PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Applicant Information:

Applicant 1

Applicant Authority ◉ Inventor ○ Legal Representative under 35 U.S.C. 117 ○ Party of Interest under 35 U.S.C. 118

Prefix	Given Name	Middle Name	Family Name	Suffix

Residence Information (Select One) ◉ US Residency ○ Non US Residency ○ Active US Military Service

City		State/Province		Country of Residence	

Citizenship under 37 CFR 1.41(b)

Mailing Address of Applicant:

Address 1	
Address 2	

City		State/Province	
Postal Code		Country	

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	20995		
Email Address	efiling@kmob.com	Add Email	Remove Email

Application Information:

Title of the Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE		
Attorney Docket Number	NETL.071A	Small Entity Status Claimed	☐
Application Type	Nonprovisional		
Subject Matter	Utility		
Suggested Class (if any)		Sub Class (if any)	
Suggested Technology Center (if any)			
Total Number of Drawing Sheets (if any)	8	Suggested Figure for Publication (if any)	

PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

Publication Information:

☐	Request Early Publication (Fee required at time of Request 37 CFR 1.219)
☐	**Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S. C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Enter either Customer Number or complete the Representative Name section below. If both sections are completed the Customer Number will be used for the Representative Information during processing.

Please Select One:	◉ Customer Number	○ US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	20995		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78(a)(2) or CFR 1.78(a)(4), and need not otherwise be made part of the specification.

Prior Application Status	Pending		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
	non provisional of	61/409893	2010-11-03

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button.

Foreign Priority Information:

This section allows for the applicant to claim benefit of foreign priority and to identify any prior foreign application for which priority is not claimed. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55(a).

Remove

Application Number	Country[1]	Parent Filing Date (YYYY-MM-DD)	Priority Claimed
			○ Yes ◉ No

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

Assignee Information:

Providing this information in the application data sheet does not substitute for compliance with any requirement of part 3 of Title 37 of the CFR to have an assignment recorded in the Office.

Assignee 1

PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

If the Assignee is an Organization check here. ☒			
Organization Name	NETLIST, INC.		
Mailing Address Information:			
Address 1	51 Discovery, #150		
Address 2			
City	Irvine	**State/Province**	CA
Country	US	Postal Code	92618
Phone Number		Fax Number	
Email Address			
Additional Assignee Data may be generated within this form by selecting the **Add** button.			

Signature:

A signature of the applicant or representative is required in accordance with 37 CFR 1.33 and 10.18. Please see 37 CFR 1.4(d) for the form of the signature.

Signature				Date (YYYY-MM-DD)	2011-11-03
First Name	Bruce	Last Name	Itchkawitz	Registration Number	47677

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

RELATED APPLICATION

[0001] This application claims the benefit of priority under 35 U.S.C. § 119(e) of U.S. Provisional Patent Application No. 61/409,893, filed on November 03, 2010, and entitled "ARCHITECTURE FOR MEMORY MODULE WITH PACKAGES OF THREE-DIMENSIONAL STACKED (3DS) MEMORY CHIPS," the disclosure of which is hereby incorporated by reference in its entirety.

BACKGROUND

Technical Field

[0002] The present disclosure relates to memory devices and memory modules. More specifically, the present disclosure relates to systems and methods for reducing the load of drivers of memory packages included on the memory modules.

Description of the Related Art

[0003] Memory modules may include a number of memory packages. Each memory package may itself include a number of array dies that are packaged together. Each array die may include an individual semiconductor chip that includes a number of memory cells. The memory cell may serve as the basic building block of computer storage representing a single bit of data.

[0004] FIGs. 1A and 1B schematically illustrate examples of existing memory package designs currently used or proposed to be used to provide the dynamic random-access memory of memory modules. FIG. 1A schematically illustrates a memory package 100 with three array dies 110 and a control die 130. The control die 130 is configured to respond to signals received by the memory package 100 by sending appropriate control signals to the array dies 110 and

includes a driver 134 for driving data signals to each of the array dies 110 via a corresponding die interconnect 120. Further, the control die 130 includes a driver 140 for driving command and/or address signals to each of the array dies 110 via another corresponding die interconnect 142. For simplicity, FIG. 1A shows only a single driver 134, die interconnect 120, driver 140, and die interconnect 142. However, additional drivers and die interconnects may be included for each bit the memory package 100 is designed to support. Thus, a 16-bit memory may include 16 pairs of drivers and die interconnects for the data signals and other similar drivers and die interconnects for the command and/or address signals. Each array die 110 also includes a chip select port 144, with the chip select ports 144 of the array dies 110 configured to receive corresponding chip select signals to enable or select the array dies for data transfer. The array dies 110 are configured to transfer data (e.g. read or write) to or from the selected memory cells identified by the command, address, and chip select signals via the die interconnects.

[0005] In some cases, the control die 130 may include memory cells and therefore, also serve as an array die. Thus, as can be seen from FIG. 1A, the control die 130 may also include a chip select port 144. Alternatively, the control die 130 and the array dies 110 may be distinct elements and the control die 130 may not include any memory cells.

[0006] FIG. 1B schematically illustrates an example of a memory package 150 that includes four array dies 160 and a control die 170 that does not include memory cells. As can be seen in FIG. 1B, each array die 160 includes a chip select port 174. However, because the control die 170 does not also serve as an array die, the control die 170 does not include a chip select port. As with memory package 100, memory package 150 includes a driver 184 that drives data signals to each of the array dies 160 along a corresponding die interconnect 182. Further, the memory package 150 includes a driver 186 for driving command and/or address signals to each of the array dies 160 via another die interconnect 188.

[0007] Generally, a load exists on each of the drivers 134, 140, 184, and 186 by virtue of the drivers being in electrical communication with the corresponding

Petitioners
Ex. 1002, p. 5

die interconnects and the corresponding circuitry of the array dies. Thus, to drive a signal along a die interconnect, a driver typically must be large enough to overcome the load on the driver. However, generally a larger driver not only consumes more space on the control die, but also consumes more power.

SUMMARY

[0008] In certain embodiments, an apparatus is provided that comprises a plurality of array dies having data ports. The apparatus further comprises at least a first die interconnect and a second die interconnect. The first die interconnect is in electrical communication with at least one data port of a first array die of the plurality of array dies and at least one data port of a second array die of the plurality of array dies and not in electrical communication with the data ports of at least a third array die of the plurality of array dies. The second die interconnect is in electrical communication with at least one data port of the third array die and not in electrical communication with the data ports of the first array die and the data ports of the second array die. In addition, the apparatus comprises a control die. The control die comprises at least a first data conduit configured to transmit a data signal to the first die interconnect and to not transmit the data signal to the second die interconnect, and at least a second data conduit configured to transmit the data signal to the second die interconnect and to not transmit the data signal to the first die interconnect.

[0009] In certain embodiments, an apparatus is provided that comprises a plurality of array dies having ports. The apparatus further comprises at least a first die interconnect and a second die interconnect. The first die interconnect is in electrical communication with at least one port of a first array die of the plurality of array dies and at least one port of a second array die of the plurality of array dies and not in electrical communication with the ports of at least a third array die of the plurality of array dies. The second die interconnect is in electrical communication with at least one port of the third array die and not in electrical communication with the ports of the first array die and the ports of the second array die. In addition, the

Petitioners
Ex. 1002, p. 6

apparatus comprises a control die. The control die comprises at least a first conduit configured to transmit a signal to the first die interconnect and to not transmit the signal to the second die interconnect, and at least a second conduit configured to transmit the signal to the second die interconnect and to not transmit the signal to the first die interconnect. Moreover, a first load on the first conduit comprises a load of the first die interconnect, a load of the first array die, and a load of the second array die. In addition, a second load on the second conduit comprises a load of the second die interconnect and a load of the third array die.

[0010] In certain embodiments, a method is provided for optimizing load in a memory package. The memory package comprises a plurality of array dies, at least a first die interconnect and a second die interconnect, and a control die. The control die comprises at least a first driver and a second driver, the first driver configured to drive a signal along the first die interconnect, and the second driver configured to drive the signal along the second die interconnect. The method comprises selecting a first subset of array dies of the plurality of array dies and a second subset of array dies of the plurality of array dies. The first subset of array dies and the second subset of array dies are exclusive of one another and are selected to balance a load on the first driver and on the second driver based at least in part on array die loads of array dies of the plurality of array dies and at least in part on die interconnect segment loads of segments of at least the first die interconnect and the second die interconnect. The method further comprises forming electrical connections between the first die interconnect and the first subset of array dies. In addition, the method comprises forming electrical connections between the second die interconnect and the second subset of array dies.

[0011] In certain embodiments, an apparatus is provided that comprises a register device configured to receive command/address signals from a memory control hub and to generate data path control signals. The apparatus further comprises a plurality of DRAM packages. Each of the DRAM packages comprises a control die. The control die comprises a plurality of command/address buffers and a data path control circuit configured to control command/address time slots and data

Petitioners
Ex. 1002, p. 7

bus time slots. The control die is configured to receive data signals from the memory control hub, the data path control signals from the register device, and command/address signals from the register device. Further, each of the DRAM packages comprises a plurality of DDR DRAM dies operatively coupled to the control die to receive the data signals from the control die. Moreover, the memory module is selectively configurable into at least two operational modes. The two operational modes comprise a first operational mode and a second operational mode. In the first operational mode the register device generates the data path control signals, and the control die uses the data path control signals to operate the data path control circuit. In the second operational mode, the control die operates the data path control circuit to provide the command/address signals to the plurality of DDR DRAM dies without decoding the command/address signals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Throughout the drawings, reference numbers are re-used to indicate correspondence between referenced elements. The drawings are provided to illustrate certain example embodiments of the inventive subject matter described herein and not to limit the scope thereof.

[0013] FIGs. 1A and 1B schematically illustrate examples of existing memory package designs.

[0014] FIG. 2 schematically illustrates an example embodiment of a memory package in accordance with the present disclosure.

[0015] FIG. 3 schematically illustrates another example embodiment of a memory package in accordance with the present disclosure.

[0016] FIG. 4 schematically illustrates an example embodiment of a driver structure of a control die in accordance with the present disclosure.

[0017] FIG. 5 presents a flowchart for an example embodiment of a load optimization process.

Petitioners
Ex. 1002, p. 8

[0018] FIGs. 6A and 6B schematically illustrate an example of a Load Reduction Dual In-line Memory Module (LRDIMM) and a HyperCloud™ Dual In-line Memory Module (HCDIMM) architecture respectively.

[0019] FIG. 7 schematically illustrates an example of a Three-Dimensional Structure Dual In-line Memory Module (3DS-DIMM) architecture.

[0020] FIG. 8 schematically illustrates an example embodiment of a memory module architecture in accordance with the present disclosure.

DETAILED DESCRIPTION OF SPECIFIC EMBODIMENTS

[0021] In addition to the below, the following U.S. patents are incorporated in their entirety by reference herein: U.S. Patent Nos. 7,289,386, 7,286,436, 7,442,050, 7,375,970, 7,254,036, 7,532,537, 7,636,274, 7,630,202, 7,619,893, 7,619,912, 7,811,097. Further, the following U.S. patent applications are incorporated in their entirety by reference herein: U.S. Patent Application Nos. 12/422,912, 12/422,853, 12/577,682, 12/629,827, 12/606,136, 12/874,900, 12/422,925, 12/504,131, 12/761,179, and 12/815,339.

[0022] Certain embodiments of the present disclosure reduce the size of drivers that are configured to drive a signal, such as a data signal, along a die interconnect to one or more array dies. Further, certain embodiments of the present disclosure reduce the power consumption of the drivers.

[0023] In certain embodiments, reducing one or both of driver size and driver power consumption may be accomplished by increasing the number of die interconnects and reducing the number of array dies that are in electrical communication with each die interconnect. For example, instead of one die interconnect in electrical communication with four array dies, there may be two die interconnects, each in electrical communication with a different pair of the four array dies.

[0024] In certain embodiments, determining the number of die interconnects and the number of array dies in electrical communication with each die interconnect is based, at least in part, on a load of each array die and a load of

the die interconnect that is in electrical communication with one or more of the array dies.

[0025] In some embodiments, the load contribution from a die interconnect may be negligible compared to the load contribution from the array dies. In such embodiments, determining the number of die interconnects and the number of array dies in electrical communication with each die interconnect may be based, at least in part, on a load of each array die without considering the load of the die interconnect. However, as the physical size of a memory package shrinks, the load of a die interconnect becomes a non-negligible value relative to the load of the array dies. Thus, as memory packages become physically smaller, it becomes more important to consider the load of the die interconnect in determining the number of die interconnects and the number of array dies in electrical communication with each die interconnect. Advantageously, certain embodiments of the present disclosure account for both the loads of the array dies and the loads of the die interconnects on a conduit (e.g., driver) in determining the number of die interconnects to be used and the number of array dies in electrical communication with each die interconnect.

[0026] FIG. 2 schematically illustrates an example embodiment of a memory package 200 in accordance with the present disclosure. One example of a memory package that includes array dies and a control die is the Hybrid Memory Cube (HMC). Examples of an HMC compatible with certain embodiments described herein are described by the IDF2011 Intel Developer Forum website, http://www.intel.com/idf/index.htm, which includes presentations and papers from the IDF2011 Intel Developer Forum including the keynote address given by Justin Rattner on September 15, 2011. Additional examples of an HMC compatible with certain embodiments described herein are described by the Hybrid Memory Cube Consortium website, http://www.hybridmemorycube.org. The memory package 200 can include any type of memory package. For example, the memory package 200 may be a DRAM package, a SDRAM package, a flash memory package, or a DDR SDRAM package (e.g., DDR3, DDR4), to name a few. A memory module (not

Petitioners
Ex. 1002, p. 10

shown) may include one or more memory packages in accordance with the memory package 200. Further, the memory package 200 may include input/output terminals (not shown) that are configured to be placed in electrical communication with circuitry of the memory module to transmit signals between the memory package 200 and a Memory Control Hub (MCH) (not shown).

[0027] The memory package 200 may include a plurality of array dies 210 (e.g., array dies 210a–210d). The plurality of array dies 210 may be sealed within the memory package 200. Further, circuitry of the array dies 210 may be in electrical communication with the input/output terminals of the memory package 200. Although generally referred to as array dies herein, the array dies 210 may also be called slave dies or slave chips. Each of the array dies 210a–210d may include circuitry (e.g., memory cells) (not shown) for storing data. Examples of array dies compatible with certain embodiments described herein are described by the existing literature regarding the Hybrid Memory Cube (e.g., as cited above). As illustrated in FIG. 2, the plurality of array dies 210 may be arranged in a stack configuration known as a three-dimensional structure (3DS). Examples of 3DS compatible with certain embodiments described herein are described by the existing literature regarding the Hybrid Memory Cube (e.g., as cited above). However, the structure or layout of the plurality of array dies 210 is not limited as such, and other structures are possible in accordance with the present disclosure. For example, the plurality of array dies 210 may be arranged in a planar structure, or in a structure that combines 3DS with a planar structure. Moreover, while the memory package 200 is illustrated as including four array dies, the memory package 200 is not limited as such and may include any number of array dies. For example, as illustrated in FIG. 3, the memory package 200 may include eight array dies. As further examples, the memory package 200 may include three or sixteen array dies.

[0028] Each of the array dies 210 may include one or more data ports (not shown). The data ports enable electrical communication and data transfer between the corresponding memory circuitry of the array dies 210 and a communication pathway (e.g., a die interconnect).

Petitioners
Ex. 1002, p. 11

[0029] In the example schematically illustrated in FIG. 2, the memory package 200 includes a plurality of die interconnects 220 (e.g., die interconnects 220a, 220b). For example, certain versions of HMC have been reported to have 512 data ports per die, with the corresponding bits of each die all connected to a single die interconnect (e.g., TSV). Examples of die interconnects include, but are not limited to, through-silicon vias (TSV), conducting rods, wire bonds, and pins. (See e.g., U.S. Patent Nos. 7,633,165 and 7,683,459.) Each of these die interconnects 220 may be coupled to, or in electrical communication with at least one data port of at least one of the array dies 210. In certain embodiments, at least one of the die interconnects 220 is in electrical communication with at least one data port from each of at least two array dies 210 without being in electrical communication with a data port from at least one array die 210, which may be in electrical communication with a different die interconnect 220.

[0030] For example, die interconnect 220a may be in electrical communication with a data port from array die 210a and a data port from array die 210b (as illustrated by the darkened circles in FIG. 2) and not in electrical communication with any data ports from array die 210c or any data ports from array die 210d. The data ports of array dies 210a and 210b in electrical communication with the die interconnect 220a can be corresponding to the same data bit (e.g., D0). Other die interconnects (not shown) can be in electrical communication with other data ports corresponding to other data bits (e.g., D1, D2, . . .) of array dies 210a and 210b. These other die interconnects can be electrically isolated from the corresponding data ports of array dies 210c and 210d.

[0031] However, continuing this example, the die interconnect 220b may be in electrical communication with a data port from array die 210c and a data port from array die 210d (as illustrated by the darkened circles in FIG. 2) (e.g., corresponding to the same data bit, e.g., D0) without being in electrical communication with any data ports from array die 210a and array die 210b. Other die interconnects (not shown) can be in electrical communication with other data ports corresponding to other data bits (e.g., D1, D2, . . .) of array dies 210c and

Petitioners
Ex. 1002, p. 12

210d. Despite not being in electrical communication with any data ports from array die 210a and 210b, in some implementations, the die interconnect 220b may pass through the array dies 210a and 210b (as illustrated by the unfilled circles) e.g., through through-holes or vias of array dies 210a and 210b. For some implementations, each of the array dies 210 may be in electrical communication with corresponding die interconnects 220, without any of the die interconnects 220 being in electrical communication with all of the array dies 210. Where existing systems may utilize a single die interconnect to be in electrical communication with the corresponding data ports of each array die (e.g., the data ports corresponding to the same bit), certain embodiments described herein utilize multiple die interconnects to provide electrical communication to the corresponding data ports of the array dies (e.g., the data ports corresponding to the same data bit) with none of the multiple die interconnects in electrical communication with data ports of all the array dies.

[0032] In addition to the plurality of array dies 210, the memory package 200 includes a control die 230, which may also be called a master die. Examples of master dies compatible with certain embodiments described herein are described by the existing literature regarding the Hybrid Memory Cube (e.g., as cited above). In some embodiments, the control die 230 may be one of the array dies 210. Alternatively, the control die 230 may be a modified version of one of the array dies 210. Thus, in some implementations, memory package 200 may include four dies or chips instead of the five illustrated in FIG. 2. Further, in some embodiments, the control die 230 may comprise a logic layer (e.g., the logic layer of an HMC).

[0033] The control die 230 may include a number of data conduits 232, which includes data conduits 232a and 232b. Each of these data conduits 232 may be configured to transmit a data signal to a single die interconnect 220. For example, the data conduit 232a may be configured to transmit a data signal to the die interconnect 220a without transmitting the data signal to the die interconnect 220b. Conversely, the data conduit 232b may be configured to transmit the data signal to the die interconnect 220b without transmitting the data signal to the die interconnect 220a (e.g., data conduit 232b is electrically isolated from die

interconnect 220a and data conduit 232a is electrically isolated from die interconnect 220b).

[0034] In some embodiments, the data conduits 232 may also include one or more drivers 234 as schematically illustrated by FIG. 2. The drivers 234 may be configured to drive the data signals along the corresponding die interconnects 220. In some embodiments, a single data conduit 232 or driver 234 may be in electrical communication with multiple die interconnects 220, each of which may be in electrical communication with different array dies 210.

[0035] Each of the data conduits 232 may be configured to receive the data signal from a common source. For instance, the data conduit 232a and the data conduit 232b may each receive a substantially similar, if not identical, data signal from the same signal source (e.g. the data signal corresponding to the same data bit). The source of the data signal may include a data path, a driver, a latch, a pin, or any other construct that may provide a data signal to a data conduit.

[0036] The data conduits 232 may each be subject to a load. Although not limited as such, this load may be measured as a capacitive load, such as a parasitic capacitance. The load on each of the data conduits 232 may include at least a load of the die interconnect 220 to which the data conduit 232 is coupled or connected as well as a load of each array die 210 with which the die interconnect 220 is in electrical communication via a data port of the die interconnect 220. Thus, for example, the load of the data conduit 232a may include loads of the die interconnect 220a, the array die 210a, and the array die 210b. Similarly, the load of the data conduit 232b may include loads of the die interconnect 220b, the array die 210c, and the array die 210d.

[0037] Generally, a load that would be on a data conduit that was in electrical communication with a die interconnect that was in electrical communication with at least one data port of each of the array dies 210 can be considered the maximum load for a data conduit. This maximum load is the load of a data conduit of a memory package that does not implement the teachings of this disclosure, but is in accordance with conventional configurations.

Petitioners
Ex. 1002, p. 14

[0038] In some implementations, the difference between the load of the data conduit 232a and the load of the data conduit 232b is less than the maximum load for a data conduit as described above. Thus, in some cases, there may exist a degree of balance or equalization between the loads of the data conduits 232a, 232b. In some implementations, the difference between the load of the data conduit 232a and the load of the data conduit 232b is zero or substantially zero. In some embodiments, the length of each die interconnect 220, and the number of array dies 210 in electrical communication with each die interconnect 220 may be selected to maintain the difference between the load of the data conduit 232a and the load of the data conduit 232b to be at or below a threshold load difference. For example, suppose that the load of each array die 210 is 1, the load of each segment of the die interconnects 220 is 0.25, and that the threshold load difference is 0.5. Using the configuration schematically illustrated in FIG. 2, the load on the data conduit 232a in this example is 2.5 and the load on the data conduit 232b in this example is 3. Thus, in this example, the difference between the load of the data conduit 232a and the load of the data conduit 232b is at the threshold load difference value of 0.5. However, an alternative configuration that places the die interconnect 220a in electrical communication with only the array die 210a, and the die interconnect 220b in electrical communication with the array dies 210b–210d would not satisfy the threshold load difference value of 0.5 of the above example. In the alternative configuration, the load on the data conduit 232a would be 1.25 and the load on the data conduit 232b would be 4. Thus, in the alternative configuration, the difference between the load of the data conduit 232a and the load of the data conduit 232b is 2.75, which is above the threshold load difference value of 0.5.

[0039] For certain embodiments, the load of each data conduit is less than the maximum load as described above. Thus, in some cases, the load of the data conduit 232a is less than the maximum load and the load of the data conduit 232b is less than the maximum load. Further, in many implementations, the combined load of the data conduit 232a and the data conduit 232b is less than the maximum load of a single data conduit. In other words, it is possible to design the data conduit

Petitioners
Ex. 1002, p. 15

232a and the data conduit 232b to reduce the overall load compared to a single data conduit that is in electrical communication with a die interconnect that is in electrical communication with at least one data port of each of the array dies 210. By reducing the overall load compared to the single data conduit, it is possible in many cases to reduce power consumption. Further, it is possible in many cases to maintain signal quality (e.g. maintain signal amplitude, maintain low signal distortion, etc.) while reducing power consumption. Advantageously, in a number of embodiments, by using multiple data conduits instead of a single data conduit, the speed of the memory package 200 can be increased. In some cases, this speed increase can include a reduced latency in accessing array dies 210 and/or operating the memory package 200 at a higher clock frequency.

[0040] Each of the die interconnects 220 may include any type of conducting material. For example, the die interconnects 220 may include copper, gold, or a conductive alloy, such as a copper/silver alloy. Further, the die interconnects 220 may include any type of structure for enabling electrical communication between the data conduits 232 and the data ports of the array dies 210. For example, the die interconnects 220 may include a wire, a conducting rod, or a conducting trace, to name a few. Moreover, the die interconnects 220 may use vias, or through-silicon vias (TSVs) to couple with or to electrically communicate with an array die. For instance, die interconnect 220a may be connected with data ports of the array dies 210a and 210b using vias (illustrated by the filled or darkened circles). Examples of TSVs which may be used with the present disclosure are described further in U.S. Patent Nos. 7,633,165 and 7,683,459.

[0041] In addition, the die interconnects 220 may use via holes to pass through an array die that is not configured to be in electrical communication with the die interconnect. For instance, die interconnect 220b may pass through array dies 210a and 210b using TSVs that do not enable electrical communication between the die interconnect 220b and data ports of the array dies 210a and 210b (illustrated by the unfilled circles). In this way, the array dies 210a, 210b are not responsive to the data signal beings transmitted by the die interconnect 220b.

Petitioners
Ex. 1002, p. 16

However, the die interconnect 220b may be connected with at least one data port of each of the array dies 210c and 210d using a via (illustrated by the filled or darkened circles). In cases where the die interconnect passes through an array die that is not configured to be in electrical communication with the die interconnect, the TSV may include an insulator or an air gap between the die interconnect and the array die circuitry that is large enough to prevent electrical communication between the die interconnect and the array die circuitry. In certain embodiments, the TSV for array dies that are configured to be in electrical communication with the die interconnect and for array dies that are not configured to be in electrical communication with the die interconnect may be configured the same. However, in such cases, electrical connections leading from the TSV of the array dies that are not configured to be in electrical communication with the die interconnect may not exist or may be stubs. These stubs are not configured to provide electrical communication with the memory cells of the array die.

[0042] Although FIG. 2 illustrates a single pair of data conduits 232 corresponding to a single pair of die interconnects 220, this is only to simplify the drawing figure. The memory package 200 may generally include as many additional data conduits and corresponding die interconnects as the number of bits the memory package 200 is designed or configured to support per memory address. Thus, if, for example, the memory package 200 is configured to be a 16-bit memory, the memory package 200 may include 16 pairs of data conduits 232 and 16 pairs of corresponding die interconnects 220. Similarly, if the memory package 200 is configured as a 32- or 64-bit memory, the memory package 200 may include 32 or 64 pairs of data conduits 232 and 32 or 64 pairs of corresponding die interconnects 220. Generally, the data conduits and die interconnects for each bit are configured identically. Thus, for the memory package 200, each data conduit is configured to be in electrical communication with a die interconnect that is configured to be in electrical communication with a pair of the array dies 210. However, it is possible that, in some embodiments, the data conduits and die interconnects of different bits may be configured differently.

Petitioners
Ex. 1002, p. 17

[0043] In certain embodiments, the same die interconnects and the same corresponding data conduits may be used to transfer data both to and from the array dies 210. In such embodiments, the die interconnects may be bi-directional. In alternative embodiments, separate die interconnects and corresponding data conduits may be used to transfer data to the array dies 210 and data from the array dies 210 to the control die 230. Thus, such embodiments may include double the number of die interconnects and data conduits as embodiments that use the same die interconnect to transfer data to and from an array die.

[0044] In some embodiments, the control die 230 may include additional command/address conduits 240 and die interconnects 242, which may be in electrical communication with at least one port of each of the array dies 210. For simplicity, FIG. 2 shows only a single such conduit 240 and die interconnect 242. The command/address conduits 240 are configured to provide corresponding signals to the die interconnects 242. These signals may be command signals, address signals, or may serve as both command and address signals (e.g., may include a memory cell address and a write command or a read command) either simultaneously or based on a determining criterion, such as the edge of a clock signal. The command die 230 may include a command/address conduit 240 and corresponding die interconnect 242 for each bit of the command/address signals that the memory package 200 is configured to support. The number of bits of the command/address signals may be the same or may be different from the number of data bits of the memory package 200.

[0045] In addition, in certain embodiments, the control die 230 may include a plurality of chip select conduits 250 (e.g., chip select conduits 250a–250d as shown in FIG. 2). Further, the control die 230 may include corresponding die interconnects 252 (e.g., die interconnects 252a–252d) with one die interconnect 252 in electrical communication with one chip select conduit 250 and one array die 210. Each of the die interconnects 252 may be in electrical communication with a different array die 210. For example, the die interconnect 252a may be in electrical communication with the array die 210a and the die interconnect 252b may be in

Petitioners
Ex. 1002, p. 18

electrical communication with the array die 210b. Each of the chip select conduits 250 may be configured to provide a chip select signal to a corresponding array die 210 via a corresponding die interconnect 252.

[0046] In some embodiments, the control die 230 may include additional drivers that are configured to drive the chip select signals along the die interconnects 252. Alternatively, the chip select signals may be driven by drivers that are external to the control die 230. For example, a register (not shown) that is part of a memory module that includes the memory package 200 may determine the chip select signals and drive the chip select signals to the array dies 210. As a second example, the chip select signals may be provided by an MCH. In some embodiments, the control die 230 may determine the array die 210 to select based on, for example, an address signal. In such embodiments, the control die may generate the chip select signals.

[0047] FIG. 3 schematically illustrates another example embodiment of a memory package 300 in accordance with the present disclosure. In certain embodiments, some or all of the embodiments described above with respect to the memory package 200 may be applicable to the memory package 300. However, for ease of illustration and to simplify discussion, certain elements are omitted in FIG. 3, such as the chip select conduits. Nevertheless, it should be understood that the memory package 300 can include the same or similar elements as described above with respect to the memory package 200, including, for example, the chip select conduits.

[0048] The memory package 300 may include a plurality of array dies 310 (e.g., array dies 310a–310h). In the implementation illustrated in FIG. 3, the memory package includes eight array dies. However, as stated earlier, the memory package 300 may include more or fewer array dies. Each of the array dies 310 may include one or more ports (not shown) that enable electrical communication between the circuitry of the array dies 310 and one or more die interconnects 320. Each of these die interconnects 320 may be coupled to, or in electrical communication with at least one port of at least one of the array dies 310. As with

Petitioners
Ex. 1002, p. 19

the memory package 200, in certain embodiments, at least one of the die interconnects 320 is in electrical communication with at least one port from each of at least two array dies 310 without being in electrical communication with a port from at least one array die 310, which may be in electrical communication with a different die interconnect 320.

[0049] In addition to the plurality of array dies 310, the memory package 300 includes a control die 330. In some embodiments, the control die 330 may include a number of conduits 332 (e.g., conduits 332a–332f). Each of these conduits 332 may be configured to transmit a signal to a single die interconnect 320.

[0050] Further, implementations of the conduits 332 may include one or more drivers 334 (e.g., drivers 334a–334f). Each of the drivers 334 may be configured to drive a signal along a corresponding die interconnect 320. For instance, the driver 334a of the conduit 332a may be configured to drive a signal along the die interconnect 320a to one or more of the array dies 310a and 310b. As a second example, the driver 334b of the conduit 332b may be configured to drive a signal along the die interconnect 320b to one or more of the array dies 310c and 310d. Although the die interconnect 320b may pass through the array dies 310a and 310b, because the die interconnect 320b, in the example illustrated in FIG. 3, is not configured to be in electrical communication with the array dies 310a and 310b, the driver 334b does not drive the signal to the array dies 310a and 310b.

[0051] In some embodiments, the signal can be a data signal, a command or address signal, a chip select signal, a supply voltage signal, or a ground voltage signal, to name a few. Further, as the signal is not limited to a data signal, in some embodiments, the conduits 332 may include conduits configured to provide signals other than data signals to the die interconnects 320. For example, the conduits may include conduits configured to provide a command or address signal, a chip select signal, a supply voltage signal, or a ground voltage signal to one or more die interconnects. Consequently, in some embodiments, the drivers 334 may be configured to drive signals other than data signals.

Petitioners
Ex. 1002, p. 20

[0052] Generally, the signal that each of the drivers 334 drive to the corresponding die interconnects 320 is from a common source. Thus, each of the drivers 334, in certain embodiments, is driving the same signal to each corresponding die interconnect 320.

[0053] The size of the drivers 334 is generally related to the load on the driver 334. In certain embodiments, the load on each driver 334 corresponds to the load of the respective conduit 332. Although not limited as such, the load may be measured as a capacitive load, such as a parasitic capacitance.

[0054] The load on each of the conduits 332 may include at least the loads of the die interconnect 320 with which the conduit 332 is coupled or connected as well as the loads of each array die 310 with which the die interconnect 320 is in electrical communication via a port of the die interconnect 320. Thus, for example, the loads of the conduit 332a may include the loads of the die interconnect 320a, the array die 310a, and the array die 310b. Similarly, the loads of the conduit 332b may include the loads of the die interconnect 320b, the array die 310c, and the array die 310d. The loads of both conduits 332a and 332b include a load of two array dies 310 because the corresponding die interconnects 320 are each configured to be in electrical communication with two array dies 310. On the other hand, the load of the conduit 332c, which may include the loads of the die interconnect 320c and the array die 310e, includes a load of one array die 310e because the corresponding die interconnect 320c is configured to be in electrical communication with only one array die 310.

[0055] As previously described with respect to FIG. 2, a load that would be on a conduit that was in electrical communication with a die interconnect that was in electrical communication with at least one port of each of the array dies 310 can be considered the maximum load for a conduit. This maximum load is the load of a conduit of a memory package that does not implement the teachings of this disclosure, but is used in accordance with conventional configurations.

[0056] In some implementations, the difference between the loads of any pair of the conduits 332 is less than the maximum load for a conduit as described

Petitioners
Ex. 1002, p. 21

above. For instance, the difference between the load of the conduit 332a and the load of the conduit 332b is less than the maximum load. As a second example, the difference between the load of the conduit 334f and any one of the conduits 334a–334e is less than the maximum load. Thus, in some cases, the load on each of the conduits 332 may be, at least partially balanced or equalized to reduce or minimize the difference between the load of any pair of the conduits 332. In some implementations, the difference between the load of a pair of the conduits 332 is zero or substantially zero. In some embodiments, the length of each die interconnect 320, and the number of array dies 310 in electrical communication with each die interconnect 320 may be selected to maintain the difference between the load of any pair of the conduits 332 to be at or below a threshold load difference.

[0057] For certain embodiments, the load of each conduit is less than the maximum load as described above. Thus, for example, the load of the conduit 332a, the load of the conduit 332b, and the load of the conduit 332c are each less than the maximum load defined above.

[0058] In certain embodiments, the load associated with each of the array dies 310 may be substantially equivalent. For example, the load of the array die 310a may be substantially equal to the load of the array die 310h. Thus, the load contribution from the array dies 310 for a specific conduit 332 may be measured as a multiple of the array dies that are in electrical communication with a die interconnect 320 corresponding to a specific conduit 332. For example, assuming that the load of each of the array dies 310 is 'L', then the contribution of the load from the array dies 310 to the conduit 332a would be 2L because the die interconnect 320a corresponding to the conduit 332a is in electrical communication with two array dies 310, array dies 310a and 310b. In alternative embodiments, the load of each of the array dies 310a may differ. For example, the load of array die 310a may be L and the load of the array die 310b may be 1.25L.

[0059] Similar to the array dies 310, in some embodiments, a die interconnect 320 can be considered to comprise a plurality of segments, with each segment of a die interconnect 320 contributing a substantially equivalent load to the

Petitioners
Ex. 1002, p. 22

load of the die interconnect 320. In this case, the segment of the die interconnect 320 may refer to the portion of the die interconnect between two successive or adjacent dies (array die-to-array die, master die-to-array die, or both) along the die interconnect 320. Further, the segment may be defined as a portion of the die interconnect 320 between the dies exclusive of a portion of the dies. For example, one segment of the die interconnect 320a may extend from the top of the array die 310a to the bottom of the array die 310b. Alternatively, the segment may be defined to include at least a portion of at least one of the array dies 310. For example, one segment of the die interconnect 320a may extend from the center of array die 310a to the center of array die 310b. As a second example, one segment of the die interconnect 320a may extend from the top of the control die 330 to the top of the array die 310a, and therefore may include a portion of the die interconnect 320a extending from the bottom of the array die 310a to the top of the array die 310a. In some implementations, the segments are substantially equal in length to each other. Moreover, the load contribution of each segment may be substantially equal to each other. Alternatively, the segments may be unequal in length and/or may each contribute a different load to the total load of a die interconnect 320. Further, in some cases, the load contribution of a segment of the die interconnect 320 that is in electrical communication with a port of an array die 310 may differ from the load contribution of a segment of the die interconnect 320 that is not in electrical communication with a port of an array die 310.

[0060] In some cases, the load contribution of each segment of the die interconnect 320 may be measured as a fraction of the load contribution from an array die 310. For example, the load of one segment of the die interconnect 320a may be equivalent to one quarter of the load of an array die 310. Thus, for example, the load of the conduit 332a may be 2.5L assuming a load contribution of L per array die 310 (two in this case) in electrical communication with the die interconnect 320a and a load contribution of 0.25L per segment (two in this case) of the die interconnect 320a. As a second example, the load of the conduit 332f may be 3L assuming the same load values as the previous example and a load

Petitioners
Ex. 1002, p. 23

contribution from one array die 310h and eight die interconnect 320f segments. Table 1 specifies the capacitive load values for each conduit 332 assuming, as with the previous two examples, that the load of each of the array dies 310 is L, and that the load of each segment of the die interconnects 320 is 0.25L. Table 1 also specifies the deviation in load from the conduits having the highest load value, which in this example are conduits 332b and 332f.

TABLE 1

Conduit	Number of Array Dies	Number of Die Interconnect Segments	Capacitive Load	Deviation from Maximum Load
332a	2	2	2.5L	0.5
332b	2	4	3L	0
332c	1	5	2.25L	0.75
332d	1	6	2.5L	0.5
332e	1	7	2.75L	0.25
332f	1	8	3L	0

[0061] As can be seen from Table 1, the maximum load of any conduit 332, using the example values previously described, is 3L, or rather three times the load of a single array die 310. Assuming the same example values, the load of a conduit in electrical communication with a die interconnect that itself is in electrical communication with each array die 310 would be 10L. Thus, certain embodiments of the present disclosure enable a reduction in the load of the conduits 332. Consequently, in some embodiments, the drivers 334 may each be smaller than a single driver that is configured to drive a signal from a conduit along a single die interconnect that is in electrical communication with a port from each of the array dies 310. Moreover, the drivers 334 may include smaller transistor sizes than a single driver that is configured to drive a signal to each of the array dies 110.

[0062] As can be seen from Table 1, conduits 332b and 332f have the largest capacitive load of the group of conduits, which is 3L. Conduit 332 has the smallest capacitive load of the group of conduits, which is 2.25L and which is a deviation of only 0.75 from the maximum load. Thus, in some embodiments, each of the drivers 334 may be substantially similar in size. In certain implementations, the drivers 334 may vary in size based on the total capacitive load on each conduit 332. Thus, the driver 334f may be larger than the driver 334e. Alternatively, each driver

Petitioners
Ex. 1002, p. 24

334 may be substantially equal and may be configured based on the drivers 334 with the largest load, which are drivers 334b and 334f in the example illustrated in FIG. 3 and Table 1.

[0063] In some embodiments, the capacitive load of each conduit 332 or driver 334 can be calculated using formula (1).

$$CL = AD + \frac{S}{M} \quad (1)$$

[0064] In formula (1), CL represents the capacitive load of a conduit 332 or driver 334, AD represents the number of array dies 310 in electrical communication with a die interconnect 320 that is in electrical communication with the conduit 332 and/or driver 334, S represents the number of die interconnect segments of the die interconnect 320, and M represents the ratio of the load of an array die 310 to the load of a segment of a die interconnect 320. Thus, using formula (1) and the example values described above, the load of the driver 332a, for example, can be calculated with the following values: AD=2, for two array dies 310; S=2, for the two segments of the die interconnect 320a; and M=4, for the ratio of the load of an array die 310, L, to the load of a segment of the die interconnect 320a, 0.25L. Therefore, as can be seen from Table 1, the capacitive load of the driver 332a is 2.5L where L is the load of a single array die.

[0065] In some cases, the load on each conduit 332 and/or driver 334 may be evenly balanced. In other words, the load on each conduit 332 and/or driver 334 may be substantially the same. To achieve a balanced load, each of the conduits 332 may be in electrical communication with a combination of array dies 310 and die interconnect 320 segments that results in a load that is substantially equivalent to the loads of the other conduits 332. In certain embodiments, the load of the conduits 332 is balanced despite each conduit 332 being in electrical communication with a different subset of the array dies 310. However, in alternative embodiments, the load of each conduit 332 and/or driver 334 may differ. This

Petitioners
Ex. 1002, p. 25

difference in the load of the conduits 332 and/or drivers 334 may be a design decision (e.g. to maintain a specific number of drivers). Alternatively, or in addition, the load difference between conduits 332 and/or drivers 334 may occur because the loads of the array dies 310 and the die interconnect segments 320 do not allow for perfect or substantially even load balancing.

[0066] As previously stated, in some embodiments the loads of the conduits 332 and/or drivers 334 may be balanced to minimize the difference between any pair of conduits 332 and/or drivers 334. For example, as illustrated in FIG. 3, a driver in electrical communication with a longer die interconnect, which may consequently include more die interconnect segments, is likely to obtain a larger load contribution from the die interconnect than a driver in electrical communication with a shorter die interconnect (e.g., compare driver 334f to 334a). Thus, the driver in electrical communication with the longer die interconnect may be in electrical communication with fewer array dies than the driver in electrical communication with the shorter die interconnect (e.g. compare driver 334f to 334a). The selection of die interconnect length and the selection of array dies to place in electrical communication with the die interconnects may therefore be dependent on the load of each segment of the die interconnects 320 and the loads of the array dies 310.

[0067] Alternatively, or in addition, the loads of the conduits 332 and/or drivers 334 may be balanced to reduce the maximum load of each conduit 332 and/or driver 334. Further, in some embodiments, the maximum load of each conduit 332 and/or driver 334 may be reduced to maintain the load of each conduit 332 and/or driver 334 to be at or below a threshold load difference.

[0068] Although not illustrated in FIG. 3, each of the conduits 332 may include one or more additional drivers configured to drive a signal received from an array die via a corresponding die interconnect 320. The additional drivers may drive the signal to a processor, a register, a latch, or any other component or device that may or may not be part of a memory module that includes the memory package 300. In some implementations, one or more of the die interconnects 320 are

Petitioners
Ex. 1002, p. 26

configured to be bi-directional, thereby enabling a signal to be driven to the array dies 310 via the die interconnect 320, and to enable a signal received from the array dies 310 along the same die interconnect 320 to be transmitted to a corresponding conduit 332 of the control die 330. Alternatively, the memory package 300 comprises one or more die interconnects 320 that are configured to enable a signal to be driven to the array dies 310 without enabling a signal to be received from the array dies 310 (e.g., the die interconnects 320 may not be bi-directional). In certain embodiments where one or more of the die interconnects 320 are not bi-directional, the memory package 300 may include additional die interconnects (not shown) that are in electrical communication with the one or more additional conduits or drivers and that are configured to enable a signal from the array dies 310 to be transmitted to the one or more additional conduits or drivers.

[0069] In addition to the drivers 334 of the control die 330, the memory package 300 may include one or more pre-drivers 340. A pre-driver 340, shown schematically in FIG. 3, may be large enough to drive a signal (e.g., a data signal) to any of the drivers 332 and subsequently, to any of the array dies 310. Thus, using the same example values as described above in relation to Table 1, a load of the pre-driver 340 may be at least 10L, which is the load of a conduit in electrical communication with a die interconnect that is in electrical communication with a port from each of the array dies 310. In some embodiments, the memory package 300 may include any number of additional drivers and/or latches for buffering and/or driving the signal to any of the array dies 310.

[0070] Just as the memory package 300 may include a pre-driver 340 for providing a signal to the conduits 332, the memory package 300 may include a post-driver (not shown) for driving an output signal from the control die 330. This post-driver may drive the output signal to additional latches and/or drivers. In some embodiments, the post-driver may drive the signal from the memory package 300 to a bus or other electrical path that is in electrical communication with the memory package 300.

Petitioners
Ex. 1002, p. 27

[0071] FIG. 4 illustrates an example embodiment of a driver structure 400 of a control die (e.g. control die 230) in accordance with the present disclosure. The driver structure 400 is configured for a single data bit. Thus, a control die for a 32-bit memory package will generally include at least 32 instances of the driver structure 400, one per bit. In some embodiments, the control die 230 may include the driver structure 400 in place of the combination of drivers 232a and 232b. Similarly, in certain embodiments, the control die 330 may include a structure similar to the driver structure 400 in place of the drivers 334. In such embodiments, the driver structure 400 would be modified to enable a signal (e.g., data signal) to be driven to the six conduits 332.

[0072] The driver structure 400 can include an input/output port 402 configured to receive or send a signal, such as a data signal. Signals received at the input/output port 402 can be provided to the drivers 404a and 404b. In turn, these drivers 404a and 404b can drive the signal to conduits 406a and 406b respectively, which are in electrical communication with corresponding die interconnects. Each of the die interconnects may be in electrical communication with different array dies in accordance with certain embodiments described herein.

[0073] In some embodiments, the driver structure 400 is bi-directional. In such embodiments, operation of the drivers 404a and 404b, and 408a and 408b may be controlled or enabled by a control signal (e.g., a directional control signal). This control signal, in some cases, may correspond to one or more of a command/address signal (e.g., read/write) and a chip select signal. In some implementations, the drivers 404a and 404b may drive a signal to the array die corresponding to the chip select signal, and not to array dies that do not correspond to a chip select signal. For example, suppose the conduit 406a is in electrical communication with array dies one and two (not shown), and that the conduit 406b is in electrical communication with array dies three and four (not shown). If a chip select signal is received corresponding to array die two, then in some implementations, driver 404a may be configured to drive a signal along a die interconnect in electrical communication with the conduit 406a to array die one and

Petitioners
Ex. 1002, p. 28

two. In this example, the driver 404b would not drive the signal because the chip select signal does not correspond to either array die three or array die four.

[0074] In some embodiments, the drivers 404a and 404b may drive a signal to all of the array dies. For example, assume the memory package that includes the driver structure 400 also includes four array dies. If the die interconnect in electrical communication with the conduit 406a is in electrical communication with two of the array dies, and if the die interconnect in electrical communication with the conduit 406b is in electrical communication with the other two array dies, then the drivers 404a and 404b may drive a signal to all four of the array dies of this example.

[0075] Similar to the input/output port 402, the conduits 406a and 406b can be configured to receive a signal from the die interconnects that are in electrical communication with the conduits 406a and 406b. In turn, the signal received at one of the conduits 406a and 406b can be provided to drivers 408a and 408b respectively. The drivers 408a and 408b can each be configured to drive a signal received from a respective data conduit 406a and 406b to the input/output port 402 and to another component that may be in electrical communication with the input/output port 402, such as a MCH.

[0076] In some embodiments, one or more chip selects, as illustrated in FIG. 2, may be used to select, determine, or enable the array die that communicates the signal to or from one or more of the conduits 406a and 406b and the drivers 408a and 408b. Similarly, in some embodiments, the chip select may select, determine, or enable the array die to receive and/or respond to the signal driven by the drivers 404a and 404b to the array dies.

[0077] It should be noted that the driver structure 400 is a non-limiting example for arranging drivers in a control die. Other driver structures are possible. For example, instead of the drivers 404a and 408a being in electrical communication with the same conduit 406a, each of the drivers 404a and 408a can be in electrical communication with a separate conduit and consequently a separate die interconnect. In such an example, the drivers 404a and 408a may still be in

Petitioners
Ex. 1002, p. 29

electrical communication with the same input/output port 402, or each driver may be in electrical communication with a separate port, the driver 404a in electrical communication with an input port, and the driver 408a in communication with an output port.

[0078] FIG. 5 presents a flowchart for an example embodiment of a load optimization process 500. In certain embodiments, the load optimization process 500 may be performed, at least in part, by one or more computing systems. Further, the process 500, in some embodiments, may be used to optimize one or more loads in a memory package (e.g. memory package 200 or memory package 300). Optimizing the loads in the memory package can include optimizing the load on one or more conduits and/or drivers. As previously described with respect to FIGs. 2 and 3, the memory package can include a plurality of array dies, a plurality of die interconnects, and a control die. Furthermore, the control die can include a plurality of drivers, each of which may be configured to drive a signal along a die interconnect.

[0079] In some implementations, the process 500 can comprise selecting a first subset of array dies and a second subset of array dies from a plurality of array dies (e.g., array dies 310), as shown in operational block 502. Generally, the first subset of array dies and the second subset of array dies may be exclusive of one another. Thus, the first subset of array dies does not include any array dies from the second subset of array dies and vice versa. However, in some cases there may be some overlap between the first subset of array dies and the second subset of array dies. Further, in some cases, at least one of the subsets of array dies includes more than one array die. For instance, the first subset of array dies may include two array dies, and the second subset of array dies may include one array die, which, depending on the embodiment, may or may not be included in the first subset of array dies.

[0080] Further, the first subset of array dies and the second subset of array dies may be selected to balance a load on a first driver and a load on a second driver based, at least in part, on the loads of the array dies 310 and on the

Petitioners
Ex. 1002, p. 30

loads of the die interconnect segments from a first and a second die interconnect. The first die interconnect may be in electrical communication with the first driver and the second die interconnect may be in electrical communication with the second driver. In some embodiments, the first subset of array dies and the second subset of array dies are selected to balance a load on a first conduit and a load on a second conduit. The load on each driver and/or conduit may be calculated using formula (1) as previously described above. In some embodiments, the first subset of array dies and the second subset of array dies may be selected to balance a load on a first driver and a load on a second driver based, at least in part, on the loads of the array dies 310 without considering the loads of the die interconnect segments from the first die interconnect and the second die interconnect.

[0081] The process 500 further comprises forming electrical connections between the first die interconnect and the first subset of array dies in an operational block 504. The process 500 further comprises forming electrical connections between the second die interconnect and the second subset of array dies in an operational block 506. Forming the electrical connections places the die interconnects in electrical communication with the respective subsets of array dies. In some embodiments, forming the electrical connections can comprise forming electrical connections between the die interconnects and at least one port from each array die of the respective subsets of array dies.

[0082] The process 500 further comprises selecting a driver size for a first driver at block 508 and selecting a driver size for a second driver at block 510. Selecting the driver size can be based, at least in part, on the calculated load on the driver. Generally, the greater the load on the driver, the larger the driver is selected to drive a signal along, for example, a die interconnect. The size of the driver may be adjusted by the selection of the transistor size and/or number of transistors included in the driver. A larger driver often consumes more power than a smaller driver. Thus, in certain embodiments, balancing the loads on the drivers to reduce the load on each driver can reduce the power consumption of a memory package.

Petitioners
Ex. 1002, p. 31

[0083] In some embodiments, the size of the first driver and the size of the second driver are both less than the size sufficient for a driver to drive a signal along a die interconnect to each of the array dies 310 (e.g., with less than a predetermined or threshold signal degradation). The threshold signal degradation can be based on any one or more characteristics of a signal. For example, the threshold signal degradation can be based on the amplitude of the signal, the frequency of the signal, the noise distortion included in or introduced into the signal, or the shape of the signal, to name a few.

[0084] The process 500 further comprises forming electrical connections between the first die interconnect and the first driver in an operational block 512. Similarly, the process 500 further comprises forming electrical connections between the second die interconnect and the second driver in an operational block 514. Forming the electrical connections places the die interconnects in electrical communication with the respective drivers. In some embodiments, forming the electrical connections can comprise forming electrical connections between the die interconnects and a data conduit that includes the respective driver.

[0085] Advantageously, certain embodiments of the present disclosure reduce the load on each conduit and on each corresponding driver. In certain embodiments, reducing the load on a driver may increase the speed of data transfer between the array dies and other components in a computer system (e.g. a MCH or a processor). Further, reducing the load on a driver may result in reduced power consumption by the driver and consequently, by the memory package. In addition, certain embodiments of the present disclosure may minimize current switching noise.

Operational Modes

[0086] A proposed three-dimensional stacking (3DS) standard for dual in-line memory modules (DIMMs) being considered by the Joint Electron Devices Engineering Council (JEDEC) addresses three major shortcomings in the current JEDEC registered DIMM (RDIMM) and JEDEC load reduced DIMM (LRDIMM)

Petitioners
Ex. 1002, p. 32

standards (an example LRDIMM structure 600 is schematically illustrated in FIG. 6A). These shortcomings include:

a) The DIMM density limitation due to the fixed number of chip-select signals received by the DIMM from the system memory controller.
b) The performance loss due to the increased load on the data bus as the DIMM density (e.g., the number of DRAM devices and number of ranks) increases.
c) The upper bound of the DIMM density due to the physical DIMM form factor.

[0087] Further, the LRDIMM structure 600 may have timing issues due to signals passing through a single memory buffer 601. In addition, the increased size of the data path of the LRDIMM architecture 600, compared to the HCDIMM architecture 602, an example of which is schematically illustrated in FIG. 6B, may result in more latency and signal integrity issues.

[0088] FIG. 7 schematically illustrates an example memory module 700 that has previously been proposed for the 3DS-DIMM standard. The proposed 3DS Dual In-line Memory Module (3DS-DIMM) schematically illustrated in FIG. 7 attempts to address the above shortcomings using two components, a 3DS register 712 and a controller die 722. In some implementations, the 3DS-DIMM 700 includes the 3DS register 712 (also known as an enhanced DDR3 register). The 3DS register 712 may include a DDR3 JEDEC standard register with a "rank multiplication" circuit, which increases the number of the output chip-select signals for selecting an array die 710 by decoding one or more higher-order row or column address bits with the incoming chip-select signals from the system controller. The 3DS Register 712 can also include a command/address buffer, a register operational code buffer (RC word buffer), and rank multiplication logic (e.g., supporting rank multiplication of 1-to-2, 1-to-4, and/or 1-to-8). The 3DS register addresses shortcoming (a) of the JEDEC RDIMM and LRDIMM listed above.

[0089] Another component of the proposed 3DS-DIMM is a 3DS dynamic random-access memory (DRAM) package 720. The 3DS DRAM package 720 includes a plurality of stacked DDR DRAM chips 724 or array dies. The 3DS DRAM package 720 has a data I/O load that is equivalent to the data I/O load of a single-die DRAM package, regardless of the actual number of DRAM dies in the 3DS

Petitioners
Ex. 1002, p. 33

DRAM package 720. The 3DS DRAM package 720 may comprise a plurality of array dies (e.g., 2, 4 or 8 DDR DRAM dies) and a controller die 722. The controller die 722 may include data buffers, a data path timing controller, a secondary command/address buffer, a programmable secondary 1-to-2 rank decoder, and a data path signal (e.g., ODT, ODT, RTT) controller. Examples of such memory packages include HMC. The 3DS DRAM package 720 addresses the shortcomings (b) and (c) of the JEDEC RDIMM and LRDIMM listed above. However, there are deficiencies in the proposed 3DS-DIMM standard as described below.

[0090] The JEDEC DDR3 RDIMM standard contains two major components: a DDR3 register and a plurality of DRAM packages each comprising one or more DRAM chips or dies. The DDR3 register serves as a command/address signal buffer and as a command/control decoder. This DDR3 register holds a set of register control (RC) words, which a computer system configures to ensure the proper operation of the RDIMM. The DDR3 register contains a phase-lock-loop (PLL), which functions as a clock synchronizer for each RDIMM. The DDR3 register outputs a set of buffered command/address signals to all the DRAM packages on the RDIMM, but the data signals are directly fed to the DRAM packages from the system memory controller.

[0091] In contrast to the JEDEC DDR3 RDIMM standard, the 3DS-DIMM proposal requires the DRAM package 720 to include the controller die 722 that controls all data paths timing and operations. This arrangement of the DRAM package reduces the load on the data bus, however, it presents three significant shortcomings:

1) The command/address buffer in the 3DS DRAM package introduces clock cycle latency since it needs to provide clock synchronous operation to the DRAM dies in the package.
2) The data path control circuit (e.g., ODT, read/write data direction) in the control die becomes very complicated to support semi-synchronous (fly-by) data path control among all 3DS DRAM packages that are on a DIMM.
3) The variations in the DRAM die timing characteristics within each 3DS DRAM package would be likely to require resynchronization of the data signals during the read/write operations, which increases the read/write latency and the read-to-write and write-to-read turn around time.

Petitioners
Ex. 1002, p. 34

[0092] The 3DS-DIMM proposal can also be compared to the HyperCloud™ (HC) DIMM architecture of Netlist, Inc. An example of the HCDIMM architecture 602 is illustrated in FIG. 6B. Further details and embodiments of the HCDIMM architecture is disclosed in U.S. Pat. Nos. 7,289,386, 7,532,537, 7,619,912, and 7,636,274, each of which is incorporated in its entirety by reference herein. One of the main topological differences between the 3DS-DIMM and HCDIMM architectures is that while the 3DS-DIMM architecture uses a control die 722 to buffer the data signals and to decode command/address signals from the 3DS register, certain configurations of the HCDIMM architecture include a plurality of isolation devices (ID), each of which includes data buffers, but no decoding capability. Unlike the HCDIMM architecture, since the command/address signal needs to pass through the control die 722 in the 3DS DRAM package 720, the 3DS-DIMM proposal presents the same shortcoming (b) of the controller die described above. The data path control signals are generated by the register device (RD) 612 in the HCDIMM architecture 602, while the data path control signals are generated by the control die 722 in the 3DS DRAM. This aspect of the 3DS-DIMM architecture creates timing critical control paths in 3DS-DIMM architecture.

[0093] In certain embodiments described herein, a memory module architecture is proposed that includes a set of device components that support JEDEC 3DS operation with the benefit of RDIMM and HCDIMM architectures (see, e.g., FIG. 8). This set of device components may comprise two components: a register device 812 (RD), which in some embodiments may be the same or similar RD component as used in HCDIMM architectures, and a plurality of array dies 824, such as a DDR DRAM Stack Package (DDSP). The DDSP may comprise a DRAM control die that can include command/address buffers and a data path control circuit. Certain embodiments of the architecture described herein differs from the 3DS-DIMM architecture in that instead of the controller die of the 3DS-DIMM (which provides a secondary address buffer, and a second rank multiplication decoder), certain embodiments described herein use an "ID+" die as a control die 822, which can provide both selective isolation and address pass-through. Selective isolation

Petitioners
Ex. 1002, p. 35

refers generally to a driver corresponding to an array die driving a signal to the array die in response to a corresponding chip select signal while additional drivers corresponding to additional array dies maintain a previous state (e.g. do not drive the signal). For example, assuming that the memory package 300 implements selective isolation, if a chip select signal is received that corresponds to a selection of array die 310b, then the driver 334a will drive a signal along the die interconnect 320a to the array die 310b, and the remaining drivers (e.g., drivers 334b–334f) will maintain their state (e.g., not drive the signal). Address pass-through is described in further detail below.

[0094] The data path control circuit of certain embodiments may have at least two operational modes: mode-C (HCDIMM/RDIMM Compatible mode) and mode-3DS (3DS DIMM compatible mode). In mode-C, the array dies 824 (e.g. a DDSP) receive the data path control signals from the register device 812, which can be configured to control a command/address time slot and a data bus time slot. In mode-3DS, the control die 822 internally generates the data path control signals to ensure the proper operation of the data path.

[0095] The control die 822 of certain embodiments enables the memory module 800 to work as either a 3DS-DIMM, a RDIMM, or a HCDIMM. However, the memory module 800 may comprise a set of optional control input pins (in addition to the package pins that are included in 3DS-DRAM packages) which in mode-C receive the data path control signals from the register device 812.

[0096] As previously mentioned, FIG. 8 schematically illustrates an example embodiment of a memory module architecture 800 in accordance with the present disclosure. Advantageously, certain embodiments of the memory module architecture 800 address the shortcomings described above without adding complexity or latency, and without causing performance loss. The memory module architecture 800 includes a memory control hub 802 (also known as a memory controller hub, a memory control handler, or a northbridge) and a memory module 810. As schematically illustrated in FIG. 8, the memory control hub 802 may communicate directly with one or more components of the memory module 810.

Petitioners
Ex. 1002, p. 36

Alternatively, the memory control hub 802 may communicate with one or more intermediary system components (not shown), which in turn communicate with one or more components of the memory module 810. In some embodiments, the memory control hub 802 may be integrated with another component of the computer system, such as a Central Processing Unit (CPU).

[0097] Further, in some embodiments, the memory control hub 802 may communicate with one or more memory modules. Each of the memory modules may be similar or substantially similar to the memory module 810. Alternatively, some of the memory modules may differ from memory module 810 in configuration, type, or both. For example, some of the memory modules may be capable of operating at different frequencies, may include a different amount of storage capability, or may be configured for different purposes (e.g. graphics versus non-graphics memory). Although in some cases a system may include memory modules capable of operating at different frequencies, each memory module may be configured to operate at the same frequency when used in a specific device. In certain implementations, the memory control hub 802 may set the operating frequency for the one or more memory modules.

[0098] The memory module 810 may include a register device 812, which is configured to receive command, address, or command and address signals from the memory control hub 802. For the purpose of simplifying discussion, and not to limit these signals, the signals will be referred to herein as command/address signals.

[0099] In some embodiments, the register device 812 receives the command/address signals via a command/address bus 814. The command/address bus 814, although illustrated as a single line, may include as many lines or signal conduits as the number of bits of the command/address signal.

[0100] Further, the register device 812 may generate data path control signals, which can be provided to a control die 822, or isolation device, of a memory package 820 via one or more data path control lines 816. In certain embodiments, the control dies 822 can include some or all of the embodiments described above

Petitioners
Ex. 1002, p. 37

with respect to the control dies 230 and 330. Moreover, in certain embodiments, the memory packages 820 can include some or all of the embodiments described above with respect to the memory package 200 and 300. In general, the memory module 812 may include one or more memory packages 820.

[0101] In certain embodiments, each control die 822, or isolation device, may be capable of address pass-through. Address pass-through, in some cases, enables the control die 822 to provide an address signal to one or more array dies 824 without decoding the address signal. This is possible, in some implementations, because the address signal received by the control die 822 is not encoded.

[0102] Some implementations of the control dies 822 include a plurality of command/address buffers (not shown). These buffers may comprise latches. In certain embodiments, the buffers are configured to hold command/address signals to control the timing of command/address signals. In some cases, controlling the timing of the command/address signals may reduce or slow signal degradation. In some implementations, the control dies 822 include a plurality of data buffers, which may control the timing of data signals to reduce or slow signal degradation. Further, the control dies 822 may include a data path control circuit (not shown) that is configured to control command/address time slots and data bus time slots. Controlling the command/address time slots and the data bus time slots enables the control dies 822 to reduce or prevent signal collisions caused by multiple memory packages 820 sharing the data path control lines 816 and the data bus 818. In some implementations, the data path control circuit may be separate from the control die 822.

[0103] Each of the control dies 822 may be configured to receive data signals from the memory control hub 802 via the data bus 818. Further, the control dies 822 may provide data signals to the memory control hub 802 via the data bus 818. The control dies 822 may also receive data path control signals and/or command/address signals from the register device 812 via the data path control lines 816.

Petitioners
Ex. 1002, p. 38

[0104] As illustrated in FIG. 8, each memory package 820 may include one or more array dies 824 operatively coupled to the control die 822. The array dies 824 may be configured to receive data signals from the control die 822. As with the array dies 210 and 310, the array dies 824 may include any type of Random Access Memory (RAM) die. For example, the array dies 824 may include DDR DRAM, SDRAM, flash memory, or SRAM, to name a few. Further, if the memory module 810 is utilized for graphics, the array dies 824 may include GDDR SDRAM or any other type of graphics memory. In addition, the array dies 824 may be configured in a stack arrangement as illustrated in FIG. 8. Alternatively, the array dies 824 may be arranged in a planar configuration or a hybrid configuration utilizing both a planar and a stack arrangement.

[0105] In some embodiments, the memory module 810 is selectively configurable into two operational modes. In the first operational mode, the register device 812 generates data path control signals provided to the control dies 822 via the data path control lines 816. The control dies 822 may decode command/address signals included with the data path control signals generated by the register device 812. In some implementations, the control dies 822 use the data path control signals to operate the data path control circuits of the control dies 822.

[0106] In the second operational mode, the control dies 822 may operate the data path control circuits to provide command/address signals to the array dies 824 without decoding the command/address signals. In this mode, the control dies 822 may use address pass-through to provide received address signals to the array dies 824.

[0107] Other operational modes may also be possible. In some embodiments, the data path control signals generated by the register device 812 may include decoded command/address signals that are decoded from command/address signals received from the memory control hub 802 via the command/address bus 814.

[0108] In some embodiments, the register device 812 may be configured to perform rank multiplication. In addition, or alternatively, the control dies 822 may

be configured to perform rank multiplication. Embodiments of rank multiplication are described in greater detail in U.S. Patent Nos. 7,289,386 and 7,532,537, each of which are incorporated in their entirety by reference herein. In such embodiments, the register device 812 can generate additional chip select signals that are provided to the array dies 824. For instance, if the memory control hub 802 is configured to recognize a single array die 824 per memory package 820, but there exists four array dies 824 per memory package 820, the memory control hub 802 may not provide the correct number of chip select signals to access a specific memory location of the plurality of array dies 824 is memory package 820. Thus, to access the specific memory location, the register device 812 can determine the array die that includes the specific memory location to be accessed based on the command/address signals received from the memory control hub 802 and can generate the correct chip select signal to access the array die that includes the specific memory location. In certain embodiments, when the memory module 810 is operating in the second operation module as described above, the memory module 810 does not perform rank multiplication.

Terminology

[0109] Unless the context clearly requires otherwise, throughout the description and the claims, the words "comprise," "comprising," and the like are to be construed in an inclusive sense, as opposed to an exclusive or exhaustive sense; that is to say, in the sense of "including, but not limited to." The term "coupled" is used to refer to the connection between two elements, the term refers to two or more elements that may be either directly connected, or connected by way of one or more intermediate elements. Additionally, the words "herein," "above," "below," and words of similar import, when used in this application, shall refer to this application as a whole and not to any particular portions of this application. Where the context permits, words in the above Detailed Description using the singular or plural number may also include the plural or singular number respectively. The word "or" in reference to a list of two or more items covers all of the following

Petitioners
Ex. 1002, p. 40

interpretations of the word: any of the items in the list, all of the items in the list, and any combination of the items in the list.

[0110] The above detailed description of embodiments of the invention is not intended to be exhaustive or to limit the invention to the precise form disclosed above. While specific embodiments of, and examples for, the invention are described above for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize. For example, while processes or blocks are presented in a given order, alternative embodiments may perform routines having steps, or employ systems having blocks, in a different order, and some processes or blocks may be deleted, moved, added, subdivided, combined, and/or modified. Each of these processes or blocks may be implemented in a variety of different ways. Also, while processes or blocks are at times shown as being performed in series, these processes or blocks may instead be performed in parallel, or may be performed at different times.

[0111] The teachings of the invention provided herein can be applied to other systems, not necessarily the system described above. The elements and acts of the various embodiments described above can be combined to provide further embodiments.

[0112] Conditional language used herein, such as, among others, "can," "might," "may," "e.g.," and the like, unless specifically stated otherwise, or otherwise understood within the context as used, is generally intended to convey that certain embodiments include, while other embodiments do not include, certain features, elements and/or states. Thus, such conditional language is not generally intended to imply that features, elements and/or states are in any way required for one or more embodiments or that one or more embodiments necessarily include logic for deciding, with or without author input or prompting, whether these features, elements and/or states are included or are to be performed in any particular embodiment.

[0113] While certain embodiments of the inventions have been described, these embodiments have been presented by way of example only, and are not

Petitioners
Ex. 1002, p. 41

intended to limit the scope of the disclosure. Indeed, the novel methods and systems described herein may be embodied in a variety of other forms; furthermore, various omissions, substitutions and changes in the form of the methods and systems described herein may be made without departing from the spirit of the disclosure. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope and spirit of the disclosure.

Petitioners
Ex. 1002, p. 42

WHAT IS CLAIMED IS:

1. A memory package, comprising:

a plurality of array dies having data ports;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with at least one data port of a first array die of the plurality of array dies and at least one data port of a second array die of the plurality of array dies and not in electrical communication with the data ports of at least a third array die of the plurality of array dies, the second die interconnect in electrical communication with at least one data port of the third array die and not in electrical communication with the data ports of the first array die and the data ports of the second array die; and

a control die comprising at least a first data conduit configured to transmit a data signal to the first die interconnect and to not transmit the data signal to the second die interconnect, and at least a second data conduit configured to transmit the data signal to the second die interconnect and to not transmit the data signal to the first die interconnect.

2. The memory package of claim 1, wherein the first data conduit and the second data conduit are configured to receive the data signal from a common source as one another.

3. The memory package of claim 1, wherein a first load on the first data conduit comprises a load of the first die interconnect, a load of the first array die, and a load of the second array die, and wherein a second load on the second data conduit comprises a load of the second die interconnect and a load of the third array die.

4. The memory package of claim 3, wherein a difference between the first load and the second load is less than a load that would be on the first data

Petitioners
Ex. 1002, p. 43

conduit upon the first die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies.

5. The memory package of claim 3, wherein:

the first load is less than a load that would be on the first data conduit upon the first die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies; and

the second load is less than a load that would be on the second data conduit upon the second die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies.

6. The memory package of claim 1, wherein the first die interconnect comprises a first conducting rod and wherein the second die interconnect comprises a second conducting rod.

7. The memory package of claim 1, wherein one of the first die interconnect and the second die interconnect is in electrical communication with at least one data port of a fourth array die of the plurality of array dies and the other one of the first die interconnect and the second die interconnect is not in electrical communication with the data ports of the fourth array die.

8. The memory package of claim 1, wherein the first die interconnect is configured to transmit signals to the first array die and the second array die, and to receive signals from the first array die and the second array die, and the second die interconnect is configured to transmit signals to the third array die and to receive signals from the third array die.

9. The memory package of claim 1, further comprising at least a third die interconnect and a fourth die interconnect, the third die interconnect in electrical

Petitioners
Ex. 1002, p. 44

communication with at least one data port of the first array die and at least one data port of the second array die and not in electrical communication with the data ports of at least the third array die, the fourth die interconnect in electrical communication with at least one data port of the third array die and not in electrical communication with the data ports of the first array die and the data ports of the second array die.

10. The memory package of claim 9, wherein the first die interconnect is configured to transmit signals to the first array die and the second array die, the second die interconnect is configured to transmit signals to the third array die, the third die interconnect is configured to receive signals from the first array die and the second array die, and the fourth die interconnect is configured to receive signals from the third array die.

11. A memory package, comprising:

a plurality of array dies having ports;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with at least one port of a first array die of the plurality of array dies and at least one port of a second array die of the plurality of array dies and not in electrical communication with the ports of at least a third array die of the plurality of array dies, the second die interconnect in electrical communication with at least one port of the third array die and not in electrical communication with the ports of the first array die and the ports of the second array die; and

a control die comprising at least a first conduit configured to transmit a signal to the first die interconnect and to not transmit the signal to the second die interconnect, and at least a second conduit configured to transmit the signal to the second die interconnect and to not transmit the signal to the first die interconnect,

wherein a first load on the first conduit comprises a load of the first die interconnect, a load of the first array die, and a load of the second array die, and

Petitioners
Ex. 1002, p. 45

wherein a second load on the second conduit comprises a load of the second die interconnect and a load of the third array die.

12. The memory package of claim 11, wherein a difference between the first load and the second load is less than a load that would be on the first conduit upon the first die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies.

13. The memory package of claim 11, wherein:

the first load is less than a load that would be on the first conduit upon the first die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies; and

the second load is less than a load that would be on the second conduit upon the second die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies.

14. The memory package of claim 11, wherein the first die interconnect comprises a first conducting rod and wherein the second die interconnect comprises a second conducting rod.

15. The memory package of claim 11, wherein:

the first conduit comprises at least a first driver, and

the second conduit comprises at least a second driver.

16. The memory package of claim 15, wherein the first driver and the second driver are configured to receive the signal from a common source as one another.

17. The memory package of claim 16, wherein the signal comprises a data signal.

Petitioners
Ex. 1002, p. 46

18. The memory package of claim 16, wherein the signal comprises at least one of the following: a command signal, an address signal, and a control signal.

19. The memory package of claim 16, further comprising a third driver and a fourth driver, wherein the third driver is configured to drive a first signal received from the first die interconnect, and wherein the fourth driver is configured to drive a second signal received from the second die interconnect.

20. A method for optimizing load in a memory package comprising a plurality of array dies, at least a first die interconnect and a second die interconnect, and a control die comprising at least a first driver and a second driver, the first driver configured to drive a signal along the first die interconnect, and the second driver configured to drive the signal along the second die interconnect, the method comprising:

selecting a first subset of array dies of the plurality of array dies and a second subset of array dies of the plurality of array dies, wherein the first subset of array dies and the second subset of array dies are exclusive of one another and are selected to balance a load on the first driver and on the second driver based at least in part on array die loads of array dies of the plurality of array dies and at least in part on die interconnect segment loads of segments of at least the first die interconnect and the second die interconnect;

forming electrical connections between the first die interconnect and the first subset of array dies; and

forming electrical connections between the second die interconnect and the second subset of array dies.

21. The method of claim 20, further comprising:

selecting a first driver size for the first driver based, at least in part, on the load on the first driver; and

Petitioners
Ex. 1002, p. 47

selecting a second driver size for the second driver based, at least in part, on the load on the second driver.

22. The method of claim 21, wherein the first driver size and the second driver size are both less than a driver size sufficient to drive the signal along a die interconnect in electrical communication with each of the plurality of array dies with a less than predetermined signal degradation.

23. The method of claim 20, wherein each array die of the plurality of array dies comprises a substantially equal array die load.

24. The method of claim 20, wherein each segment of the first die interconnect and the second die interconnect comprises a substantially equal die interconnect segment load.

25. The method of claim 20, wherein the segments of the die interconnect comprise one or more portions of a die interconnect between two adjacent array dies.

26. The method of claim 20, wherein the segments of the die interconnect comprise one or more portions of a die interconnect between the control die and an adjacent array die.

27. The method of claim 20, wherein the first subset of array dies comprises at least two array dies.

28. The method of claim 20, wherein balancing the load on the first driver and the load on the second driver comprises minimizing a load difference between the first driver and the second driver.

29. A memory module comprising:

a register device configured to receive command/address signals from a memory control hub and to generate data path control signals; and

Petitioners
Ex. 1002, p. 48

a plurality of DRAM packages, each DRAM package comprising:

a control die comprising a plurality of command/address buffers and a data path control circuit configured to control command/address time slots and data bus time slots, the control die configured to receive data signals from the memory control hub, the data path control signals from the register device, and command/address signals from the register device; and

a plurality of DDR DRAM dies operatively coupled to the control die to receive the data signals from the control die,

wherein the memory module is selectively configurable into at least two operational modes comprising:

a first operational mode in which the register device generates the data path control signals, and the control die uses the data path control signals to operate the data path control circuit; and

a second operational mode in which the control die operates the data path control circuit to provide the command/address signals to the plurality of DDR DRAM dies without decoding the command/address signals.

30. The memory module of claim 29, wherein the register device is further configured to perform rank multiplication.

31. The memory module of claim 29, wherein the memory module does not perform rank multiplication while operating in the second operational mode.

32. The memory module of claim 29, wherein the control die is further configured to perform rank multiplication.

33. The memory module of claim 29, wherein at least one of the plurality of DRAM packages comprises the memory package of claim 1.

Petitioners
Ex. 1002, p. 49

34. The memory module of claim 29, wherein at least one of the plurality of DRAM packages comprises the memory package of claim 11.

Petitioners
Ex. 1002, p. 50

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

ABSTRACT

An apparatus is provided that includes a plurality of array dies and at least two die interconnects. The first die interconnect is in electrical communication with a data port of a first array die and a data port of a second array die and not in electrical communication with data ports of a third array die. The second die interconnect is in electrical communication with a data port of the third array die and not in electrical communication with data ports of the first array die and the second array die. The apparatus includes a control die that includes a first data conduit configured to transmit a data signal to the first die interconnect and not to the second die interconnect, and at least a second data conduit configured to transmit the data signal to the second die interconnect and not to the first die interconnect.

12136053
102011

Petitioners
Ex. 1002, p. 51

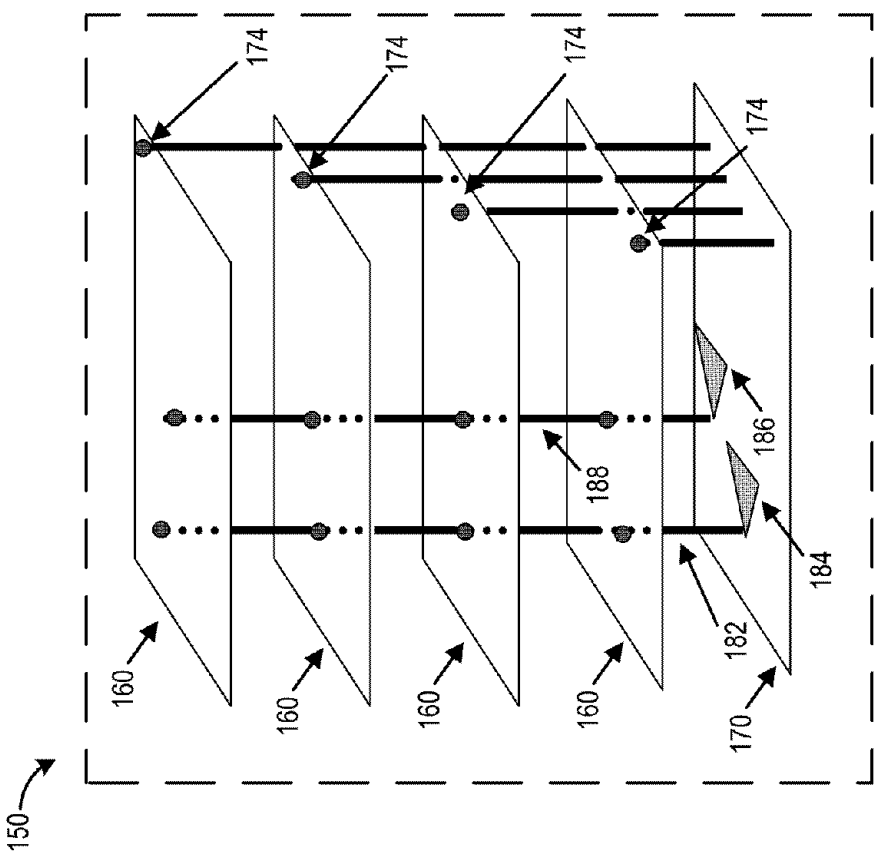

150
160
160
160
160
170
174
174
174
174
182
184
186
188

FIG. 1B

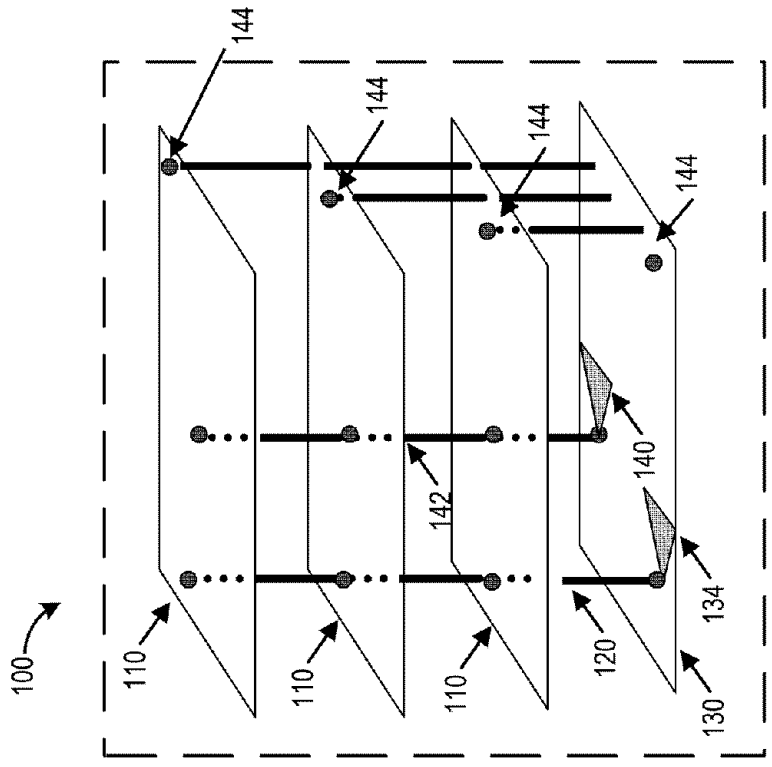

100
110
110
110
120
130
134
140
142
144
144
144
144

FIG. 1A

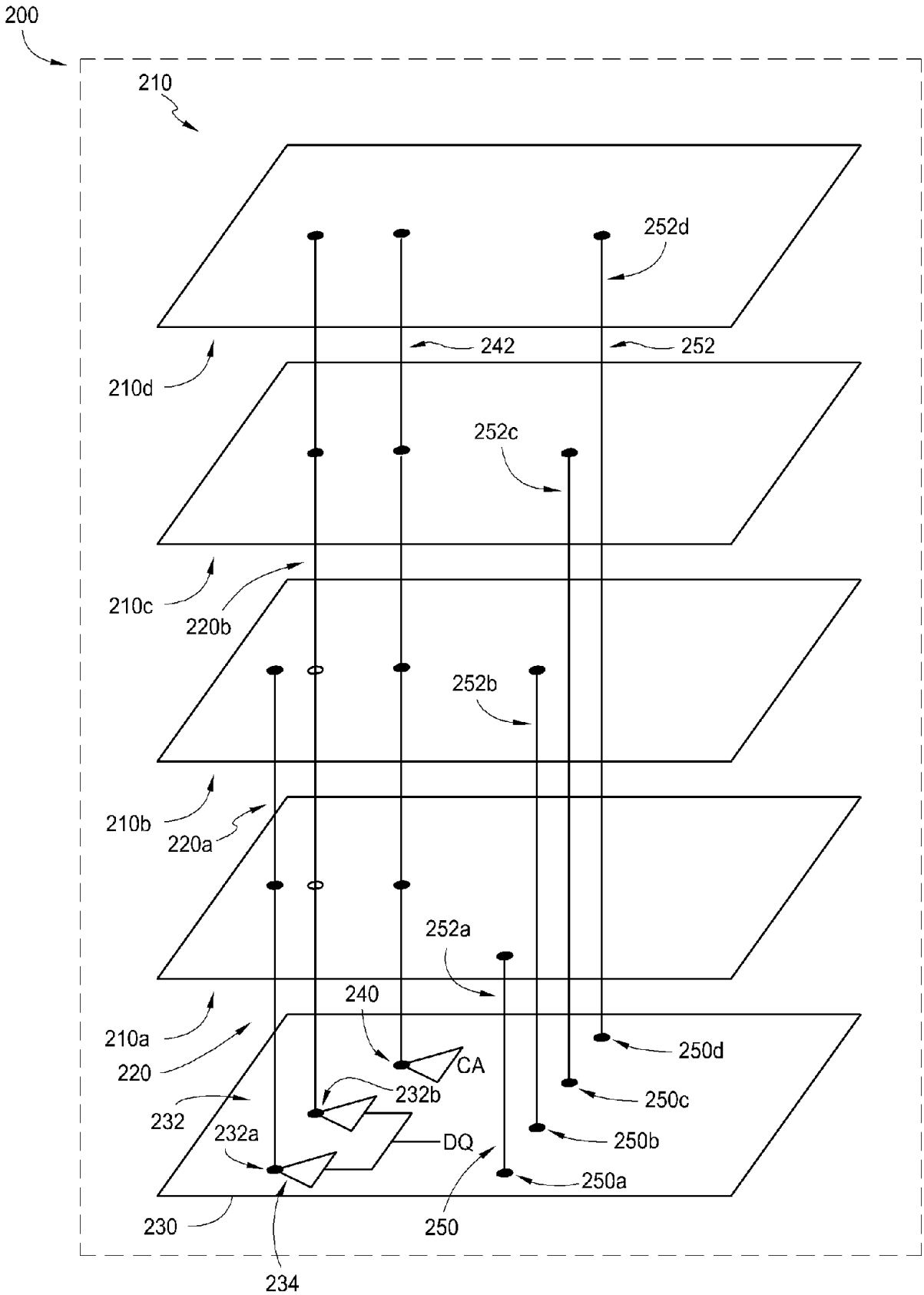

200
210
252d
242
252
210d
252c
210c
220b
252b
210b
220a
252a
240
210a
220
CA
250d
232b
250c
232
232a
DQ
250b
250a
230
250
234

FIG. 2

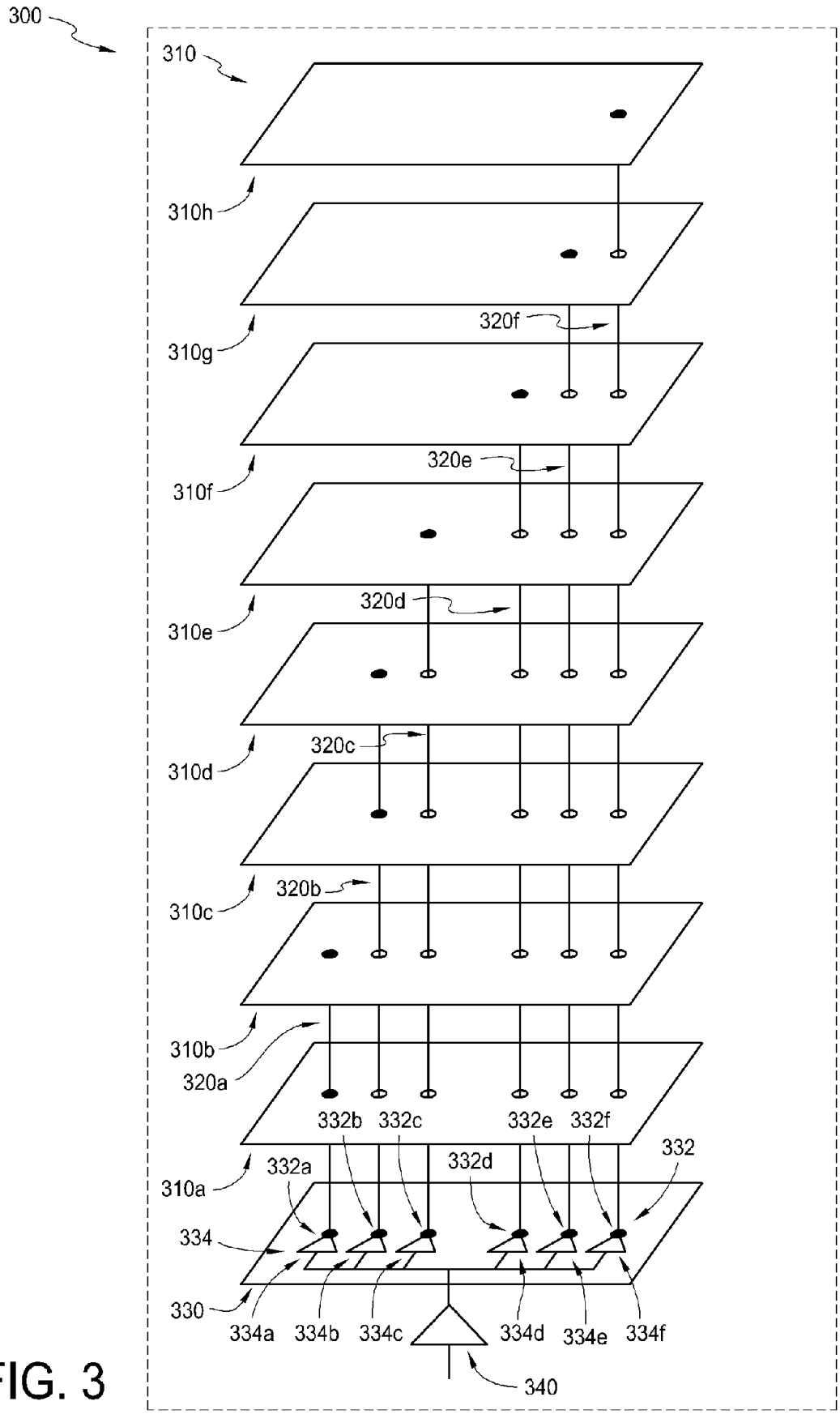

300
310
310h
320f
310g
320e
310f
320d
310e
320c
310d
320b
310c
310b
320a
332b
332c
332e
332f
332
332a
332d
310a
334
330
334a
334b
334c
334d
334e
334f
340

FIG. 3

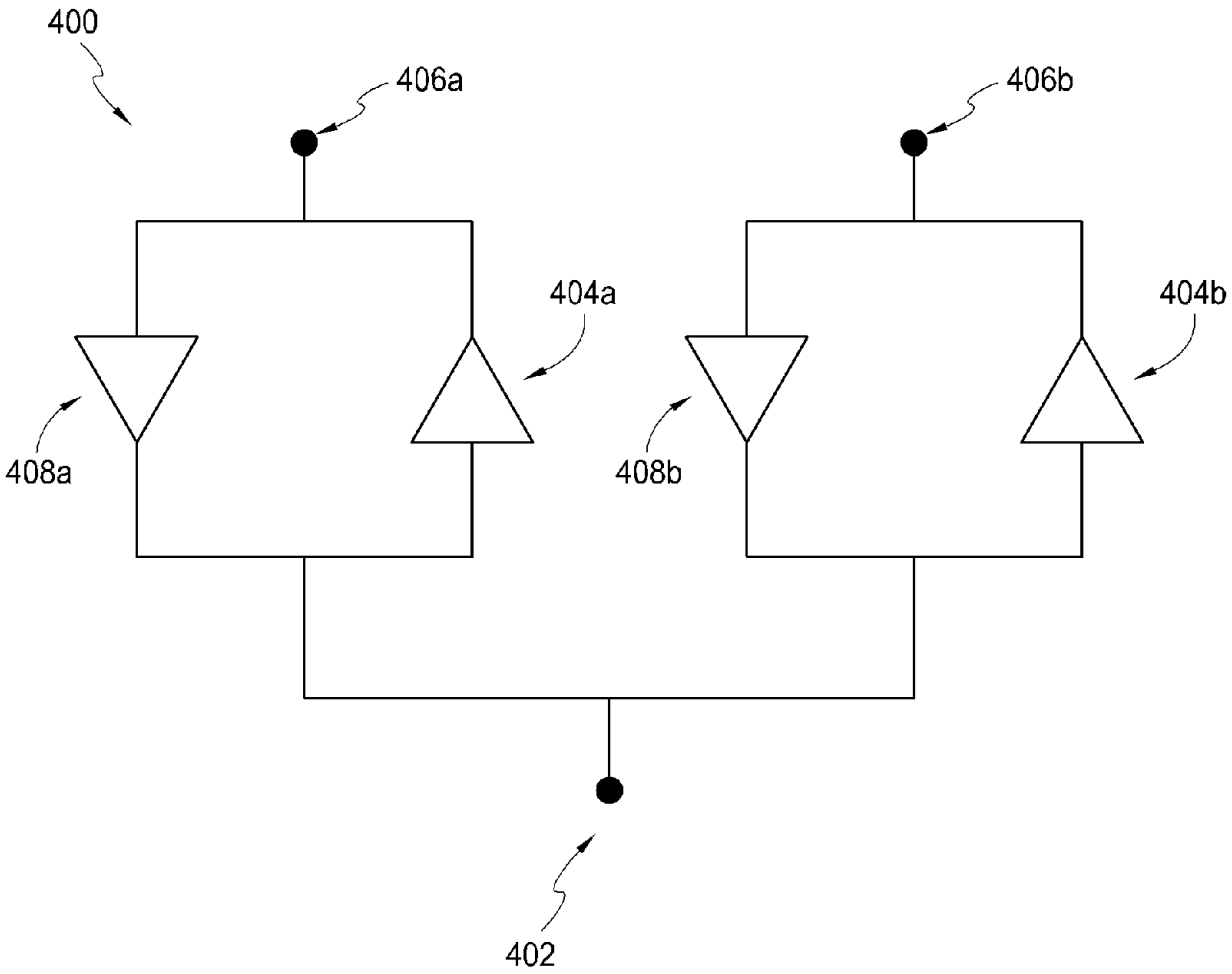


FIG. 4

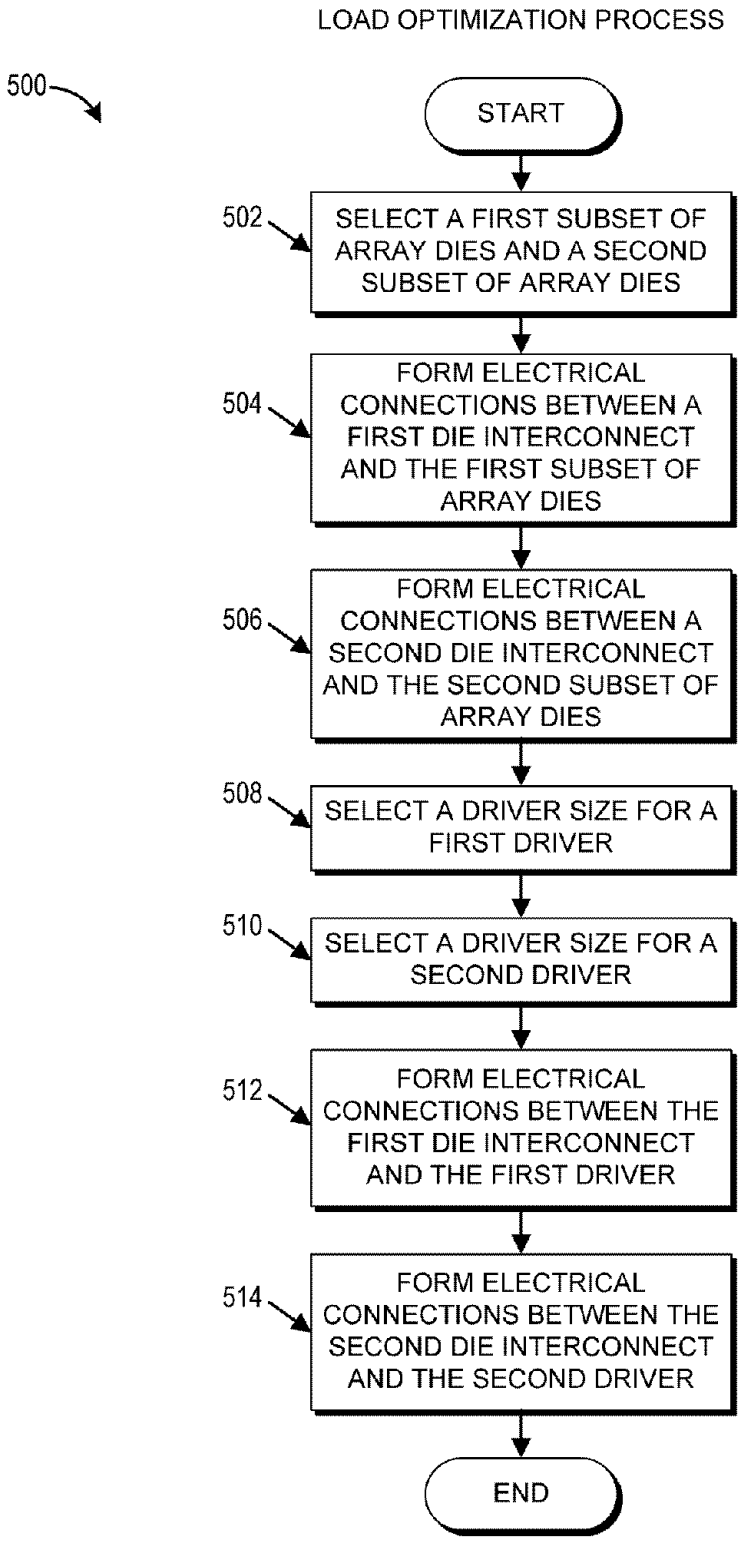

LOAD OPTIMIZATION PROCESS
500
START
502
SELECT A FIRST SUBSET OF ARRAY DIES AND A SECOND SUBSET OF ARRAY DIES
504
FORM ELECTRICAL CONNECTIONS BETWEEN A FIRST DIE INTERCONNECT AND THE FIRST SUBSET OF ARRAY DIES
506
FORM ELECTRICAL CONNECTIONS BETWEEN A SECOND DIE INTERCONNECT AND THE SECOND SUBSET OF ARRAY DIES
508
SELECT A DRIVER SIZE FOR A FIRST DRIVER
510
SELECT A DRIVER SIZE FOR A SECOND DRIVER
512
FORM ELECTRICAL CONNECTIONS BETWEEN THE FIRST DIE INTERCONNECT AND THE FIRST DRIVER
514
FORM ELECTRICAL CONNECTIONS BETWEEN THE SECOND DIE INTERCONNECT AND THE SECOND DRIVER
END

FIG. 5

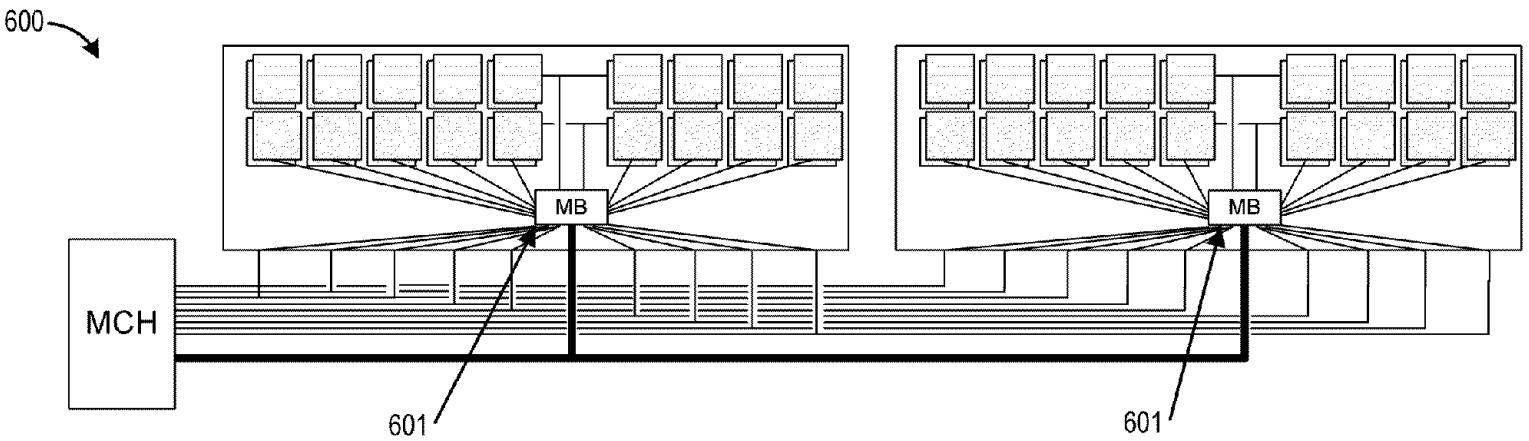


FIG. 6A

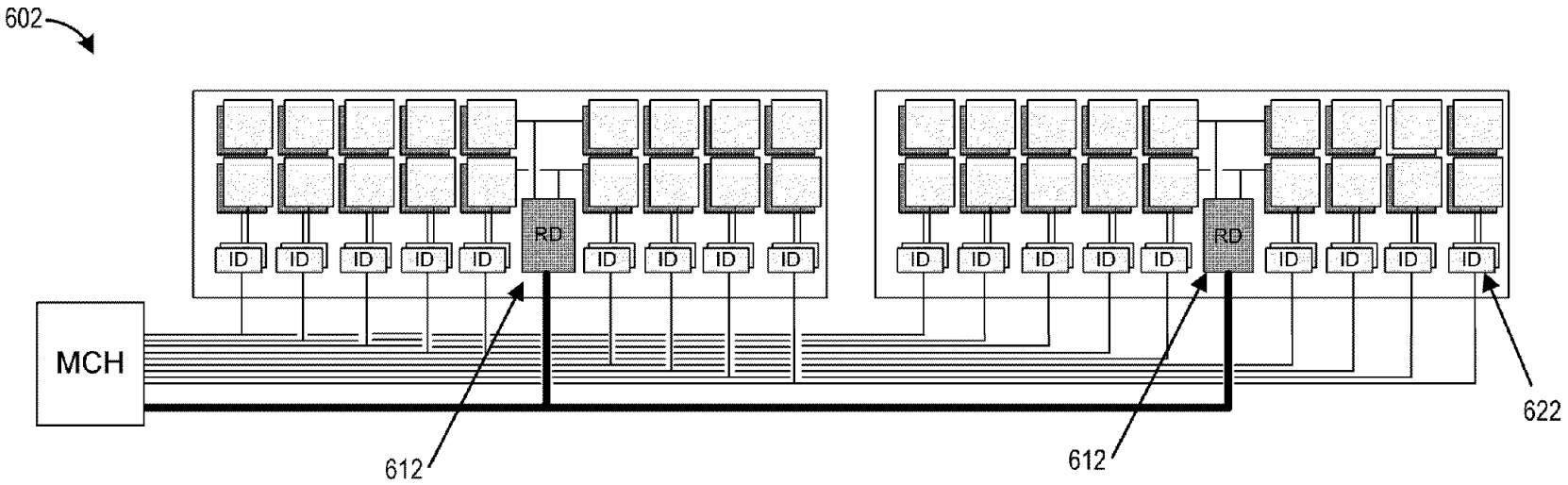


FIG. 6B

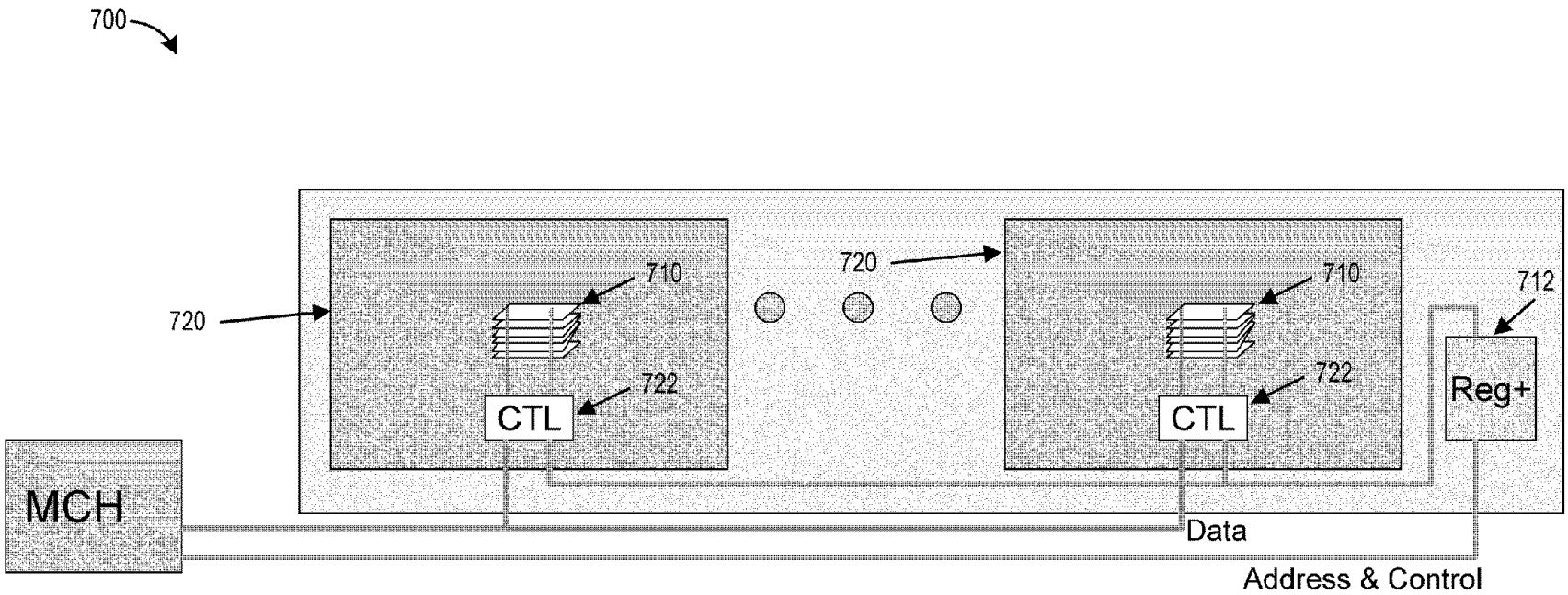

700
720
710
722
CTL
720
710
722
CTL
712
Reg+
MCH
Data
Address & Control

FIG. 7

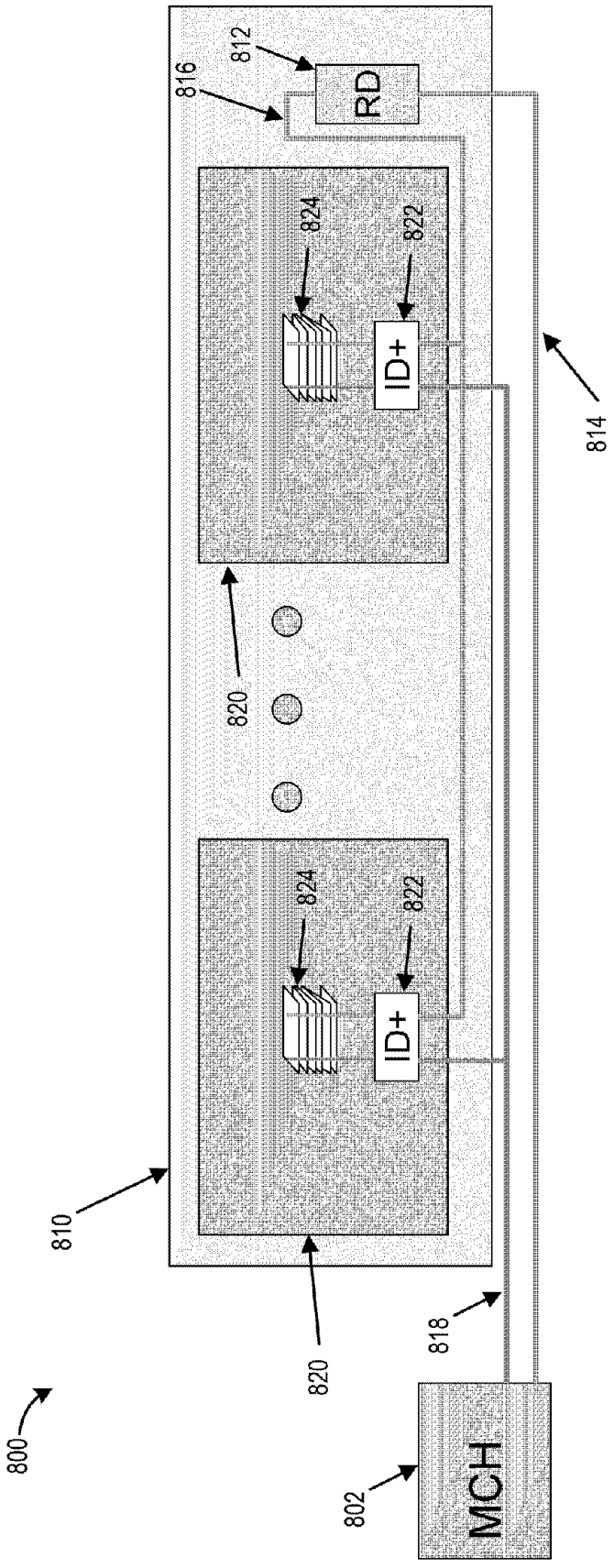


FIG. 8

Electronic Acknowledgement Receipt

EFS ID:	11333501
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	. .
Customer Number:	20995
Filer:	Bruce S. Itchkawitz/Sabrina Jacob
Filer Authorized By:	Bruce S. Itchkawitz
Attorney Docket Number:	NETL.071A
Receipt Date:	03-NOV-2011
Filing Date:	
Time Stamp:	19:18:38
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	ADS_NETL071A.pdf	174605 2c3f5c9174404d841b47d4bc505fd36507350f08	no	3

Warnings:	
Information:	

This is not an USPTO supplied ADS fillable form

2		Spec_NETL071A.pdf	2233814 c845c8c916c01c6c331f0c659fcd86bdf8256518	yes	48

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Specification	1	39
Claims	40	47
Abstract	48	48

Warnings:

Information:

3	Drawings-only black and white line drawings	Drawings_NETL071A.pdf	673727 9ebca21121abf287d1cd4768054119b91b034b9a	no	8

Warnings:

Information:

Total Files Size (in bytes):	3082146

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/288,850	11/03/2011		NETL.071A

CONFIRMATION NO. 2466

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

FORMALITIES LETTER

OC000000050966737

Date Mailed: 11/18/2011

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The statutory basic filing fee is missing.
  *Applicant must submit $**380** to complete the basic filing fee for a non-small entity. If appropriate, applicant may make a written assertion of entitlement to small entity status and pay the small entity filing fee (37 CFR 1.27).*
- The oath or declaration is missing.
  A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.
  Note: If a petition under 37 CFR 1.47 is being filed, an oath or declaration in compliance with 37 CFR 1.63 signed by all available joint inventors, or if no inventor is available by a party with sufficient proprietary interest, is required.

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- Additional claim fees of $**1090** as a non-small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
- A surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of $**130** for a non-small entity, must be submitted.

SUMMARY OF FEES DUE:

Total fee(s) required within **TWO MONTHS** from the date of this Notice is $**2470** for a non-small entity
- $**380** Statutory basic filing fee.
- $**130** Surcharge.
- The application search fee has not been paid. Applicant must submit $**620** to complete the search fee.

- The application examination fee has not been paid. Applicant must submit $**250** to complete the examination fee for a non-small entity.
- Total additional claim fee(s) for this application is $**1090**
  - $**250** for **1** independent claims over 3.
  - $**840** for **14** total claims over 20.

Replies should be mailed to:

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

Registered users of EFS-Web may alternatively submit their reply to this notice via EFS-Web.
https://sportal.uspto.gov/authenticate/AuthenticateUserLocalEPF.html

For more information about EFS-Web please call the USPTO Electronic Business Center at **1-866-217-9197** or visit our website at http://www.uspto.gov/ebc.

If you are not using EFS-Web to submit your reply, you must include a copy of this notice.

/mhteklu/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
13/288,850	11/03/2011	2811	0.00	NETL.071A	34	4

CONFIRMATION NO. 2466

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

FILING RECEIPT

OC00000050966736

Date Mailed: 11/18/2011

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)
Assignment For Published Patent Application
NETLIST, INC., Irvine, CA
Power of Attorney: None

Domestic Priority data as claimed by applicant
This appln claims benefit of 61/409,893 11/03/2010

Foreign Applications (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

If Required, Foreign Filing License Granted: 11/14/2011

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/288,850**

Projected Publication Date: To Be Determined - pending completion of Missing Parts

Non-Publication Request: No

Early Publication Request: No

Petitioners
Ex. 1002, p. 64

Title

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

Preliminary Class

257

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as

Petitioners
Ex. 1002, p. 65

set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage, facilitate, and accelerate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 13/288,850

APPLICATION AS FILED - PART I

	(Column 1)	(Column 2)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE($)	FEE($)		RATE($)	FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	380
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	620
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	250
TOTAL CLAIMS (37 CFR 1.16(i))	34 minus 20 =	* 14			OR	x 60 =	840
INDEPENDENT CLAIMS (37 CFR 1.16(h))	4 minus 3 =	* 1				x 250 =	250
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						0.00
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							0.00
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	2340

APPLICATION AS AMENDED - PART II

		(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1)		(Column 2)	(Column 3)					
		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

MULTIPLE DEPENDENT CLAIM FEE CALCULATION SHEET

Substitute for Form PTO-1360
(For use with Form PTO/SB/06)

Application Number	Filing Date
13288850	

Applicant(s)

* May be used for additional claims or amendments

CLAIMS	AS FILED		AFTER FIRST AMENDMENT		AFTER SECOND AMENDMENT			*		*		*	
	Indep	Depend	Indep	Depend	Indep	Depend		Indep	Depend	Indep	Depend	Indep	Depend
1	1						51						
2		1					52						
3		1					53						
4		1					54						
5		1					55						
6		1					56						
7		1					57						
8		1					58						
9		1					59						
10		1					60						
11	1						61						
12		1					62						
13		1					63						
14		1					64						
15		1					65						
16		1					66						
17		1					67						
18		1					68						
19		1					69						
20	1						70						
21		1					71						
22		1					72						
23		1					73						
24		1					74						
25		1					75						
26		1					76						
27		1					77						
28		1					78						
29	1						79						
30		1					80						
31		1					81						
32		1					82						
33		(1)					83						
34		(1)					84						
35							85						
36							86						
37							87						
38							88						
39							89						
40							90						
41							91						
42							92						
43							93						
44							94						
45							95						
46							96						
47							97						
48							98						
49							99						
50							100						
Total Indep	4		0		0								
Total Depend	30	↵	0	↵	0	↵							
Total Claims	34		0		0								

Please Direct All Correspondence to Customer Number 20995

RESPONSE TO FILE MISSING PARTS

Applicant	:	Hyun Lee
App. No.	:	13/288,850
Filed	:	November 3, 2011
For	:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
Art Unit	:	2811
Conf No.	:	2466

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

The above-captioned application was filed without a Declaration and/or filing fees. Enclosed in compliance with 37 CFR 1.53(f) are the following.

(X) A Declaration in 1 page.

(X) Fees will be paid via EFS Web. Extension of time is requested by payment of any extension fee.

The Commissioner is hereby authorized to charge any additional fees which may be required, now or in the future, or credit any overpayment, to Account No. 11-1410.

Bruce S. Itchkawitz
Registration No. 47,677
Attorney of Record
Customer No. 20995
(949) 760-0404

12506819

DECLARATION FOR UTILITY OR DESIGN APPLICATION UNDER 37 CFR 1.63

Docket No.: NETL.071A

Title: METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

Inventors: Hyun Lee

Please Direct All Correspondence to Customer Number 20995

This Declaration is directed to the invention described in the application that:

Was filed as Application No. 13/288,850 filed on November 3, 2011

As a below named inventor:

I believe the inventor(s) named below to be the original and first inventor(s) of the subject matter which is described and claimed and for which a patent is sought;

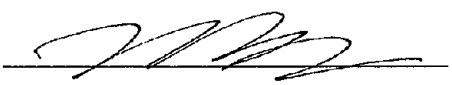
I have reviewed and understand the contents of the above-identified application, including the claims, and any amendment filed herewith or identified above;

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56;

This Declaration is being submitted with an Application Data Sheet (ADS). The ADS sets forth any applicable Foreign Priority Claims under 35 USC § 119, and sets forth the full mailing address and residence of each inventor whose signature appears below as allowed under 37 CFR 1.63(c). The ADS also sets forth any Domestic Priority Claims under 35 USC §§ 119(e), 120, 121, and 365.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful, false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of **first** inventor: Hyun Lee

Signature:  Date: 12/15/2011

Citizenship: U.S.

Send Correspondence To:
KNOBBE, MARTENS, OLSON & BEAR, LLP
Customer No. 20995

12339133
112811

Electronic Patent Application Fee Transmittal

Application Number:	13288850
Filing Date:	03-Nov-2011
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	
Filer:	Bruce S. Itchkawitz/Nicole Jevsevar
Attorney Docket Number:	NETL.071A

Filed as Large Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Utility application filing	1011	1	380	380
Utility Search Fee	1111	1	620	620
Utility Examination Fee	1311	1	250	250
Pages:				
Claims:				
Claims in excess of 20	1202	14	60	840
Independent claims in excess of 3	1201	1	250	250
Miscellaneous-Filing:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Late filing fee for oath or declaration	1051	1	130	130
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**2470**

Electronic Acknowledgement Receipt

EFS ID:	11745939
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	
Customer Number:	20995
Filer:	Bruce S. Itchkawitz
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	03-JAN-2012
Filing Date:	03-NOV-2011
Time Stamp:	14:29:05
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$2470
RAM confirmation Number	704
Deposit Account	111410
Authorized User	KNOBBE MARTENS OLSON AND BEAR

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Applicant Response to Pre-Exam Formalities Notice	RESPONSE.pdf	32475 b5a21dfa5021992daa770638f983debc9a2d9e1a	no	1
Warnings:					
Information:					
2	Oath or Declaration filed	declaration.pdf	52711 1c4d126c6df950d1577987b1dc89bb5d744fcced	no	1
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	40417 08d309b26aecb80615040bda92b3c31b994a2a75	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	125603		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number: 13/288,850

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	380
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	620
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	250
TOTAL CLAIMS (37 CFR 1.16(i))	34 minus 20 =	* 14			OR	x 60 =	840
INDEPENDENT CLAIMS (37 CFR 1.16(h))	4 minus 3 =	* 1				x 250 =	250
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						0.00
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							0.00
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	2340

APPLICATION AS AMENDED - PART II

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/288,850	11/03/2011	Hyun Lee	NETL.071A

CONFIRMATION NO. 2466

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

NOTICE



Date Mailed: 01/12/2012

NOTICE OF INFORMAL APPLICATION

This application is considered to be informal since it does not comply with the regulations for the reason(s) indicated below. The period within to correct the informalities noted below and avoid abandonment is set in the accompanying Office action.

Items Required To Avoid Processing Delays:

The item(s) indicated below are also required and should be submitted with any reply to this notice to avoid further processing delays.

A new oath or declaration, identifying this application number, or, if appropriate, an application data sheet (37 CFR 1.76), is required. The oath or declaration does not comply with 37 CFR 1.63 in that it:

- does not identify the residence (e.g., city and either state or foreign country) of each inventor.
- does not identify the complete mailing or post office address of each inventor.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
13/288,850	11/03/2011	2811	2470	NETL.071A	34	4

CONFIRMATION NO. 2466

UPDATED FILING RECEIPT

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

OC00000051873698

Date Mailed: 01/12/2012

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)

Hyun Lee, Residence Not Provided;

Assignment For Published Patent Application

NETLIST, INC., Irvine, CA

Power of Attorney: None

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/409,893 11/03/2010

Foreign Applications (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

If Required, Foreign Filing License Granted: 11/14/2011

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/288,850**

Projected Publication Date: 05/03/2012

Non-Publication Request: No

Early Publication Request: No

Title

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

Preliminary Class

257

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as

set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage, facilitate, and accelerate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/288,850	11/03/2011	Hyun Lee	NETL.071A

CONFIRMATION NO. 2466

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

PUBLICATION NOTICE



Title:METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

Publication No.US-2012-0106228-A1
Publication Date:05/03/2012

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 1 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT DOCUMENTS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	6,243,283 B1	06-05-2001	Bertin et al.	
	2	6,551,857 B2	04-22-2003	Leedy	
	3	7,098,541 B2	08-29-2006	Adelmann	
	4	7,200,021 B2	04-03-2007	Raghuram	
	5	7,254,036 B2	08-07-2007	Pauley et al.	
	6	7,269,042 B2	09-11-2007	Kinsley et al.	
	7	7,286,436 B2	10-23-2007	Bhakta et al.	
	8	7,289,386 B2	10-30-2007	Bhakta et al.	
	9	7,375,970 B2	05-20-2008	Pauley et al.	
	10	7,442,050 B1	10-28-2008	Bhakta et al.	
	11	7,532,537 B2	05-12-2009	Solomon et al.	
	12	7,619,893 B1	11-17-2009	Yu	
	13	7,619,912 B2	11-17-2009	Bhakta et al.	
	14	7,630,202 B2	12-08-2009	Pauley et al.	
	15	7,633,165 B2	12-15-2009	Hsu et al.	
	16	7,636,274 B2	12-22-2009	Solomon et al.	
	17	7,683,459 B2	03-23-2010	Ma et al.	
	18	7,811,097 B1	10-12-2010	Bhakta et al.	
	19	7,827,348 B2	11-02-2010	Lee et al.	
	20	7,830,692 B2	11-09-2010	Chung et al.	
	21	7,839,645 B2	11-23-2010	Pauley et al.	
	22	7,864,627 B2	01-04-2011	Bhakta et al.	
	23	7,881,150 B2	02-01-2011	Solomon et al.	
	24	7,894,229 B2	02-22-2011	Lahtinen et al.	
	25	7,894,230 B2	02-22-2011	Kim	
	26	7,978,721 B2	07-12-2011	Jeddeloh et al.	
	27	7,990,171 B2	08-02-2011	Chung et al.	
	28	7,999,367 B2	08-16-2011	Kang et al.	

Examiner Signature	Date Considered
*__Examiner:__ Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 2 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT DOCUMENTS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	29	8,001,434 B1	08-16-2011	Lee et al	
	30	8,019,589 B2	09-13-2011	Rajan et al.	
	31	8,033,836 B1	10-11-2011	Bhakta et al.	
	32	2006/0233012 A1	10-19-2006	Sekiguchi et al.	
	33	2007/0096332 A1	05-03-2007	Satoh et al.	
	34	2008/0025123 A1	01-31-2008	Rajan et al.	
	35	2008/0094808 A1	01-24-2008	Kanapathippillai et al.	
	36	2008/0025134 A1	01-31-2008	Rajan et al.	
	37	2008/0253085 A1	10-16-2008	Soffer	
	38	2008/0296779 A1	12-04-2008	Matsui et al.	
	39	2009/0070727 A1	03-12-2009	Solomon	
	40	2009/0103345 A1	04-23-2009	McLaren et al.	
	41	2009/0290442 A1	11-26-2009	Rajan	
	42	2010/0020583 A1	01-28-2010	Kang et al.	
	43	2010/0090338 A1	04-15-2010	Lee et al.	
	44	2010/0174858 A1	07-08-2010	Chen et al.	
	45	2011/0006360 A1	01-13-2011	Ikebuchi	
	46	2011/0016250 A1	01-20-2011	Lee et al.	
	47	2011/0016269 A1	01-20-2011	Lee et al.	
	48	2011/0050320 A1	03-03-2011	Gillingham	
	49	2011/0108888 A1	05-12-2011	Or-Bach et al.	
	50	2011/0125982 A1	05-26-2011	Choi et al.	
	51	2011/0156232 A1	06-30-2011	Youn et al.	
	52	2011/0169171 A1	07-14-2011	Marcoux	

Examiner Signature	Date Considered

*__Examiner:__ Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

__T[1]__ - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 3 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT APPLICATIONS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Filing Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	53	12/422,912 (Attorney Docket No.NETL.044A)	04-13-2009	Hyun Lee, et al.	
	54	12/422,853 (Attorney Docket No.NETL.045A)	04-13-2009	Hyun Lee, et al.	
	55	12/815,339 (Attorney Docket No.NETL.062A)	06-14-2010	Hyun Lee	

FOREIGN PATENT DOCUMENTS						
Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	56	WO 2010-138480	12-02-2010	Rambus Inc.		
	57	WO 2011-049710	04-28-2011	Rambus Inc.		
	58	WO 2011-094437	08-04-2011	Hewlett Packard Development Co.		

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	59	AHMAD et al., "Modeling of peak-to-peak switching noise along a vertical chain of power distribution TSV pairs in a 3D stack of ICs interconnected through TSVs," Norchip Conference, Article Number 5669473, Nov. 15-16, 2010, IEEE Computer Society	
	60	BLACK et al., "Die Stacking (3D) Microarchitecture," MICRO-39, 39th Annual IEEE/ACM International Symposium on, Dec. 2006, 469-479, Orlando, FL	
	61	DANESHTALAB et al., "CMIT – A novel cluster-based topology for 3D stacked architectures," 3D Systems Integration Conference (3DIC), 2010 IEEE International, Nov. 16-18, 2010, pp. 1-5	
	62	FUNAYA et al., "Cache partitioning strategies for 3-D stacked vector processors," IEEE 3D System Integration Conference, article no. 5751453, Nov. 16-18, 2010, IEEE Computer Society	
	63	GHOSH et al., "Smart Refresh: An Enhancement Memory Controller Design for Reducing Energy in Conventional and 3D Die-Stacked DRAMs," Microarchitecture, 2007, 40th Annual IEEE/ACM International Symposium, Dec. 1-5, 2007, pp. 134-145.	
	64	KANG et al., "8 Gb 3-D DDR3 DRAM using through-silicon-via technology," IEEE Journal of Solid-State Circuits, v 45, n 1, 111-19, Jan. 2010, IEEE, USA	

Examiner Signature	Date Considered
***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 4 OF 4	Attorney Docket No.	NETL.071A

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	65	KANG et al., "Signal integrity and reliability of a new Multi-Stack Package using a Pressure Conductive Rubber," Electrical Design of Advanced Packaging and Systems Symposium, 214-17, Dec. 2008, Seoul, South Korea	
	66	KAWANO, "A 3D Packaging Technology for High-Density Stacked DRAM," VLSI Technology, Systems and Applications, 2007, 23-25 April 2007, pp. 1-2	
	67	KIM et al., "A Quantitative Analysis of Performance Benefits of 3D Die Stacking on Mobile and Embedded SoC	
	68	KURITA et al., "A 3-D packaging technology with highly-parallel memory/logic interconnect," IEICE Transactions on Electronics, v E92-C, no. 12, pp. 1512-1522, 2009, Maruzen Col, Ltd.	
	69	KURITA et al., "Vertical Integration of Stacked DRAM and High-Speed Logic Device Using SMAFTI Technology," Advanced Packaging, IEEE Transactions, Aug. 2009, Volume 32 Issue 3, pp. 657-665	
	70	"Posts Tagged '3D Stacking'," from Chip Design Mag., http://chipdesignmag.com/lpd/blog/tag/3d-stacking/ (Printed October 13, 2011).	
	71	RUSSELL, Gill, "Intel Micron Hybrid Memory Cube: The Future of Exascale Computing," Bright Side of News. Sept. 19, 2011, <http://www.brightsideofnews.com/news/2011/9/19/intel-micron-hybrid-memory-cube-the-future-of-exascale-computing.aspx> Printed October 13, 2011 in 8 pages.	
	72	LOH, "3D-Stacked Memory Architectures for Multi-core Processors," Proceedings of the 35th Annual International Symposium on Computer Architecture (ISCA '08), 453-464, IEEE Computer Society, Washington DC, USA	
	73	VAL, "The 3D interconnection applications for mass memories and microprocessors," Proceedings of the Technical Conference, 1991 International Electronic Packaging Conference, 851-60, vol. 2, 1991, Int. Electron. Packaging Soc, Wheaton, IL, USA, 2008	
	74	WEIS et al., "Design space exploration for 3D-stacked DRAMs," Proceedings – Design, Automation and Test in Europe Conference and Exhibition, 2011	
	75	ZHANG et al. "A Customized Design of DRAM Controller for On-Chip 3D DRAM Stacking" Custom Integrated Circuits Conference (CICC), 2010 IEEE, Sept. 19-22 2010, pp. 1-4	

13081580

Examiner Signature	Date Considered
*Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 December 2010 (02.12.2010)

PCT

(10) International Publication Number
WO 2010/138480 A2

(51) International Patent Classification:
H01L 23/60 (2006.01)

(21) International Application Number:
PCT/US2010/036020

(22) International Filing Date:
25 May 2010 (25.05.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/181,251 26 May 2009 (26.05.2009) US

(71) Applicant *(for all designated States except US)*: **RAMBUS INC.** [US/US]; 4440 El Camino Real, Los Altos, CA 94022 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **BEST, Scott, C.** [US/US]; 4283 Ponce Drive, Palo Alto, CA 94306 (US).

(74) Agents: **BREGMAN, Dion, M.** et al.; Morgan Lewis & Bockius LLP, 2 Palo Alto Square, 3000 El Camino Real, Suite 700, Palo Alto, CA 94306 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: STACKED SEMICONDUCTOR DEVICE ASSEMBLY

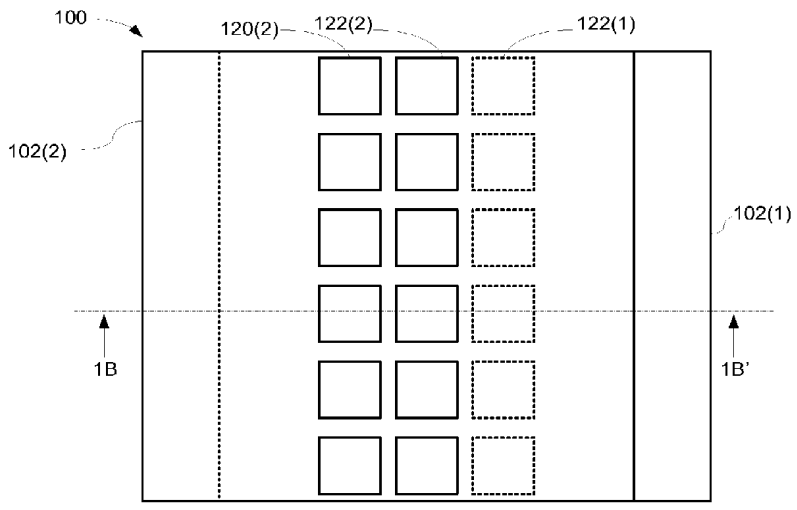


FIG. 1A

(57) **Abstract**: The semiconductor device system includes multiple stacked substantially identical semiconductor devices each including a first side and an opposing second side. First and second pads are disposed at the first side of the semiconductor device, while third and fourth pads are disposed at the second side of the semiconductor device. First interface circuit is electrically coupled to the first pad and the third pad, while second interface circuit is electrically coupled to the second pad and the fourth pad. The second interface circuit is separate and distinct from the first interface circuit. At least one first semiconductor device of the multiple semiconductor devices is offset from other of the multiple semiconductor devices such that the fourth pad on the first semiconductor device is aligned with, and electrically connected to, the first pad on an adjacent one of the multiple semiconductor devices. In some embodiments, the first pad is associated with a first capacitance, while the second pad is associated with a second capacitance that is smaller than the first capacitance.

WO 2010/138480 A2

STACKED SEMICONDUCTOR DEVICE ASSEMBLY

TECHNICAL FIELD

[0001] The embodiments disclosed herein relate to semiconductor devices, and in particular to a stacked semiconductor device assembly with at least one of the semiconductor devices offset from the remainder of the semiconductor devices.

BACKGROUND

[0002] As computer systems evolve, so does the demand for increased memory for such systems. To increase memory density, some memory packages consist of several integrated circuit (IC) dies stacked one on top of the other. These stacked multi-die packages increase the capacity of each memory device without requiring additional space on the underlying circuit board or memory module. Furthermore, Thru-Silicon Via (TSV) technology is emerging as a popular solution for enabling the largest number of die-to-die interconnect in multi-die packages.

[0003] There are, however, a number of drawbacks associated with stacked multi-die packages, especially when TSV's are used as the interconnection means. For one, it is often desirable from a manufacturing cost perspective that all of the dies in the stack are derived from the same mask set, *i.e.,* that all of the dies in a stack are substantially identical. When identical dies are stacked in a package and interconnected by TSV's, however, it is often difficult to selectively enable different modes of operation to a subset of the dies within the stack. For example, if the TSV's interconnect the stacked devices at the input/output ("IO") pad of each die, there will be a large aggregation of capacitance at that point. In particular, each die contributes capacitance associated with that die's IO pad metal, IO device loading, IO devices, and electrostatic discharge (ESD) device loading. With such a large aggregated capacitance, the stack of die would become severely limited in operational speed compared to one of the die on its own.

[0004] In other stacked multi-die packages, the TSV's may interconnect the stacked dies internally, "behind" the IO system, using some sort of multi-drop data bus topology. In these packages, one of the dies acts as the "bus master", while others act as "slave devices." If the dies are substantially identical, it is difficult to designate one of the devices as the "bus master" without utilizing bus-multiplexing circuitry that is non-optimal from a cost and performance perspective.

Petitioners
Ex. 1002, p. 86

[0005] In yet other stacked multi-die packages, it may be possible to add one or more manufacturing steps to modify some features or connections on some of the devices to facilitate die stacking. Performing such modifications, however, adds significant manufacturing and inventory control costs. As a result, the complexity and manufacturing and assembly costs for the stacked semiconductor device assembly are increased substantially.

[0006] As such, it would be highly desirable to provide a stacked die assembly that utilizes substantially identical dies and that enables selective modes of operation for at least a subset of the dies in the stack.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] For a better understanding of the disclosure herein, reference should be made to the following detailed description taken in conjunction with the accompanying drawings, in which:

[0008] Figure 1A is a plan view of an exemplary stacked semiconductor device assembly;

[0009] Figure 1B is a cross sectional view of the stacked semiconductor device assembly shown in Figure 1A, as viewed along line 1B-1B'of Figure 1A;

[0010] Figure 2A is a schematic diagram of a semiconductor device in the stacked semiconductor device assembly shown in Figures 1A and 1B;

[0011] Figure 2B is a schematic circuit diagram 210 of one embodiment of a portion of the semiconductor device shown in Figure 2A;

[0012] Figure 2C is a schematic circuit diagram of a portion of a stacked semiconductor die assembly using devices like the one shown in Figure 2B;

[0013] Figure 2D is a schematic diagram of another stacked semiconductor die assembly with mode enable circuitry;

[0014] Figure 2E is a schematic circuit diagram of the mode enable circuitry of Figure 2D;

[0015] Figure 3 is a schematic cross-sectional side view (cross-hatching removed for clarity) of another exemplary stacked semiconductor device assembly;

Petitioners
Ex. 1002, p. 87

[0016] Figure 4 is a schematic cross-sectional side view (cross-hatching removed for clarity) of another exemplary stacked semiconductor device assembly;

[0017] Figure 5 is a plan view of another exemplary stacked semiconductor device assembly; and

[0018] Figure 6 is a schematic of a computer system using a stacked memory semiconductor device assembly.

[0019] Like reference numerals refer to the same or similar components throughout the several views of the drawings.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0020] The following describes a number of exemplary assemblies that address the issues associated with multiple semiconductor devices being stacked together to form a semiconductor packaging assembly (or package), in which at least one semiconductor device in the stack performs different functions or electrical operations as compared to the other devices in the stack. Stated differently, an interface semiconductor device, while being substantially identical to the other devices in the assembly, performs different functions from those performed by the other semiconductor devices in the stack; or the interface semiconductor device operates in a different mode of operation from the mode of operation of the other devices. In one embodiment, the interface device connects to circuitry external to the stacked semiconductor device assembly, and is conveniently placed at the bottom of the stack of semiconductor devices in the assembly. For example, the interface semiconductor device may include all of the active external-IO circuitry (such as the ESD circuitry), while the remainder of the semiconductor device in the stack have lower-power "intra-stack" IO circuitry. In another example, the interface device may perform master read/write operations while the other devices perform slave read/write operations.

[0021] According to some embodiments there is provided a semiconductor device assembly that includes multiple stacked substantially identical semiconductor devices each including a first side and an opposing second side. First and second pads are disposed at the first side of the semiconductor device, while third and fourth pads are disposed at the second side of the semiconductor device. First interface circuit is electrically coupled to the first pad and the third pad, while second interface circuit is electrically coupled to the second pad and

Petitioners
Ex. 1002, p. 88

the fourth pad. The second interface circuit is separate and distinct from the first interface circuit. At least one first semiconductor device of the multiple semiconductor devices is offset from other of the multiple semiconductor devices such that the fourth pad on the first semiconductor device is aligned with, and electrically connected to, the first pad on an adjacent one of the multiple semiconductor devices. The semiconductor device assembly may be a multi-die memory chip, such as a multi-die DRAM chip, which may then be attached to a memory module, such as a DRAM memory module

[0022] In some embodiments, the first side and the second side are substantially planar and parallel to one another. Also in some embodiments, the first and third pads are electrically coupled to each other by a first via through the semiconductor device, and the second and fourth pads are electrically coupled to each other by a second via through the semiconductor device. These vias may be through-silicon-vias (TSVs). In these embodiments, the first and third pads are aligned with one another along an axis substantially perpendicular to the first and second sides of the semiconductor device, and the second and fourth pads are aligned with one another along an axis substantially perpendicular to the first and second sides of the semiconductor device.

[0023] In use, and in the context of the semiconductor device being a memory device (such as a DRAM, an SRAM, a Flash memory IC, or other types of memory chip), the semiconductor device communicates data over an internal signal line to or from core interface circuit using either the first interface circuit or the second interface circuit depending on whether the first and third pads, or the second and fourth pads, are coupled to an external signal line. The core interface circuit may include a data transmitter or a data receiver. The internal signal line is either unidirectional or bidirectional.

[0024] In some embodiments, the first interface circuit and the second interface circuit have different capacitances, *e.g.,* the first interface circuit includes electrostatic discharge (ESD) circuitry. In other embodiments, the first interface circuit is a read channel master and the second interface circuit is a read channel slave, *e.g.,* the first interface circuit is a write channel master and the second interface circuit is a write channel slave.

[0025] According to some embodiments there is provided a method for selectively enabling a mode of semiconductor device in a stacked semiconductor device assembly. Multiple semiconductor devices that are substantially identical to one another are provided,

Petitioners
Ex. 1002, p. 89

as described above. The multiple semiconductor devices are then stacked one on top of the other. Finally, at least a first semiconductor device of the multiple semiconductor devices is spatially offset from the remainder of the multiple semiconductor devices such that the fourth pad on the first semiconductor device is aligned with, and electrically connected to, the first pad on an adjacent one of the multiple semiconductor devices.

Finally, according to some embodiments there is provided a computing assembly that includes a bus;a processor coupled to the bus; communication circuitry coupled to the bus; semiconductor devices, as described above, forming multi-die memory chips that are coupled to the bus, where the multi-die memory chips include stacked semiconductors one on top of the other, and spatially offsetting at least a first semiconductor device of the multiple semiconductor devices from the remainder of the multiple semiconductor devices such that the fourth pad on the first semiconductor device is aligned with, and electrically connected to, the first pad on an adjacent one of the multiple semiconductor devices.

[0026] These embodiments provide a stacked semiconductor device assembly that allows for one or more semiconductor devices in the stack of identical semiconductor devices to have a different mode of operation to the remainder of the semiconductor devices in the stack.

[0027] Figure 1A is a plan view of an exemplary stacked semiconductor device assembly 100, while Figure 1B is a cross sectional view of the stacked semiconductor device assembly shown in Figure 1A, as viewed along line 1B-1B'of Figure 1A. The stacked semiconductor device assembly includes at least two semiconductor devices 102(1) and 102(2). In some embodiments, the semiconductor device is a die, *i.e.*, an unpackaged bare integrated circuit or chip. In some of these embodiments, the semiconductor device is a memory die (e.g., DRAM, SRAM, Flash, RRAM, FeRAM, MRAM, etc.). In other embodiments, the semiconductor device is any packaged or unpackaged integrated circuit or the like. The stacked semiconductor device assembly 100 is particularly well suited to memory because of the needs for multiple substantially identical die that, in aggregate, satisfy the ever increasing demands on memory capacity and density.

[0028] The semiconductor devices 102(1), 102(2) are substantially identical to one another, so they can be fabricated using a same or substantially same set of semiconductor fabrication process steps. Each of the semiconductor devices 102(1), 102(2) has a substantially planar

Petitioners
Ex. 1002, p. 90

first side 106(1), 106(2) and an opposing substantially planar second side 108(1), 108(2). Each semiconductor device 102(1), 102(2) also includes active circuitry 110(1), 110(2), including interface circuit, modal select circuitry, core interface circuit, a core, and/or the like, as described below. In some embodiments, the active circuitry is formed near or at the first side 106(1), 106(2) of each semiconductor device.

[0029] Each of the semiconductor devices 102(1), 102(2) has a first group of connection pads 112(1), 112(2); a second group of connection pads 114(1), 114(2); a third group of connection pads 120(1), 120(2); and a fourth group of connection pads 122(1), 122(2). The connection pads on a semiconductor device are for connection circuits formed on the device to circuits external to the device. In one example, the connection pad may be input/output (I/O) pads for connection to circuits external the assembly (e.g., other chips in the computer system). In another example, the connection pads may be contact pads for connecting to one or more intra-assembly buses. The connection pads may be of any suitable shape, such as square, round or the like. In some embodiments each group of connection pads is a stripe of connection pads that extend substantially across a respective side of each semiconductor device, *i.e.,* each stripe may form a column (or row, depending on orientation) across a side of each semiconductor device. In some embodiments, each group of connection pads is located near the middle of a respective side of a semiconductor device, however, in other embodiments, the connection pads may be located near an edge or anywhere else on the semiconductor device. In other embodiments, each of the groups of connection pads may include multiple stripes of connection pads, *e.g.*, each group includes two stripes of connection pads. In some embodiments, first group of connection pads 112(1), 112(2) and the second group of connection pads 114(1), 114(2) are disposed at the first surface 106(1), 106(2) of each of the respective semiconductor devices. Similarly, in some embodiments, the third group of connection pads 120(1), 120 (2) and the fourth group of connection pads 122(1), 122 (2) are disposed at the second surface 108(1), 108(2) of each of the respective semiconductor devices.

[0030] In some embodiments, each connection pad of the first group of connection pads 112 is aligned with and electrically connected to a corresponding connection pad of the third group of connection pads 120. Similarly, each connection pad of the second group of connection pads 114 is aligned with and electrically connected to a corresponding connection pad of the fourth group of connection pads 122. By aligned with it is meant that the

connection pads are aligned along axes that are substantially perpendicular to both the first side 106 and the second side 108 of each semiconductor device (*e.g.,* along axes parallel to a Z-axis perpendicular to both of the X and Y axes of Figure 5). In some embodiments, the electrical connections between the connection pads of the first group of connection pads 112 and the connection pads of the third group of connection pads 120 is through a series of vias, two of which are shown by reference numerals 116(1), 116(2) (Figure 1B) that extend through each semiconductor device from the connection pads at the first side 106 to the connection pads at the second side 108 of each of the respective semiconductor devices. Similarly, the electrical connections between the connection pads of the second group of connection pads 114 and the connection pads of the fourth group of connection pads 122 is also through a series of vias, one of which are shown by reference numerals 118(2) (Figure 1B) that extend through each semiconductor device from the connection pads at the first side 106(1), 106(2) to the connection pads at the second side 108(1), 108(2) of each of the respective semiconductor devices. In other words, in some embodiments, each connection pad on the first side is connected to a corresponding pad at the second side through a single via.

[0031] In some embodiments, each of the vias may be referred to as a through chip via ("TCV") or through silicon via ("TSV"). Also in some embodiments, the vias 116 and 118 are substantially aligned with their respective pads along the axes defined above. The vias may be formed by any suitable technique, such as by thinning a wafer that contains a deep metal via formed during semiconductor processing, or by electroplating through-holes formed by laser drilling subsequent to semiconductor manufacturing, or by other techniques well known in the industry. In some embodiments (not shown), the vias are filled with a conductive material. In other embodiments, the electrical connections between the connection pads on the first side 106(1), 106(2) and the connection pads on the second side 108(1), 108(2) can be made through any other suitable electrical connections including traces, redistribution layers, different types and sizes of vias, leads, tape, wires, or any combination of the aforementioned.

[0032] As will be explained in more detail below, during assembly, the second semiconductor device 102(2) (typically an upper semiconductor device) is stacked adjacent to the first semiconductor device 102(1) (typically a bottommost semiconductor device) such that the second group of connection pads 114(2) of the second semiconductor device 102(2)

Petitioners
Ex. 1002, p. 92

is adjacent to and aligned with the third group of connection pads 120(1) of the first semiconductor device 102(1). Stated differently, the second semiconductor device is offset (stepped or moved along the X and/or Y axes of Figure 5) from the first semiconductor device by a distance equal to an integer multiple of the pitch between the respective groups of connection pads (120 and 122, and/or 112 and 114) on each side of each semiconductor device. Figure 2B illustrates an offset distance where the integer multiple is "1", but other values may be used. Each of the connection pads of the second group of connection pads 114(2) of the second semiconductor device 102(2) is then electrically connected to a corresponding connection pad of the third group of connection pads 120(1) of the first semiconductor device 102(1). These electrical connections are made by any suitable means, such as by a solder reflow process, as depicted by solder contact 128. Thermosonic flip chip bonding, thermocompression bonding, or other techniques common in multi-die bonding may also be used. In some embodiments, these electrical connections are made through "solder bumps", "micro-bumps", or "C4 bumps" at each connection pad.

[0033] In some embodiments, the first semiconductor device 102(1) is electrically and/or mechanically connected to a package substrate 124, such as is used in flip-chip (FC) packaging. The FC package substrate is then usually electrically coupled to a printed circuit board (PCB) via a ball-grid array (BGA) connection technology (not shown). In particular, the connection pads of the first group of connection pads 112(1) on the first semiconductor device 102(1) are electrically connected to the substrate 124. In some embodiments, these electrical connections are made by any suitable means, such as by soldering connection pads to one another, as depicted by reference numeral 126. In some embodiments, these electrical connections are made through "solder bumps" or "C4 bumps" at each connection pad.

[0034] Once assembled, the stacked semiconductor device assembly 100 allows electrical signals to be communicated between electrical connections on the substrate 124 and those on the second semiconductor device 102(2) through the first semiconductor device 102(1). Here, the first semiconductor device 102(1) acts as the primary interface semiconductor device between the substrate 124 and the remainder of the semiconductor devices (here semiconductor device 102(2)) in the stacked semiconductor device assembly 100, and it may have active interface circuitry (e.g., ESD protection circuitry) that is different from the remainder of the semiconductor devices in the stack.

Petitioners
Ex. 1002, p. 93

[0035] Figure 2A is a schematic diagram of a semiconductor device 102 in the stacked semiconductor device assembly shown in Figures 1A and 1B. The semiconductor device 102 shows a first connection pad 112 from the first group of connection pads 112(1), 112(2) of Figures 1A and 1B, and a second connection pad 114 from the second group of connection pads 114(1), 114(2) of Figures 1A and 1B.

[0036] The first connection pad 112 is electrically connected to first interface circuit 220, and the second connection pad 114 is electrically connected to second interface circuit 222.

[0037] In some embodiments, the first interface circuit 220 and the second interface circuit 222 electrically connect to a single signal path 221 made of, for example, a signal conducting line. This signal path 221 may be unidirectional or bidirectional. The signal line 221 is in turn electrically connected to core interface circuit 212, which may include a serializer or deserializer, and which may either generate and transmit signals to the signal line 221, or receive signals from the signal line 221. The core interface circuit 212 electrically connects to the core 111 (such as memory cells or memory banks) through multiple connections. The core 111 and core interface circuit 212, are well understood in the art.

[0038] The ability to access the core 111 via either the first connection pads 112 or the second connection pads 114 provides the designer of the stacked semiconductor device assembly the ability to have a particular semiconductor device in the stack operate in a different mode from the remainder of the semiconductor devices in the stacked semiconductor assembly, *i.e.*, perform functional or electrical operations differently to other semiconductor devices in the stacked semiconductor assembly. The ability to have different modes of operation within a stacked semiconductor assembly is provided despite the stacked semiconductor assembly containing substantially identical semiconductor devices.

[0039] Figure 2B is a schematic circuit diagram 210 of one embodiment of a portion of the semiconductor device shown in Figure 2A. In this embodiment, interface circuits 220 and 222 include output drivers 226 and 224, respectively, for transmitting signals from the semiconductor device. One skilled in the art, however, will appreciate that a similar circuit may be included in interface circuits 226 or 224 for receiving signals into the semiconductor device. The first connection pad 112 connects to the first interface circuit 220, while the second connection pad 114 connects to the second interface circuit 222. The first interface circuit 220 includes first driver 226 and ESD circuitry 228. The second interface circuit

Petitioners
Ex. 1002, p. 94

includes second driver 224. The first and second drivers 226, 224, respectively, electrically connect to the single signal line 221 that is in turn connected to a data generator 232, which forms part of the core interface circuit 212. The first and second driver circuitry 226, 224, respectively, also electrically connects to modal select circuitry 234, which may or may not form part of the core interface circuit 212. The modal select circuitry enables either the first driver 226 or the second driver 224. Suitable modal select circuitry is described in relation to Figure 2E. In other embodiments, no modal select circuitry 234 is provided. In these embodiments, the desired circuitry or mode is enabled (*e.g.*, powered-up, connected to an enable signal, connected to an output pad, *etc.*) by virtue of either making or not making a physical electrical connection at the interface semiconductor device.

[0040] In some embodiments, such as the embodiment described above in relation to Figure 2B, the first interface circuit 220 has a first capacitance (C+), and the second interface circuit 222 has a second capacitance (C-) that is lower than the first capacitance. In some embodiments, these capacitances are representative of the aggregated capacitances of the active circuitry itself, the ESD circuitry 228, driver circuitry 224 or 226, the pad metal, and all electrical connections, such as traces *etc.*, that are coupled to each of the connection pads 112 and 114. In use, the modal select circuitry can enable either the signal path through the first interface circuit 220 and the first connection pad 112, or the signal path through the second interface circuit 222 and the second connection pad 114.

[0041] Figure 2C is a schematic circuit diagram of a portion of a stacked semiconductor die assembly using devices like the one shown in Figure 2B. By offsetting a semiconductor device, as described in relation to Figures 1A and 1B, and in some embodiments controlling the modal select circuitry 234, different semiconductor devices in the stacked semiconductor assembly operate in different modes. An interface semiconductor device 102(1) is offset (see Figure 1B) from the remainder of the device in the stacked semiconductor assembly, so that the connection pad 112 and the interface circuit 220 is enabled. The connection pad 112 on the interface device 102(1) is connected to the connection pads 114 on the remainder of the semiconductor devices 102(2), 102(3), and 102(4) in the stacked semiconductor die assembly. The connections between the connection pads 112 and 114 may be through vias and other connection pads as described above in relation to Figures 1A and 1B. As the connection pad 112 and the first interface circuit 220 is enabled on the interface semiconductor device 102(1) (and not the second connection pad 114 and second interface

Petitioners
Ex. 1002, p. 95

circuit 222), the interface semiconductor device 102(2) operates in a different mode to the remainder of the semiconductor devices in the stacked semiconductor assembly. For example, the interface semiconductor device has the driver connected to ESD circuitry enabled, while the remainders of the devices in the stacked semiconductor assembly do not. In embodiments where no modal select circuitry 234 is provided, the first interface circuit 220 is used if the signal path is through the first connection pad 112, and the second interface circuit 222 is used if the signal path is through the second connection pad 114.

[0042] There are many advantages of the above described stacked semiconductor assembly. For example, as the higher capacitance is disabled for all except the interface device 102(1), the resulting capacitive loading on the external interface is drastically reduced, and, therefore, the overall performance of the stack is increased, as compared to an embodiment having identical semiconductor devices all having high capacitance loading. The above described embodiment also eliminates having to downsize the ESD protection circuitry in each of the semiconductor devices in the stack so that the overall aggregated ESD protection of the stack is at a desired level. Downsizing the ESD protection in each semiconductor device restricts the stack count to a number or amount of semiconductor devices that have a collective ESD protection amount. For example, if each semiconductor device has 500V HBM ESD tolerance, where the desired ESD protection amount of the stack is 2000V HBM, then no less than four semiconductor devices can be used in the stack. Instead, in the above described embodiments, the ESD protection circuitry is only enabled in the interface semiconductor device 102(1), thereby eliminating: (i) having to downsize the ESD protection circuitry in each of the semiconductor devices, and (ii) any restriction on the number of semiconductor devices that can be used in any one stacked semiconductor assembly. It should also not be overlooked that the stacked semiconductor assemblies described herein all make use of substantially identical semiconductor devices. This drastically reduces the complexity and cost of manufacturing and assembling such stacked semiconductor assemblies.

[0043] Figure 2D is a schematic diagram of another stacked semiconductor device assembly 240 that includes a plurality of substantially identical semiconductor devices 236(1)-236(n), wherein the spatial offset technique is used to enable different modes of operation for the substantially identical semiconductor devices. Each of the semiconductor devices 236 includes a number of connection pads, including: a master-in connection pad (M

Petitioners
Ex. 1002, p. 96

In), a master enable connection pad (or mode enable or select connection pad, labeled "M/E"), a first master connection pad (M1), a second master connection pad (M2), a first slave connection pad (S1), and a second slave connection pad (S2). In some embodiments, the master-in connection pad (M in), master enable connection pad (M/E), first master connection pad (M1), and first slave connection pad (S1) are disposed at a first side of each semiconductor device, while the second master connection pad (M2) and the second slave connection pad (S2) are disposed at the second side of the semiconductor device. The first and the second sides are opposite to one another and may be substantially planar. The first master connection pad (M1) is connected to the second master connection pad (M2) through a via, such as a TSV. Similarly, the first slave connection pad (S2) is connected to the second slave connection pad (S2) through a via, such as a TSV.

[0044] In some embodiments, each semiconductor device 236 also includes circuitry, including: an I/O buffer 244 for receiving and buffering input signals received at the master-in connection pad (M In); a read and/or write channel master circuit 246; a read and/or write channel slave circuit 248; a core interface circuit 250; a core 252 (e.g., a DRAM memory core, a flash memory core, a processor array core, etc.); and/or a mode select circuit 234. In some embodiments, the mode select circuit is located at or near the first side of each semiconductor device.

[0045] In some embodiments, the read and/or write channel master circuit 246 is electrically connected to the I/O buffer 244; the mode select circuitry 234; the first master connection pad (M1), which is in turn connected to the second master connection pad (M2) through a TSV; and the read and/or write channel slave circuit 248. In these embodiments, the read and/or write channel slave circuit 248 is electrically connected to the read and/or write channel master circuit 246; the core interface circuit 250; the first slave connection pad (S1), which is in turn connected to the second slave connection pad (S2) through a TSV; and/or mode select circuitry 234. The core interface circuit 250 is connected to the core 252.

[0046] In use, the first semiconductor device 236(1) can be spatially offset from the remainder of the semiconductor devices in the stack, such that the second master connection pad (M2) of the first semiconductor device 236 (1) is connected to and aligned with the first slave connection pad (S1) of the second semiconductor device 236(2) in the stack. As with the embodiment described in Figures 1A and 1B, the second master connection pad (M2) is aligned with the first slave connection pad (S1) along an axis that is substantially

Petitioners
Ex. 1002, p. 97

perpendicular to the opposing two sides of the semiconductor device (*i.e.,* along an axis substantially parallel to the Z-axis and substantially perpendicular to the X and Y axes of Figure 5).

[0047] While any suitable mode enable or select circuitry 234 may be used, Figure 2E demonstrates how a resistor to ground can provide the functionality of one such suitable circuitry. The schematic 234 includes a multiplexer (Mux) connected to inputs associated with, for example, data path (0) and data path (1). In this embodiment, the first path (0) is the default mode, while the second mode (1) is the mode enabled if the M/E signal is electrically attached to VDD.

[0048] In Figure 2E, the select input of the Mux is connected to ground through a resistor. If the M/E signal pad is left unconnected, the resistor will "pull down" the select input to ground, thereby defaulting the Mux to the first mode (0). A similar technique can be used to create a global "Slave Enable" for the entire die. On the other hand, if a VDD connection is made to the master enable connection pad (M/E), the mode select signal will driven to VDD, thereby selecting the Mux to the second mode (1). A similar technique can be used to create a global "Master Enable" for the entire die. In other embodiments, separate mode select or enable circuitry may be provided for each of the read and/or write channel master circuit 246 (Figure 2D) and the read and/or write channel slave circuit 248 (Figure 2D).

[0049] Returning to Figure 2D, a VDD signal is attached to the master enable connection pad (M/E) of the first semiconductor device 236(1) only, thereby enabling the read and/or write channel master circuit 246. The master enable connection pad (M/E) of the remainder of the semiconductor devices 236(2)-236(n) is not connected, thereby, enabling the default read and/or write channel slave circuit 248 on the remainder of the semiconductor devices 236(2)-(n).

[0050] In use, write signals received on the master-in connection pad (M In) are received and pass through the I/O buffer 244 to the read and/or write channel master circuit 246. Typically, a read interface of master circuit 246 includes data receivers, while a write interface of master circuit 246 includes data transmitters. The signal then passes from the read and/or write channel master circuit 246 to both the first master connection pad (M1), through the TSV to the second master connection pad (M2) and to the first slave connection pad (S2) on the adjacent semiconductor device. The signal also passes from the read and/or

Petitioners
Ex. 1002, p. 98

write channel master circuit 246 to the read and/or write channel slave circuit 248 of the same device, whereby it is passed onto the core interface circuit 250 and the core 252.

[0051] Write signals received at the second and subsequent semiconductor devices travel from the first slave connection pad (S1) to the second slave connection pad (S2) through a TSV and onto subsequent semiconductor devices in the stack. The signal also travels from the first slave connection pad (S1) of the second and subsequent devices 236(2)-(n) to the read and/or write channel slave circuit 248. The read and/or write channel slave circuit 248 may include read and/or write interfaces (not shown). The read interface of slave circuit 248 includes data transmitters, while the write interface of slave circuit 248 includes data receivers. In one embodiment, duplicate read and/write interfaces or duplicate transmitters and/or receivers may be provided in slave circuit 248 to communicate data with the master circuit 246 on the same device as the slave circuit 248 and the master circuit on another device, respectively. As shown in Figure 2D, slave circuit 248 may also use the signal generated by select circuitry 234 to determine which of its receivers and/or transmitters need to be activated (*i.e.*, whether it's interfacing with a master interface circuit 246 on the same die as itself or on another die in the stack). Finally, the signal travels from the read and/or write channel slave circuit 248 to the core interface circuit 250 and the core 252. Accordingly, by spatially offsetting the first of the substantially identical semiconductor devices in the stack, the first semiconductor device is master enabled, while the remainder of the devices in the stack are slave enabled. Additionally, transmitters in the stack's slave circuit 248 communicate to receivers in the bottom die's master circuit 246, while transmitters in the bottom die's master circuit 246 communicate to receivers in the stack's slave circuit 248.

[0052] Read signals should travel in opposite direction as the write signals. For example, a read signal originated from device 236(4) would travel from a read channel slave circuit 248 to the S1 pad of device 236(4), through the TSV's between pads S2 and S1 in devices 236(3) and 236(2), and through the TSV between pads M2 and M1 in device 236(1), to a read channel master circuit 246, and output by I/O buffer 244 through the "M In" pad of device 236(1).

[0053] Figure 3 is a schematic cross-sectional side view (cross-hatching removed for clarity) of another exemplary stacked semiconductor device assembly 300. As shown, the second group of connection pads of the interface semiconductor device (bottommost first

Petitioners
Ex. 1002, p. 99

semiconductor device) 302 is electrically connected (such as through "solder bumps" or "C4 bumps") to the substrate 312 of the stacked semiconductor device assembly 300. The fourth group of connection pads of the first semiconductor device 302 is then electrically coupled to the first group of connection pads of the second semiconductor device 304. The third group of connection pads of the second semiconductor device 304 is connected to the first group of connection pads of the third semiconductor device 306; the third group of connection pads of the third semiconductor device 306 is connected to the first group of connection pads of the fourth semiconductor device 308; and the third group of connection pads of the fourth semiconductor device 308 is connected to the first group of connection pads of the fifth semiconductor device 310. Overmolding 316 is them provided over the entire stack of semiconductor devices to insulate and protect the semiconductor devices. The stacked semiconductor device assembly may connect to other electronic circuitry via any suitable electrical connections, such as micro-BGA balls 314.

[0054] The stacked semiconductor device assembly allows any number of identical semiconductor devices to be stacked one on top of the other. Offsetting the bottom semiconductor device from the remainder of the semiconductor devices in the stack allows only the interface (or bottom) semiconductor device to perform functional or electrical operations differently as compared to the other semiconductor devices in the stack. For example, as shown in Figures 2B and 2C, only the connection pad on the interface semiconductor device is connected to ESD circuitry, where the ESD circuitry on the remainder of the semiconductor devices is not used or needed. In another example, as shown in Figure 2D, the bottom die may act as a "stack Master", buffering IO data to multiple "stack Slaves" which are substantially identical but have disabled master IO portions.

[0055] As all of the semiconductor devices in the stack are connected to one another in series, in some embodiments, it is desirable to select a particular semiconductor device for receiving the transmitted signal. Selecting the appropriate semiconductor device may be achieved using any suitable a chip select mechanism. One such suitable chip select mechanism is described in co-pending U.S. Patent Application No. 11/402,393, which is hereby incorporated by reference in its entirety.

[0056] Figure 4 is a schematic cross-sectional side view (cross-hatching removed for clarity) of another exemplary stacked semiconductor device assembly 400. This embodiment includes a stacked semiconductor device assembly connected to a circuit board 408. The

Petitioners
Ex. 1002, p. 100

stacked semiconductor device assembly 400 includes multiple stacks of semiconductor devices 402 and 404. One of the stacks of semiconductor devices 402 may include semiconductor devices that are offset in two opposing directions. The stacked semiconductor device assembly 400 may also include other circuitry 422 that is not stacked.

[0057] Figure 5 is a plan view of another stacked semiconductor device assembly 500. Unlike the stacked semiconductor device assembly 100 shown in Figure 1A, this stacked semiconductor device assembly 400 includes one or more semiconductor devices that are offset from one another in more than one dimension. For example, the uppermost semiconductor device 102(2) is offset from the bottommost semiconductor device 102(1) along both of the orthogonal axes Y and X.

[0058] Figure 6 is a schematic of a computer system using a stacked memory semiconductor device assembly 626. The system 600 includes a plurality of components, such as at least one central processing unit (CPU) 602; a power source 606, such as a power transformer, power supply, or batteries; input and/or output devices, such as a keyboard and mouse 610 and a monitor 608; communication circuitry 612; a BIOS 620; a level two (L2) cache 622; Mass Storage (MS) 624, such as a hard-drive; Random Access Memory (RAM) 626; and at least one bus 614 that connects the aforementioned components. These components are at least partially housed within a housing 616. Some components may be consolidated together, such as the L2 cache 622 and the CPU 602. The RAM 626 includes one or more stacked semiconductor device assemblies 300 described above.

[0059] While the foregoing description and drawings represent the preferred embodiments of the present invention, it will be understood that various additions, modifications and substitutions may be made therein without departing from the spirit and scope of the present invention as defined in the accompanying claims. In particular, it will be clear to those skilled in the art that the present invention may be embodied in other specific forms, structures, arrangements, proportions, and with other elements, materials, and components, without departing from the spirit or essential characteristics thereof. The presently disclosed embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims, and not limited to the foregoing description.

Petitioners
Ex. 1002, p. 101

What is claimed is:

1. A stacked semiconductor device assembly comprising:

first and second substantially identical integrated circuit (IC) devices, each of the first and second IC devices including:

first interface circuits each having a first capacitance;

second interface circuits each having a second capacitance that is substantially lower than the first capacitance;

a first set of through vias coupled to respective ones of the first interface circuits; and

a second set of through vias coupled to respective ones of the second interface circuits; and

wherein the first and second IC devices are stacked one on top of another and offset from each other such that the first set of through vias of the first IC device are aligned with and electrically connected to respective ones of the second set of through vias of the second IC device.

2. The stacked semiconductor device assembly of claim 1, wherein each of the first and second IC devices comprises core circuitry, wherein the stacked semiconductor device assembly is configured to communicate data between the core circuitry and the first interface circuits on the first IC device and to communicate data between the core circuitry and the second interface circuits on the second IC device.

3. The stacked semiconductor device assembly of claim 2, wherein the stacked semiconductor device assembly is further configured to communicate data between the core circuitry of the second IC device and the first interface circuits of the first IC device.

4. The stacked semiconductor device assembly of claim 1, wherein each of the first interface circuits on a respective IC device includes an I/O pad connected to electrostatic discharge circuitry on the respective IC device.

Petitioners
Ex. 1002, p. 102

5. The stacked semiconductor device assembly of claim 1, wherein each of the second interface circuits on a respective IC device includes an I/O pad that is not connected to electrostatic discharge circuitry on the respective IC device.

6. The stacked semiconductor device assembly of claim 1, wherein the stacked semiconductor device assembly is a memory module.

7. The stacked semiconductor device assembly of claim 1, wherein the first interface circuits are a read channel master and the second interface circuits are a read channel slave.

8. The stacked semiconductor device assembly of claim 1, wherein the first interface circuits are a write channel master and the second interface circuits are a write channel slave.

9. The stacked semiconductor device assembly of claim 1, wherein each IC device is configured to communicate data using either the first interface circuits or the second interface circuits depending on whether the first and third pads or the second and fourth pads are coupled to an external signal line.

10. The stacked semiconductor device assembly of claim 1, wherein each semiconductor device of the semiconductor devices is configured to operate in a different mode depending on whether the first set of through vias or the second set of through vias are coupled to an external signal line.

11. The stacked semiconductor device assembly of claim 1, wherein the stacked semiconductor device assembly is a DRAM module.

12. The stacked semiconductor device assembly of claim 1, wherein the stacked semiconductor device assembly includes multiple distinct groups of stacked IC devices.

13. The stacked semiconductor device assembly of claim 1, wherein at least one of the IC device is offset from at least another one of the multiple IC devices along two axes that are

Petitioners
Ex. 1002, p. 103

perpendicular to one another and that are substantially parallel with substantially planer sides of the IC devices.

14. A integrated circuit (IC) device, comprising:

I/O buffer circuits;

read/write channel master circuits coupled to the I/O buffer circuits;

a first set of through vias coupled to respective ones of the read/write channel master circuits;

read/write channel slave circuits;

a second set of through vias coupled to respective ones of the read/write channel slave circuits;

mode select circuitry to set the IC device as either a slave device that communicates data using the read/write channel slave circuits and the second set of through vias or as a master device that communicates data using the I/O buffer and the read/write channel master circuits.

15. The IC device of claim 14 in a stacked semiconductor device assembly including the IC device and at least one other IC device, wherein the IC device and the one other IC device are stacked one on top of another and offset from each other such that the first set of through vias of the IC device are aligned with and electrically connected to respective ones of the second set of through vias of the other IC device, wherein the IC device is configured to function as a master device and the other IC device is configured to function as a slave device.

16. A stacked semiconductor device assembly comprising:

first and second substantially identical semiconductor devices, each of the first and second semiconductor devices including:

a first interface circuit having a first capacitance;

a second interface circuit having a second capacitance that is substantially lower than the first capacitance;

a first set of through vias coupled to the first interface circuit; and

Petitioners
Ex. 1002, p. 104

a second set of through vias coupled to the second interface circuit; and

wherein the first and second semiconductor devices are stacked one on top of another and offset from each other such that the first set of through vias of the first semiconductor device are aligned with and electrically connected to respective ones of the second set of through vias of the second semiconductor device.

17. The stacked semiconductor device assembly of claim 16, wherein each of the first and second semiconductor devices comprises core circuitry, wherein the stacked semiconductor device assembly is configured to communicate data between the core circuitry and the first interface circuit on the first semiconductor device and to communicate data between the core circuitry and the second interface circuit on the second semiconductor device.

18. The stacked semiconductor device assembly of claim 17, wherein the stacked semiconductor device assembly is further configured to communicate data between the core circuitry of the second semiconductor device and the first interface circuit of the first semiconductor device.

19. The stacked semiconductor device assembly of claim 16, wherein each of the first interface circuit on a respective semiconductor device includes an I/O pad connected to electrostatic discharge circuitry on the respective semiconductor device.

20. The stacked semiconductor device assembly of claim 16, wherein the second interface circuit on a respective semiconductor device includes an I/O pad that is not connected to electrostatic discharge circuitry on the respective semiconductor device.

Petitioners
Ex. 1002, p. 105

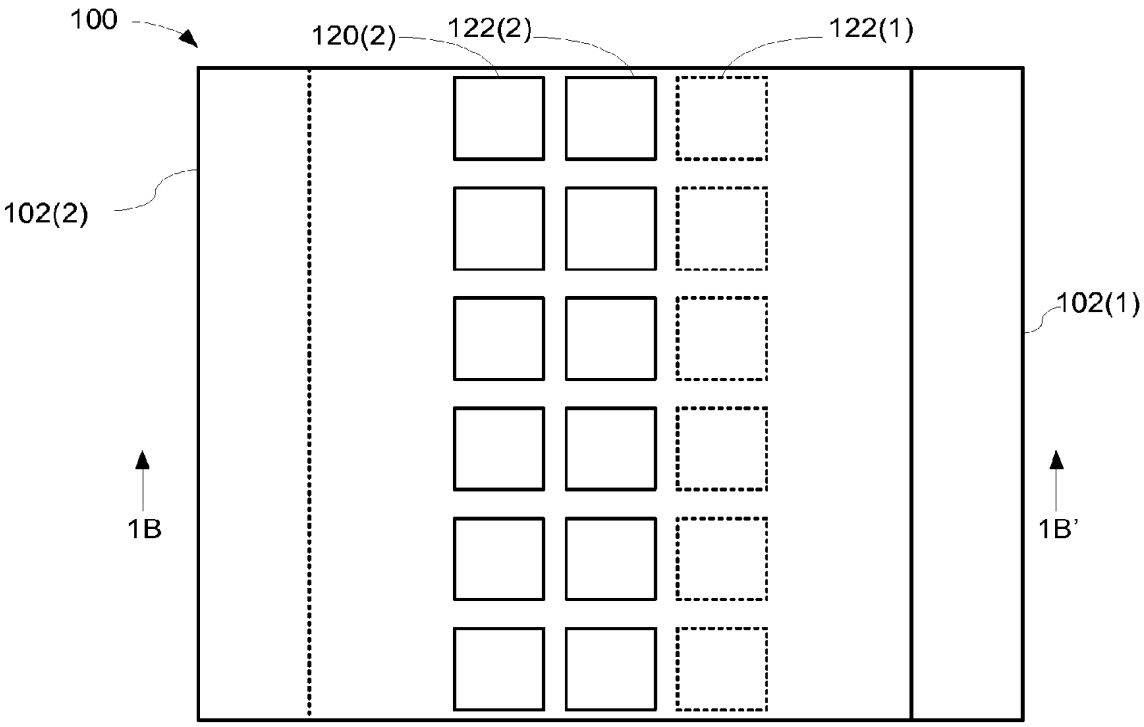

100
120(2)
122(2)
122(1)
102(2)
102(1)
1B
1B'

FIG. 1A

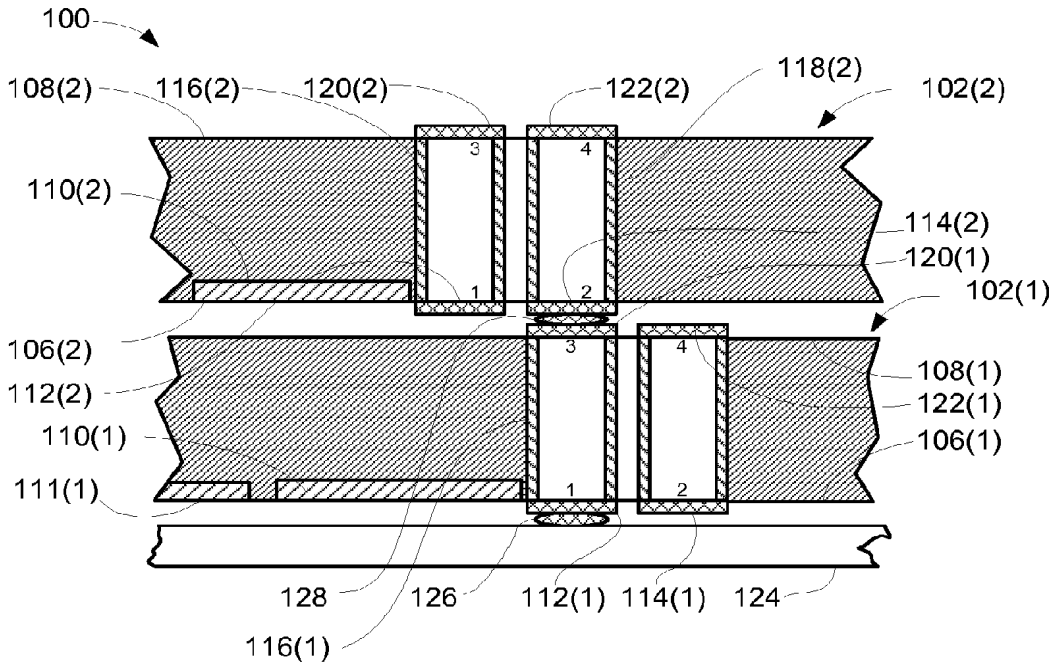

100
108(2)
116(2)
120(2)
122(2)
118(2)
102(2)
110(2)
114(2)
120(1)
102(1)
106(2)
112(2)
110(1)
111(1)
108(1)
122(1)
106(1)
128
126
112(1)
114(1)
124
116(1)
1
2
3
4

FIG. 1B

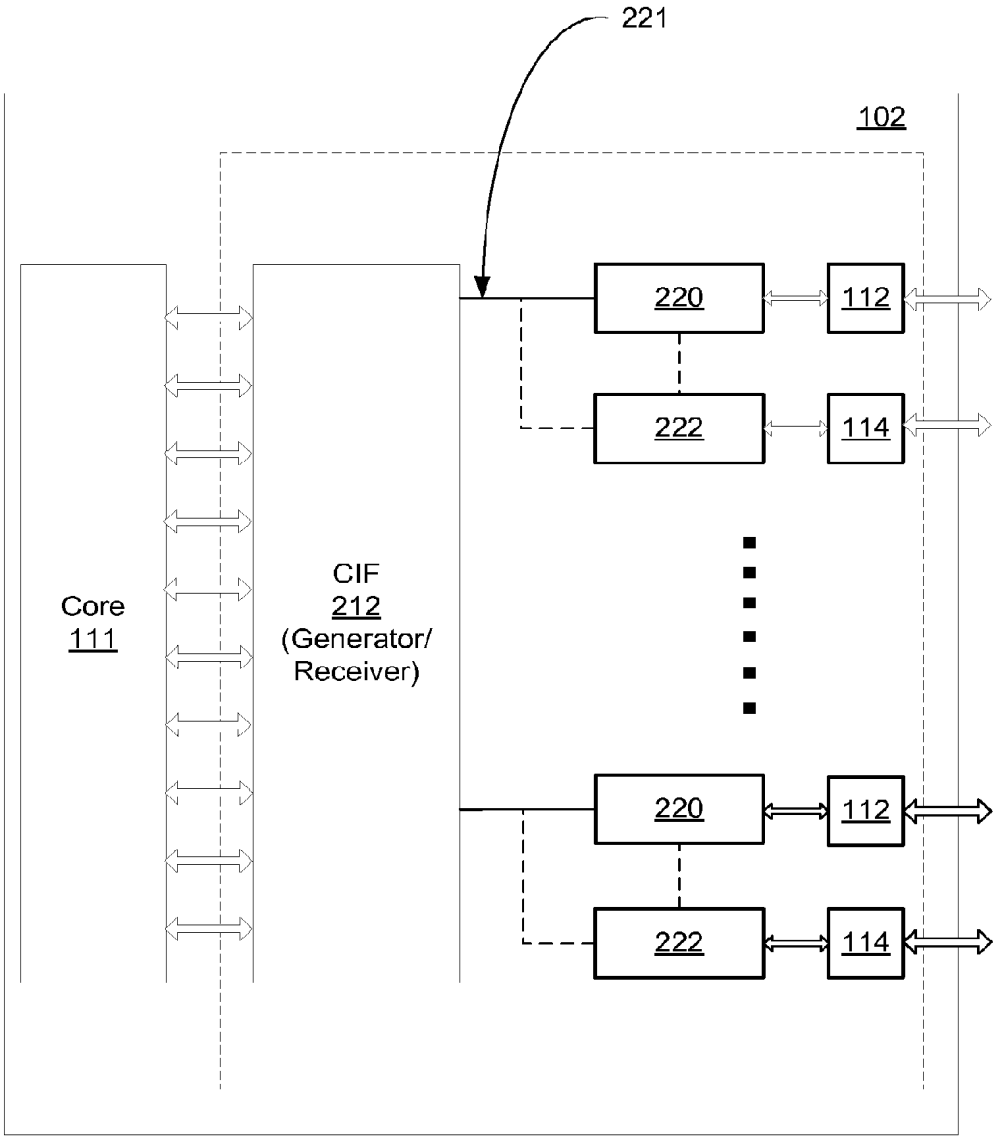

221
102
220
112
222
114
Core
111
CIF
212
(Generator/
Receiver)
220
112
222
114

FIG. 2A

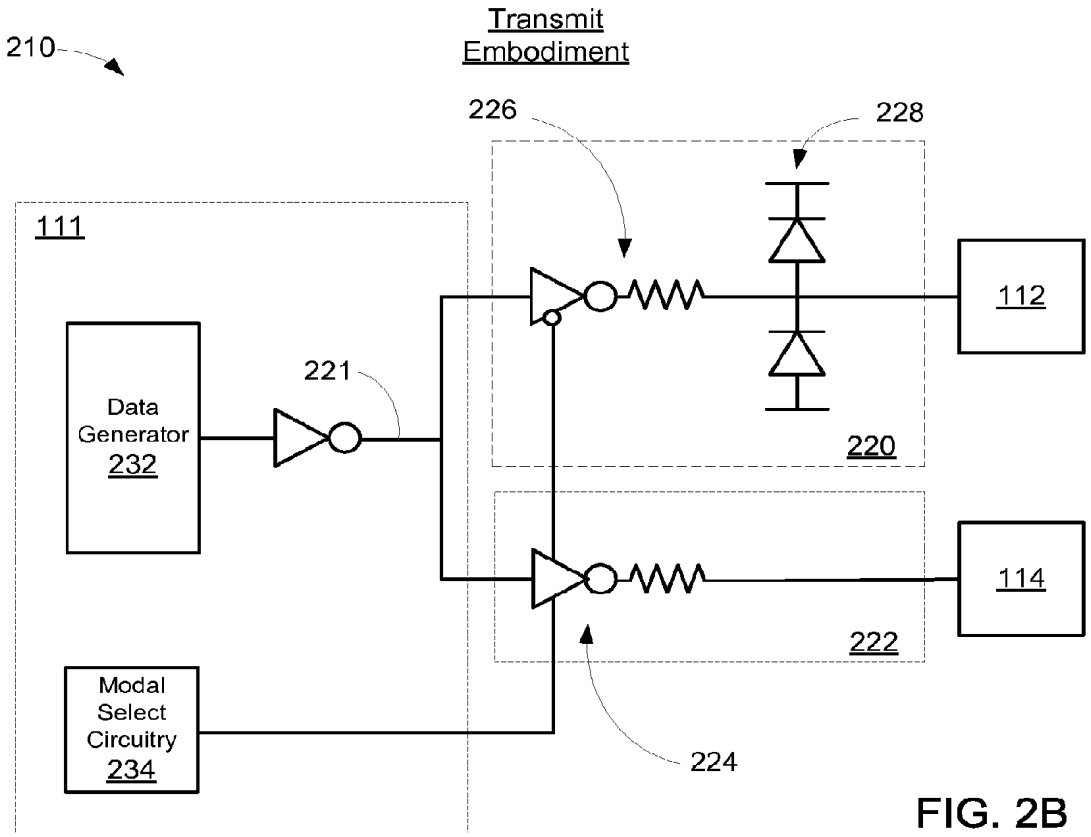

210
Transmit
Embodiment
226
228
111
112
Data
Generator
232
221
220
114
222
Modal
Select
Circuitry
234
224

FIG. 2B

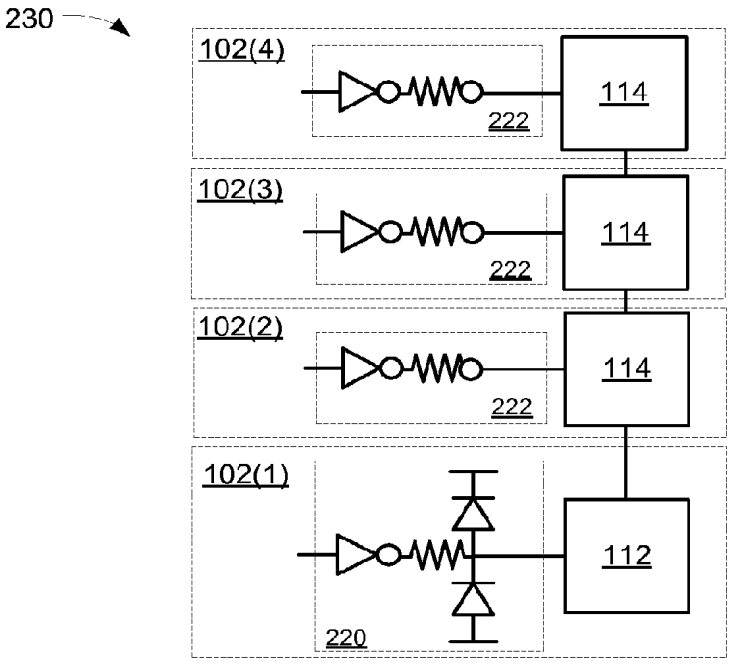

230
102(4)
114
222
102(3)
114
222
102(2)
114
222
102(1)
112
220

FIG. 2C

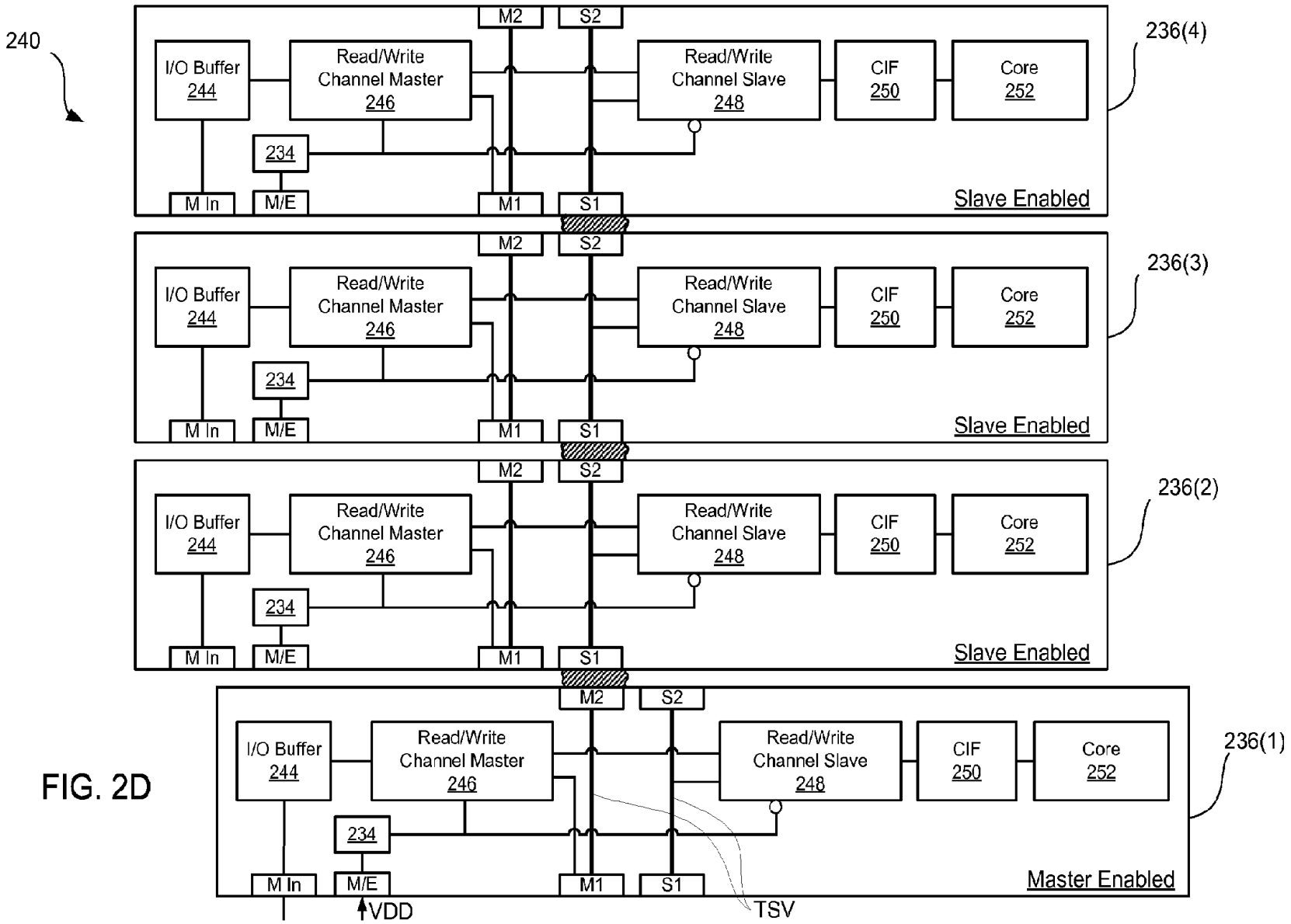

240
236(4)
236(3)
236(2)
236(1)
I/O Buffer 244
Read/Write Channel Master 246
Read/Write Channel Slave 248
CIF 250
Core 252
234
M In
M/E
M1
M2
S1
S2
Slave Enabled
Master Enabled
VDD
TSV

FIG. 2D

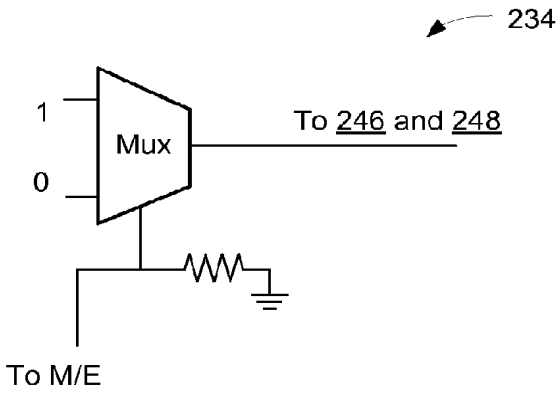


FIG. 2E

Petitioners
Ex. 1002, p. 110

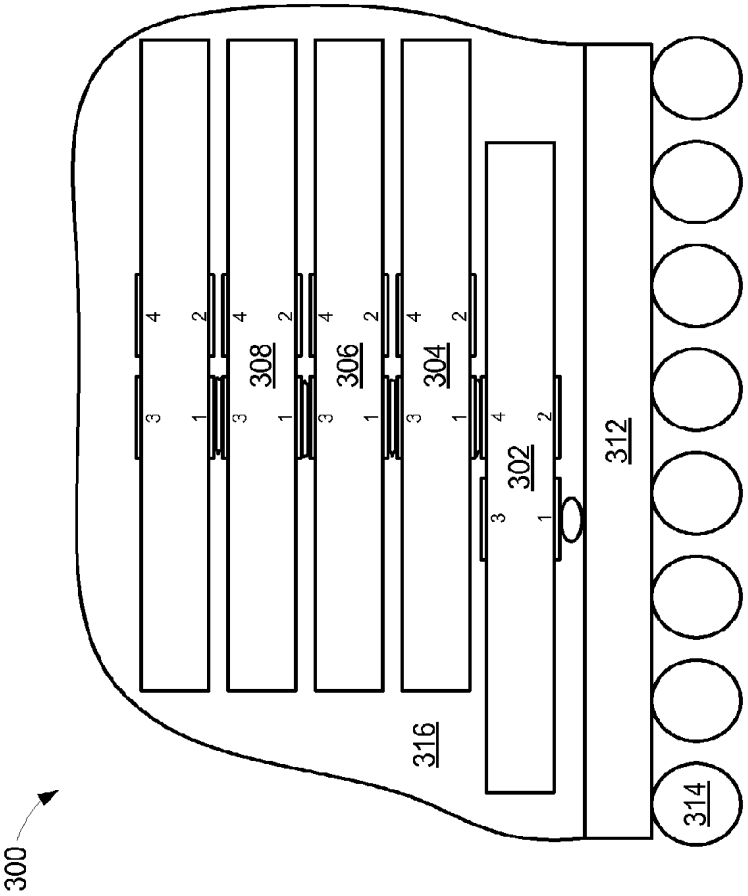


FIG. 3

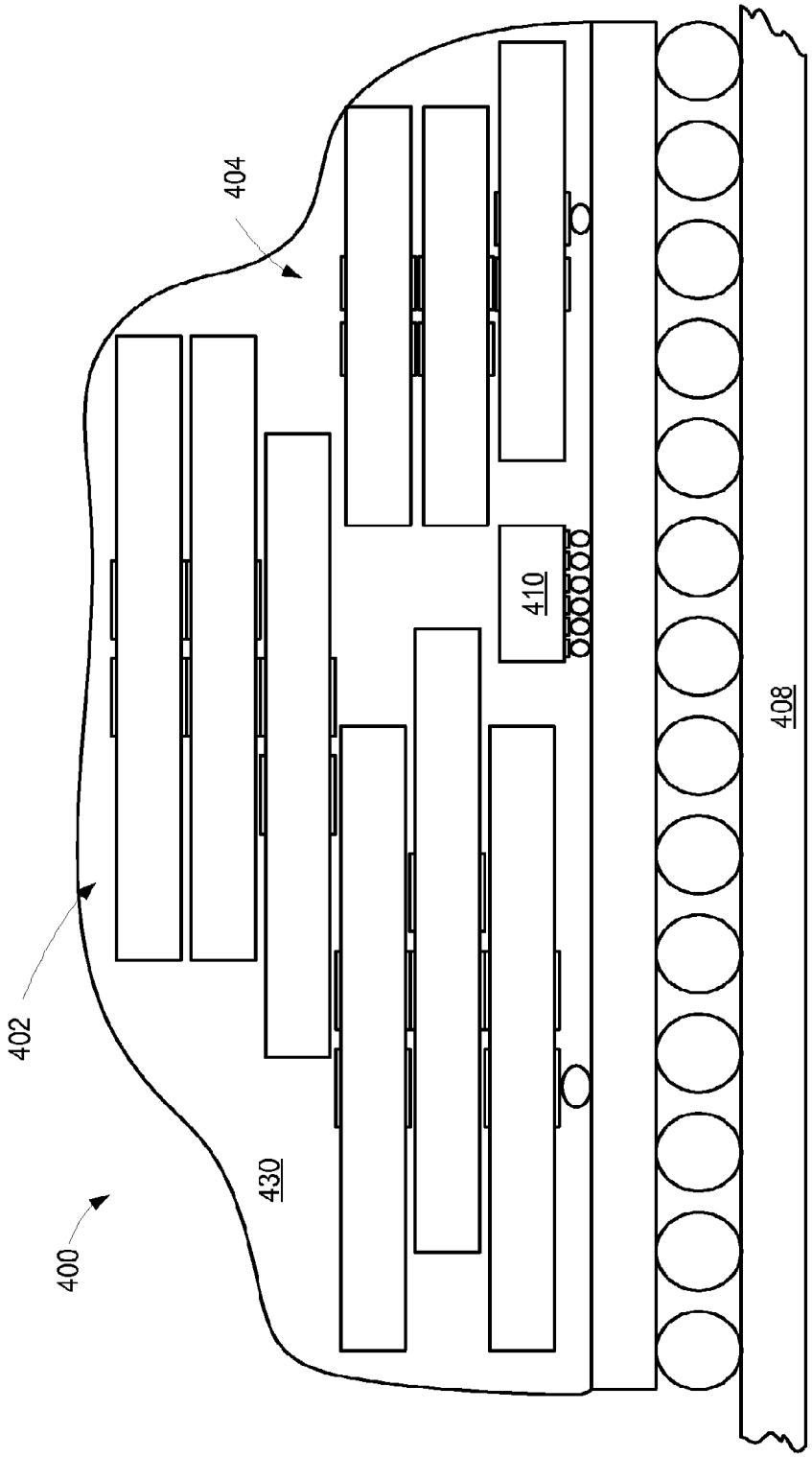


FIG. 4

Petitioners
Ex. 1002, p. 112

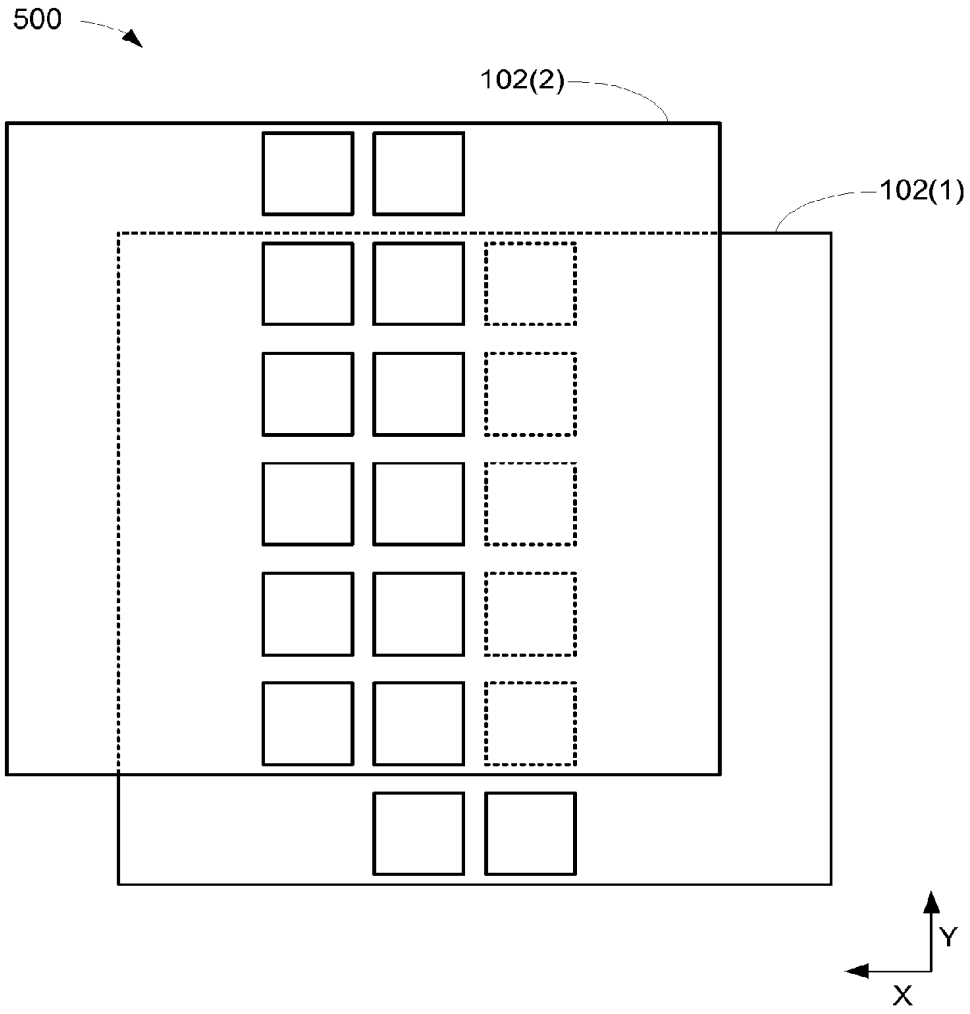


FIG. 5

Petitioners
Ex. 1002, p. 113

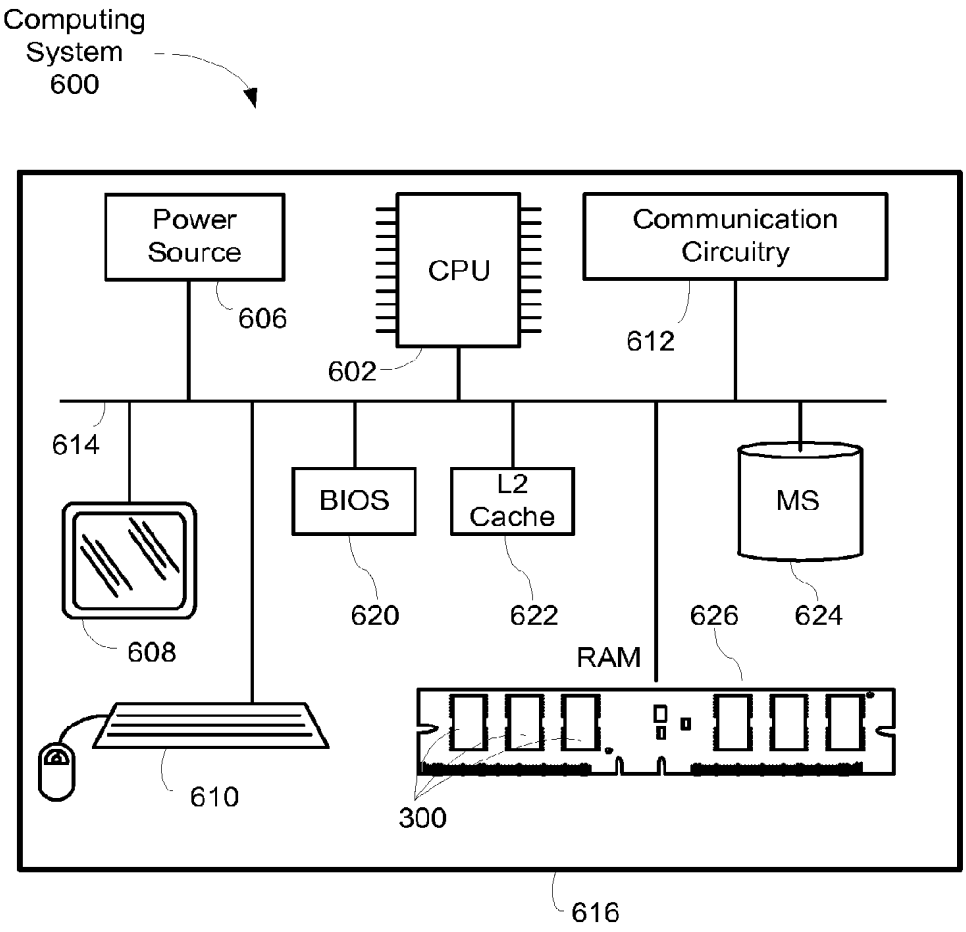

Computing System 600
Power Source
606
CPU
602
Communication Circuitry
612
614
BIOS
L2 Cache
MS
620
622
626
624
608
RAM
610
300
616

FIG. 6

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 April 2011 (28.04.2011)

PCT

(10) International Publication Number
WO 2011/049710 A2

(51) International Patent Classification:
H01L 23/48 (2006.01) *H01L 23/055* (2006.01)
H01L 21/60 (2006.01) *H01L 23/045* (2006.01)

(21) International Application Number:
PCT/US2010/050273

(22) International Filing Date:
24 September 2010 (24.09.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/254,519 23 October 2009 (23.10.2009) US

(71) Applicant *(for all designated States except US)*: **RAMBUS INC.** [US/US]; 4440 El Camino Real, Los Altos, CA 94022 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **WARE, Frederick, A.** [US/US]; 13961 Fremont Pines Lane, Los Altos Hills, CA 94022 (US).

(74) Agent: **KREISMAN, Lance, M.**; Mahamedi Paradice Kreisman LLP, 550 South Winchester Blvd., Suite 605, San Jose, CA 95128 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

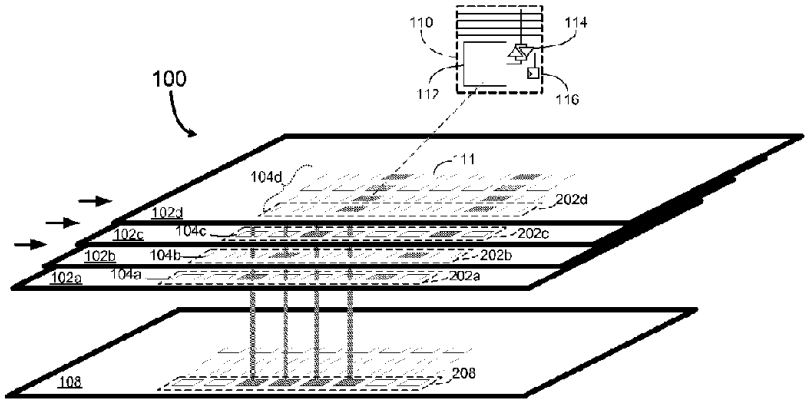
— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: STACKED SEMICONDUCTOR DEVICE

FIG. 1



(57) Abstract: A stacked semiconductor device is disclosed that includes a plurality of semiconductor dies. Each die has oppositely disposed first and second surfaces, with pads formed on each of the surfaces. A plurality of through- vias connect respective pads on the first surface to respective pads on the second surface. The through- vias include a first group of through- vias coupled to respective I/O circuitry on the semiconductor die and a second group of through- vias not coupled to I/O circuitry on the semiconductor die. The plurality of semiconductor dies are stacked such that the first group of through- vias in a first one of the plurality of semiconductor dies are aligned with respective ones of at least a portion of the second group of through- vias in a second one of the plurality of semiconductor dies.

WO 2011/049710 A2

STACKED SEMICONDUCTOR DEVICE

TECHNICAL FIELD

[0001] The disclosure herein relates to semiconductor devices and more particularly to stacking configurations and methods for semiconductor devices.

BACKGROUND

[0002] Semiconductor packaging often involves assembling and electrically interconnecting groups of semiconductor devices together in space-efficient, yet inexpensive ways. One packaging approach involves stacking devices to reduce the horizontal footprint of the assembled system. Conventional stacking schemes typically rely on signal redistribution layers and/or wire-bond resources to provide electrical access to and from each device. Moreover, conventional stacking arrangements often overlook power efficiency concerns.

[0003] While conventional stacking methods work well for their intended applications, the need exists for power and cost efficient stacked die arrangements and methods. Embodiments of packaged semiconductor systems and methods described herein satisfy these needs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] Embodiments of the disclosure are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

Figure 1 illustrates a three-dimensional exploded view of a packaged semiconductor system employing stacked devices in accordance with one embodiment of the disclosure herein;

Figure 2 illustrates a two-dimensional partial side view of the stacked device system of Figure 1 with the pad and I/O circuitry for each device shown as if from a top plan view;

Petitioners
Ex. 1002, p. 116

Figure 3 illustrates a partial cross-sectional view of two semiconductor devices showing formed vias employed in the system of Figure 1;

Figure 4 illustrates a generalized top plan view of an alignment technique for offsetting the stacked semiconductor devices of Figure 1;

Figure 5 illustrates a three-dimensional exploded view similar to Figure 1 for an alternative embodiment;

Figure 6 shows a two-dimensional partial side view similar to Figure 2 for the alternative embodiment of Figure 5;

Figure 7 illustrates the routing of a serial bus used in one embodiment of the stacked semiconductor system of Figure 5 to initialize one or more of the stacked semiconductor devices; and

Figure 8 shows one embodiment of a memory system employing point-to-point and multi-drop interconnection paths.

DETAILED DESCRIPTION

[0005] Embodiments of systems and methods are described herein where a plurality of semiconductor devices employing through-silicon-vias are stacked to allow access of at least one of the semiconductor devices by way of through-silicon-vias in at least one other of the semiconductor devices. The through-silicon-vias are designed and device-aligned to reduce input/output (I/O) loading along a given via path. By reducing I/O loading on via paths, the system enables optimal point-to-point connections between I/O circuitry in a base device, and I/O circuitry in one of the plurality of devices. This, in turn, improves system speed and power efficiency. In one embodiment, the through-silicon-vias in each of the plurality of semiconductor devices are arranged in an array. Each semiconductor device has two opposing surfaces and a corresponding pad array on each of the opposing surfaces. The reduced I/O loading is achieved by staggering I/O circuits among a pad array and offsetting adjacent devices by at least one pad

Petitioners
Ex. 1002, p. 117

location. In other words, the two adjacent devices are offset from each other by at least the distance between two adjacent columns of pads such that the first column of the through-silicon-via array in a first one of the adjacent devices are aligned with respective ones of the second column of the through-silicon-via array in the other one of the adjacent devices, the second column of the through-silicon-via array in the first one of the adjacent devices are aligned with respective ones of the third column of the through-silicon-via array in the other one of the adjacent devices, etc. The cooperation between the staggered I/O circuits and the offsetting between adjacent devices minimizes the number of I/O circuits coupled to each via path, resulting in reduced I/O loading thereof. Alternatively, the alignment is accomplished through a straight-up stack without offsetting. In such an embodiment, the I/O circuitry is not staggered, but rather selectively enabled to reduce the I/O loadings.

[0006] Figures 1-3 illustrate respective three-dimensional and two-dimensional views for one specific embodiment of a semiconductor die stack 100 that facilitates point-to-point interconnections while reducing power dissipation and cost. A plurality of semiconductor dies 102a – 102d are provided for stacking. Each of the semiconductor dies 102a – 102d has oppositely disposed first and second surfaces 304 and 306, a plurality of pads 308 formed on the first surface, a plurality of pads 308 formed on the second surface, and a plurality of through-vias 302 connecting respective pads on the first surface to respective pads on the second surface. The pads and vias are arranged in respective arrays, such as the array of pads 104a – 104d on the first surfaces. The arrays of pads 104a – 104d include I/O pads 112 (in magnified callout 110), which are pads coupled to corresponding I/O circuitry on the same die, and non-I/O pads, which are pads 111 not coupled to any I/O circuits on the same die. Accordingly, the array of through-vias in each semiconductor die would include a first group of through-vias coupled to respective I/O circuitry on the same semiconductor die and a second group of through-vias not coupled to I/O circuitry on the same semiconductor die.

Petitioners
Ex. 1002, p. 118

[0007] With continued reference to Figures 1-3, a processor die or device 108 preferably forms the base of the stack 100, but may be remotely disposed therefrom as the application dictates. The processor device 108 accesses a die higher in the stack, such as die 102c, by way of some of the through-vias in the same die and some of the through-vias in the dies lower in the stack, such as dies 102a and 102b, i.e., these vias form via paths for accessing the die 102c. The processor device 108 optionally employs a via array (shown in phantom in Figure 3) that may or may not correspond to the stacked die via arrays depending on whether the active circuitry on the processor device is facing the stack.

[0008] Further referring to Figure 1, and more specifically the magnified callout, each of the I/O pads 110 may take one of several forms. Generally, each pad represents the area occupied by the via structure (the via post diameter in addition to the surrounding metal pad) that couples to a corresponding via structure on an adjacent silicon die. In one embodiment, each I/O pads 112 electrically couples to transmit and/or receiver circuits such as a driver/receiver pair (or transceiver) 114, and possibly electro-static-discharge (ESD) circuitry (not shown) associated with the transmit and/or receive circuits, on the same die. An optional programmable mode register 116 provides a selectable enable/disable state for the driver/receiver circuitry. The capacitive coupling of this I/O circuit type may have the ability to load a via path with capacitance on the order of hundreds of femptofarads.

[0009] In an effort to minimize the effects of I/O pad capacitance on via paths, and referring still to Figure 1, the non-I/O pads 111 are employed. In one embodiment, the non-I/O pads are pads not having any associated I/O circuitry (e.g., driver/transceiver pair and ESD circuitry). In one embodiment, illustrated more clearly in Figure 2, only one in every four pad locations (such as at 210a) in a given row of pads 212 (in phantom) is formed with I/O circuitry, defining a staggered I/O circuit pattern. This allows for point-to-point connections between processor die 208 and up to four semiconductor dies 202a – 202d stacked above the processor die 208. For ease of reference in the discussion that follows, the shaded I/O pad location for an I/O pad on

Petitioners
Ex. 1002, p. 119

each stacked device of Figure 2 will be referred to as "Location1", with subsequent adjacent locations for non-I/O pads referred to in order as Locations 2 through 4.

[0010] In one embodiment, the point-to-point connectivity enables sizing of the I/O transceivers on the processor device for driving/receiving signals to/from one load (as opposed to multiple loads for a multi-drop configuration). The sizing may be in terms of current or voltage driving capability or the like. This, in turn, provides significant power savings, especially as the number of I/Os scale upwards for a given device stack.

[0011] With continued reference to Figure 2, from an overall stack perspective, representations of end-to-end vias from a cross-sectional view are shown as vertically interconnected paths 214, 216, 218, and 220. From an individual device perspective, the via array generally reflects the array shape, size, and pitch exhibited by the pad array.

[0012] Referring now to Figure 3, each TSV 300 is formed to provide a conductive path 302 oriented orthogonally between the active surface of the die 304 and the backside surface of the die 306. A raised pad structure 308 caps each via end. Pads disposed on the active surface of the die are generally formed coextensive with, and may actually define the I/O contact pads. Thus, for a given die, depending on the regular "stagger" of the I/O circuit pattern described above, a first set or group of vias will be coupled to I/O pads, while at least a portion of the rest (a second group) will be coupled to non-I/O pads on that die (no active circuitry). Note that the base die 108 in a given stack may optionally employ the via array 314 (in phantom) since its active circuitry lies on the top surface, at 312. In one embodiment, the plurality of semiconductor dies are stacked such that the first group of through-vias in a first one of the plurality of semiconductor dies (e.g., die 102a) are aligned with respective ones of at least a portion of the second group of through-vias in a second one of the plurality of semiconductor dies (i.e., die 102b).

Petitioners
Ex. 1002, p. 120

[0013] Preferably, the semiconductor dies are identical in terms of structural form, but may differ somewhat in performance. Employing identical devices in the stack simplifies die stacking, alignment, and reduces the cost per unit through the benefits of volume manufacturing.

[0014] Referring back to Figures 1 and 2, in one embodiment, the plurality of semiconductor dies are substantially identical and are offset from each other to achieve alignment of the first group of through-vias in a first one of the plurality of semiconductor dies with respective ones of at least a portion of the second group of through-vias in a second one of the plurality of semiconductor dies. For example, the stacked semiconductor dies can be aligned in such a way that the TSV arrays of adjacent dies are offset by at least one pad location. The die offsets create a "stairstep" arrangement for the stack, beginning at one edge of the pad array, and progressing by at least one pad location for each subsequently stacked device. The offsets allow the vias to electrically cooperate with the I/O circuit pattern formations on each die (preferably structurally identical). This is done such that each via path 214, 216, 218, and 220 (connected at one end to processor die I/O circuits 222, 224, 226, and 228) couples to only one I/O circuit ("location 1") in only one of the stacked dies.

[0015] With continued reference to Figure 2, the couplings from the base die I/O circuits 222, 224, 226, and 228 to each stacked die "Location1" are illustrated at 210a, 210b, 210c, and 210d. For the specific instance where four dies are to be stacked, and one of every four pad locations (in a row of pads) for each die include I/O circuitry, a single pad offset in a staircase manner results in an alignment where groups of vias not coupled to I/O circuits on their own respective die (pad locations 2, 3, and 4 for each set of four I/O locations) make electrical contact to I/O circuitry in different devices. Conversely, the alignment enables vias coupled to the I/O circuitry for pad Location 1 (for the lower-most stacked device 202a) to interface with groups of vias not coupled to I/O circuitry (pad locations 4, 3 and 2 of stacked devices 202b, 202c and 202d, respectively). In this specific manner, each via path forms a true point-to-point connection

Petitioners
Ex. 1002, p. 121

between the processor I/O circuit and one I/O circuit formed in one of the stacked semiconductor devices.

[0016] In some instances, there may be a desire to add capacity and stack additional die over and above the originally designed stack. Each additional die will preferably be formed with the appropriate TSV array, but may lack the staggered I/O pad array described above to cooperate properly with the TSV aligned offset scheme. Instead, each I/O location on the additional device may employ conventional pad I/O circuitry, including a programmable register to enable the driver/receiver circuitry. For enabled I/O circuitry, the result on the corresponding electrically coupled end-to-end via path will be a multi-drop situation. For non-enabled I/O circuits on the additional device, a capacitive loading will exist on the corresponding via (even though the active circuitry is "off"). While the added device presents a slight overall capacitive load to each via, the overall power benefit achieved by the original stack more than makes up for the loading.

[0017] In some applications, both point-to-point and multi-drop paths may need to be realized in the stacked die. Multi-drop paths are often employed in memory systems to transfer common signals such as address, command and control signals from a memory controller to memory dies. This may be done in a variety of different ways. For example, a group of adjacent I/O circuits on the processor (for example, pad locations 1 – 4 of the processor die) may be tied or grouped together electrically to form a super-pad such that a commonly transmitted signal is transferred to the same I/O circuit location on all of the stacked die. Another approach involves populating a stretch of via sites with I/O circuitry similar to Figure 6, although modified such that four I/O sites are coupled from the processor 604 to sixteen I/O sites in semiconductor devices 602a-602d. Generally, the approach unscrambles the four received signals (along each row of four I/O sites) from the processor in each of the four semiconductor devices 602a-602d by rotating the four bits by 0, 1, 2, or 3 bit positions, depending on the device location in the stack.

[0018] Manufacturing the stacked semiconductor devices in accordance with the embodiment described above involves relatively straightforward fabrication methods including

Petitioners
Ex. 1002, p. 122

staggering the I/O pad pattern (with the I/O circuitry), fabricating the TSV array, and aligning the devices in a stacked relationship.

Formation of the I/O circuitry and pad array for each device may be accomplished during wafer fabrication, involving generally understood principles of integrated circuit device processing. Of interest is the notion that for each device, pad locations are preferably reserved for I/O circuitry, even if omitted. This runs counter to the trend of electrically utilizing every possible piece of silicon real estate. However, this cost is more than offset by the advantages in power savings that result. By designing the device such that the overall chip layout is identical among devices, fabrication costs due to stacked-die-specific applications may be spread out or amortized due to volume manufacturing.

[0019] Fabricating the TSV array for each device may occur at the wafer (device) level, or stacked die level. Referring again to Figure 3, the two stacked die 102a and 108 are illustrated that each exhibit a pair of TSVs 300. Openings 301 to initially define the via shapes may be formed at the device level through a plasma or chemical etching or mechanical drilling process. At the stacked level, the via arrays may be efficiently formed by a suitable drilling method where openings are formed through all of the stacked devices in a single drilling step. Following formation of the via openings, a suitable conductor such as aluminum is deposited therein to form an electrically conductive path 302. Each end of the via is capped off with a raised pad contact 308 suitable for engaging a confronting pad contact, such as that at 309 from an adjacent stacked die. Pads formed at the active surface 304 of the die 102a electrically couple to the pad I/O locations. If I/O circuitry for a given pad was masked during I/O fabrication, then the via will not be coupled to I/O circuitry.

[0020] Following formation of the TSV array for each die, singulated die are aligned to each other in a stacked relationship such that the TSV arrays for adjacent devices (corresponding to the pad arrays) are offset by at least one pad location. This is accomplished during assembly through well-known registration techniques, such as that shown in Figure 4, wherein each die

Petitioners
Ex. 1002, p. 123

400a – 400d includes a pair of identically located registration reticles 402 and 404. The reticles in-turn are aligned with a pair of die pads (not shown). By aligning the leftmost reticle 402 of one die to the rightmost reticle 404 of an adjacent die (note that the dies are generally thinned and transparent during processing to allow visual alignment), a die-to-die offset of one pad may be straightforwardly achieved.

[0021] As each die is stacked in offset relationship to the die beneath it, pads from the lower surface of the upper device engage aligned pads disposed on the upper surface of the lower device. Once the stack is completely assembled, electrically conductive paths will extend orthogonally from the lowest device (typically a processor chip) up through the entire stack of identically formed dies.

[0022] In operation, the via signal paths define channels that are susceptible to various loading effects. Specifically, capacitance associated with each via affects current usage and related power efficiency parameters. However, by employing point-to-point signaling paths with the via and I/O circuit arrangement described above, each end-to-end via path electrically couples to (preferably) one I/O circuit. With each I/O circuit having the potential to load a via by a few hundred femptofarads, eliminating capacitance associated with three out of every four potential pad locations significantly improves power efficiency.

[0023] Figures 5 and 6 illustrate an alternative embodiment of a low power and low cost stacked die solution that avoids offsetting adjacent die. Like the previous embodiment illustrated in Figures 1 – 4, a plurality of semiconductor dies 502a – 502d are provided and stacked. Each die includes a formed TSV array 504 that corresponds to a pad array formed on each of the opposing die surfaces, with each via oriented orthogonally to the planar die surfaces and extending from one surface to the other.

[0024] Unlike the previously described embodiment, I/O circuitry 508 may be formed at each pad location 510 (phantom), but with circuitry 512 that enables selective coupling to a given via or pad. In one embodiment, a programmable fuse or anti-fuse may be employed to

Petitioners
Ex. 1002, p. 124

electrically decouple the I/O circuitry from the I/O pad. In so doing, the capacitive loading that would normally take place with a coupled I/O circuit (even when disabled) is avoided.

[0025] Thus, for each semiconductor device, a group of vias in a given semiconductor device will electrically couple to I/O circuitry on that device, while at least a portion of the remaining vias will be electrically decoupled from I/O circuits disposed adjacent their respective via pads.

[0026] Figure 6 illustrates the assembled alignment of stacked semiconductor dies 602a – 602d with respect to a base semiconductor die 604. While each I/O pad location includes I/O circuitry, for the first stacked device 602a, respective fuses have decoupled the I/O circuitry at Locations 2, 3 and 4 from the respective I/O pads (represented by "open" connections between each pad and I/O circuitry, such as at 606). Similar de-fusings are carried out with the remaining stacked dies.

[0027] Referring back to Figures 5 and 6, employing programmable fuses enables a device manufacturer to avoid masking out selected I/O locations, although at a cost of creating the fuses. With all of the I/O pad locations exhibiting identical circuitry, effecting point-to-point paths with the stacked die merely involves programming the fuses appropriately, such that a single via path coupling multiple devices couples to less than all of the I/O circuits adjacent that via path. The net result of using fuses allows the die stack to take on a true vertically aligned structure, with no offsets or process steps to precisely align offsets.

[0028] One additional process step in assembling the stack in accordance with the fuse embodiment described above involves initializing the fuses to arrive at the desired coupled/decoupled structures. Figure 7 generally illustrates how a serial bus 700 may be used in a daisy-chained manner to initialize selected fuses 702 on each die 704. Other ways to accomplish this include providing horizontal offset paths between adjacent devices to direct enable signals to the separate devices; enabling each device to measure a resistance of a given via path so that it can determine where it lies in the stack; or providing wirebond connections to drive device selection inputs.

Petitioners
Ex. 1002, p. 125

[0029] In the embodiments described above, an emphasis on limiting the via to pad I/O electrical coupling to one circuit was applied. In practice, as long as at least one I/O circuit is decoupled from the via path, or less than all of the I/O circuits are coupled to the via path, power savings may be realized. Moreover, although point-to-point paths for the via connections were emphasized above, multi-drop connections that still provide power and cost savings may be beneficial.

[0030] The stacked die structures described above are particularly well-suited for low-cost master-slave applications such as memory systems, where a master processing device, such as a memory controller, interfaces with a plurality of slave devices, such as memory devices. Figure 8 illustrates one general topology for interconnecting master and slave dies in a memory system 800. The system employs an integrated circuit (IC) memory controller 802 that interfaces with a plurality of IC memory devices 804a – 804d via point-to-point data buses 806 and a multi-drop address/control bus 808. The point-to-point paths generally interconnect data I/O circuitry on the controller with corresponding data I/O circuitry on each memory device. The multi-drop address/control bus routes address and control signals from the memory controller to all of the devices.

[0031] With continued reference to Figure 8, the memory devices 804a – 804d generally load the multi-drop bus in aggregate, while loading each point-to-point path individually. Preferably, dynamic random access memory (DRAM) devices are employed in the die stack with a corresponding DRAM controller. Non-volatile memory devices are also well suited to benefit from the unique die stacking arrangements disclosed herein. In addition to the system topology shown in Figure 8, various other topologies are suitable for the embodiments described herein, such as where all of the interconnections comprise either point-to-point paths, or multi-drop paths.

[0032] It should be noted that the various circuits disclosed herein may be described using computer aided design tools and expressed (or represented), as data and/or instructions embodied

Petitioners
Ex. 1002, p. 126

in various computer-readable media, in terms of their behavioral, register transfer, logic component, transistor, layout geometries, and/or other characteristics. Formats of files and other objects in which such circuit expressions may be implemented include, but are not limited to, formats supporting behavioral languages such as C, Verilog, and VHDL, formats supporting register level description languages like RTL, and formats supporting geometry description languages such as GDSII, GDSIII, GDSIV, CIF, MEBES and any other suitable formats and languages. Computer-readable media in which such formatted data and/or instructions may be embodied include, but are not limited to, non-volatile storage media in various forms (e.g., optical, magnetic or semiconductor storage media) and carrier waves that may be used to transfer such formatted data and/or instructions through wireless, optical, or wired signaling media or any combination thereof. Examples of transfers of such formatted data and/or instructions by carrier waves include, but are not limited to, transfers (uploads, downloads, e-mail, etc.) over the Internet and/or other computer networks via one or more data transfer protocols (e.g., HTTP, FTP, SMTP, etc.).

[0033] When received within a computer system via one or more computer-readable media, such data and/or instruction-based expressions of the above described circuits may be processed by a processing entity (e.g., one or more processors) within the computer system in conjunction with execution of one or more other computer programs including, without limitation, net-list generation programs, place and route programs and the like, to generate a representation or image of a physical manifestation of such circuits. Such representation or image may thereafter be used in device fabrication, for example, by enabling generation of one or more masks that are used to form various components of the circuits in a device fabrication process.

[0034] In the foregoing description and in the accompanying drawings, specific terminology and drawing symbols have been set forth to provide a thorough understanding of the present invention. In some instances, the terminology and symbols may imply specific details that are not required to practice the invention. For example, any of the specific numbers of bits, signal

Petitioners
Ex. 1002, p. 127

path widths, signaling or operating frequencies, component circuits or devices and the like may be different from those described above in alternative embodiments. Also, the interconnection between circuit elements or circuit blocks shown or described as multi-conductor signal links may alternatively be single-conductor signal links, and single conductor signal links may alternatively be multi-conductor signal links. Signals and signaling paths shown or described as being single-ended may also be differential, and vice-versa. Similarly, signals described or depicted as having active-high or active-low logic levels may have opposite logic levels in alternative embodiments. Component circuitry within integrated circuit devices may be implemented using metal oxide semiconductor (MOS) technology, bipolar technology or any other technology in which logical and analog circuits may be implemented. With respect to terminology, a signal is said to be "asserted" when the signal is driven to a low or high logic state (or charged to a high logic state or discharged to a low logic state) to indicate a particular condition. Conversely, a signal is said to be "deasserted" to indicate that the signal is driven (or charged or discharged) to a state other than the asserted state (including a high or low logic state, or the floating state that may occur when the signal driving circuit is transitioned to a high impedance condition, such as an open drain or open collector condition). A signal driving circuit is said to "output" a signal to a signal receiving circuit when the signal driving circuit asserts (or deasserts, if explicitly stated or indicated by context) the signal on a signal line coupled between the signal driving and signal receiving circuits. A signal line is said to be "activated" when a signal is asserted on the signal line, and "deactivated" when the signal is deasserted. Additionally, the prefix symbol "/" attached to signal names indicates that the signal is an active low signal (i.e., the asserted state is a logic low state). A line over a signal name (e.g., '$\overline{<\text{signal name}>}$') is also used to indicate an active low signal. The term "coupled" is used herein to express a direct connection as well as a connection through one or more intervening circuits or structures. Integrated circuit device "programming" may include, for example and without limitation, loading a control value into a register or other storage circuit within the

Petitioners
Ex. 1002, p. 128

device in response to a host instruction and thus controlling an operational aspect of the device, establishing a device configuration or controlling an operational aspect of the device through a one-time programming operation (e.g., blowing fuses within a configuration circuit during device production), and/or connecting one or more selected pins or other contact structures of the device to reference voltage lines (also referred to as strapping) to establish a particular device configuration or operation aspect of the device. The term "exemplary" is used to express an example, not a preference or requirement.

[0035] While the invention has been described with reference to specific embodiments thereof, it will be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. For example, features or aspects of any of the embodiments may be applied, at least where practicable, in combination with any other of the embodiments or in place of counterpart features or aspects thereof. Accordingly, the specification and drawings are to be regarded in an illustrative rather than a restrictive sense.

Petitioners
Ex. 1002, p. 129

CLAIMS

We Claim:

1. A stacked semiconductor system comprising:

a plurality of semiconductor dies, each having oppositely disposed first and second surfaces, a plurality of pads formed on the first surface, a plurality of pads formed on the second surface, and a plurality of through-vias connecting respective pads on the first surface to respective pads on the second surface, the plurality of through-vias in each semiconductor die including a first group of through-vias coupled to respective I/O circuitry on the semiconductor die and a second group of through-vias not coupled to I/O circuitry on the semiconductor die;

wherein the plurality of semiconductor dies are stacked such that the first group of through-vias in a first one of the plurality of semiconductor dies are aligned with respective ones of at least a portion of the second group of through-vias in a second one of the plurality of semiconductor dies.

2. The stacked semiconductor system according to claim 1 wherein first and second ones of the plurality of semiconductor dies are substantially identical and are offset from each other to achieve alignment of the first group of through-vias in the first one of the plurality of semiconductor dies with respective ones of at least a portion of the second group of through-vias in the second one of the plurality of semiconductor dies.

3. The stacked semiconductor system according to claim 2 wherein the first and second ones of the plurality of semiconductor dies comprise integrated circuit memory devices.

4. The stacked semiconductor system according to claim 3 wherein a third one of the plurality

Petitioners
Ex. 1002, p. 130

of semiconductor dies is stacked beneath the first one of the dies such that a portion of the first group of through-vias from the second one of the plurality of semiconductor dies electrically couples to at least a portion of the first group of through-vias of the third semiconductor die, the electrical coupling forming a plurality of conductive paths between the second semiconductor die and the third semiconductor die.

5. The stacked semiconductor system according to claim 4 wherein the third one of the plurality of semiconductor dies comprises a memory controller.

6. The stacked semiconductor system according to claim 4 wherein selected conductive paths of the plurality of conductive paths define point-to-point connections from I/O circuitry formed on the third semiconductor die to selected I/O circuits on one of the first or second semiconductor dies.

7. The stacked semiconductor system according to claim 4 wherein selected conductive paths of the plurality of conductive paths define multi-drop connections between the third semiconductor die and a plurality of I/O circuits on one or both of the first and second semiconductor dies.

8. The stacked semiconductor system according to claim 1 wherein the second group of through-vias are coupled to I/O pad not connected to I/O circuitry on a same die.

9. The stacked semiconductor system according to claim 1 wherein the plurality of pads define respective pad locations and the plurality of semiconductor dies are aligned such that each die is laterally offset from an adjacent die by at least one pad location.

Petitioners
Ex. 1002, p. 131

10. The stacked semiconductor system according to claim 1 wherein the plurality of pads define respective pad locations, the pad locations formed with respective I/O circuits, wherein at least one of the pad locations is programmably decoupled from its associated I/O circuit.

11. The stacked semiconductor system according to claim 10 wherein at least one of the pad locations is programmably decoupled from respective I/O circuitry by blowing an anti-fuse.

12. The stacked semiconductor system according to claim 1 wherein each I/O circuit includes transceiver circuitry.

13. The stacked semiconductor system according to claim 12 wherein each input/output circuit includes an electro-static-discharge device.

14. The stacked semiconductor system according to claim 1 wherein each through-via couples a pad location disposed on the first surface of a given die to a pad location disposed on the second surface with a conductive path formed orthogonal to the first and second surfaces.

15. A method of making a stacked semiconductor system comprising:

selecting a plurality of semiconductor dies, each having: oppositely disposed first and second surfaces, a plurality of pads formed on the first surface, a plurality of pads formed on the second surface, and a plurality of through-vias connecting respective pads on the first surface to respective pads on the second surface, the plurality of through-vias including a first group of through-vias coupled to respective I/O circuitry on the semiconductor die and a second group of through-vias not coupled to I/O circuitry on the semiconductor die;

stacking the plurality of semiconductor dies such that the first group of through-

Petitioners
Ex. 1002, p. 132

vias in a first one of the plurality of semiconductor dies is aligned with respective ones of at least a portion of the second group of through-vias in a second one of the plurality of semiconductor dies; and

electrically coupling the through-vias to form orthogonal paths between the plurality of semiconductor die.

16. The method of making a stacked semiconductor system according to claim 15 wherein the plurality of pads define respective pad locations, and stacking comprises offsetting each of the plurality of semiconductor dies with adjacent die by at least one pad location.

17. The method of making a stacked semiconductor system according to claim 16 wherein the pad locations are formed with respective I/O circuits and at least one of the pad locations is programmably decoupled from its associated input/output circuits.

18. A memory system comprising:

a first memory die having oppositely disposed first and second surfaces, and a first pad pattern defining a row of pad locations dedicated to conducting input/output signals, the pad locations formed on the first surface, each pad location formed with a conductive pad and an adjacent area, wherein less than all of the adjacent areas include respective input/output circuits for coupling to respective pads, the first memory die including an array of through-vias corresponding to the row of pad locations and coupling each pad location to the second surface;

a second memory die formed with the first pad pattern and the first via pattern, and stacked beneath the first memory die such that the first via pattern of the first memory die is offset by at least one pad location with respect to the via pattern of the second memory die and electrically couples to at least a portion of the first pad pattern of the second memory

Petitioners
Ex. 1002, p. 133

die, the electrical coupling forming vertical paths between the first memory die and the second memory die; and

a memory controller die formed with the first pad pattern and stacked beneath the second memory die such that the first via pattern of the second memory die electrically couples to at least a portion of the first pad pattern of the memory controller die, the electrical coupling forming conductive paths between the second memory die and the memory controller die.

19. The memory system according to claim 18 wherein selected conductive paths define point-to-point paths from the memory controller die to selected I/O circuits on one of the first or second memory dies.

20. The memory system according to claim 18 wherein selected conductive paths define multi-drop paths between the memory controller die and a plurality of I/O circuits on one or both of the first and second memory dies.

21. A stacked semiconductor system comprising:

a plurality of semiconductor dies, each having oppositely disposed first and second surfaces, a plurality of pads formed on the first surface, a plurality of pads formed on the second surface, and a plurality of through-vias connecting respective pads on the first surface to respective pads on the second surface;

wherein the plurality of pads define respective pad locations and the plurality of semiconductor dies are substantially identical and aligned such that each die is laterally offset from an adjacent die by at least one pad location.

22. A stacked semiconductor system comprising:

Petitioners
Ex. 1002, p. 134

a plurality of semiconductor dies, each having oppositely disposed first and second surfaces, a plurality of pads formed on the first surface, a plurality of pads formed on the second surface, and a plurality of through-vias connecting respective pads on the first surface to respective pads on the second surface, the plurality of through-vias in each semiconductor die including a first group of through-vias coupled to respective I/O circuitry on the semiconductor die and a second group of through-vias not coupled to I/O circuitry on the semiconductor die;

wherein the plurality of semiconductor dies are stacked such that through-vias in a first one of the plurality of semiconductor dies are aligned with respective ones of through-vias in a second one of the plurality of semiconductor dies and the I/O circuitry is sized to drive a load presented by one transmitter and one receiver.

Petitioners
Ex. 1002, p. 135

FIG. 1

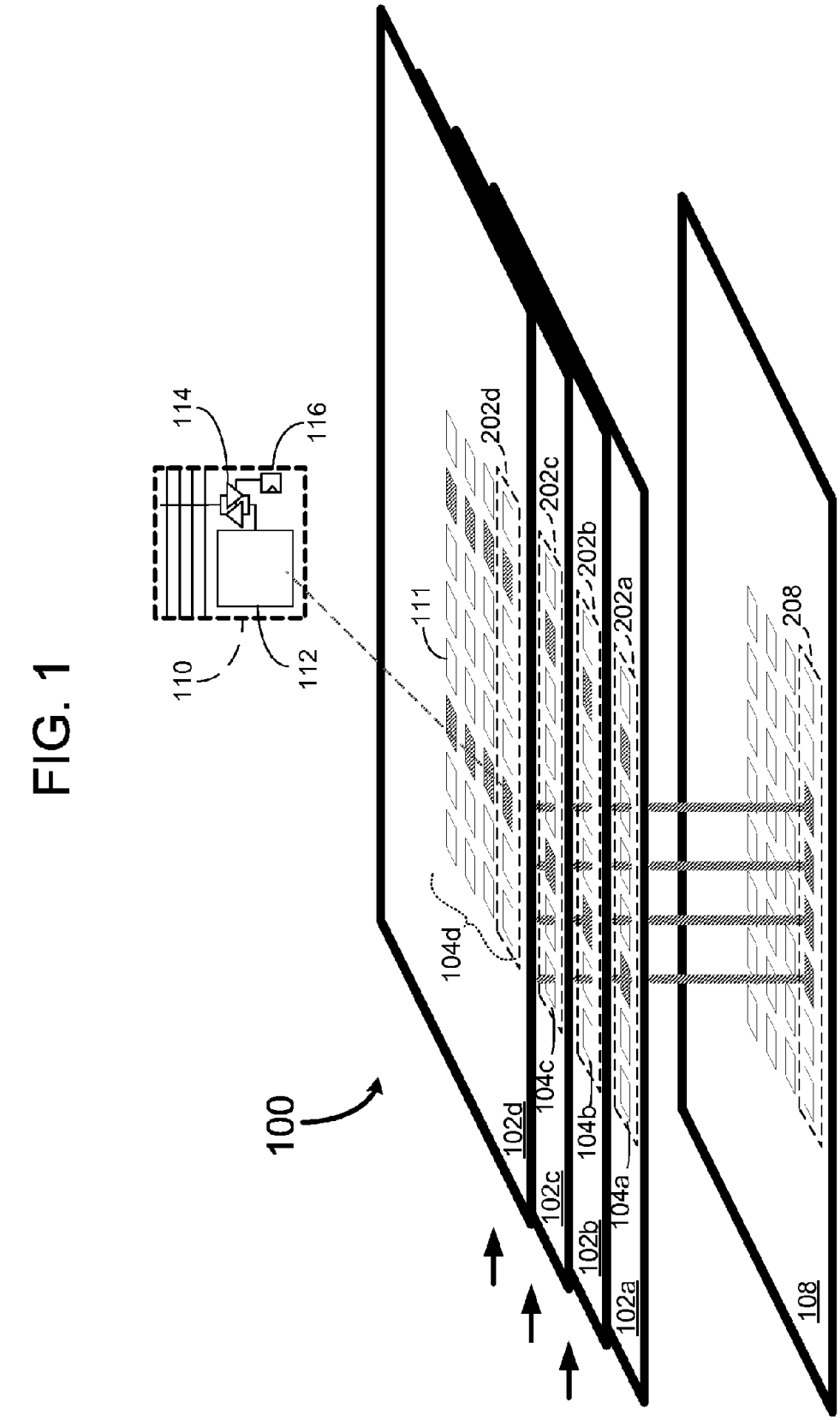

100
102d
102c
102b
102a
104d
104c
104b
104a
108
110
111
112
114
116
202d
202c
202b
202a
208

Fig. 2

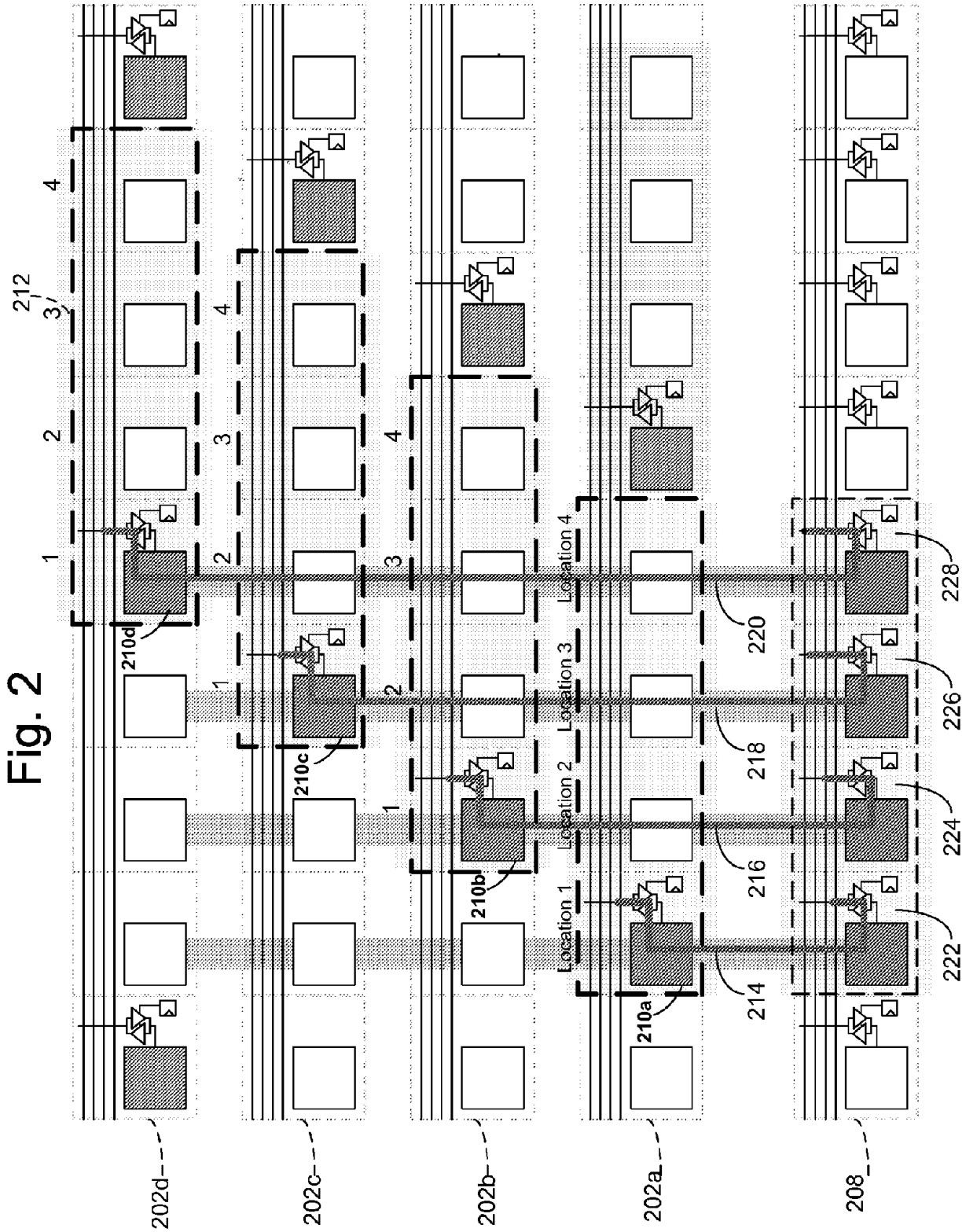

202d
202c
202b
202a
208
212
210a
210b
210c
210d
214
216
218
220
222
224
226
228
Location 1
Location 2
Location 3
Location 4
1
2
3
4

Fig. 3

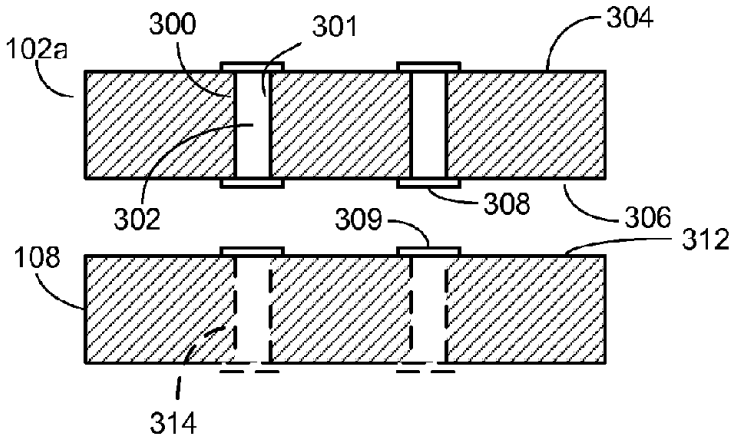

300
301
304
102a
302
308
306
309
312
108
314

Fig. 4

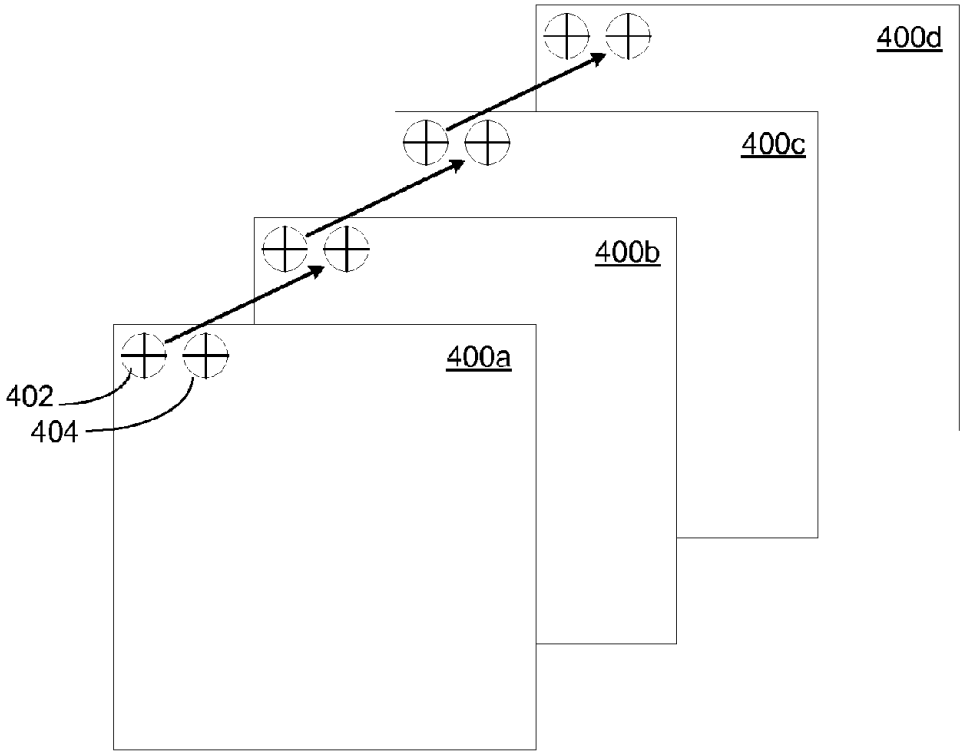

400d
400c
400b
400a
402
404

FIG. 5

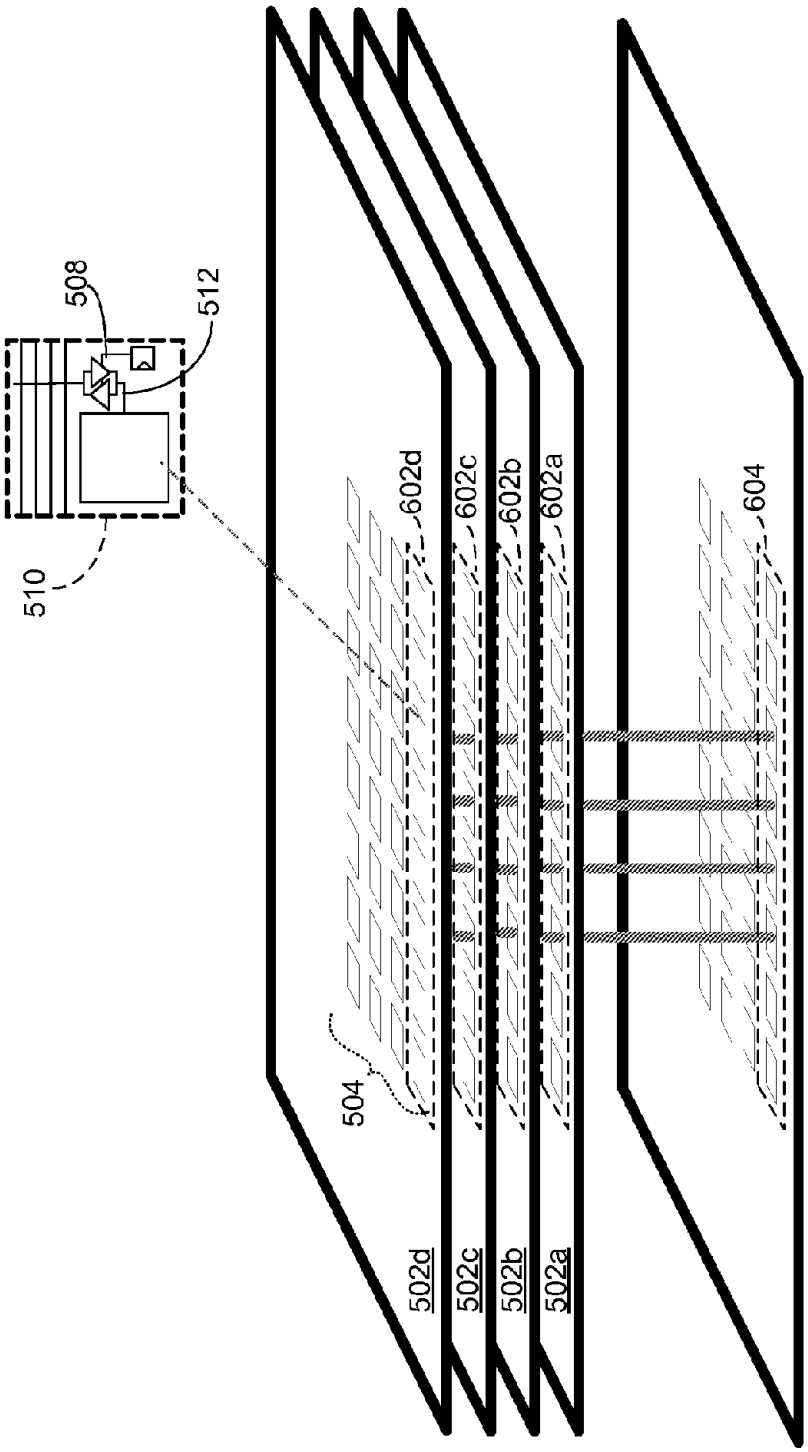


Fig. 6

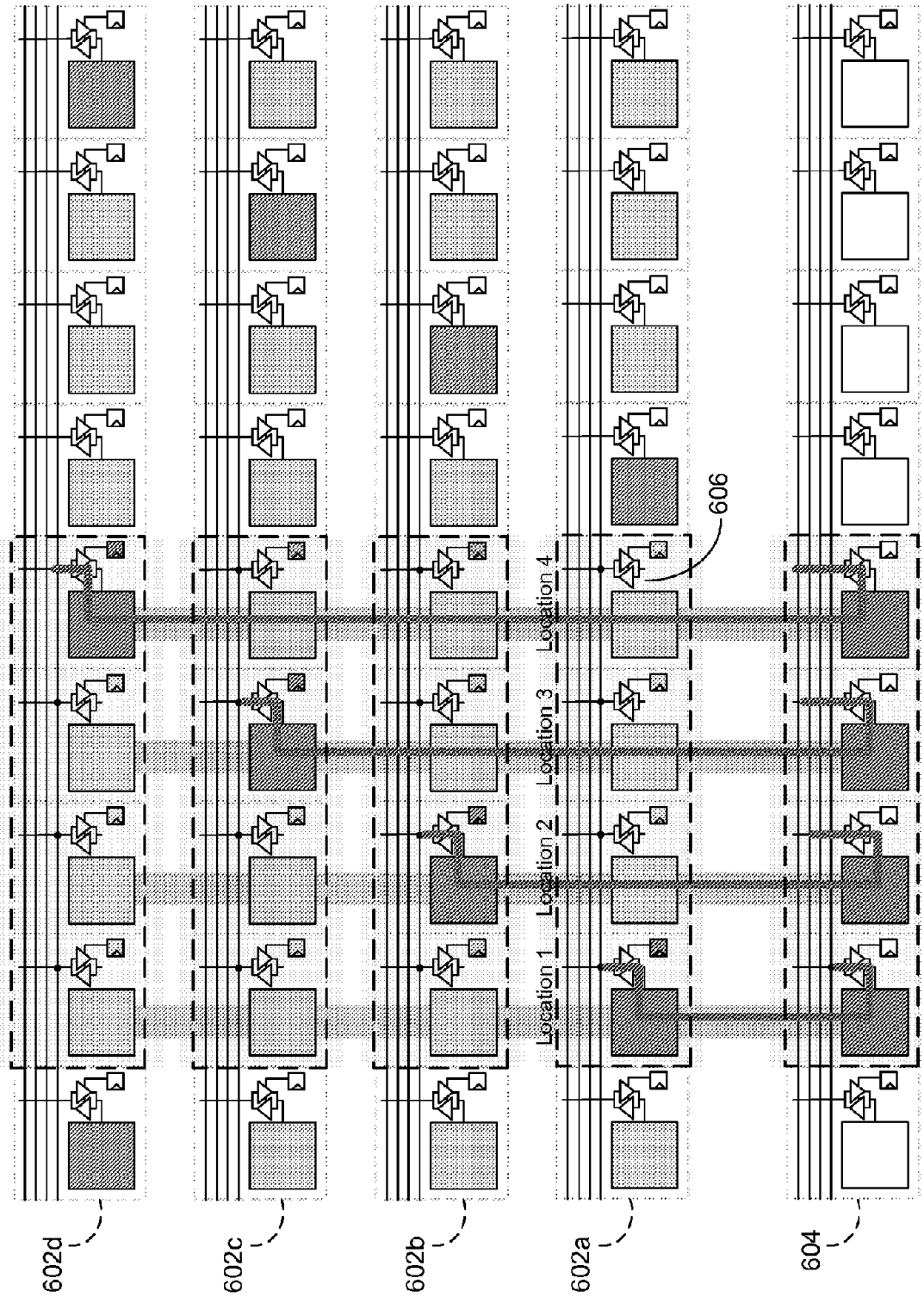

606
Location 4
Location 3
Location 2
Location 1
602d
602c
602b
602a
604

Fig. 7

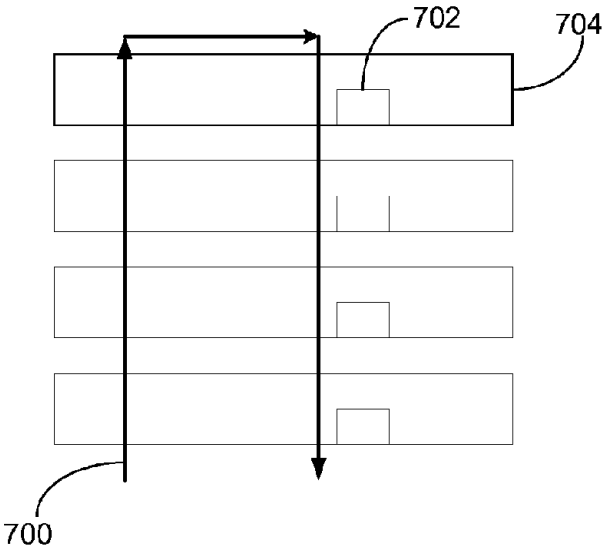
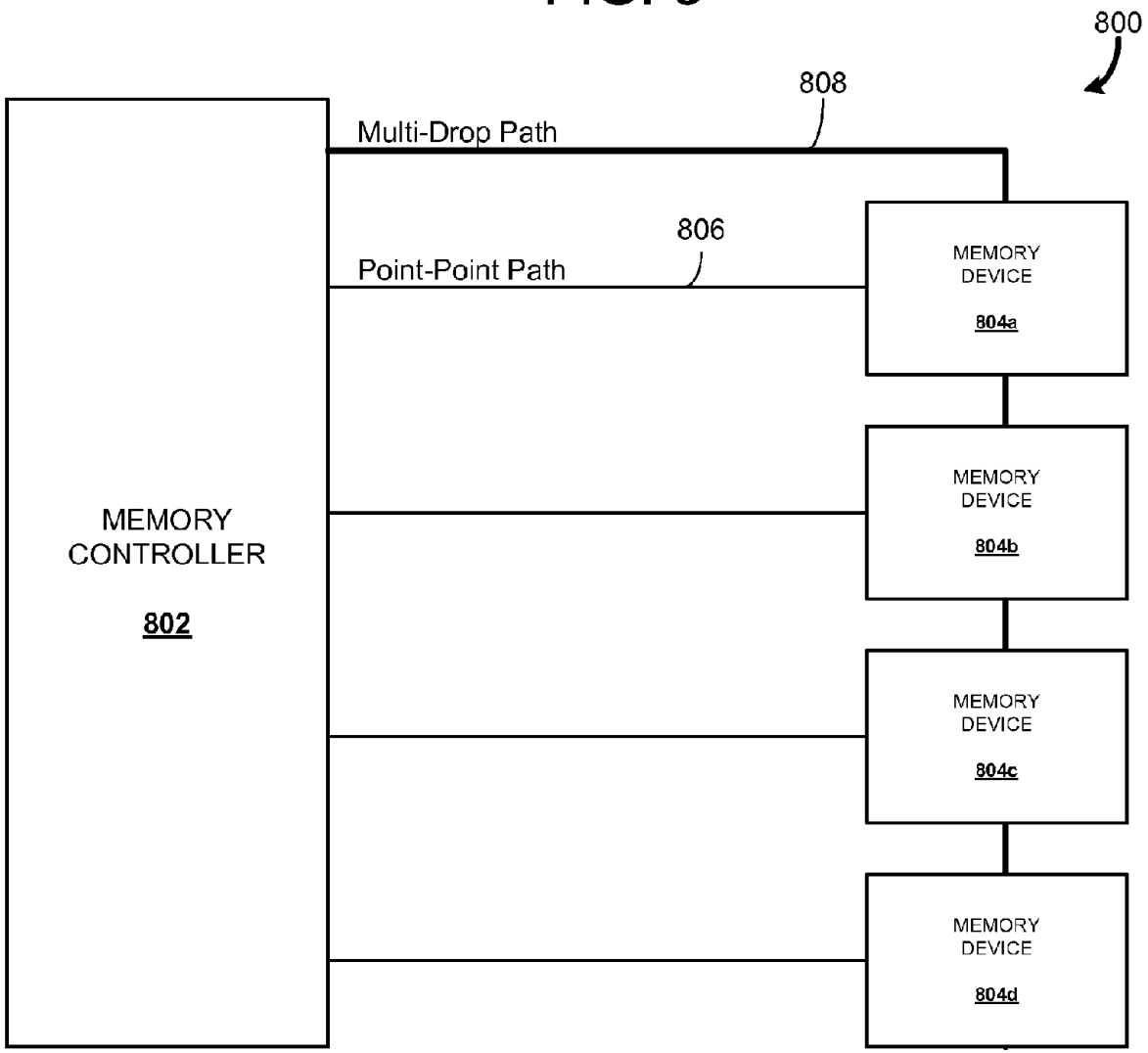


FIG. 8


800
808
Multi-Drop Path
806
Point-Point Path
MEMORY CONTROLLER
802
MEMORY DEVICE
804a
MEMORY DEVICE
804b
MEMORY DEVICE
804c
MEMORY DEVICE
804d

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 August 2011 (04.08.2011)

PCT

(10) International Publication Number
WO 2011/094437 A2

(51) International Patent Classification:
G11C 11/4063 (2006.01) *G06F 12/00* (2006.01)
G11C 7/10 (2006.01)

(21) International Application Number:
PCT/US2011/022763

(22) International Filing Date:
27 January 2011 (27.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/299,155 28 January 2010 (28.01.2010) US

(71) Applicants *(for all designated States except US)*: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, Texas 77070 (US). **UNIVERSITY OF UTAH RESEARCH FOUNDATION** [US/US]; 615 Arapeen Drive, Suite 310, Salt Lake City, Utah 84108 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MURALIMANOHAR, Naveen** [IN/US]; Hewlett-Packard Co., 1501 Page Mill Rd., Mail Stop 1183, Palo Alto, California 94304 (US). **UDIPI, Aniruddha** [IN/US]; University of Utah, School of Computing, 50 S Central Campus Dr. 3190, Salt Lake City, Utah 84112 (US) **CHATTERJEE, Niladrish** [IN/US]; University of Utah, School of Computing, 50 S Central Campus Dr. 3190, Salt Lake City, Utah 84112 (US). **BALASUBRAMONIAN, Rajeev** [IN/US]; University of Utah, School of Computing, 50 S Central Campus Dr. 3190, Salt Lake City, Utah 84112 (US). **DAVIS, Alan Lynn** [US/US]; Hewlett-Packard Co., 1501 Page Mill Rd., B-3 3U H35, Mail Stop 1177, Palo Alto, California 94304 (US). **JOUPPI, Norman Paul** [US/US]; Hewlett-Packard Co., 1501 Page Mill Rd., B-3 3U G37, Mail Stop 1181, Palo Alto, California 94304 (US).

(74) Agents: **CHANG, Marcia R.** et al.; Hewlett-Packard Company, Intellectual Property Administration, 3404 E. Harmony Road, Mail Stop 35, Fort Collins, Colorado 80528 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

[Continued on next page]

(54) Title: MEMORY ACCESS METHODS AND APPARATUS

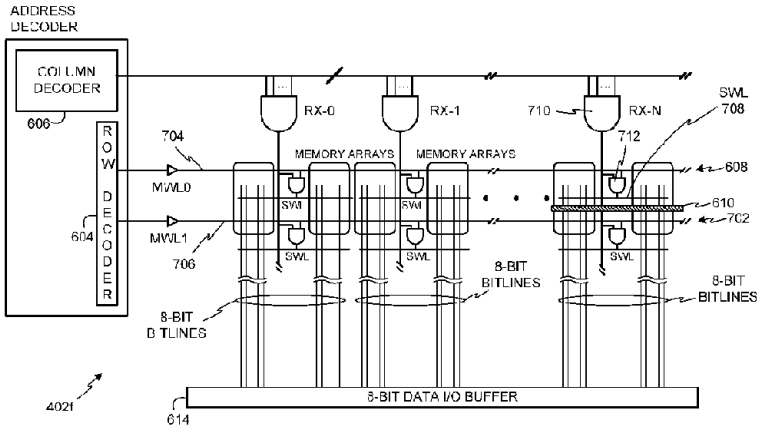


FIG. 7

(57) Abstract: A disclosed example apparatus includes a row address register (412) to store a row address corresponding to a row (608) in a memory array (602). The example apparatus also includes a row decoder (604) coupled to the row address register to assert a signal on a wordline (704) of the row after the memory receives a column address. In addition, the example apparatus includes a column decoder (606) to selectively activate a portion of the row based on the column address and the signal asserted on the wordline.

WO 2011/094437 A2

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

MEMORY ACCESS METHODS AND APPARATUS

RELATED APPLICATIONS

[0001] This is an International Patent Cooperation Treaty patent application that claims the benefit of U.S. Provisional Patent Application No. 61/299,155, filed on January 28, 2010, which is hereby incorporated herein by reference in its entirety.

BACKGROUND

[0002] Traditionally, memories such as dynamic random access memories (DRAMs) have been designed for low-cost and fast access time. However, gains in cost and access speed have been made via tradeoff decisions resulting in increased power-consumption. A DRAM is made of thousands of transistors organized to form bit cells that store bit-level information. When accessed together, combinations of bit cells can store meaningful and/or useful information. DRAM architectures include signal busses used to activate different combinations of bit cells for writing and/or reading information at addressable storage locations.

[0003] Some traditional DRAM architectures are structured such that a DRAM can be operated to quickly retrieve data stored at sequential address locations in response to a single data request and address provided by a processor or memory controller. For example, processors are typically configured to retrieve one or more entire cache lines from DRAM based on a single read request. In a

traditional DRAM module, a single conventional read operation requires pre-charging an entire row of bitlines in each DRAM chip of the DRAM module. Each pre-charged row typically corresponds to multiple cache lines.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] FIG. 1 depicts an interface configuration between a memory controller and a memory.

[0005] FIG. 2 depicts an example printed circuit board (PCB) in-line memory module having a plurality of memory chips.

[0006] FIG. 3 depicts an example chip stack memory module.

[0007] FIG. 4 is a dynamic random access memory (DRAM) configuration to selectively activate particular segments of rows by delaying use of a row address to access information stored therein.

[0008] FIG. 5 is a diagram of known DRAM chip.

[0009] FIG. 6 is a diagram of the DRAM configuration of FIG. 4 showing row and column address decoders working cooperatively to selectively activate specific segments of rows in DRAM bit cell array banks.

[0010] FIG. 7 is a detailed diagram of one of the DRAM chips of the DRAM configuration of FIGS. 4 and 6 showing row and column address decoders of the DRAM chip working cooperatively to selectively activate particular bitlines of the DRAM chip.

[0011] FIG. 8 is a diagram of a DRAM configuration to selectively activate a particular DRAM chip to access an entire cache line from that DRAM chip.

[0012] FIG. 9 is a detailed diagram of the DRAM module of FIG. 8.

[0013] FIG. 10 is a diagram of another DRAM memory module that can alternatively be used to implement the DRAM module of FIG. 8 to provide data access to only a single DRAM chip at a time.

[0014] FIG. 11 is a diagram of another DRAM memory module that can alternatively be used to implement the DRAM module of FIG. 8 to selectively activate particular DRAM chips per memory access.

[0015] FIG. 12 is an example cache line storage configuration that can be used in connection with the example DRAM memory modules of FIGS. 8-11.

[0016] FIG. 13 depicts a detailed diagram of a memory controller of FIG. 8.

[0017] FIG. 14 depicts a flow diagram of an example process that can be used to access memory devices as described in connection with FIGS. 4, 6, and 7.

[0018] FIG. 15 depicts a flow diagram of an example process that can be used to access memory devices as described in connection with FIGS. 8-10, 12, and 13.

[0019] FIG. 16 depicts a flow diagram of an example process that can be used to access memory devices as described in connection with FIG. 11.

[0020] FIG. 17 depicts an example memory module configured to store error-checking and parity information to enable error-checking and data correction operations when accessing data stored therein.

[0021] FIG. 18 depicts a flow diagram of an example process that can be used to perform error-checking and data correction operations when accessing data in a memory.

[0022] FIG. 19 is a table showing example test bench parameters of an example processor system model used to analyze the performance of the example methods, apparatus, and articles of manufacture disclosed herein.

[0023] FIG. 20 is a table showing example DRAM energy parameters of the example memory access techniques and structures disclosed herein and known memory access techniques and structures that were used to implement the performance analyses.

[0024] FIG. 21 is a graph showing example row buffer hit rates in an example code execution environment as a function of quantities of traditional memory arrays.

[0025] FIG. 22 is a graph showing example row re-use counts in an example code execution environment based on an 8-array traditional memory.

[0026] FIG. 23 is a graph comparing example normalized DRAM energy consumption for row/wordline accesses using known memory access techniques/structures and the example memory access techniques/structures disclosed herein.

[0027] FIG. 24 is a graph comparing example DRAM energy consumed by different example memory subsystems when accessing rows/wordlines using

known memory access techniques/structures and the memory access techniques/structures disclosed herein.

[0028] FIG. 25 is a graph showing example memory latency impact of using low-power states in connection with known memory access techniques/structures.

[0029] FIG. 26 is a graph showing example energy reduction impact of using low-power states in connection with known memory access techniques/structures.

[0030] FIG. 27 is a graph comparing example average main memory latency when accessing memory using known memory access techniques/structures and the memory access techniques/structures disclosed herein.

[0031] FIG. 28 is a graph comparing example memory latencies contributed by different memory subsystems when accessing memory using known memory access techniques/structures and the memory access techniques/structures disclosed herein.

[0032] FIG. 29 is a graph comparing example normalized instructions per cycle for known memory access techniques/structures and the memory access techniques/structures disclosed herein.

[0033] FIG. 30 is a graph comparing example differences in memory access performance for use and non-use of chipkill error detection and correction techniques.

DETAILED DESCRIPTION

[0034] Example methods, apparatus, and articles of manufacture disclosed herein can be used to implement electronic memory devices that consume less power than traditional memory devices. Example methods, apparatus, and articles of manufacture are disclosed in connection with dynamic random access memory (DRAM) architectures, but may be implemented in connection with other types of memories including memristor memory devices, static RAM (SRAM) devices, Ferroelectric RAM (FRAM) devices, etc.

[0035] Traditionally, DRAM has been designed for low-cost and fast access time. However, gains in cost and access speed have been made via tradeoff

decisions resulting in increased power-consumption. A DRAM is made of thousands of transistors organized to form bit cells that store bit-level information. When accessed together, combinations of bit cells can store meaningful and/or useful information. DRAM architectures include signal busses used to activate different combinations of bit cells for writing and/or reading information at addressable storage locations. As discussed in greater detail below, bitlines are selectively enabled or activated to selectively read or write information in different bit cells.

[0036] To increase performance, while maintaining or decreasing physical space requirements, some traditional DRAM architectures are structured such that a DRAM can be operated to quickly retrieve data stored at sequential address locations in response to a single data request and address provided by a processor or memory controller. However, such data access techniques typically result in retrieving more information from a memory than is necessary and, thus, unnecessarily consume more power by activating portions of memory that may not otherwise be needed. For example, processors are typically configured to retrieve one or more entire cache lines from DRAM based on a single read request. A cache line refers to the number of bits or bytes that make up the width of a processor's cache memory (e.g., a 32-byte cache line, a 64-byte cache line, a 128-byte cache line, etc.). In a traditional DRAM module, a single conventional read operation requires pre-charging an entire row of bitlines in each DRAM chip of the DRAM module. Each pre-charged row typically corresponds to multiple cache lines. Thus, even when a processor requests a single cache line, internal DRAM logic must pre-charge all of the bitlines of the DRAM chip row(s) that store(s) the requested cache line. Pre-charging an entire row activates thousands of bitlines even though only a subset of those activated bitlines store the single relevant cache line requested by the requesting processor. Such known access techniques, achieve fast access performance when there is high locality of adjacently stored information that is needed by the requesting processor in a consecutive manner. That is, by pre-charging an entire row, every cache line stored in that row can be read-out to the requesting processor during consecutive low-latency read cycles (e.g.,

DRAM burst-mode reads or page-mode reads). While there is some initial overhead delay in pre-charging to read-out the first cache line, each subsequent consecutive cache line can be read out with much lower latency because the row storing those cache lines has already been pre-charged. However, if the requesting processor requires only one cache line from the pre-charged row, significant energy is wasted using these known techniques, especially when memory access locality is reduced as core, thread, and socket counts increase in processor operations.

[0037] Unlike traditional DRAM architectures and operations, example methods, apparatus, and articles of manufacture disclosed herein facilitate accessing memory devices, while consuming relatively lower power by activating relatively less DRAM circuitry per memory access than known DRAM devices. An example implementation referred to herein as Selective Bitline Activation (SBA) involves configuring DRAM control logic to selectively access single cache lines by applying address information differently from the manner in which traditional DRAM architectures use such address information. Using SBA, an internal DRAM architecture can operate while maintaining compatibility with industry-adopted external interface device standards used to ensure compatible interoperability between different memory controllers and different memory device manufacturers. The JEDEC Solid State Technology Association (formerly the Joint Electron Devices Engineering Council (JEDEC)) publishes and maintains several DRAM interface standards that are used to establish interface timing and control requirements to which memory device manufacturers must adhere.

[0038] Another example implementation referred to herein as Single Subarray Access (SSA) organizes the layout of DRAM chips and the mapping of data to DRAM chips so that an entire cache line can be fetched from a single DRAM memory chip of a memory module. In contrast, known memory modules, such as known single in-line memory module (SIMM) or dual in-line memory module (DIMM) architectures, fetch a cache line from separate cache line portions or cache line segments distributed among multiple DRAM chips of a memory module. (A cache line segment is a sub-group of bits that form a portion of an

entire cache line.) Thus, unlike traditional DRAM configurations, in which multiple chips in a memory module must be activated or pre-charged to read out a single cache line, example SSA techniques disclosed herein enable reading out an entire cache line by activating only a single memory chip of a memory module. By eliminating the need to activate and pre-charge multiple memory chips of a memory module to access a single cache line, relatively less power is used per memory access because the non-activated memory chips in the memory module can remain in a low-power consumption mode throughout the fetch operation. Example preliminary studies show that memory designs configured in accordance with the SSA memory access technique can yield memory energy savings of 6X, while incurring an area cost of only 4.5%, and improving performance (e.g., read latency) by up to 15%.

[0039] Example methods and apparatus are also disclosed herein to provide error detection and correction capabilities through the use of checksum information. Such error detection and correction capabilities involve the use of chipkill techniques and may be advantageously used in connection with the SSA memory access technique disclosed herein, which relies on a single memory chip to provide the entire contents of a single cache line. The error detection and correction capabilities are not, however, committed for use only with the SSA techniques. Example preliminary studies show that the chipkill implementation described herein has significantly lower capacity and energy overheads than known chipkill solutions, while incurring only an 11% performance penalty over a comparable memory implementation without chipkill error detection and correction capabilities.

[0040] In the following description and corresponding figures, FIG. 1 shows a memory interface representative of interface line configurations that can be used in either of the SBA or SSA memory access techniques. The diagrams of FIGS. 4, 6, 7, and 14 are used, by way of example, to describe example memory implementations using SBA memory configurations. The diagrams of FIGS. 8-13, 15, and 16 are used, by way of example, to describe example memory implementations using SSA memory configurations. The chipkill error

detection and correction capabilities are described in connection with FIGS. 17 and 18.

[0041] Turning now to FIG. 1, an example memory interface configuration 100 shows a memory controller 102 operatively coupled to a memory 104. The memory controller 102 can be a standalone memory controller integrated circuit (IC) or an embedded memory controller implemented in a processor chip (e.g., fabricated on the same die or located in the same chip package as a processor core). In the illustrated example, the memory 104 is a DRAM memory. The DRAM memory 104 can be a single DRAM memory IC or a memory module including multiple DRAM memory ICs. In some example implementations, the DRAM memory 104 can be embedded memory implemented in a processor chip. As shown, the memory interface configuration 100 includes a data bus 106, an address bus 108, and control lines 110.

[0042] The data bus 106 may be any number of bits wide and is used to communicate data between the memory controller 102 and the DRAM memory 104. When implemented as a memory module (e.g., the memory modules of FIGS. 2 and 3), the lines of the data bus 106 can be formed of separate data line groups corresponding to data buses of respective memory chips.

[0043] Turning briefly to FIGS. 2 and 3, the example memory modules 200 and 300 depicted therein can be used to implement any of the memory configurations disclosed herein. In the example illustrated in FIG. 2, a DRAM PCB in-line memory module 200 (e.g., a dual in-line memory module (DIMM)) is implemented as a multi-chip memory module including four memory chips 202a-d mounted on a PCB 204. The DRAM PCB in-line memory module 200 may be advantageously used in optical interface systems in which the memory module 200 is connected to other subsystems (e.g., other memory modules and/or memory controllers) via optical lines. Alternatively, the memory module 200 may also be used in electrical interface systems. A data bus 208a-d of each of the memory chips 202a-d interfaces with a respective, separate data line group 206a-d of the DRAM memory module 200. Each data line of the memory module 200 can be implemented using an electrical contact pad formed (e.g., etched) on the PCB 204 and connected to a respective one of the memory

chips 202a-d via an electrically conductive trace formed in and/or on the PCB 204.

[0044] In some examples, one or more of the memory chips 202a-d are implemented using a 3D chip stack in which multiple memory dies (e.g., of homogeneous or heterogeneous memory technology types) are stacked on top of one another (e.g., similar to the 3D stack structure shown in FIG. 3).

[0045] An example 3D chip stack memory module 300 is shown in FIG. 3. The 3D stacked memory module 300 may be advantageously used in electrical interface systems in which the memory module 300 is connected to other subsystems (e.g., other memory modules and/or memory controllers) via electrical lines. Alternatively, the memory module 300 may also be used in optical interface systems. The example 3D chip stack memory module 300 of FIG. 3 includes a first IC die 302 stacked on a second IC die 304, which is stacked on a third IC die 305. The IC dies 302, 304, and 305 are carried on a ball grid array (BGA) chip package 306. In the illustrated example, the first IC die 302 can be a SDRAM memory and the second IC die 304 can be another SDRAM memory or any other type of memory (e.g., memristor memory, SRAM, flash memory, PCRAM, etc.) or IC (e.g., a processor, a controller, etc.). In example implementations in which an SDRAM die (or any other memory technology die) is stacked on a processor or controller die, address, control, and data lines of the SDRAM die can be routed directly to the processor or controller die internal to the chip stack package. In such implementations, memory access external from the chip stack package might not be necessary. Alternatively or additionally, to enable external memory access, address, control, and data lines of the memory IC dies can be routed to external chip interfaces (e.g., BGA pads, surface mount pads, chip leads, etc.). In some examples, a memory module buffer may be stacked with multiple memory die. For example, in the illustrated example of FIG. 3, the IC die 305 may be a memory module buffer (e.g., similar or identical to memory module buffers 404 of FIG. 4, 908 of FIG. 9, 1008 of FIG. 10, and/or 1104 of FIG. 11). Although the 3D chip stack memory module 300 is shown as a BGA package, other types of packages may be used.

[0046] Returning to FIG. 1, the address bus 108 is used to communicate address information from the memory controller 102 to the DRAM memory 104. In DRAM architectures, an address for each memory controller-initiated memory access may be communicated in two parts during two consecutive address communications. During a first address part communication, the memory controller 102 may output a row address on the address bus 108. During a second address part communication, the memory controller 102 may output a column address on the address bus 108. To provide access to the desired information addressed by the memory controller 102, the DRAM memory 104 latches the row and column addresses to activate corresponding bit cells for read or write access.

[0047] In the illustrated example, the control lines 110 can include control lines suitable for implementing DRAM or synchronous DRAM (SDRAM) type interfaces. Such control lines can include a row address strobe (RAS) line, a column address strobe (CAS) line, a write-enable (WE) line, a chip select (CS) line, a clock (CLK) line, a clock enable (CE) line, bank select lines, etc. In some example implementations, bank select lines can be connected to lines of the address bus 108. In some example implementations, SDRAM-type interfaces can be used in connection with non-SDRAM type memory technologies. For example, an SDRAM interface may be used to interface to a memristor IC die or a SRAM IC die so that interface compatibility can be maintained with SDRAM memory controllers even though different memory technologies are employed.

[0048] Also shown in FIG. 1 is a processor 112 in communication with the memory controller 102. The processor 112 includes a cache memory 114 that functions as a temporary quick access storage area for a processor core of the processor 112. The cache memory 114 is formed of a plurality of cache lines, one of which is denoted as cache line 116. In some example processor system implementations, a size of a cache line (e.g., a 64-byte cache line, a 128-byte cache line, etc.) dictates the number of bits that may be read from an external memory (e.g., a DRAM) per memory access. For example, if the cache line 116 is 64 bytes wide and the DRAM memory 104 is a memory module (e.g., the PCB in-line memory module 200 of FIG. 2) with a 64-bit data bus, the memory

controller 102 can retrieve the 64-byte wide cache line from the DRAM memory 104 using eight read cycles. Alternatively, if the DRAM memory 104 is a single DRAM chip having an 8-bit data bus, the memory controller 102 can retrieve the 64-byte wide cache line using 64 8-bit read cycles.

[0049] Turning now to FIG. 4, in accordance with SBA memory access, a DRAM configuration selectively activates particular segments of rows in respective DRAM chips by delaying use of a row address to access information stored therein. As shown in FIG. 4, a multi-chip memory module 400 includes a plurality of DRAM chips 402a-h and a memory module buffer 404. In the illustrated example, the memory module buffer 404 includes an address buffer 406 to receive and buffer row and column addresses and RAS/CAS buffers 408 to receive and buffer RAS and CAS signals from a memory controller. Although the RAS/CAS buffers 408 are shown, in other example implementations, the RAS/CAS buffers 408 may be omitted and the memory module buffer 404 can dynamically (or on-the-fly) generate the RAS and CAS signals as needed when communicating address information (e.g., row and column addresses) to the DRAM chips 402a-h. In some examples, the memory module buffer 404 may be omitted and memory module buffers substantially similar or identical to the memory module buffer 404 may be implemented in one or more of the DRAM chips 402a-h such that the memory module 400 would operate the same as if it were provided with the memory module buffer 404 as shown in FIG. 4.

[0050] In the illustrated example, signal timings required by the memory module 400 are the same as signal timings of traditional memory modules. As a result, the memory module 400 can be used in connection with known or legacy memory controllers. However, as explained below, the operation of each of the DRAM chips 402a-h is different from the manner in which traditional memory chips operate to provide access to stored memory contents.

[0051] Upon receiving a row address in the address buffer 406 and a RAS signal at the RAS/CAS buffers 408, the address buffer 406 and the RAS/CAS buffers 408 of the illustrated example communicate the row address and the RAS signal in broadcast fashion to all of the DRAM chips 402a-h. In addition, upon receiving a column address in the address buffer 406 and a CAS signal at

the RAS/CAS buffers 408, the address buffer 406 and the RAS/CAS buffers 408 of the illustrated example communicate the column address and the CAS signal in broadcast fashion to all of the DRAM chips 402a-h. In this manner, the memory module 400 is backwards compatible with known memory controllers.

[0052] However, unlike known DRAM memory chips, the DRAM chips 402a-h delay use of the received row address to enable selectively activating relatively less circuitry of a DRAM storage area per memory access than is activated by known DRAM technologies. In this manner, SBA memory accesses use less power per memory access by activating relatively less circuitry. To enable SBA memory accesses, each SDRAM chip 402a-h is provided with a row address register 412 and a column address register 414, which are shown by way of example in connection with SDRAM chip 402f. The row address register 412 enables use of a posted RAS by storing the row address to postpone its use in decoding an address of requested data until a CAS signal and column address are received.

[0053] The SDRAM chip 402f is provided with a demultiplexer 416 to select one of the row address register 412 or the column address register 414. A RAS/CAS select input to the demultiplexer 416 is connected to the RAS/CAS buffers 408 to control the demultiplexer 416 to select either the row address register 412 or the column address register 414 to store a received address. For example, under traditional timings, the memory module 400 receives a row address in connection with an assertion of a RAS control signal. That is, the row address is latched into the address buffer 406 and communicated to the DRAM chip 402f (and to the other DRAM chips 402a-e, 402g, and 402h), while the RAS assertion signal is latched into the RAS/CAS buffers 408 and also communicated to the DRAM chip 402f (and to the other DRAM chips 402a-e, 402g, and 402h). Similarly, under traditional timings, the memory module 400 receives a column address in connection with an assertion of a CAS control signal. That is, the column address is latched into the address buffer 406 and communicated to the DRAM chip 402f (and to the other DRAM chips 402a-e, 402g, and 402h), while the CAS assertion signal is latched into the RAS/CAS buffers 408 and also communicated to the DRAM chip 402f (and to the other

DRAM chips 402a-e, 402g, and 402h). Thus, when the demultiplexer 416 senses an assertion of the RAS signal, the address bits at its address inputs correspond to the row address, and the demultiplexer 416 selects the row address register 412 to store the address bits. Similarly, when the demultiplexer 416 senses an assertion of the CAS signal, the address bits at its address inputs correspond to the column address, and the demultiplexer 416 selects the column address register 414 to store the address bits.

[0054] Unlike known techniques that act on a row address by activating an entire row of a bit cell array bank before receiving the column address, the DRAM chip 402f of the illustrated example waits until both of the row address and column address are received before decoding the row and column addresses and before activating any of the bit cell access circuitry. In this manner, the DRAM chip 402f can use the row and column addresses concurrently in combination to activate only the necessary storage area corresponding to the requested memory access and, thus, retrieve only a relevant cache line into a row buffer. Thus, in the illustrated example, when both the row address and the column address for a particular memory access are stored in the registers 412 and 414, row and column decoders activate only the portion of a row that stores the specific information that is requested and addressed by the combination of the row and column addresses. This concept is illustrated in greater detail in FIG. 6.

[0055] Before continuing to FIG. 6, a brief discussion of the operation of known DRAM chips may be instructive. Unlike the operation of the DRAM chip 402f described in connection with FIGS. 4, 6, and 7, known DRAM chips activate an entire row per memory access as shown in FIG. 5. Turning to FIG. 5, upon receipt of a row address and before receiving a column address, a known DRAM chip 500 activates an entire row 502 of a cell array bank 504 based on the row address loaded into a row decoder 506. In known DRAM chips, activating an entire row is referred to as opening a page such that a page-mode access is enabled for consecutively reading out bursts of adjacently stored information. When the entire row 502 is active, the DRAM chip 500 loads all of the bits of the active row 502 into a row buffer 508. When the

DRAM chip 500 receives a column address, a column decoder 510 decodes the column address and selects a particular portion 512 of the row buffer 508 and outputs the row buffer portion 512 via a data I/O buffer 514. Thus, known DRAM chips activate an entire row and load all of the bits of the active row into a row buffer before receiving a column address. Activating an entire row in this manner for each memory access results in consuming more power than necessary when only a portion of the information stored in that row is requested. The example SBA techniques described herein reduce power consumption by only activating a portion of a row in a cell array bank storing relevant information as discussed below in connection with FIGS. 6 and 7.

[0056] Turning to FIG. 6, the DRAM chip 402f is shown with a plurality of bit cell array banks 602 that store information, a row decoder 604, and a column decoder 606. Each bit cell array bank 602 is arranged as rows of bit cells that store information. A row 608 is shown in one of the banks 602 and is also referred to herein as a wordline. In the illustrated example, the row decoder 604 receives the row address from the row address register 412 (FIG. 4), and the column decoder 606 receives the column address from the column address register 414 (FIG. 4). The row decoder 604 and the column decoder 606 then concurrently decode their respective address information to activate only a portion of the row 608 (or wordline) within a particular one of the cell array banks 602 containing the requested information. The activated portion of the row 608 is denoted herein as a wordline segment 610, which is a portion of the row 608 that is activatable without requiring activation of other portions of the same row 608. In the illustrated example of FIG. 6, separate wordline segments of respective ones of the DRAM chips 402a-h can be activated and read out as discussed below with reference to the DRAM chip 402f. In this manner, the bits read out from the separate SDRAM chips 402a-h can be concatenated to form a cache line requested by a memory controller (e.g., the memory controller 102 of FIG. 1). In the illustrated example of FIG. 6, a 64-byte cache line is formed by concatenating 8-bit data outputs from each of the DRAM chips 402a-h to form eight 64-bit cache line portions or segments that are subsequently concatenated to form the entire cache line.

[0057] When the wordline segment 610 is active, its bits are loaded into a row buffer 612 for output via data I/O buffer 614. By activating only a relevant wordline segment (e.g., the wordline segment 610) in a row, the SBA technique imposes a close-page policy, which prevents burst accesses associated with known DRAM chips that allow opening a row (or page) once and reading out two or more row segments (e.g., multiple 8-bit data readouts) based on the single row opening. Instead, the close-page policy of SBA memory accesses requires decoding a row address for each memory access. However, the close-page policy is more energy-efficient than known DRAM chip access techniques when used in connection with applications in which relevant data is not stored with high locality (e.g., not often stored in the same row), because while known DRAM chip accesses will open an entire row even when only a single datum (e.g., a cache line) in that row is relevant (which is often the case for information stored with low locality), the SBA techniques described herein only open a single wordline segment of a row per memory access.

[0058] The wordline segment 610 can include a number of bits equal to or greater than the width of the data I/O bus of the DRAM chip 402f. In the illustrated example, each of the DRAM chips 402a-h has an 8-bit data I/O bus. Thus, for a cache line that is 64 bytes wide, each activated wordline segment (the wordline segment 610) of each DRAM chip 402a-h stores a portion or segment of a requested 64-byte cache line. When each of the DRAM chips 402a-h retrieves 8 bits from its respective wordline segment, the 8 bits from each of the DRAM chips 402a-h are concatenated to form 64 bits output by the memory module 400. This process can be repeated until the DRAM chips 402a-h have provided the entire 64-byte cache line. In such an example implementation, each of the DRAM chips 402a-h stores a respective 8 bytes of the 64-byte cache line.

[0059] FIG. 7 is a detailed diagram of the DRAM chip 402f of the example memory module 400 of FIGS. 4 and 6 showing the row decoder 604 and the column decoder 606 of the DRAM chip 402f working cooperatively to selectively activate particular bitlines of the DRAM chip 402f per memory access. The illustrated example of FIG. 7 shows how hierarchical wordlines can be used to

selectively activate portions of rows or wordlines without activating portions of the row that do not include data of interest (e.g., selectively activate the wordline segment 610 of the row 608 as shown in FIG. 6) to perform SBA memory accesses.

[0060] In the illustrated example of FIG. 7, two rows (or wordlines) 608 and 702 of the DRAM chip 402f are shown. Each row 608 and 702 includes a main wordline (MWL) and controlling sub-wordlines (SWL). In some example implementations, the MWL's can be formed in first-level metal and the SWLs can be formed in a poly layer. In the illustrated example, access to the row 608 is enabled through assertion of a MWL0 line 704 and access to the row 702 is enabled through assertion of a MWL1 line 706. The SWL's connect to memory bit cells in different memory cell arrays. In the illustrated example, a SWL 708 connects to the wordline segment 610 of FIG. 6 to enable selectively activating the wordline segment 610 to transfer its bits to the data I/O buffer 614.

[0061] To enable activating the SWLs, the DRAM chip 402f is provided with region select (RX) 'AND' gates, one of which is denoted by reference numeral 710. In addition, the DRAM chip 402f is provided with SWL 'AND' gates, one of which is denoted by reference numeral 712. The RX 'AND' gates (e.g., RX-0, RX-1, RX-N) and the SWL 'AND' gates enable selecting particular SWLs within respective columns of memory arrays. Although 'AND' gates are shown in the illustrated example, 'NOR' gates can be used instead, and the signal polarities of MWL and RX signal lines can be inverted to operate with the 'NOR' gates. In the illustrated example, a first input of the SWL 'AND' gate 712 is connected to the RX-N 'AND' gate 710 and a second input of the SWL 'AND' gate 712 is connected to the MWL0 line 704. In addition, the output of the SWL 'AND' gate 712 is connected to the SWL 708 to activate the SWL 708 upon assertion of the MWL0 line 704 and the output of the RX-N 'AND' gate 710. The other SWL 'AND' gates of the DRAM chip 402f are connected in an analogous manner so that any one of the SWLs of the DRAM chip 402f can be selected for outputting based on selecting a corresponding combination of an RX 'AND' gate (e.g., the RX-N 'AND' gate 710) and a MWL (e.g., the MWL0 line 704).

[0062] As shown in FIG. 7, the MWL0 line 704 and the MWL1 line 706 are connected to the row decoder 604, and the RX 'AND' gates are connected to the column decoder 606. In an example implementation that uses 15 bits for row addresses, the row decoder 604 can decode the selection of up to 32,768 MWLs. In an example implementation that uses 11 bits for column addresses, the inputs of the RX 'AND' gates can be connected to the least significant 7 bits of the column address interface of the column decoder 606 to enable individual selection of any one of up to 128 SWLs. In the illustrated example, when the row decoder 604 asserts a high signal on the MWL0 line 704 and the address output by the column decoder 606 selects the RX-N 'AND' gate 710, the inputs of the SWL 'AND' gate 712 are both high, causing the output of the SWL 'AND' gate 712 to activate the wordline segment 610 for outputting its bits onto the data I/O buffer 614.

[0063] In the illustrated example of FIG. 7, the MWL0 line 704 and the MWL1 line 706 are loaded only by a few SWL 'AND' gates, which significantly reduces the capacitance on the MWLs and, thus, their signal propagation delays. When only a subset of the CAS address is used to trigger RX signals associated with the RX 'AND' gates as discussed above, the activation area and wordline/bitline energy is advantageously reduced (e.g., less address lines are routed). In addition, energy is also advantageously conserved because the MWLs are not directly connected to the memory cells. As a result of not being directly connected to the memory cells, activation of an MWL across a memory array does not result in destruction of data, because only a small subset of memory cells (e.g., the wordline segment 610) connected to the active SWL read their data out.

[0064] In some example implementations, it may be die-area prohibitive or cost-prohibitive to provide the quantity of 'AND' gates (e.g., SWL 'AND' gates and RX 'AND' gates) to enable a single cache line per access. In such example implementations, less 'AND' gates can be provided with the tradeoff in energy required to activate and read out more cache lines per access. For example, the DRAM chip 402f depicted in FIG. 7 can be implemented with less 'AND' gates to instead read out 16 cache lines per access (e.g., the SWLs become 16

times longer between 'AND' gates). Such an example implementation would still produce relatively high energy savings over known DRAM chips.

[0065] As discussed above, the SBA technique advantageously reduces energy consumption during memory accesses while enabling interfacing memory controllers with memories in accordance with current (and likely future) industry standards for interfacing with DRAM memories. Another example memory access technique referred to herein as Single Subarray Access (SSA) organizes the layout of DRAM chips and the mapping of data to the DRAM chips so that an entire processor cache line can be fetched from a single DRAM chip (e.g., one of the memory chips 202a-d of FIG. 2) of a memory module (e.g., the DRAM PCB in-line memory module 200 of FIG. 2). The SSA memory access techniques are described below in connection with FIGS. 8-13.

[0066] Unlike known techniques for accessing cache lines from a DRAM memory module, which involve activating every DRAM chip of the memory module and concatenating the data output from each DRAM chip to form a single cache line, the example SSA memory access techniques illustrated in FIGS. 8-13, 15, and 16 involve reading an entire requested cache line from a single DRAM chip (e.g., one of the memory chips 202a-d of FIG. 2) of a memory module (e.g., the DRAM PCB in-line memory module 200 of FIG. 2). In this manner, the SSA technique eliminates the need to activate every DRAM chip in a memory module when accessing a requested cache line.

[0067] Turning to FIG. 8, a memory controller 802 is shown interfacing with a multi-chip DRAM memory module 804. In the illustrated example, the memory controller 802 is configured to access contents of the DRAM module 804 in accordance with the SSA technique described below. As shown, the DRAM module 804 includes eight DRAM chips 806a-h, each having an 8-bit output data bus. The DRAM chip 806g is shown in detail. The other DRAM chips 806a-f and 806h are substantially similar to the DRAM chip 806g and, thus, are not illustrated in detail. The DRAM chip 806g is provided with a plurality of banks, each bank having a plurality of subarrays and bitlines to access individual bit cells in the subarrays. For example, a bank 808 includes subarrays 810a and 810b, bitlines 812, and a row buffer 814. In the illustrated

example, each subarray of the DRAM chip 806g is 64 bytes wide. Thus, each row of the subarrays 810a-b can store a contiguous 64-byte wide cache line of data. In contrast, known DRAM memory modules store a cache line as several cache line portions or segments among several DRAM chips on the same memory module. For example, a known memory module having eight DRAM chips stores a 64-byte cache line by dividing the cache line into eight segments (e.g., eight 8-byte segments). Each of the eight cache line segments is then stored in a separate, respective DRAM chip such that access to the 64-byte cache line requires enabling all eight DRAM chips at the same time. Unlike those known DRAM memory modules, each of the SDRAM chips 806a-h of the example of FIG. 8 can store whole, contiguous cache lines. In this manner, when a particular cache line is requested by the memory controller 802, only the DRAM chip that stores the entire requested cache line needs to be activated while the other DRAM chips can remain in a sleep mode or other low-power mode to prevent excess power consumption (e.g., prevent energy leakage).

[0068] As shown in FIG. 8, the DRAM chip 806g is provided with a shared 8-bit external data I/O bus 816. Thus, a 64-byte wide cache line in a subarray is accessed using multiple consecutive row accesses to the DRAM chip 806g. Each of the row accesses causes the bitlines 812 to perform 8-bit read outs (or write ins) of the subarray to/from the shared 8-bit external data I/O bus 816 until the entire cache line has been accessed. In the illustrated example, the addressing organization used by the memory controller 802 to access the contents of the DRAM chips 806a-h enables selectively enabling specific ones of the DRAM chips 806a-h and specifying particular rows of the enabled DRAM chips 806a-h. For example, as described in greater detail below in connection with FIGS. 9 and 10, a portion of an address communicated by the memory controller 802 function as chip selects to enable selecting a specific one of the DRAM chips 806a-h per memory access.

[0069] In the illustrated example of FIG. 8, the SSA architecture is used in combination with the SBA technique of activating only necessary circuitry in a DRAM chip per memory access discussed above. Thus, because entire cache lines are stored in respective banks and subarrays (e.g., the subarrays 810a-b

of the bank 808), a memory access for a particular cache line requires a DRAM chip to activate only that portion of its circuitry that stores the requested cache line. For example, if the memory controller 802 requests a cache line stored in the subarray 810a, the DRAM chip 806g need only activate the circuitry for the bank 808 and the subarray 810a to retrieve the requested cache line, while maintaining the circuitry for the other banks and subarrays of the DRAM chip 806b (as well as the other DRAM chips 806a-f and 806h of the module 804) in a low-power mode. In this manner, a large portion of the accessed DRAM chip 802g as well as larger portions of the DRAM chips 806a-f and 806h can remain in a low-power mode while only some of the circuitry of the addressed chip (e.g., the DRAM chip 806g) is activated to retrieve a requested cache line. This approach realizes a significant power savings over known DRAM memory devices that require activating all banks during a memory access.

[0070] The SSA architecture can be used to dynamically reduce energy by activating a single chip and a single subarray (e.g., with short wordlines and the required quantity of bitlines) when accessing a cache line of memory. In addition, background energy can be reduced because a large fraction of on-chip die area is inactive at any point in time and can be placed in low-power modes, even when other portions (e.g., other subarrays and banks) of the memory are active. Also, the SSA architecture can be used to favorably impact performance by having relatively more banks per memory module and, thus, higher concurrency. For example, if each memory chip of a memory module has eight banks, the entire memory module has 64 separately accessible banks, each of which can store cache lines in their entireties. As such, performance can be positively impacted by adopting close-page access policies when there is low locality of stored data. In instances where there is high locality, the SSA architecture disclosed herein can be used in combination with probability techniques for approximating or estimating how row buffer hits can be used to issue prefetches to consecutive lines that may reside in other memory chips or banks.

[0071] In some examples, the SSA architecture of FIG. 8 is also configured to finish pre-charging a bank (e.g., the bank 808) when a corresponding DRAM

chip (e.g., the DRAM chip 806g) closes a row or subarray in that bank. In this manner, any subsequent access to the same bank will begin faster than if the bank were not pre-charged in this manner.

[0072] FIG. 9 is a detailed diagram of the DRAM module 804 of FIG. 8 implemented as a PCB in-line DRAM memory module that can provide concurrent data access to all of its DRAM chips 806a-h. In the illustrated example of FIG. 9, the DRAM module 804 is provided with a 64-bit segmented data bus 902. The segmented data bus 902 is physically divided into eight separate 8-bit data buses 904a-h, each of which corresponds to a respective one of the DRAM chips 806a-h. Using the example configuration of FIG. 9, two or more of the DRAM chips 806a-h can concurrently read out separate cache lines to the memory controller 802 (FIG. 8).

[0073] To select specific ones of the DRAM chips 806a-h and specify storage locations therein, the DRAM memory module 804 is provided with a memory module buffer 908 having an address buffer 910 and a demultiplexer 912. In the illustrated example, the address buffer 910 is connected to an address bus 914, and the demultiplexer 912 is connected to at least three of the most significant bits (MSB) 916 of the address bus 914. In some example implementations, three bits other than the MSB bits may be used instead. The three MSBs 916 of the address bus 914 are used by the demultiplexer 912 as chip selects to enable a specified one of the eight DRAM chips 806a-h for each memory access. During a memory access, the address buffer 910 latches address bits communicated by the memory controller 802, and a portion of the address bits (e.g., the three MSBs 916) are latched by the demultiplexer 912. The demultiplexer 912 decodes the three MSBs 916 to enable a specified one of the DRAM chips 806a-h (e.g., the DRAM chips identified by the three MSBs 916) and communicates the address bits from the address buffer 910 to the specified one of the DRAM chips 806a-h. The specified one of the DRAM chips 806a-h can then provide access (e.g., read/write access) to an entire cache line stored therein at the address location specified by the address latched by the address buffer 910. For example, if a read operation accesses the DRAM chip

806g, the DRAM chip 806g reads out a 64-byte cache line by outputting 8-bits per read cycle on its 8-bit data bus 904g over multiple read cycles.

[0074] In the illustrated example, the memory module buffer 908 is also provided with a control buffer 918 to latch, buffer, or otherwise store one or more memory access control signals from the memory controller 802. An example control signal is a write enable signal. The one or more buffered control signals can also be communicated to respective ones of the DRAM chips 806a-h via the demultiplexer 912.

[0075] In the illustrated example of FIG. 9, because each of the DRAM chips 806a-h interfaces to a separate one of the data buses 904a-h, multiple ones of the DRAM chips 806a-h can output respective cache lines simultaneously. For example, multiple concurrent memory access requests communicated by the memory controller 802 (FIG. 8) to the memory module buffer 908 can be pipelined in the memory module buffer 908 and consecutively communicated to respective ones of the DRAM chips 806a-h. In this manner, each of the selected one of the DRAM chips 806a-h can concurrently read out a requested cache line via its respective one of the data buses 904a-h. In such example implementations, the memory controller 802 is configured to organize concurrently received cache lines and to communicate them to a processor's cache memory (e.g., the cache memory 114 of the processor 112 of FIG. 1) in a coherent manner.

[0076] FIG. 10 is a diagram of an alternative multi-chip memory module structure 1000 that can be used to implement the DRAM module 804 of FIG. 8 to provide data access to only a single DRAM chip at a time (e.g., a single one of the DRAM chips 806a-h per cache line readout). In the illustrated example of FIG. 10, the DRAM module 1000 is provided with a shared one-chip wide data bus 1004, which is shown by way of example as an 8-bit wide data bus. The shared one-chip wide data bus 1004 is used by each of the DRAM chips 806a-h, one chip at a time. Using the example configuration of FIG. 10, only one of the DRAM chips 806a-h can read out data to the memory controller 802 (FIG. 8) per read out cycle.

[0077] To select specific ones of the DRAM chips 806a-h and specify storage locations therein, the DRAM memory module 1000 is provided with a memory module buffer 1008 having an address buffer 1010 and a demultiplexer 1012. In the illustrated example, the address buffer 1010 is connected to an address bus 1014, and the demultiplexer 1012 is connected to at least three of the most significant bits (MSB) 1016 of the address bus 1014. The three MSBs 1016 of the address bus 1014 are used by the demultiplexer 1012 as chip selects to enable a specified one of the eight DRAM chips 806a-h for each memory access. In addition, the memory module buffer 1008 is also provided with a control buffer 1018 to latch, buffer, or otherwise store one or more memory access control signals (e.g., a write enable signal) from the memory controller 802.

[0078] During a memory access, the address buffer 1010 latches address bits communicated by the memory controller 802, and a portion of the address bits (e.g., the three MSBs 1016) are used by the demultiplexer 1012. In addition, the control buffer 1018 latches one or more control signals. The demultiplexer 1012 decodes the three MSBs 1016 (or other three selected bits instead of the three MSBs) to enable a specified one of the DRAM chips 806a-h and communicates the address bits from the address buffer 1010 and the control signal(s) from the control buffer 1018 to the specified one of the DRAM chips 806a-h. The three MSBs1016 can, thus, be thought of as selecting the output of the demultiplexer 1012 to which the remaining address bits are routed. The specified one of the DRAM chips 806a-h can then provide access (e.g., read/write access) to an entire cache line stored therein at the address location specified by the address latched by the address buffer 1010. For example, if a read operation accesses the DRAM chip 806g, the DRAM chip 806g reads out a 64-byte cache line by outputting 8-bits per read cycle on the shared one-chip wide data bus 1004 over multiple read cycles.

[0079] FIG. 11 is a diagram of another multi-chip DRAM memory module 1100 that can be used to implement the DRAM module 804 of FIG. 8 to selectively activate particular DRAM chips 1102a-h per memory access in accordance with SSA techniques. The DRAM memory module 1100 can be

implemented as a PCB in-line memory module or any other type of memory module. In addition, a data bus (not shown) of the DRAM memory module 1100 can be implemented as a segmented data bus such as the segmented data bus 902 of FIG. 9 to enable concurrent memory access in separate ones of the DRAM chips 1102a-h, or the data bus may be implemented as a shared one-chip wide data bus such as the shared one-chip data bus 1004 of FIG. 10, or the data bus may be implemented using some combination thereof.

[0080] In the illustrated example of FIG. 11, the DRAM memory module 1100 is provided with a memory module buffer 1104 having an address buffer 1106 and a control buffer 1108. Upon receiving an address from a memory controller (e.g., the memory controller 802 of FIG. 8), the address buffer 1106 latches the address bits and communicates the address in broadcast fashion to all of the DRAM chips 1102a-h, unlike the example implementations of FIGS. 9 and 10, in which the respective memory module buffers 908 and 1008 use chip selects and demultiplexers to selectively communicate an address to only a relevant DRAM chip. In addition, the control buffer 1108 latches one or more control signal(s) (e.g., write enable control signals) and also communicates the control signal(s) in broadcast fashion to all of the DRAM chips 1102a-h. In the illustrated example of FIG. 11, each of the DRAM chips 1102a-h is configured to determine whether received addresses specify their storage locations or whether addresses can be ignored. When a DRAM chip determines that it can ignore an address, it can remain in a sleep mode or other low-power mode. When a DRAM chip determines that an address is relevant, the DRAM chip is activated and provides access to the specified memory contents.

[0081] In some examples the memory module buffer 1104 may be omitted and memory module buffers substantially similar or identical to the memory module buffer 1104 may be implemented in each of the DRAM chips 402a-h such that the memory module 1100 operates the same as if it were provided with the memory module buffer 1104 as shown in FIG. 11.

[0082] In the illustrated example of FIG. 11, DRAM chip 1102f is shown in detail. The other DRAM chips 1102a-e, 1102g, and 1102h are substantially similar in structure and operation to the DRAM chip 1102f and, thus, to avoid

redundancy, are not illustrated. To determine whether to decode or ignore received addresses, the DRAM chip 1102f is provided with a tri-state buffer 1110 that receives address bits from the address buffer 1106 and uses a portion of the address bits (e.g., the MSBs) as a chip select input. The DRAM chip 1102f uses the chip select input of the tri-state buffer 1110 to determine whether received addresses are relevant. By providing the tri-state buffer 1110, the DRAM chip 1102f can be selectively activated to provide access to an entire requested cache line stored therein.

[0083] If the chip select input of the tri-state buffer 1110 enables the DRAM chip 1102f, the tri-state buffer 1110 communicates the address to a demultiplexer 1112, which is configured to identify whether received address bits correspond to row addresses or column addresses based on assertions of RAS/CAS signals. The DRAM chip 1102f is configured to store received row addresses in a row address register 1114 and received column addresses in a column address register 1116 before decoding the addresses and activating a corresponding storage area. In the illustrated example of FIG. 11, once the DRAM chip 1102f is activated or selected based on the provided address information, the DRAM chip 1102f is configured to act on one or more control signal(s) from the control buffer 1108 and decode the row and column addresses to selectively activate specific storage areas (e.g., SWLs) as discussed above in connection with FIGS. 4, 6, and 7 to provide access to the requested data.

[0084] FIG. 12 illustrates an example cache line storage configuration that can be used in connection with any of the example DRAM memory modules of FIGS. 8-11. In the illustrated example, four DRAM chips 1202a-d of a multi-chip memory module 1200 are shown. In the illustrated example, to distribute load and increase concurrency, data is interleaved between the DRAM chips 1202a-d such that consecutive cache lines are first placed in different ones of the DRAM chips 1202a-d and then in different banks of each DRAM chip. In the illustrated example, each of the DRAM chips 1202a-d includes four banks. The DRAM chip 1202a is shown as having banks 0-3. Cache lines C0-C15 are distributed among the different DRAM chips 1202a-d and different ones of the

respective banks of the DRAM chips 1202a-d in a chip-bank-interleaved fashion using the least significant bits (LSBs) of a cache line address as bank select lines. For example, cache line C0 is stored in bank 0 of the DRAM chip 1202a and cache line C1 is stored in bank 0 of the DRAM chip 1202b. When a last DRAM chip is reached (e.g., the DRAM chip 1202d), the next sequential cache line is stored in a sequentially next bank of the first DRAM chip. For instance, as shown in FIG. 12, cache line C3 is stored in bank 0 of the DRAM chip 1202d and cache line C4 is stored in bank 1 of the DRAM chip 1202a.

[0085] FIG. 13 is a detailed diagram illustrating the example memory controller 802 of FIG. 8. The memory controller 802 is configured to be connected between a processor (e.g., the processor 112 of FIG. 1) and one or more memory devices (e.g., the memory modules of FIGS. 8-11). In some example implementations, the memory controller 802 can be implemented as a standalone IC chip separate from a processor chip. Alternatively, the memory controller 802 may be formed on the same die or located in the same IC package as a processor. In the illustrated example, a single address bus and a single command line bus are used to communicate with all DRAM chips on a memory module.

[0086] The memory controller 802 can be configured to provide an address in two transfers. This is particularly advantageous when pin limitations arise for a memory module or an IC package of the memory controller 802. Such a two-transfer address is similar to transferring separate row and column addresses in a conventional manner. However, the address transfers of the memory controller 802 need not necessarily correspond to row and column addresses. As discussed above in connection with FIG. 8, cache lines are stored in DRAM chips using a bank/subarray arrangement, which is a different organization from row/column arrangements in known DRAM chips. Thus, the cache lines can be addressed using a single address that is transferred by the memory controller 802 to a memory module in two separate address cycles. In the illustrated example, the address bits from both transfers identify a unique subarray (e.g., one of the subarrays 810a or 810b) and row (i.e., a cache line) within that subarray. Therefore, an entire address is required before a subarray can be

identified or accessed. In addition, a portion of the address (e.g., the MSBs of the address) identifies the DRAM chip that stores the requested cache line, and that portion of the address can be used as a chip select as discussed above in connection with FIGS. 9-11 to activate a corresponding DRAM chip storing the requested cache line.

[0087] In the illustrated example of FIG. 13, the memory controller 802 includes a processer address interface 1302 and a processor data interface 1304, each of which is configured to communicate with a processor (e.g., the processor 112 of FIG. 1). To translate addresses received from a processor into address and control bits to access memory contents in a memory module (e.g., the memory modules of FIGS. 9-12), the memory controller 802 is provided with an address translator 1306. For example, the address translator 1306 can split an address into two separately transferrable portions so that a lengthy address can be transferred to a memory module using two separate address transfers via a memory address interface 1308. To facilitate a two-transfer address, the address translator 1306 can cause memory interface control lines 1310 to assert corresponding control signal lines to inform a memory module of the ordering of the address transfers. For example, a first address transfer signal (not shown) of the memory interface control lines 1310 can be asserted to indicate that a first portion of an address is being output by the memory address interface 1308, while a second address transfer signal can be asserted to indicate that a second portion of an address is being output by the memory address interface 1308.

[0088] In some example implementations, in which virtual-to-physical address memory translation is performed as part of memory controller operations, the address translator 1306 can also translate addresses between a virtual memory map used by the processor and a physical memory map corresponding to the addressing scheme of physical memory modules connected to the memory controller 802. In some example implementations, the address translator 1306 may also update a physical memory map when information is stored in or deleted from memory. If virtual-to-physical memory address translation is not

performed as part of memory controller operations, such memory address translations can be performed in a processor.

[0089] To transfer data between the memory controller 802 and memory modules, the memory controller 802 is provided with one or more data buses. In the illustrated example, the memory controller 802 is provided with eight separate, narrow channel, data bus interfaces 1312a-h. In this manner, the memory controller 802 can access entire cache lines from respective individual DRAM chips as discussed above in connection with FIGS. 8-12 using respective data bus interfaces 1312a-h. In the illustrated example, each of the data bus interfaces 1312a-h includes eight data lines. In some example implementations, the memory controller 802 may be provided with a single narrow channel data bus when configured to interface with memory modules having a corresponding single narrow channel data bus such as the shared one-chip data bus 1004 of FIG. 10.

[0090] To concatenate cache line segments to form entire cache lines, the memory controller 802 is provided with data concatenators 1314a-h. Each of the data concatenators 1314a-h may be implemented using a shift register that receives a number of bits at a time equal to the width of one of the data bus interfaces 1312a-h and shifts the bits per read access until a complete cache line is formed. The data concatenators 1314a-h can communicate their cache lines to the processor data interface 1304 for communication to a processor. In the illustrated example, the data concatenators 1314a-h can operate as queues for data to be communicated to a processor, and the processor data interface 1304 can arbitrate communications from the data concatenators 1314a-h to communicate cache lines to the processor in an order expected by the processor.

[0091] In some examples, the memory controller 802 is provided with separate data queues for each bank in a DRAM chip (e.g., each of the DRAM chips 1102a-h of FIG. 11). For example, as shown in FIG.13, the data concatenator 0 1314a is connected to a data queue array 1324 having separate data queues for each bank of a DRAM chip corresponding to the data concatenator 0 1314a. In particular, the data queue array 1324 includes a B0

queue 1326a for a first bank of an SDRAM chip, a B1 queue 1326b for a second bank of the SDRAM chip, a B2 queue 1326c for a third bank of the SDRAM chip, and a B3 queue 1326d for a fourth bank of the SDRAM chip. In the illustrated example, the bank queues 1326a-d are configured to hold or queue data received from an SDRAM chip for subsequently forwarding to a processor. In the illustrated example, the bank queues 1326a-d enable buffering data that is received faster than it can be communicated to a requesting processor. In this manner, requested data can be quickly received at the memory controller 802 from a memory so that the memory can be freed-up to service other memory requests. Although not shown in FIG. 13, each of the other data concatenators 1314b-h is also provided with a respective data queue array substantially similar or identical to the data queue array 1324. That is, in the illustrated example, the quantity of bank queues (substantially similar or identical to the bank queues 1326a-d) provided in the memory controller 802 is equal to the number of banks in a single DRAM chip multiplied by the number of DRAM chips in a memory module (e.g., the memory module 1100 of FIG. 11).

[0092] To enable error detection and correction in connection with memory accesses, the memory controller is provided with an error detector 1316, a checksum generator 1318, a parity generator 1320, and a data reconstructor 1322. The memory controller 802 of the illustrated example uses the error detector 1316 to verify checksums stored in association with cache lines in memory chips as discussed below in connection with FIGS. 17 and 18 and detect errors in stored information. That is, when an error exists in a data storage location storing some or all of a cache line (e.g., a cache line 1704 of FIG. 17), the error is exposed by checksum information (e.g., a checksum 1706 of FIG. 17) corresponding to the cache line and the error detector 1316 is configured to detect the error based on the checksum information. The memory controller 802 of the illustrated example uses the checksum generator 1318 to generate new checksums each time the memory controller 802 writes data to a memory. The memory controller 802 of the illustrated example uses the parity generator 1320 to generate updated parity values each time the memory controller 802 writes data to a memory. The memory controller 802 of the

illustrated example uses the data reconstructor 1322 to reconstruct uncorrupted data when the error detector 1316 detects an error in a cache line based on corresponding checksum information exposing the error. The storage locations of checksum and parity values may be stored in a physical memory map in association with storage locations of their respective cache lines.

[0093] FIG. 14 is a flow diagram illustrating an example process that can be used to perform SBA memory accesses as described above in connection with FIGS. 4, 6, and 7. The example process of FIG. 14 is described below as performed by a single DRAM chip (e.g., the DRAM chip 402f of FIGS. 4, 6, and 7). However, the same process may be implemented in similar manner by any of the DRAM chips 402a-h of the memory module 400 (FIG. 4) or DRAM chips of other memory structures. Initially, the DRAM chip 402f receives a row address (block 1402), selects the row address register 412 (FIG. 4) (block 1404), and stores the row address (block 1406) in the row address register 412. For example, the DRAM chip 402f can select the row address register 412 based on an assertion of a RAS signal received at the demultiplexer 416.

[0094] The DRAM chip 402f receives a column address (block 1408), selects the column address register 414 (FIG. 4) (block 1410), and stores the column address (block 1412) in the column address register 414. The row decoder 604 and column decoder 606 (FIGS. 6 and 7) concurrently decode the row and column addresses (block 1414). The row decoder 604 activates a MWL (e.g., one of the MWL0 line 704 or MWL1 line 706 of FIG. 7) corresponding to the row address (block 1416). The column decoder 606 enables a RX line (e.g., one of the RX-0, RX-1, ... RX-N 'AND' gates) corresponding to the column address (or a portion of the column address) (block 1418). By activating the MWL at block 1416 and one of the RX lines at block 1418, a SWL 'AND' gate (e.g., the SWL 'AND' gate 712 of FIG. 7) outputs a high signal to enable a requested wordline segment of data (e.g., the wordline segment 610 of FIGS. 6 and 7). The row buffer 612 (FIG. 6) loads the activated wordline segment data 610 (block 1420), and the data I/O buffer 614 (FIGS. 6 and 7) outputs the wordline segment data 610 (block 1422). The example process of FIG. 14 is then ended.

[0095] FIG. 15 is a flow diagram illustrating an example process that can be used to perform SSA memory accesses as described above in connection with FIGS. 8-10, 12, and 13. The example process of FIG. 15 is described below as performed by the memory module 800 of FIG. 8. However, the same process may be implemented in similar manner by any of the other memory modules depicted in FIGS. 9-10 or other memory structures. Initially, the memory module 800 receives an address (block 1502), and the address buffer 910 (FIG. 9) stores the address (block 1504). For example, as discussed above, the address can be received in two address transfers and concatenated by the address buffer 910 for communication to a corresponding one of the DRAM chips 806a-h.

[0096] The demultiplexer 912 then selects one of the DRAM chips 806a-h corresponding to the received address (block 1506). For example, the dumultiplexer 912 decodes a portion (e.g., the MSBs) of the address bits communicated across the address bus 914 to select a corresponding one of the DRAM chips 806a-h. The demultiplexer 912 then communicates the address from the address buffer 910 to the selected one of the DRAM chips 806a-h (block 1508). In some example implementations, the address from the address buffer 910 can be communicated in two transfers to a selected DRAM chip if the number of address bits exceeds the pin count of the address interface of the DRAM chip.

[0097] The selected one of the DRAM chips (e.g., the DRAM chip 806g) decodes the received address (block 1510) and activates a bank (e.g., the bank 808 of FIG. 8) (block 1512) specified by the address and a subarray (e.g., one of the subarrays 810a or 810b) in the active bank that is also specified by the address (block 1514). The DRAM chip 806g then activates a row (block 1516) also specified by the address and selects a first cache line segment (block 1518) of the cache line stored in the active row. A cache line segment is a sub-group of bits that form a portion of an entire cache line. Thus, if a cache line is 64 bytes wide, a cache line segment can be 8 bytes wide so that the 64-byte cache line can be read out using multiple 8-bit read cycles.

[0098] The row buffer 814 (FIG. 8) loads the selected cache line segment data (block 1520), and the shared 8-bit external data I/O bus 816 outputs the selected cache line segment data (block 1522). If there is another cache line segment remaining (e.g., another segment of a 64-byte wide cache line) (block 1524), the DRAM chip 806g selects the next cache line segment (block 1526) and control returns to block 1520. When no other cache line segments remain to be output (block 1524), this signifies that the DRAM chip 806g has completed reading out an entire cache line. In some examples, the DRAM chip 806g closes the active row (or subarray) and finishes pre-charging the bank in which the row (or subarray) is located to enable subsequent accesses to the same bank to begin faster than if the bank were not pre-charged in this manner. The example process of FIG. 15 is then ended.

[0099] FIG. 16 is a flow diagram illustrating an example process that can be used to perform SSA memory accesses as described above in connection with FIG. 11. The example process of FIG. 16 is described below as performed by the memory module 1100 of FIG. 11; however, other memory structures could be used. Initially, the memory module 1100 receives an address (block 1602), and the address buffer 1106 (FIG. 11) stores the address (block 1604). For example, as discussed above, the address can be received in two address transfers and concatenated by the address buffer 1106 for communication to the DRAM chips 1102a-h.

[00100] The memory module buffer 1104 (FIG. 11) communicates the address from the address buffer 1106 to all of the DRAM chips 1102a-h (block 1606). In some example implementations, the address from the address buffer 1106 can be communicated in two transfers to the DRAM chips 1102a-h if the number of address bits exceeds the pin count of the address interface of the DRAM chips 1102a-h.

[00101] Each of the DRAM chips 1102a-h determines whether to decode or ignore the received address (block 1608). For example, as discussed above in connection with FIG. 11, the tri-state buffer 1110 of the DRAM chip 1102f can decode a portion (e.g., the MSBs) of the address as if it were a chip select to

determine whether the DRAM chip 1102f is intended to act upon the received address.

[00102] The DRAM chip specified by the address (e.g., the DRAM chip 1102f) decodes the address (block 1610). The specified DRAM chip activates a bank (e.g., the bank 808 of FIG. 8) (block 1612) specified by the address and a subarray (e.g., one of the subarrays 810a or 810b) in the active bank that is also specified by the address (block 1614). The addressed DRAM chip then activates a row (block 1616) also specified by the address and selects a first cache line segment (block 1618) of the cache line stored in the active row. A cache line segment is a sub-group of bits that form part of an entire cache line. Thus, if a cache line is 64 bytes wide, a cache line segment can be 8 bytes wide so that the 64-byte cache line can be read out using multiple 8-bit read cycles.

[00103] A row buffer (e.g., the row buffer 814 of FIG. 8) of the addressed DRAM chip loads the selected cache line segment data (block 1620). A shared 8-bit external data I/O bus of the addressed DRAM chip outputs the selected cache line segment data (block 1622). If there is another cache line segment remaining (e.g., another segment of a 64-byte wide cache line) (block 1624), the addressed DRAM chip selects the next cache line segment (block 1626) and control returns to block 1620. When no other cache line segments remain to be output (block 1624) (i.e., the addressed DRAM chip has completed reading out an entire requested cache line), the example process of FIG. 16 is ended. In some examples, the addressed DRAM chip closes the active row (or subarray) and finishes pre-charging the bank in which the row (or subarray) is located to enable subsequent accesses to the same bank to begin faster than if the bank were not pre-charged in this manner.

[00104] FIG. 17 depicts an example multi-chip memory module 1700 configured to store error-checking and parity information to enable error-checking and data correction operations when accessing data stored therein. In the illustrated example, the memory module 1700 includes nine memory chips 1702a-i. The memory module 1700 can be advantageously used to implement memory modules that enable SSA memory accesses as described above in connection with FIGS. 8-13, 15, and 16. That is, the memory structures

described above in connection with FIGS. 8-12 enable retrieving an entire cache line from a single DRAM chip. To facilitate detecting corrupted cache lines, the memory module 1700 stores checksums in association with each stored cache line. In addition, to facilitate reconstructing corrupted cache lines, the memory module 1700 stores parity information among all of the memory chips 1702a-i.

[00105] In the illustrated example of FIG. 17, the architecture of the memory module 1700 is adapted so that each of the memory chips 1702a-i can store entire cache lines and a corresponding checksum for each cache line. In FIG. 7, a cache line L0 1704 is stored by the first memory module 1702a in association with a corresponding checksum (C) 1706. Storing a respective checksum for each cache line in this manner requires additional storage space of the memory module 1700. For example, storing an 8-byte checksum for each 64-byte cache line incurs a 1.625% storage overhead. To retrieve a cache line (e.g., the cache line L0 1704) and its corresponding checksum (e.g., the checksum (C) 1706), a larger burst length can be used. For example, nine 8-byte data reads can be used to retrieve a 64-byte cache line and its 8-byte checksum rather than using only eight 8-byte data reads to retrieve only the 64-byte cache line. Similarly, nine 8-byte data writes can be used to write a 64-byte cache lines and its checksum to a corresponding storage location.

[00106] In the illustrated example of FIG. 17, the memory module 1700 stores a parity value corresponding to eight cache lines. To accommodate the additional storage required for parity information, the memory module 1700 is provided with nine memory modules 1702a-i instead of only eight memory modules. The parity values are stored among different ones of the memory chips 1702a-i. For example, as shown in FIG. 17, a parity value P0 1708 corresponding to cache lines L0-L7 is stored in the memory module 1702i. Further, the parity value P0 1708 is stored in association with a checksum (C) 1710 corresponding to the parity value P0 1708. When an error is detected in a retrieved cache line, an uncorrupted cache line can be reconstructed by reading a corresponding parity value and the other cache lines corresponding to that parity value. For example, if an error is detected upon reading the cache line L0

1704, an uncorrupted instance can be reconstructed (e.g., by the data reconstructor 1322 of FIG. 13) based on the parity value P0 1708 and cache lines L1-L7 using, for example, known reconstruction techniques associated with the redundant array of inexpensive disks (RAID) techniques of RAID-4 and/or RAID-5 data reconstruction processes.

[00107] During a read access, checksum verification can be performed by a memory controller such as the memory controller 802 of FIG. 13. For example, a checksum verification function can be performed by the error detector 1316 (FIG. 13) using bit inversion to ensure detection of stuck-at-zero faults. If an error is detected, the data reconstructor 1322 (FIG. 13) can reconstruct the requested cache line based on the stored parity information. To enable the checksum verification, during a read access of data at a storage location an additional burst is employed (e.g., during the same read cycle as the read access of the data) by the memory module 1700 to also read out a corresponding checksum value to enable the error detector 1316 (FIG. 13) of the memory controller 802 to perform the checksum verification. For example, while servicing a read request of the cache line L0 1704, the memory chip 1702a reads out the requested cache line L0 1704 and uses an additional burst to also read out the checksum value 1706 corresponding to the cache line L0 1704. In this manner, the memory chip 1702a provides the requested data and the corresponding stored checksum to the requesting memory controller 802 to detect any errors that may exist in the requested data. In the illustrated example, the memory controller 802 reads in cache lines (data), checksums, and parity values via the data bus interfaces 1312a-g (FIG. 13).

[00108] During a write access, the memory controller 802 writes a cache line and corresponding checksum to their designated addressed location in a memory chip (e.g., one of the memory chips 1702a-i). In addition, the memory controller 802 writes an updated parity value corresponding to the written cache line. To perform such a write operation, the memory controller 802 must first read out the old cache line (e.g., the cache line L0 1704) and checksum (e.g., the checksum 1706) from the memory chip (e.g., the memory chip 1702a), read out the old corresponding parity value (e.g., the parity value P0 1708) from the

memory chip, write the new cache line and checksum (in replacement of the old cache line and checksum), and write the corresponding updated parity value (in replacement of the old parity value). For each write operation, the memory controller 802 can use the checksum generator 1318 (FIG. 13) to generate new checksum values and use the parity generator 1320 (FIG. 13) to generate new parity values. In the illustrated example, the memory controller 802 reads in and writes out cache lines (data), checksums, and parity values via the data bus interfaces 1312a-g (FIG. 13).

[00109] FIG. 18 is a flow diagram illustrating an example process that can be used to perform error-checking and data correction operations when accessing data in a memory. The example process of FIG. 18 is described in connection with the example memory controller 802 of FIG. 13 and the memory module 1700 of FIG. 17, but other processing devices and/or memories may be used in addition to or instead of the memory controller 802. Memory accesses performed in connection with the example process of FIG. 18 can be performed using SSA memory accesses as described above in connection with FIGS. 8-12.

[00110] Initially, the memory controller 802 determines whether it is to perform a read or write operation (block 1802). If the memory controller 802 is to perform a read operation, the memory address interface 1308 communicates an address to the memory module 1700 (block 1804). The address can be communicated as two separate address portions in two transfer cycles. The memory controller 802 then receives a cache line (e.g., the cache line L0 1704 of FIG. 17) and corresponding checksum (e.g., the checksum (C) 1706 of FIG. 17) (block 1806) via one of the data bus interfaces 1312a-h (FIG. 13). For example, the cache line and checksum can be latched into a corresponding one of the data concatenators 1314a-h (FIG. 13).

[00111] The error detector 1316 performs a data verification operation on the received cache line based on the received checksum value (block 1808). If an error is detected (block 1810), the data reconstructor 1322 performs a data reconstruction operation (block 1812). For example, the data reconstructor 1322 can cause the memory controller 802 to retrieve a parity value (e.g., the

parity value 1708 of FIG. 17) corresponding to the corrupted cache line detected at block 1808 and all of the other cache lines (e.g., the cache lines L1-L7 of FIG. 7) corresponding to the retrieved parity value. The data reconstructor 1322 can then reconstruct an uncorrupt version of the requested cache line based on the other cache lines and the retrieved parity values using, for example, known reconstruction techniques associated with RAID-4 and/or RAID-5 data reconstruction processes.

[00112] If an error is not detected (block 1810) or after the data reconstructor 1322 reconstructs an uncorrupted cache line, the processor data interface 1304 (FIG. 13) communicates the requested cache line to a requesting processor (block 1814) and the process of FIG. 18 is ended.

[00113] If at block 1802 the memory controller 802 determines that it needs to perform a write operation, the processor data interface 1304 receives the data to be written from a processor (block 1816). The checksum generator 1318 generates a checksum value for the data to be written based on the data (block 1818), and one of the data concatenators 1314a-h concatenates the data to be written with the corresponding checksum value (block 1820). The parity generator 1320 (FIG. 13) causes the memory controller 802 to read the original data and checksum stored at the data location to be written and to read the parity value from the memory module 1700 (block 1822). For example, if the write operation is to write a value to the cache line L0 1704 (FIG. 17), the memory controller 802 reads the original value of the cache line L0 1704, the checksum value 1706, and the parity value 1708 and its checksum 1710.

[00114] The parity generator 1320 generates a new parity value based on the data to be written and its checksum value generated at block 1818 (block 1824). The checksum generator 1318 generates a new checksum value for the new parity value based on the parity value (block 1826). A corresponding one of the data concatenators 1314a-h concatenates the updated parity value with its new checksum value (block 1828). The memory controller 802 then writes the new cache line value and its checksum value to the memory module 1700 (block 1830) and writes the updated parity value and its checksum value to the

memory module 1700 (block 1832). The example process of FIG. 18 is then ended.

[00115] In some example implementations, one or more of the example processes of FIGS. 14-16 and/or 18 may be implemented using machine readable instructions that, when executed, cause a device (e.g., a programmable controller or other programmable machine or integrated circuit) to perform the operations shown in FIGS. 14-16 and/or 18. For instance, the example processes of FIGS. 14-16 and/or 18 may be performed using a processor, a controller, and/or any other suitable processing device. For example, the example processes of FIGS. 14-16 and/or 18 may be implemented in coded instructions stored on a tangible machine readable medium such as a flash memory, a read-only memory (ROM), and/or a random-access memory (RAM) associated with a processor or controller. As used herein, the term tangible computer readable medium is expressly defined to include any type of computer readable storage and to exclude propagating signals. Additionally or alternatively, the example processes of FIGS. 14-16 and/or 18 may be implemented using coded instructions (e.g., computer readable instructions) stored on a non-transitory computer readable medium such as a flash memory, a read-only memory (ROM), a random-access memory (RAM), a cache, or any other storage media in which information is stored for any duration (e.g., for extended time periods, permanently, brief instances, for temporarily buffering, and/or for caching of the information). As used herein, the term non-transitory computer readable medium is expressly defined to include any type of computer readable medium and to exclude propagating signals.

[00116] Alternatively, the example processes of FIGS. 14-16 and/or 18 may be implemented using any combination(s) of application specific integrated circuit(s) (ASIC(s)), programmable logic device(s) (PLD(s)), field programmable logic device(s) (FPLD(s)), discrete logic, hardware, firmware, etc. Also, the example processes of FIGS. 14-16 and/or 18 may be implemented as any combination(s) of any of the foregoing techniques, for example, any combination of firmware, software, discrete logic, and/or hardware. Further, although the example processes of FIGS. 14-16 and/or 18 are described with

reference to the flow diagrams of FIGS. 14-16 and/or 18, other methods of implementing the processes of FIGS. 14-16 and/or 18 may be employed. For example, the order of execution of the blocks may be changed, and/or some of the blocks described may be changed, eliminated, sub-divided, or combined. Additionally, any or all of the example processes of FIGS. 14-16 and/or 18 may be performed sequentially and/or in parallel by, for example, separate processing threads, processors, devices, discrete logic, circuits, etc.

[00117] FIGS. 19-20 depict example test settings and example memory energy parameters associated with analyzing the energy consumption and memory access performance of memory access techniques and structures disclosed herein relative to known memory access techniques and structures. As discussed below, the memory access techniques and structures disclosed herein advantageously impact the energy consumption and memory access performance of memory subsystems over known memory access techniques and structures.

[00118] FIG. 19 illustrated an example test bench parameters table 1900 showing example test bench parameters of an example processor system model used to analyze the performance of the methods, apparatus, and articles of manufacture disclosed herein. The example performance analyses discussed below were conducted by modeling a baseline 8-core out-of-order processor with private L1 caches and a shared L2 cache as indicated in the test bench parameters table 1900. For purposes of the performance analyses, baseline structures were implemented or modeled in accordance with a Micron SDRAM memory device (Micron part MT47H256M8).

[00119] For purposes of the performance analyses, the main memory capacity was set to 4 gigabytes and organized as shown in the test bench parameters table 1900. A VIRTUTECH® Simics full-system simulator was used to implement a simulation infrastructure with out-of-order timing supported by Simics' 'ooo-micro-arch' module. (Simics is a full-system simulator used to run production binaries of target hardware at high-performance speeds.) The 'trans-staller' module was heavily modified to accurately capture DRAM device timing information including multiple channels, ranks, banks and open rows in

each bank. Both open-row and close-row page management policies with first-come-first-serve (FCFS) and first-ready-first come-first-serve (FR-FCFS) scheduling with appropriate queuing delays were modeled. In addition, overlapped processing of commands by the memory controller were modeled to hide pre-charge and activation delays. Also, bus models were used to model transfers of data between a memory controller and DIMMs. Address mapping policies were adopted from the DRAMSim framework (a memory system simulator). DRAM timing information was obtained from a Micron datasheet for part MT47H256M8.

[00120] Area, latency, and energy numbers of DRAM banks were obtained from CACTI 6.5 and were modified to include accurate models for commodity DRAM, both for the baseline memory design and with hierarchical wordlines. CACTI is a model that integrates models for cache and memory access time, cycle time, area, leakage, and dynamic power. By integrating all of these models together, tradeoffs between time, power, and area are all based on the same assumptions and, hence, are mutually consistent.

[00121] By default CACTI divides a large DRAM array into a number of mats and employs an H-tree to connect various mats. Such an organization incurs low latency but requires large area. However, traditional DRAM banks are heavily optimized for area to reduce cost and employ very large arrays with minimal peripheral circuitry overhead. In addition, a read or write operation is done using long bitlines spanning the array instead of using an H-tree interconnect. For the performance analyses, CACTI was modified to reflect a typical commodity DRAM implementation. While modeling hierarchical wordlines along with hierarchical bitlines, the number of MWL to SWL connectors can be reduced and, thus, the array area can also be reduced by trading off for increased bitline energy. For example, performing selective bitline activation in a second level bitline and reading all bitlines in a first level will increase energy but can significantly reduce the array area.

[00122] For purposes of the SBA performance analyses, a particular analytical model for hierarchical wordlines was incorporated into CACTI 6.5 to quantify the overhead of required die array area. A high area overhead was incurred when

enough SWLs were introduced to activate exactly one cache line in a bank. This is because the area overhead introduced by the SWL 'AND' gate (e.g., the SWL 'AND' gate 712 of FIG. 7) and RX 'AND' gate (e.g., the RX 'AND' gate 710 of FIG. 7) signals was needed every few memory cells in the array. While this results in activating a minimum number of bitlines, the die-area cost may be prohibitive. However, energy can be traded off for lower cost by not being as selective. For example, 16 cache lines were read out instead of a single cache line per access, the SWLs become 16 times longer (e.g., 16 times less 'AND' gates). This still leads to relatively higher energy savings over the baseline, and an area overhead of only 12% over the baseline. The 16 cache line read outs are used in a close-page policy approach.

[00123] The example DRAM energy parameters used to conduct the performance analyses are shown in FIG. 20. In particular, FIG. 20 shows a table 2000 including example DRAM energy parameters of the memory access techniques and structures disclosed herein and known memory access techniques and structures. The baseline energy values shown in table 2000 are based on the Micron datasheet for part MT47H256M8 and Micron's System Power Calculator (http://www.micron.com/support/part info/powercalc). The SBA and SSA memory access techniques and structures disclosed herein were evaluated using subsets of the known multi-threaded PARSEC (Princeton Application Repository for Shared-Memory Computers), NAS (NASA Advanced Supercomputing), and STREAM (a simple synthetic benchmark program that measures sustainable memory bandwidth (in megabytes (MB)/s) and the corresponding computation rate for simple vector kernels) benchmark suites. During the analyses, applications were run for one million DRAM accesses and reported total energy consumption and instructions per cycle (IPC).

[00124] FIGS. 21 and 22 show performance graphs of row access operations for traditional memory devices in a code execution environment. In particular, graph 2100 of FIG. 21 shows row buffer hit rates in a code execution environment as a function of quantities of chips in traditional memories and graph 2200 of FIG. 22 shows row re-use counts in a code execution environment based on an 8-chip traditional memory. As shown in graph 2100,

row-buffer hit rates steeply drop as the number of threads that simultaneously access memory increases. For instance, graph 2100 shows average rates drop from 67% for a 1 core system to 34% for a 16 core system. In addition, as shown in graph 2200 of FIG. 22, whenever a row is fetched into a row buffer (e.g., the row buffer 508 of FIG. 5), the number of times it is used before being closed due to a conflict is often just one or two times. This indicates that even on benchmarks with high locality and relatively good average row buffer hit rates (e.g., the NAS conjugate gradient (cg) benchmark), a large number of pages still exhibit relatively low reuse in the row-buffer. This means that the energy costs of activating an entire 8 kilobyte (KB) row (as in traditional memory access techniques) is amortized over only very few accesses, which wastes a significant amount of energy.

[00125] Turning now to FIG. 23, a graph 2300 compares example normalized DRAM energy consumption for row/wordline accesses using known memory access techniques and structures and the memory access techniques and structures disclosed herein. In particular, graph 2300 shows the energy consumption using a close-row baseline, SBA memory accesses, and SSA memory accesses, normalized to the open-row baseline. As shown, the close-row baseline is relatively worse in terms of energy consumption than the open-row baseline. For instance, graph 2100 shows an average increase in energy consumption by 73%, with numbers for individual benchmarks varying based on their respective row-buffer hit rates. This is due to the fact that even accesses that were potentially row-buffer hits (thus not incurring the energy of activating the entire row again) need to perform an entire activate-read-precharge cycle.

[00126] In FIG. 24, graph 2400 shows that in the baseline organizations (both open and close row), the total energy consumption in the baseline memory device is dominated by energy in the bitlines. This is because every access to a new row of the baseline memory device results in a large number of bitlines being activated twice (once to read data out of the cells into the row-buffer and once to pre-charge the array).

[00127] Using the SBA or SSA structures and memory access techniques disclosed herein eliminates a relatively large portion of the energy consumption

components as shown in the graphs 2300 and 2400. By waiting for the CAS signal and only activating/pre-charging the exact cache line needed to serve a memory access request, bitline energy goes down by a factor of 128. This results in a dramatic energy reduction on every access. However, prohibitive area overheads (due to additional logic (e.g., the RX 'AND' gates 710 and corresponding lines and the SWL 'AND' gates 712 and corresponding lines of FIG. 7)) may sometimes dictate using a coarser grained selection in SBA, which may in some cases lead to slightly larger energy consumption compared to SSA. Compared to a baseline open-page system, SBA exhibits an average dynamic memory energy savings of 3X and SSA exhibits an average dynamic memory energy savings of over 6.4X. The results of FIGS. 23 and 24 show that SBA and SSA provide the most energy reduction in the bitlines. The energy overhead due to other components such as decoder, pre-decoder, inter-bank bus, bus termination, etc. remains relatively the same. Thus, their contribution to the total energy increases as bitline energy decreases.

[00128] In an open page policy, open row hits can sometimes have a relatively high energy efficiency because they do not involve any additional activation for subsequent accesses to that row when it is open. An analytical model was constructed to estimate how many cache lines had to be read out of a row buffer for the open-page baseline to match the energy per access of the SBA/SSA structures and memory access techniques. The results showed that the average level of row reuse needs to be 23. However, because the typical row reuse levels of known applications is relatively lower (as shown in graph 2200), energy performance of SBA/SSA is better than the baseline memory using an open-page policy.

[00129] Turning to FIGS. 25 and 26, graphs 2500 and 2600 show example impacts on memory latency and energy reduction, respectively, when using low-power states in connection with the known memory access techniques and structures. Localizing and managing DRAM accesses at a granularity as fine as a subarray (e.g., using SSA) allows more opportunity to place relatively larger portions of a DRAM into low-power states. Known DRAM devices support multiple levels of power-down (e.g., by turning off different levels of circuitry),

each power-down level having a different wake-up latency penalty. For purposes of performance analyses, two simple low-power modes were evaluated with *P* (Power savings factor) and *W* (Wakeup) values calculated based on the energy numbers shown in table 2000 of FIG. 20 (obtained from the Micron datasheet and Micron's System Power Calculator). For the deepest sleep mode (i.e., self refresh mode) of a baseline memory device, *P* is 23 and *W* is 200 memory cycles. In a less deep sleep mode (i.e., power down mode) of the baseline memory device, *P* is lower at 7.5, and *W* is also lower at just 3 memory cycles. These measurements were conducted while varying the idle cycle threshold (*I*) as multiples of the wake-up time *W*.

[00130] The graphs 2500 and 2600 show an example impact of these low-power states on performance and energy consumption. As shown in graph 2600, the more latency-expensive low-power mode (self refresh mode) achieves relatively less energy savings compared to the more latency-efficient mode (power down mode). As shown in graph 2500, decreasing the frequency of transitions to low-power states (i.e., increase in idle cycle threshold (*I*)) causes a decrease in the average memory latency penalty from just under 4% to just over 0.5% for the power down mode. Energy savings correspondingly drop from 86% to 74% as shown in graph 2600. As shown in graph 2500, performance impacts are relatively larger for the latency-expensive self refresh mode, changing from over 400% at a very aggressive idle cycle threshold (*I*) to under 20% in the least aggressive idle cycle threshold (*I*) case. Corresponding energy savings range from 90% to 20% as shown in graph 2600. These power down modes can be applied to the baseline memory architecture as well. However, the granularity at which this can be done is much coarser due to the baseline memory being a DIMM bank at best. This means that there are fewer opportunities for the baseline memory to move into low-power states.

[00131] As a comparison, the application of the low-overhead power down state is analyzed relative to the baseline memory. On average, even with an aggressive sleep threshold, baseline memory banks can only be put in this low-overhead power down mode about 80% of the time, while incurring a penalty of 16% in terms of added memory latency. Unlike the SBA and SSA techniques

disclosed herein, being less aggressive significantly impacts the ability to power down a baseline memory, with banks going into sleep mode only 17% of the time with a minimal 3% latency penalty. As another comparison point, comparisons are made for the percentage of time that subarrays or banks can be put in the deepest sleep mode (self refresh mode) in SSA vs. the baseline, for a constant 10% latency overhead. The results show that subarrays in SSA can go into deep sleep nearly 18% of the time whereas banks in a baseline memory can only go into deep sleep about 5% of the time.

[00132] FIG. 27 is a graph 2700 comparing example average main memory latency when accessing memory using known memory access techniques and structures and the memory access techniques and structures disclosed herein. FIG. 28 is a graph 2800 comparing example memory latencies contributed by different memory subsystems when accessing memory using known memory access techniques and structures and the memory access techniques and structures disclosed herein. Graph 2700 shows that employing either the SBA or SSA structures and memory access techniques has positive impacts on memory access latency. Graph 2800 breaks down the example latencies of graph 2700 down into the average contributions of the various memory components. One of the primary factors affecting latency is the page management policy. Moving to a close-page policy from an open-page baseline results in a drop in average memory latency by about 16% for a majority (8 of 13) of the benchmarks. This exhibits beneficial advantages of using a close-page policy in connection with the SBA and SSA structures and memory access techniques. The remaining benchmarks show an increase in memory latency by about 17% on average when moving to close-page. Employing the posted RAS technique described above in connection with the SBA structures and memory access techniques results in an additional small latency of just under 8% on average.

[00133] In the example graph 2800 the queuing delay is the dominant contributor to total memory access latency. However, this is also true in known DRAM systems. Thus, the additional latency introduced by the posted RAS of SBA does not significantly change average memory access latency.

[00134] For the example SSA structure and memory access techniques illustrated herein, each cache line return is serialized over, for example, eight lines to a memory controller. The delay introduced by such serialization becomes a dominant factor in the total access time. However, this is compensated for by a large increase in parallelism in the system. That is, each memory chip in a memory module can service independent sets of requests, significantly reducing the queuing delay. As a result, memory latency is not significantly increased, even though the data transfers are serialized. In 8 of the 12 benchmarks depicted in graph 2700, latency increases just over 40%. The other four benchmarks can exploit the parallelism much better, which compensates for the serialization latency with average access time decreasing by about 30%. These are also the applications with the highest memory latencies, and as a result SSA outperforms the baseline model.

[00135] FIG. 29 is a graph 2900 comparing example normalized IPCs for known memory access techniques and structures and the memory access techniques and structures disclosed herein. A majority of the example benchmarks show better performance with a close-row policy than with an open-row policy. In particular, there are performance improvements of just over 5% for 9 of the 12 benchmarks. The other three example benchmarks show performance decreases of about 20% on average. These were the benchmarks with relatively higher miss rates (on the order of 10 for every 1000 instructions). Employing the posted RAS of SBA results in a marginal decrease in IPC over close-row baseline (about 3% on average), which is largely influenced by the three highly sensitive benchmark applications (stream, cg, is).

[00136] The example SSA techniques illustrated herein exhibit a performance decrease of 5.5% on average compared to the open-page baseline on the five benchmarks that exhibited a memory latency increase in graph 2700 of FIG. 27. The other seven benchmarks with decreased memory access latency exhibit performance gains of 42% on average for SSA. These high numbers are observed because these applications are limited by bank contention and SSA addresses that bottleneck. In sum, in addition to significantly lowered DRAM access energies, SSA can boost performance, while yielding only minor

decreases in performance. As the quantity of memory chips increases per memory subsystem and higher queuing pressures are exerted on memory controllers, SSA can be used to achieve even higher energy and performance improvements.

[00137] FIG. 30 is a graph 3000 comparing example differences in memory access performance for use and non-use of chipkill error detection and correction techniques described above in connection with FIGS. 13, 17, and 18. In particular, graph 3000 shows the decrease in IPC performance resulting from using SSA with the chipkill approach disclosed herein. This is caused by the increased bank contention during writes. On average, the decrease in memory latency performance is 40%, resulting in an 11% decrease in IPC performance. Compared to the non-chipkill SSA implementation, there is also additional energy consumption on every write for the chipkill SSA implementation. This results in a 2.2X increase in dynamic energy to retain chipkill-level reliability, which is still significantly lower than the dynamic energy used by a baseline memory structure.

[00138] Although certain methods, apparatus, and articles of manufacture have been described herein, the scope of coverage of this patent is not limited thereto. To the contrary, this patent covers all methods, apparatus, and articles of manufacture fairly falling within the scope of the appended claims either literally or under the doctrine of equivalents.

What is Claimed is:

1. A memory, comprising:

a row address register (412) to store a row address corresponding to a row (608) in a memory array (602);

a row decoder (604) coupled to the row address register to assert a signal on a wordline (704) of the row after the memory receives a column address; and

a column decoder (606) to selectively activate a portion of the row based on the column address and the signal asserted on the wordline.

2. A memory as defined in claim 1, further comprising an address buffer (406) to latch the row address received from a memory controller (102) and send the row address to the row address register and to latch the column address and send the column address to a column address register (414) of the memory to facilitate concurrently decoding the row address and the column address at the memory.

3. A memory as defined in claim 2, wherein the memory is a first memory, the address buffer is located external to the memory on a memory module (400) containing the memory and a second memory, the address buffer to send the row and column addresses to the first memory and the second memory.

4. A memory as defined in claim 1, wherein the portion of the row includes a cache line segment (610) that forms a cache line addressed by the row and column addresses.

5. A memory as defined in claim 4, wherein the row and column decoders selectively activate a second portion of the memory array based on the same row and column addresses, the second portion containing a second cache line segment to form the cache line, and the row and column decoders selectively activating additional portions until the entire cache line is output from a memory module (400) containing the memory.

6. A memory as defined in claim 5, wherein the portion and the second portion are located in a corresponding row across the memory and a second memory of the memory module.

7. A memory as defined in claim 4, wherein the memory is a first memory, the first memory is located on a memory module containing a second memory, and a second cache line segment comprising part of the cache line is activated in the second memory and output by the memory module in response to the row and column addresses.

8. A memory as defined in claim 1, wherein the row decoder is to assert a first signal on the wordline connected to a first input of a logic gate (712), the column decoder is to assert a second signal at a second input of the logic gate, of the first and second signals to cause the logic gate to activate the portion of the row without activating a second portion of the row.

9. A memory as defined in claim 1, wherein the memory is a dynamic random access memory chip.

10. A memory as defined in claim 1, wherein the memory is located on one of a dual inline memory module (DIMM) (200) containing a second memory or a 3D stack memory module (300) containing the second memory.

11. A memory module, comprising:

a first memory (402f) having a first row (608);

a second memory (402a) having a second row; and

an address buffer (406) to latch a row address and a column address and forward the row and column addresses to the first and second memories, the first memory to selectively activate a first portion of the first row and the second memory to selectively activate a second portion of the second row in response to the row and column addresses, the first portion and the second portion activated without activating other portions of the first and second rows.

12. A memory module as defined in claim 11, wherein the first portion of the first row is a first cache line segment (610) and the second portion of the second row is a second cache line segment, the first and second memories to output the first and second portions to form a cache line.

13. A memory module as defined in claim 11, wherein the first and second memories selectively activate portions of third and fourth rows respectively located in the first and second memories based on the same row and column addresses, the portions respectively containing third and fourth cache line segments to form the cache line, and the first and second memories selectively activating the portions of the third and fourth rows until an entirety of the cache line is output from the memory module (400).

14. A memory module as defined in claim 11 structured as one of a dual inline memory module (DIMM) or a 3D stack memory module (300) including the first and second memories.

15. A memory module as defined in claim 11, wherein the first memory includes a row decoder (604) and a column decoder (606), the row decoder to assert a first signal on a main wordline (704) connected to a first input of a logic gate (712), the column decoder to assert a second signal at a second input of the logic gate, the first and second signals to cause the logic gate to activate the first portion of the first row exclusive of other portions of the first row.

16. A memory, comprising:

a data interface (816, 902, 1004); and

an integrated circuit memory die (806g) in communication with the data interface, the integrated circuit memory die comprising:

a plurality of banks (808);

a plurality of subarrays (810a, 810b) in each bank; and

bitlines (812) in each bank, first ones of the bitlines corresponding to a first one of the banks being addressable to access an entire cache line in a single one of the subarrays of the first bank.

17. A memory as defined in claim 16, further comprising an address buffer (910, 1010, 1106) to latch row and column addresses to facilitate concurrently decoding the row and column addresses to read out the entire cache line from the single one of the subarrays.

18. A memory as defined in claim 17, wherein the address buffer is external to the integrated circuit memory die and located on a memory module (200, 300) containing the integrated circuit memory die and a second integrated circuit memory die, the address buffer to send the row and column addresses to the integrated circuit memory die and the second integrated circuit memory die.

19. A memory as defined in claim 16, wherein the first ones of the bitlines activate a portion of the single one of the subarrays to access the cache line after receiving both row and column addresses associated with the cache line.

20. A memory as defined in claim 16, wherein the integrated circuit memory die is a first die located on a memory module (200, 300) containing at least a second integrated circuit memory die, the second integrated circuit memory die in communication with the data interface, the first integrated circuit memory die to be activated without activating the second integrated circuit memory die when an address is received at the memory module to retrieve the entire cache line from the first integrated circuit memory die.

21. A memory as defined in claim 16, wherein the integrated circuit memory die is located on a memory module (200, 300) containing the data interface, the data interface of the memory module having a larger bit length than a second data interface of the integrated circuit memory die, the readout of the entire cache line from the memory module to be performed using a burst size based on the bit length of the second data interface of the integrated circuit memory die.

22. A memory as defined in claim 16, wherein the integrated circuit memory die is located on a memory module (200, 300) containing at least a second integrated circuit memory die in communication with the data interface, the integrated circuit memory die to output the cache line via the data interface while the second integrated circuit memory die concurrently outputs a second cache line via the data interface.

23. A memory as defined in claim 16, wherein the integrated circuit memory die is to pre-charge the first bank after the first bitlines close the single one of the subarray, the pre-charge of the first bank to facilitate faster access to the first bank during a subsequent access of the first bank.

24. A memory as defined in claim 16, wherein the data interface is in communication with a memory controller (102), the memory controller including a separate data queue for each bank of the integrated circuit memory die.

25. A memory module, comprising:

a first memory (806g) having a first subarray (810a);

a second memory (806a) having a second subarray;

first bitlines (812) in the first memory to enable access to an entire first cache line in the first subarray; and

second bitlines in the second memory to enable access to an entire second cache line in the second subarray.

26. A memory module as defined in claim 25, further comprising an address buffer (910, 1010, 1106) to latch row and column addresses received from a memory controller (102) to facilitate concurrently decoding the row and column addresses to activate the entire first cache line for reading out from the first subarray.

27. A memory module as defined in claim 26, wherein the address buffer is external to the first and second memories and located on a memory module (200, 300) containing the first and second memories, the address buffer to send the row and column addresses to the first and second memories.

28. A memory module as defined in claim 25, wherein the first bitlines activate a portion of the first subarray to access the first cache line after receiving both row and column addresses associated with the cache line.

29. A memory module as defined in claim 25, wherein the first memory is to be activated to access the first cache line without activating the second memory, and the second memory is to be activated to access the second cache line without activating the first memory.

30. A memory module as defined in claim 25, wherein the first and second memories are both in communication with a data interface (902), the first memory to output the first cache line via the data interface while the second memory is to concurrently output the second cache line via the data interface.

31. A memory, comprising:

a plurality of memory chips (1702a, 1702i); and

a plurality of data storage locations and parity storage locations in each memory chip, each of the data storage locations to store a respective cache line (1704) and a corresponding checksum (1706), and each parity storage location to store a parity value (1708) for a corresponding subset of the data storage locations, the cache lines of a subset of the storage locations being stored in different memory chips.

32. A memory as defined in claim 31, wherein upon an occurrence of an error in at least one of the data storage locations exposed by the checksum, an uncorrupted instance of the least one of the data storage locations is reconstructable based on the parity value.

33. A memory as defined in claim 32, wherein the uncorrupted instance of the least one of the data storage locations is reconstructable using a reconstruction technique selected from a redundant array of inexpensive disks (RAID)-4 data reconstruction process or a RAID-5 data reconstruction process.

34. A memory as defined in claim 31, wherein a first one of the memory chips is to service a read request by outputting a requested cache line and using an additional burst in a same read cycle to output a checksum corresponding to the requested cache line.

35. A memory as defined in claim 34, wherein the checksum is stored in the first memory chip in association with the requested cache line.

36. A memory as defined in claim 31, wherein a first one of the memory chips is to service a write request by storing a new cache line and its corresponding checksum to a first one of the memory chips.

37. A memory as defined in claim 36, wherein the first memory chip is to further service the write request by storing a corresponding parity value associated with the second cache line.

38. A memory as defined in claim 36, wherein to service the write request, the memory chip outputs an old cache line and a corresponding old checksum to a memory controller (802) prior to receiving the second cache line and the corresponding second checksum from the memory controller to replace the old cache line and the corresponding old checksum.

39. A memory controller, comprising:

a data interface (1312a) to be in communication with a plurality of memory chips (1702a, 1702i) and exchange information with a plurality of data storage locations and parity storage locations in the at least one memory chip, the data interface to access a first cache line (1704) and a corresponding first checksum (1706) from a first subset of the data storage locations, and the data interface to access parity values (1708) for a corresponding subset of the data storage locations; and

a checksum generator (1318) to generate the first checksum for storing in association with the first cache line.

40. A memory controller as defined in claim 39, further comprising:

an error detector (1316) to detect an error in at least one of the data storage locations based on the first checksum; and

a data reconstructor (1322) to reconstruct an uncorrupted instance of the least one of the data storage locations based on the parity value.

41. A memory controller as defined in claim 40, wherein the data reconstructor is to reconstruct the uncorrupted instance of the at least one of the data storage locations using a reconstruction technique selected from a redundant array of inexpensive disks (RAID)-4 data reconstruction process or a RAID-5 data reconstruction process.

42. A memory controller as defined in claim 39, wherein the data interface is to perform a read operation in a read cycle by reading a requested cache line from at least one of the memory chips and by initiating an additional burst in the same read cycle to read a checksum corresponding to the requested cache line.

43. A memory controller as defined in claim 39, further comprising a parity generator (1320), the data interface is to follow a write operation with a read operation of a cache line written using the write operation and a parity value corresponding to the written cache line, and the parity generator is to calculate a new parity value for the written cache line based on the parity value.

44. A memory controller as defined in claim 39, wherein the data interface is to perform a write operation by writing a new cache line and its corresponding checksum to a first one of the memory chips.

45. A memory controller as defined in claim 44, wherein to perform the write operation, the data interface is to read an old cache line and a corresponding old checksum from the at least one memory chip prior to writing the second cache line and the corresponding second checksum to the at least one memory chip to replace the old cache line and the corresponding old checksum.

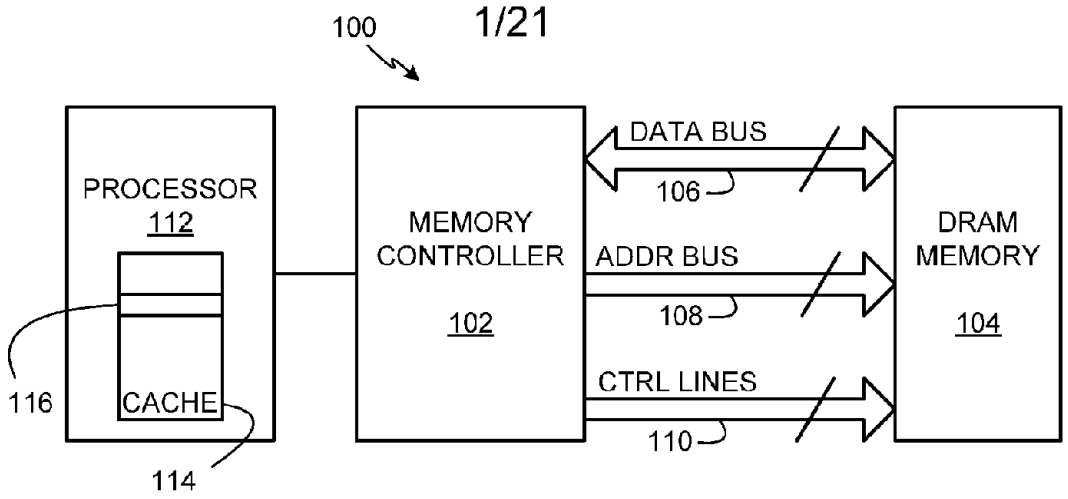


FIG. 1

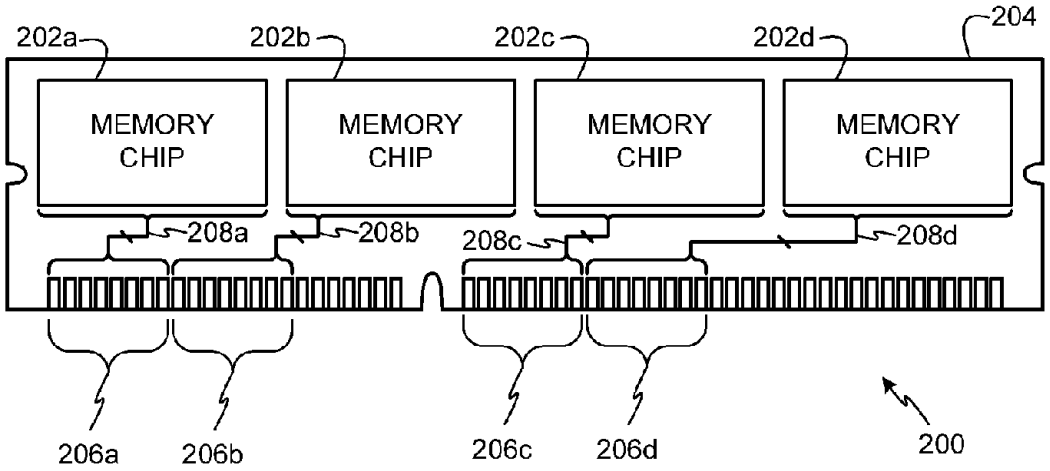


FIG. 2

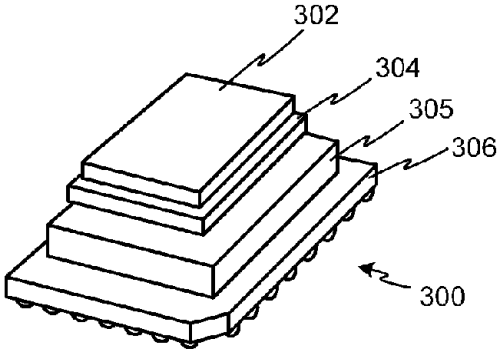


FIG. 3

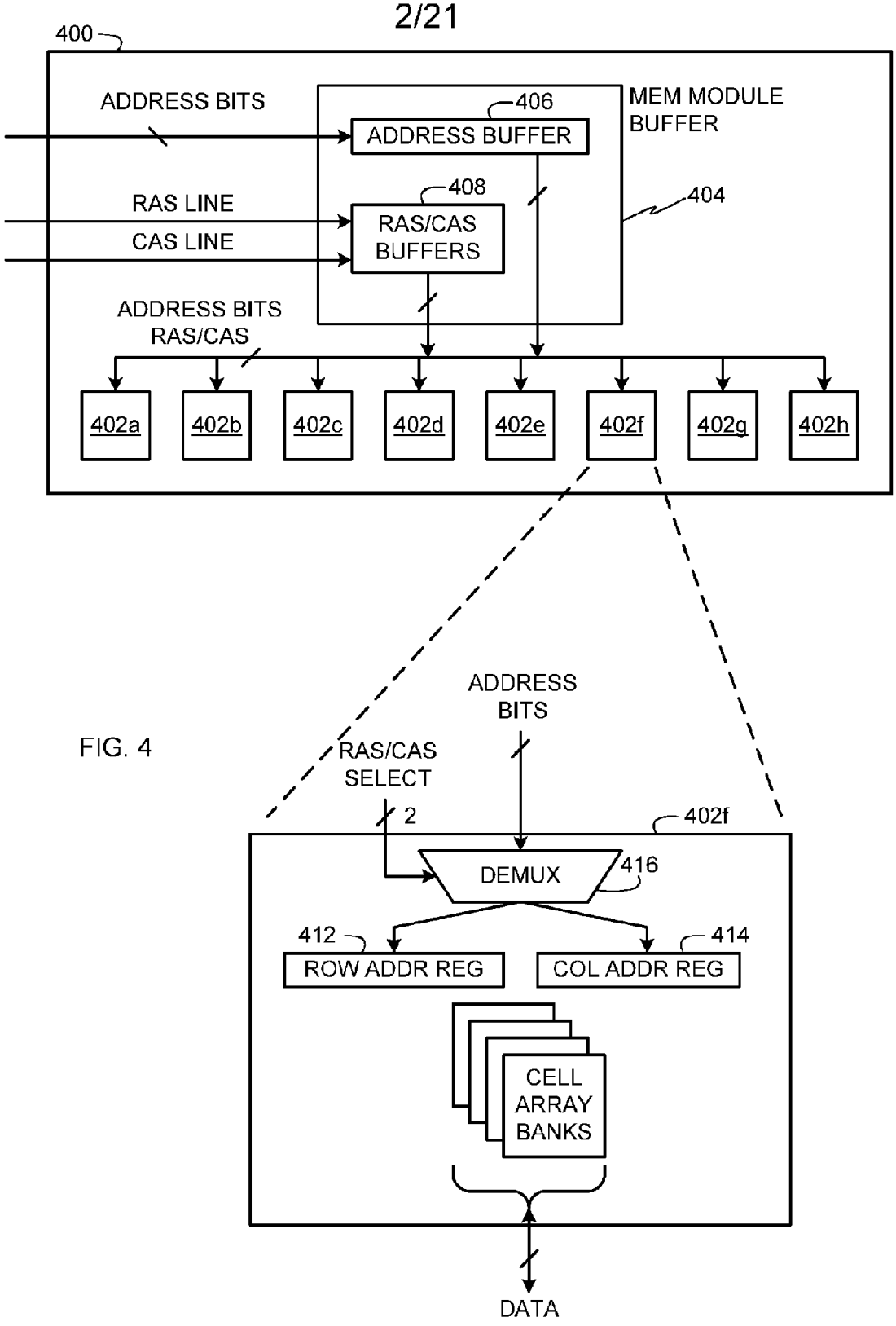

400
ADDRESS BITS
406
ADDRESS BUFFER
MEM MODULE BUFFER
404
RAS LINE
CAS LINE
408
RAS/CAS BUFFERS
ADDRESS BITS RAS/CAS
402a
402b
402c
402d
402e
402f
402g
402h
ADDRESS BITS
RAS/CAS SELECT
2
402f
DEMUX
416
412
414
ROW ADDR REG
COL ADDR REG
CELL ARRAY BANKS
DATA

FIG. 4

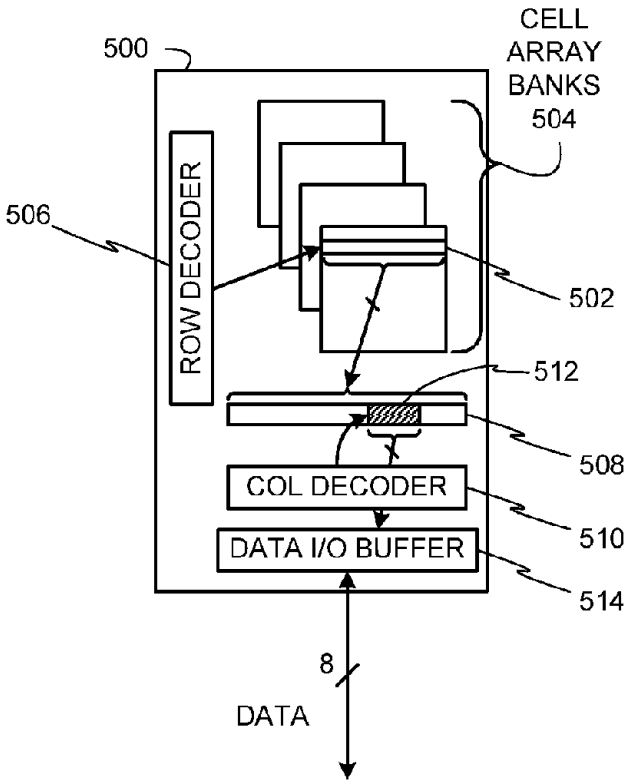


FIG. 5
(PRIOR ART)

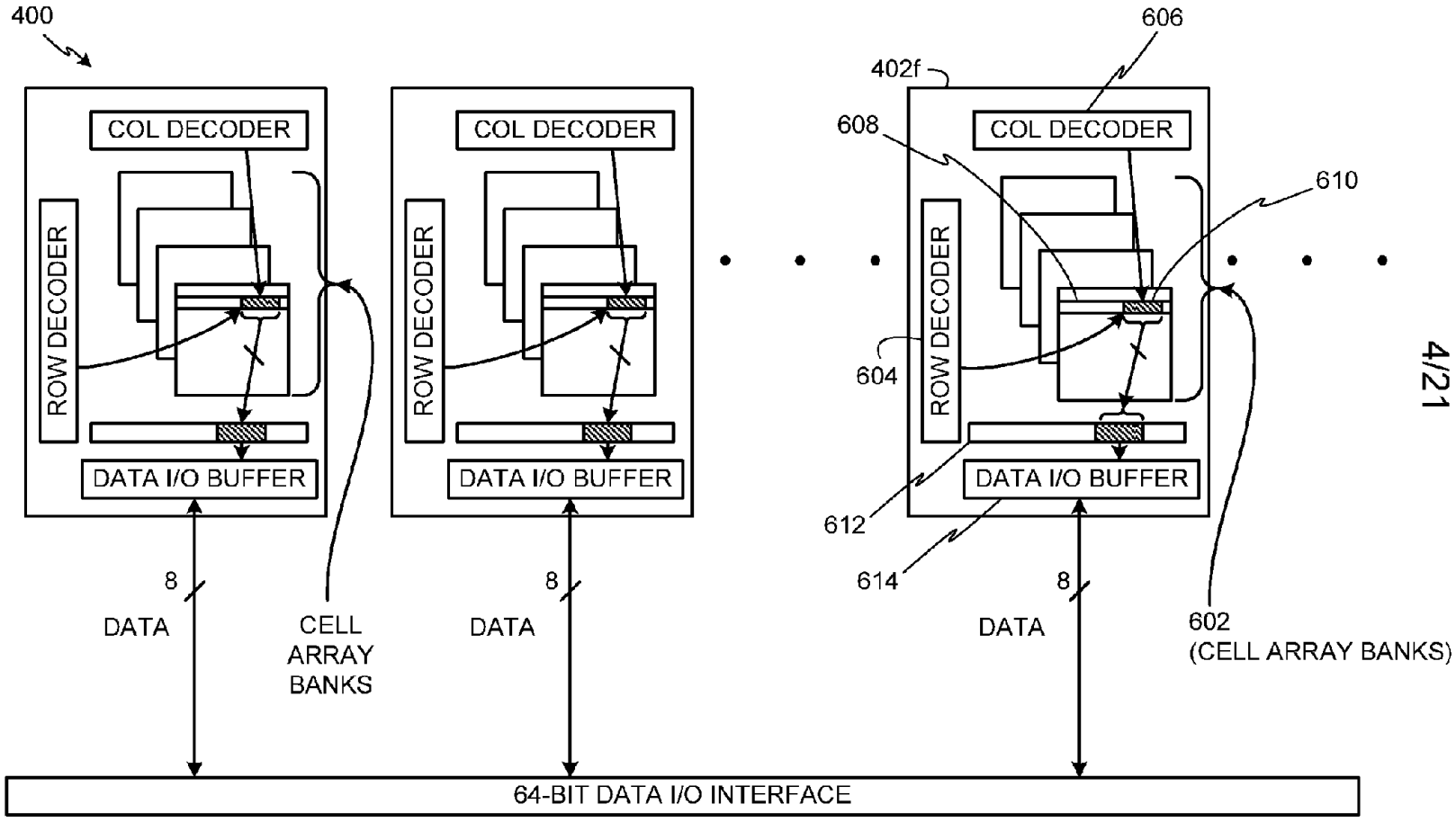

400
COL DECODER
ROW DECODER
DATA I/O BUFFER
8
DATA
CELL ARRAY BANKS
606
402f
608
610
604
612
614
602 (CELL ARRAY BANKS)
64-BIT DATA I/O INTERFACE

FIG. 6

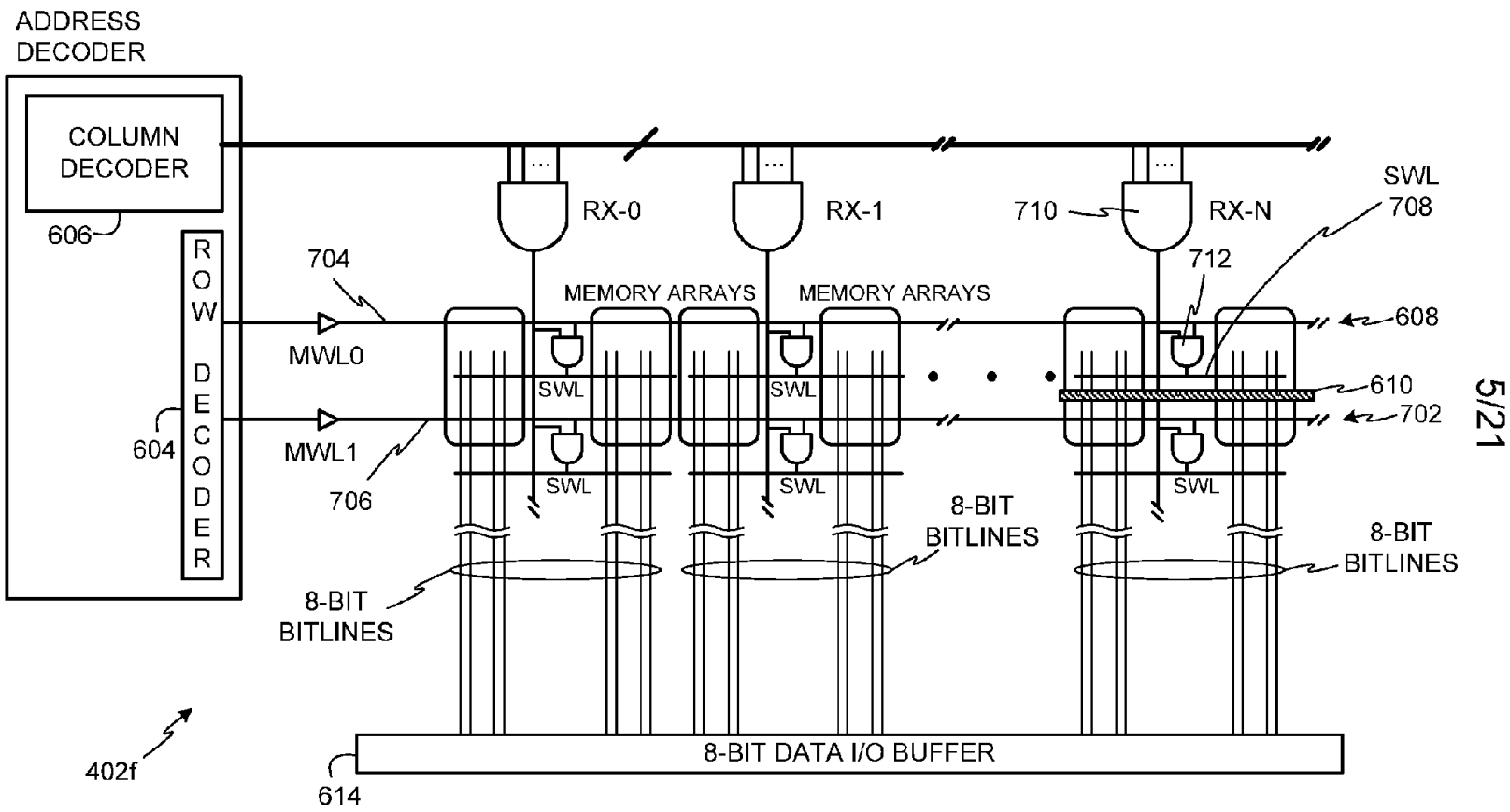

ADDRESS DECODER
COLUMN DECODER
606
ROW DECODER
604
704
MWL0
MWL1
706
RX-0
RX-1
710
RX-N
SWL
708
712
MEMORY ARRAYS
MEMORY ARRAYS
608
610
702
SWL
8-BIT BITLINES
8-BIT BITLINES
8-BIT BITLINES
402f
8-BIT DATA I/O BUFFER
614

FIG. 7

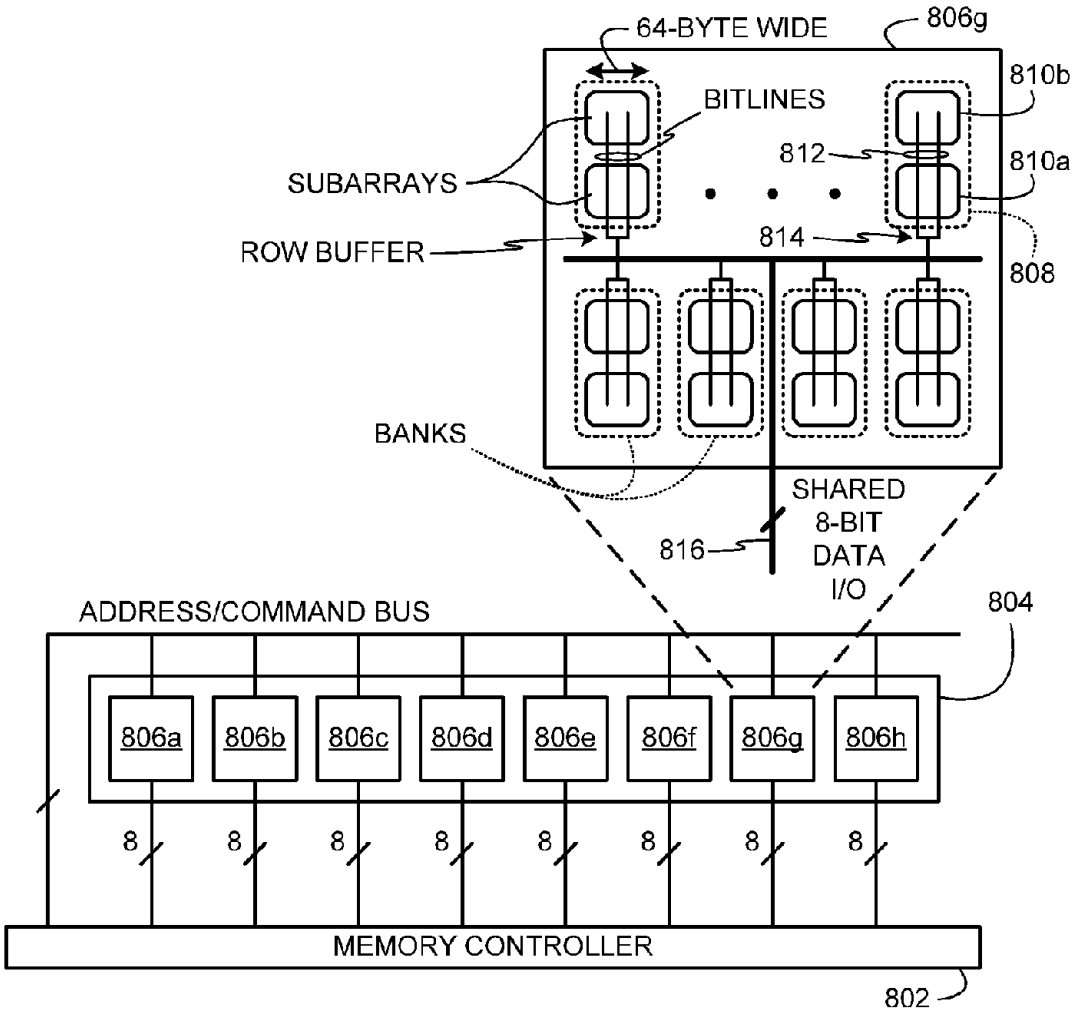

64-BYTE WIDE
806g
810b
BITLINES
812
810a
SUBARRAYS
ROW BUFFER
814
808
BANKS
SHARED
8-BIT
DATA
I/O
816
ADDRESS/COMMAND BUS
804
806a
806b
806c
806d
806e
806f
806g
806h
8
MEMORY CONTROLLER
802

FIG. 8

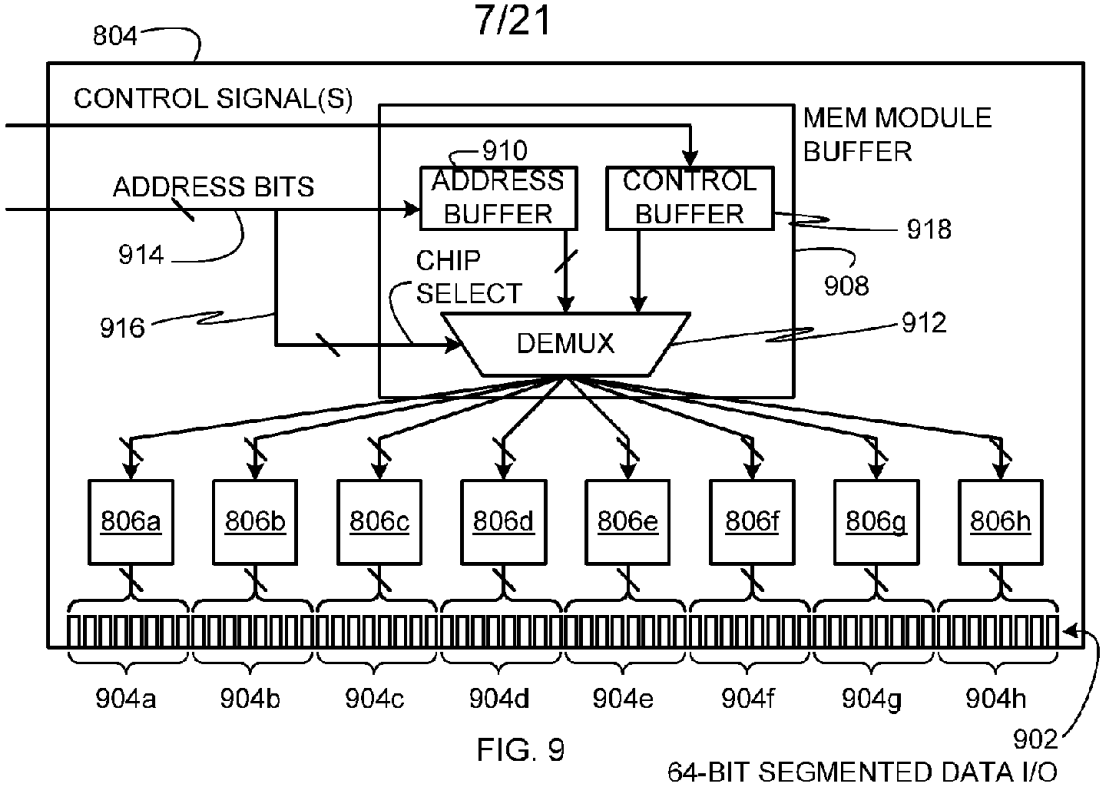

804
CONTROL SIGNAL(S)
MEM MODULE BUFFER
910
ADDRESS BUFFER
CONTROL BUFFER
ADDRESS BITS
914
916
CHIP SELECT
DEMUX
918
908
912
806a
806b
806c
806d
806e
806f
806g
806h
904a
904b
904c
904d
904e
904f
904g
904h
902
64-BIT SEGMENTED DATA I/O

FIG. 9

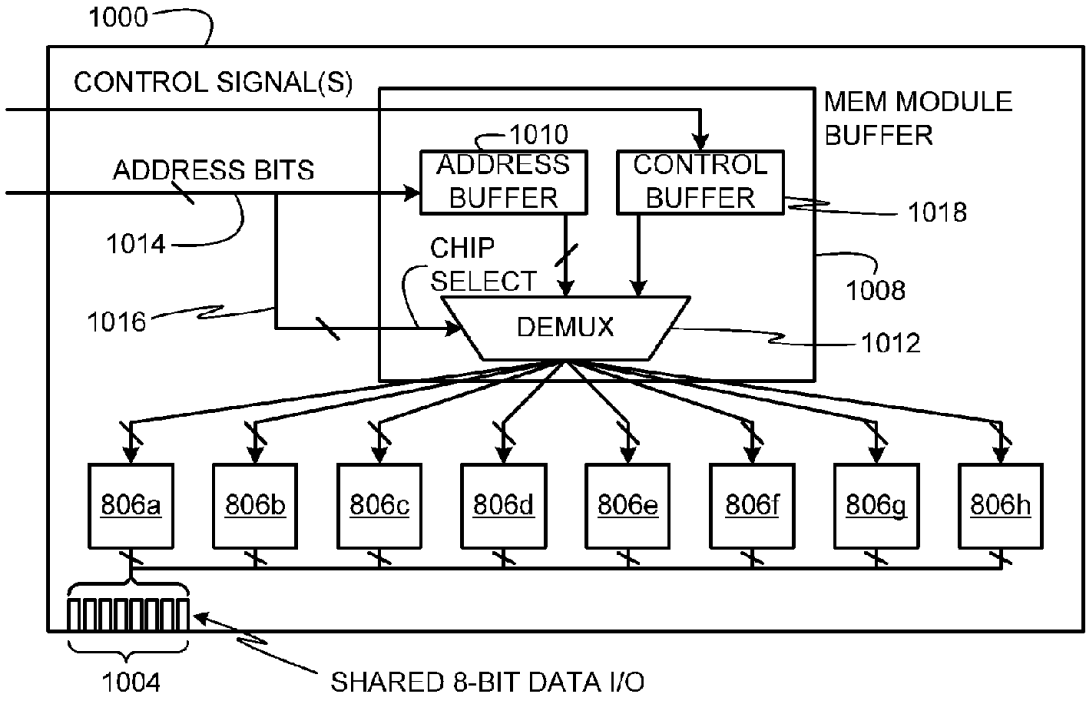

1000
CONTROL SIGNAL(S)
MEM MODULE BUFFER
1010
ADDRESS BUFFER
CONTROL BUFFER
ADDRESS BITS
1014
1016
CHIP SELECT
DEMUX
1018
1008
1012
806a
806b
806c
806d
806e
806f
806g
806h
1004
SHARED 8-BIT DATA I/O

FIG. 10

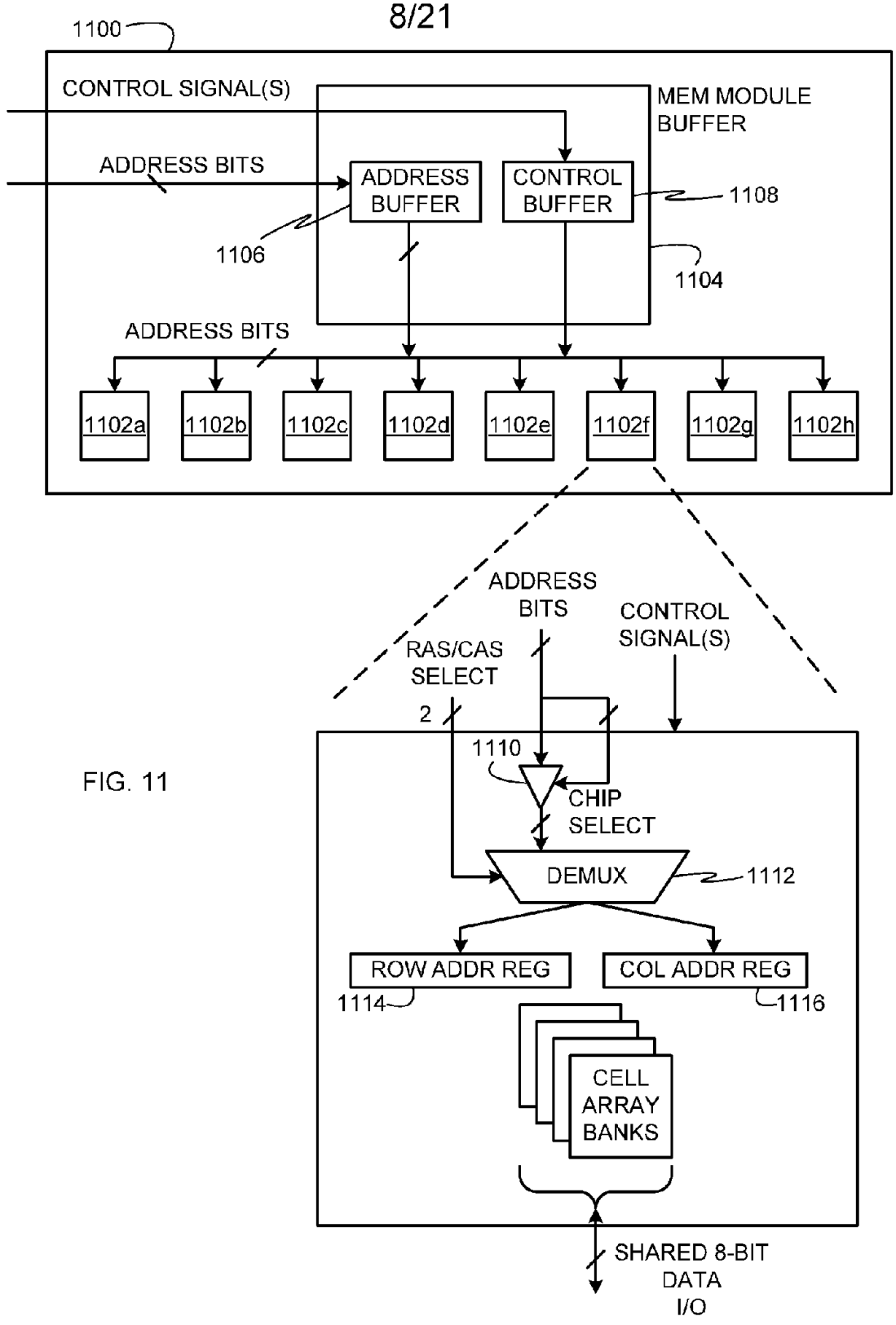

1100
CONTROL SIGNAL(S)
MEM MODULE BUFFER
ADDRESS BITS
ADDRESS BUFFER
CONTROL BUFFER
1108
1106
1104
ADDRESS BITS
1102a
1102b
1102c
1102d
1102e
1102f
1102g
1102h
ADDRESS BITS
CONTROL SIGNAL(S)
RAS/CAS SELECT
2
1110
CHIP SELECT
DEMUX
1112
ROW ADDR REG
COL ADDR REG
1114
1116
CELL ARRAY BANKS
SHARED 8-BIT DATA I/O

FIG. 11

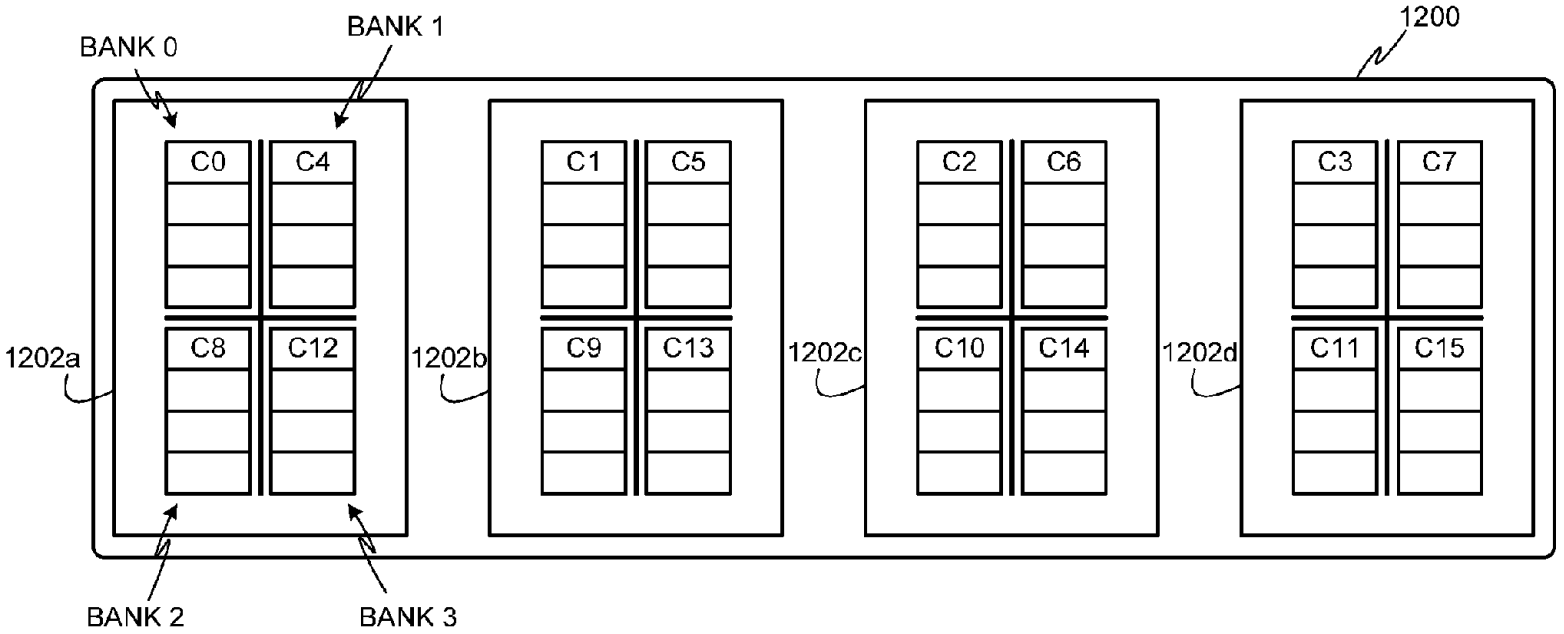

BANK 0
BANK 1
1200
C0
C4
C1
C5
C2
C6
C3
C7
C8
C12
C9
C13
C10
C14
C11
C15
1202a
1202b
1202c
1202d
BANK 2
BANK 3

FIG. 12

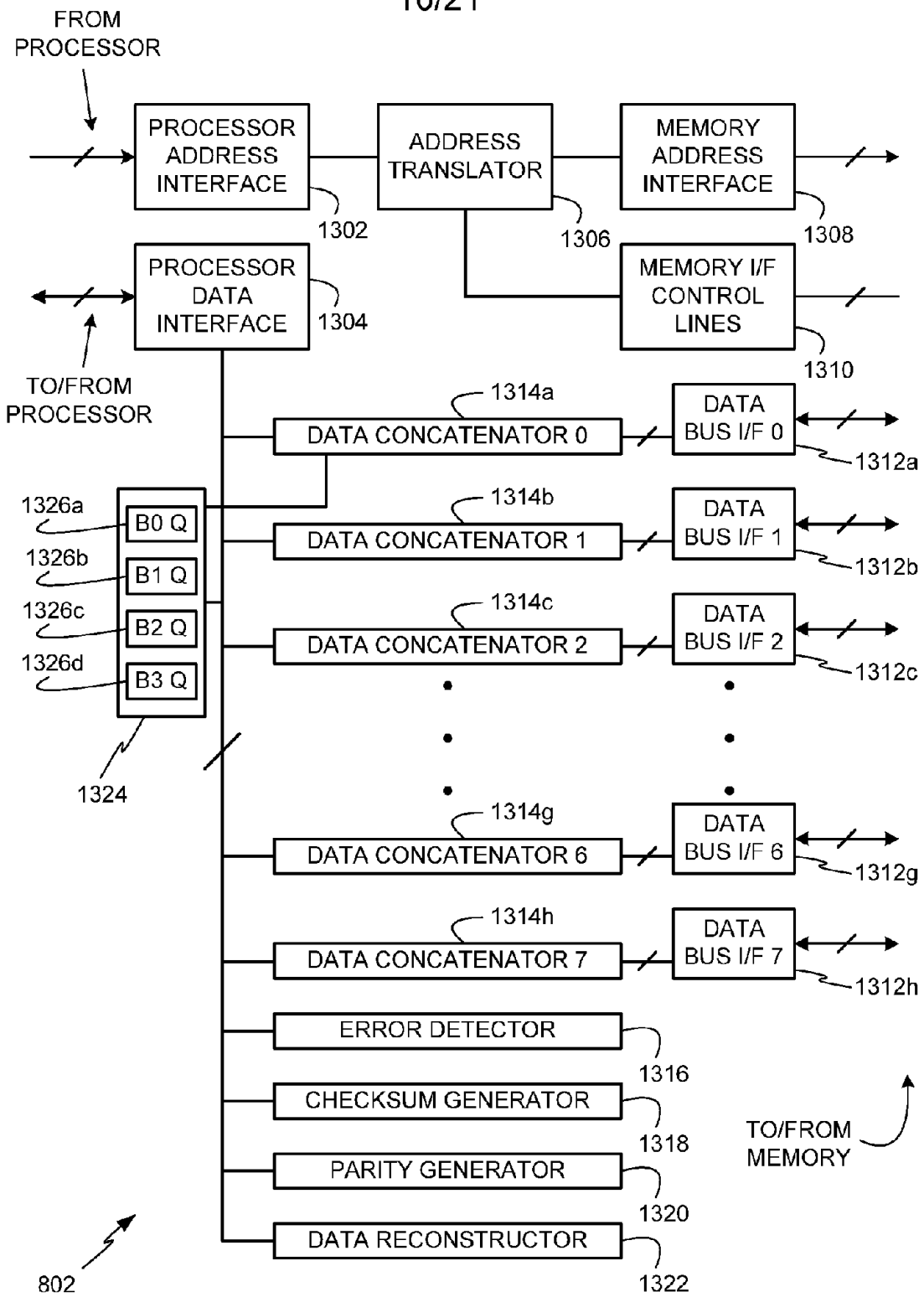

FROM PROCESSOR
PROCESSOR ADDRESS INTERFACE
1302
ADDRESS TRANSLATOR
1306
MEMORY ADDRESS INTERFACE
1308
PROCESSOR DATA INTERFACE
1304
MEMORY I/F CONTROL LINES
1310
TO/FROM PROCESSOR
1314a
DATA CONCATENATOR 0
DATA BUS I/F 0
1312a
1326a
B0 Q
1314b
DATA CONCATENATOR 1
DATA BUS I/F 1
1312b
1326b
B1 Q
1326c
B2 Q
1314c
DATA CONCATENATOR 2
DATA BUS I/F 2
1312c
1326d
B3 Q
1324
1314g
DATA CONCATENATOR 6
DATA BUS I/F 6
1312g
1314h
DATA CONCATENATOR 7
DATA BUS I/F 7
1312h
ERROR DETECTOR
1316
CHECKSUM GENERATOR
1318
TO/FROM MEMORY
PARITY GENERATOR
1320
DATA RECONSTRUCTOR
1322
802
MEMORY CONTROLLER

FIG. 13

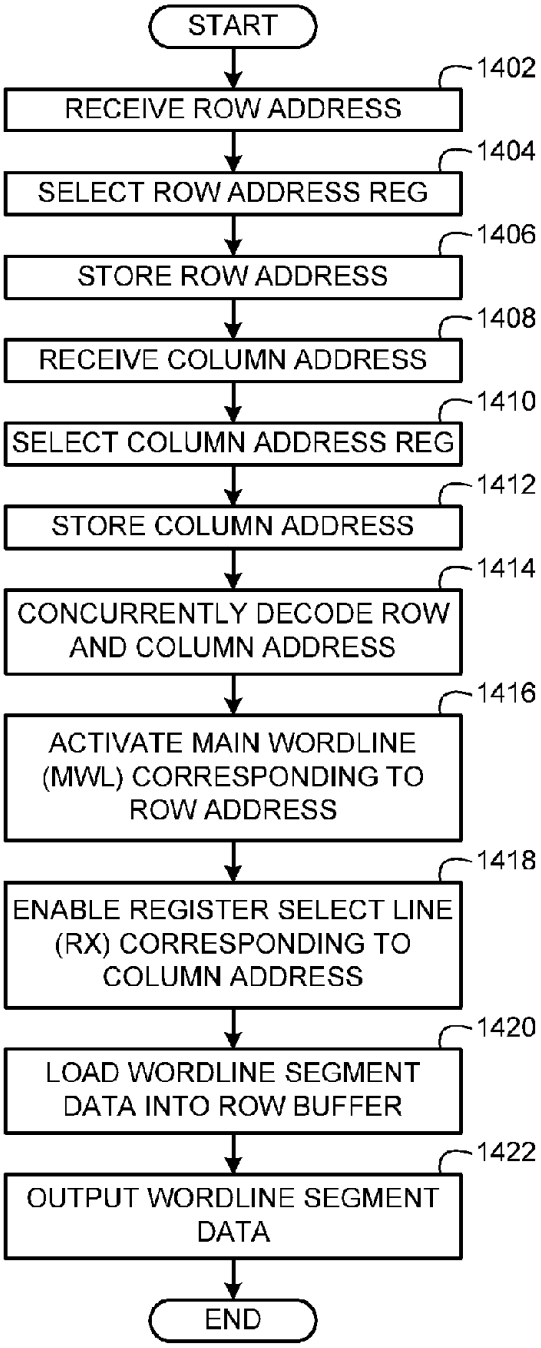

START
1402
RECEIVE ROW ADDRESS
1404
SELECT ROW ADDRESS REG
1406
STORE ROW ADDRESS
1408
RECEIVE COLUMN ADDRESS
1410
SELECT COLUMN ADDRESS REG
1412
STORE COLUMN ADDRESS
1414
CONCURRENTLY DECODE ROW AND COLUMN ADDRESS
1416
ACTIVATE MAIN WORDLINE (MWL) CORRESPONDING TO ROW ADDRESS
1418
ENABLE REGISTER SELECT LINE (RX) CORRESPONDING TO COLUMN ADDRESS
1420
LOAD WORDLINE SEGMENT DATA INTO ROW BUFFER
1422
OUTPUT WORDLINE SEGMENT DATA
END

FIG. 14

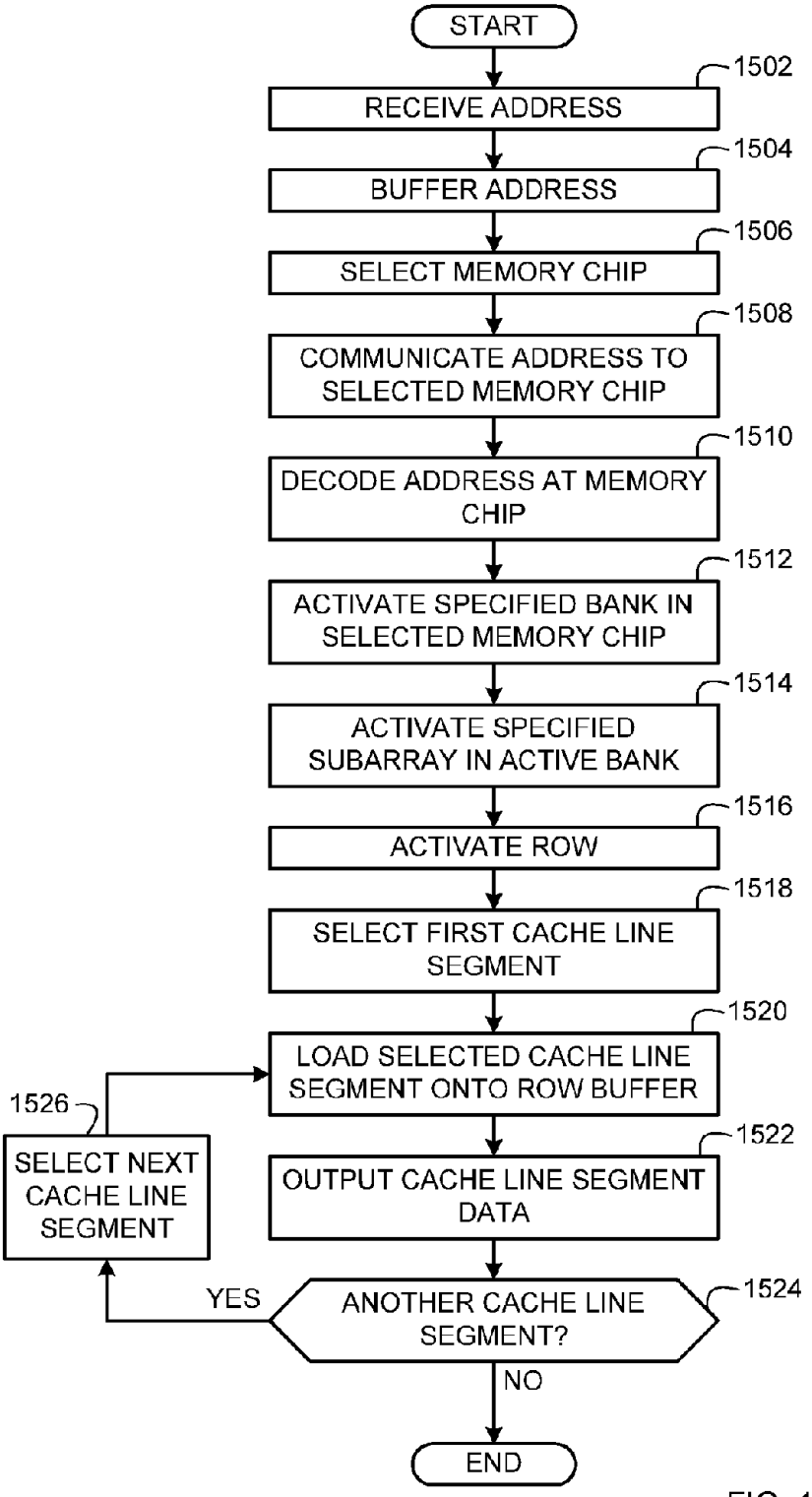

START
1502
RECEIVE ADDRESS
1504
BUFFER ADDRESS
1506
SELECT MEMORY CHIP
1508
COMMUNICATE ADDRESS TO SELECTED MEMORY CHIP
1510
DECODE ADDRESS AT MEMORY CHIP
1512
ACTIVATE SPECIFIED BANK IN SELECTED MEMORY CHIP
1514
ACTIVATE SPECIFIED SUBARRAY IN ACTIVE BANK
1516
ACTIVATE ROW
1518
SELECT FIRST CACHE LINE SEGMENT
1520
LOAD SELECTED CACHE LINE SEGMENT ONTO ROW BUFFER
1522
OUTPUT CACHE LINE SEGMENT DATA
1524
ANOTHER CACHE LINE SEGMENT?
YES
NO
1526
SELECT NEXT CACHE LINE SEGMENT
END

FIG. 15

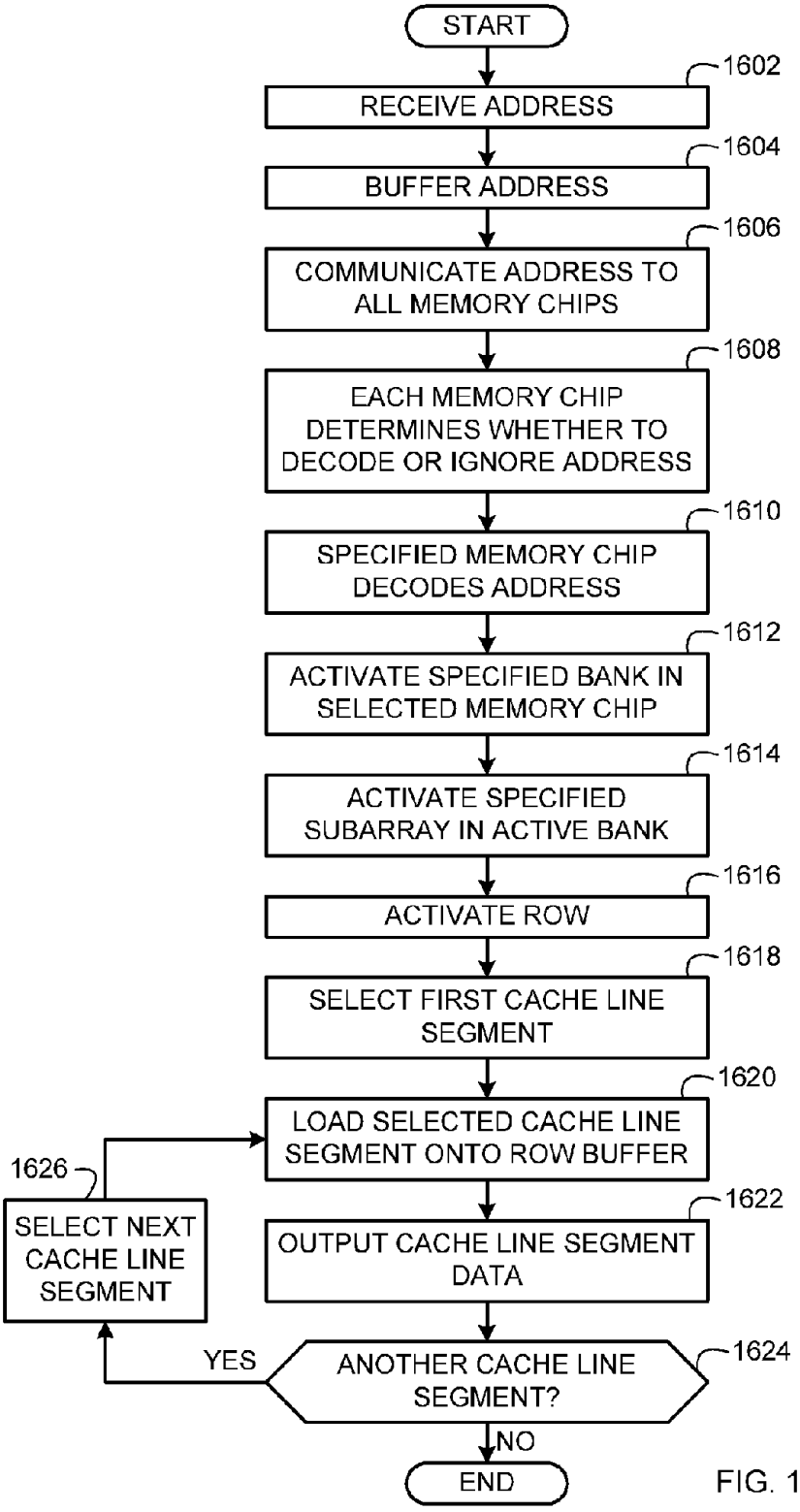

START
1602
RECEIVE ADDRESS
1604
BUFFER ADDRESS
1606
COMMUNICATE ADDRESS TO ALL MEMORY CHIPS
1608
EACH MEMORY CHIP DETERMINES WHETHER TO DECODE OR IGNORE ADDRESS
1610
SPECIFIED MEMORY CHIP DECODES ADDRESS
1612
ACTIVATE SPECIFIED BANK IN SELECTED MEMORY CHIP
1614
ACTIVATE SPECIFIED SUBARRAY IN ACTIVE BANK
1616
ACTIVATE ROW
1618
SELECT FIRST CACHE LINE SEGMENT
1620
LOAD SELECTED CACHE LINE SEGMENT ONTO ROW BUFFER
1622
OUTPUT CACHE LINE SEGMENT DATA
1624
ANOTHER CACHE LINE SEGMENT?
YES
NO
1626
SELECT NEXT CACHE LINE SEGMENT
END

FIG. 16

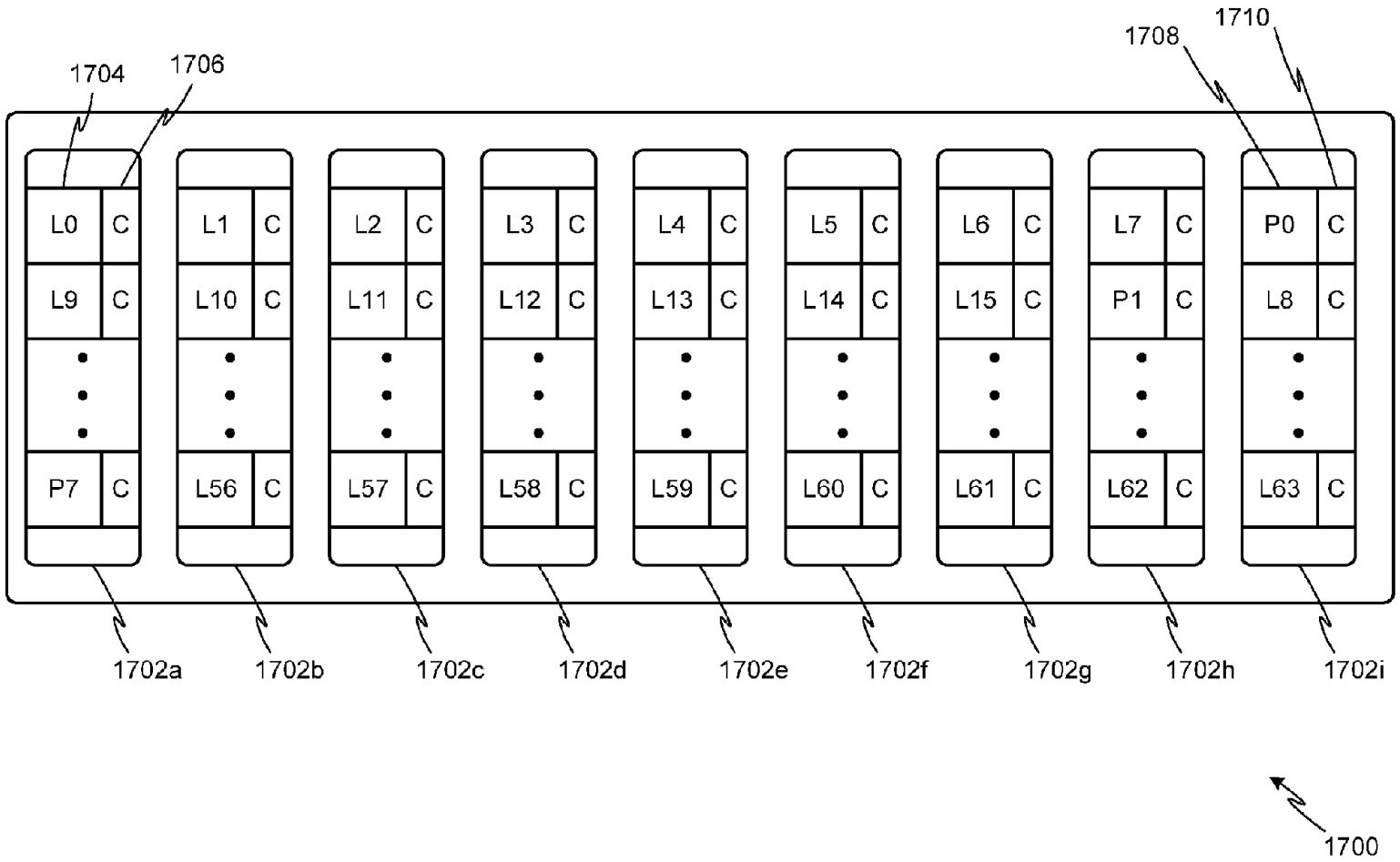

1704
1706
1708
1710
L0
C
L1
C
L2
C
L3
C
L4
C
L5
C
L6
C
L7
C
P0
C
L9
C
L10
C
L11
C
L12
C
L13
C
L14
C
L15
C
P1
C
L8
C
P7
C
L56
C
L57
C
L58
C
L59
C
L60
C
L61
C
L62
C
L63
C
1702a
1702b
1702c
1702d
1702e
1702f
1702g
1702h
1702i
1700

FIG. 17

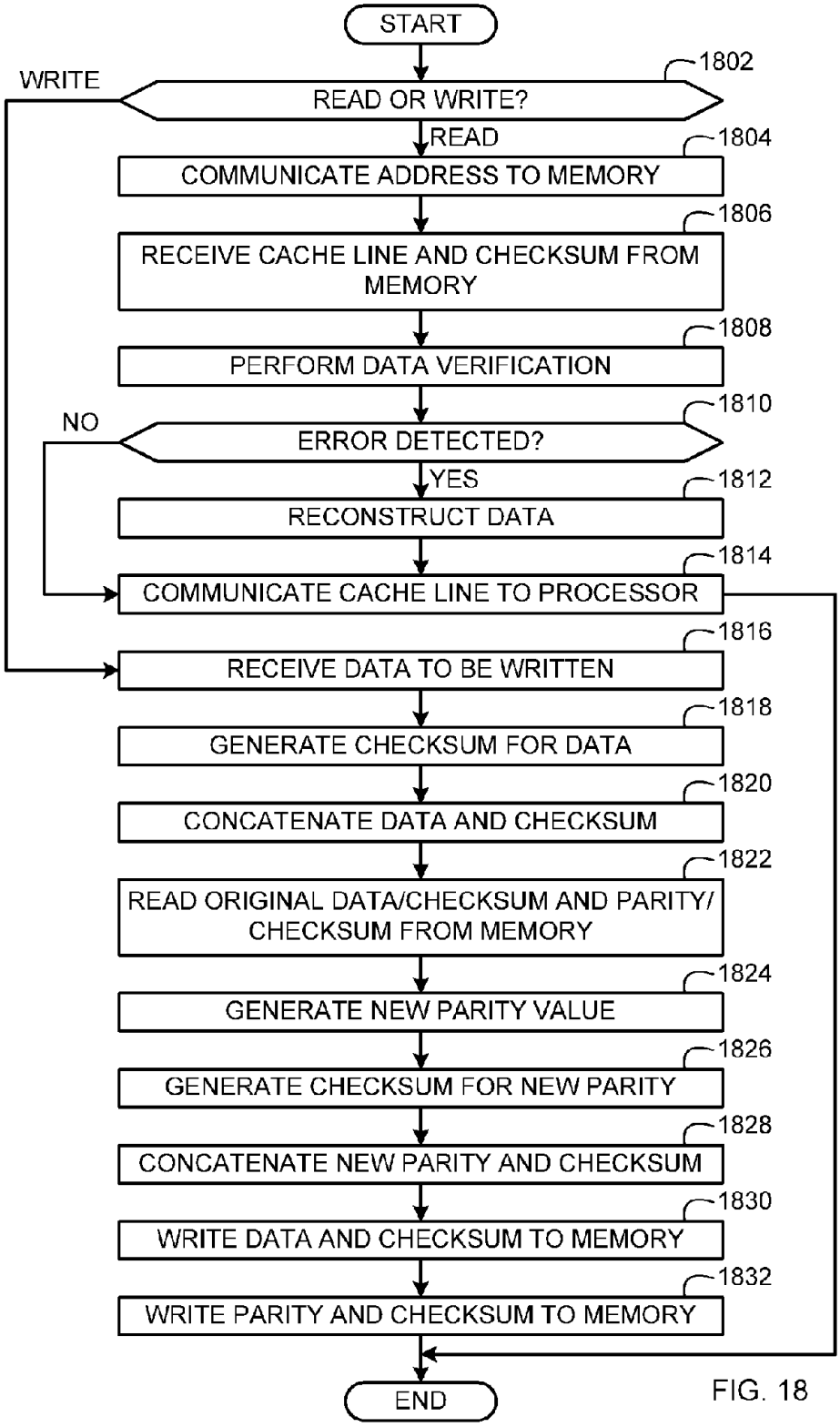

START
1802
WRITE
READ OR WRITE?
READ
1804
COMMUNICATE ADDRESS TO MEMORY
1806
RECEIVE CACHE LINE AND CHECKSUM FROM MEMORY
1808
PERFORM DATA VERIFICATION
1810
NO
ERROR DETECTED?
YES
1812
RECONSTRUCT DATA
1814
COMMUNICATE CACHE LINE TO PROCESSOR
1816
RECEIVE DATA TO BE WRITTEN
1818
GENERATE CHECKSUM FOR DATA
1820
CONCATENATE DATA AND CHECKSUM
1822
READ ORIGINAL DATA/CHECKSUM AND PARITY/ CHECKSUM FROM MEMORY
1824
GENERATE NEW PARITY VALUE
1826
GENERATE CHECKSUM FOR NEW PARITY
1828
CONCATENATE NEW PARITY AND CHECKSUM
1830
WRITE DATA AND CHECKSUM TO MEMORY
1832
WRITE PARITY AND CHECKSUM TO MEMORY
END

FIG. 18

1900

Processor	8-Core OOO, 2GHz
L1 cache	Fully Private, 3 cycle 2-way, 32 KB each I and D
L2 cache	Fully shared, 10 cycle 8-way, 2 MB, 64B Cache lines
Row-buffer size	8 KB
DRAM Frequency	400 MHz
DRAM Part	512MB, x8
Devices per DIMM	8
Channels	1
Ranks	2
Banks	4
T-rcd, T-cas, T-rp	5 cyc

TEST BENCH PARAMETERS

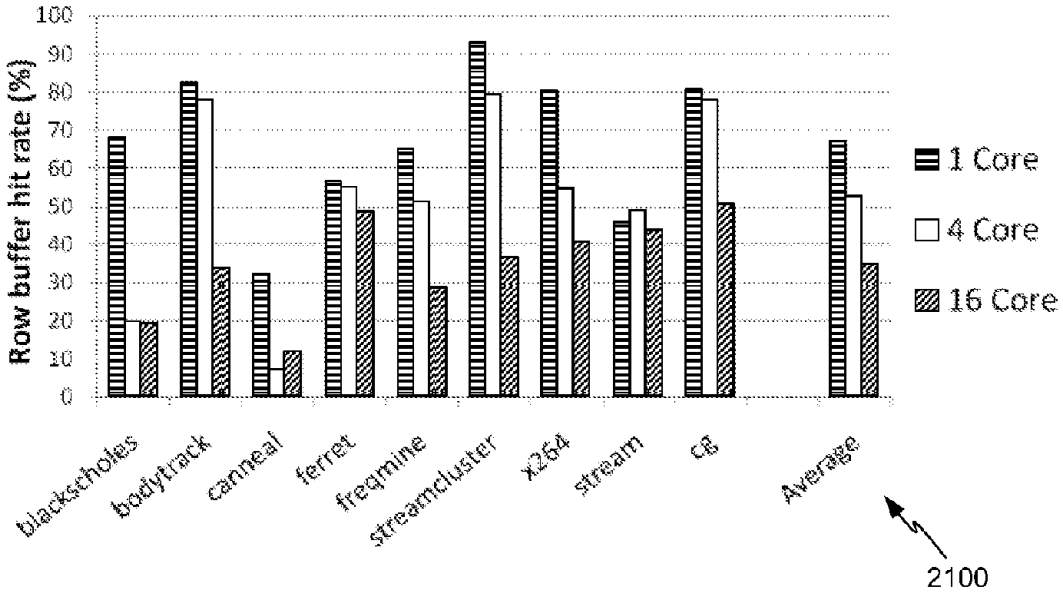
FIG. 19

2000

Component	**Dynamic Energy (nJ)**
Decoder/Wordline - Baseline	1.429
Decoder/Wordline - SBA	0.024
Decoder/Wordline - SSA	0.013
Bitlines - Baseline	19.282
Bitlines - SBA/SSA	0.151
Termination Resistors Baseline/SBA/SSA	14.569
Output Drivers	2.185
Others - Baseline/SBA/SSA	1.143
Low-power mode	**Background power (mW)**
Active	142.5
Power Down (3 mem. cyc)	19
Self Refresh (200 mem. cyc)	6.1

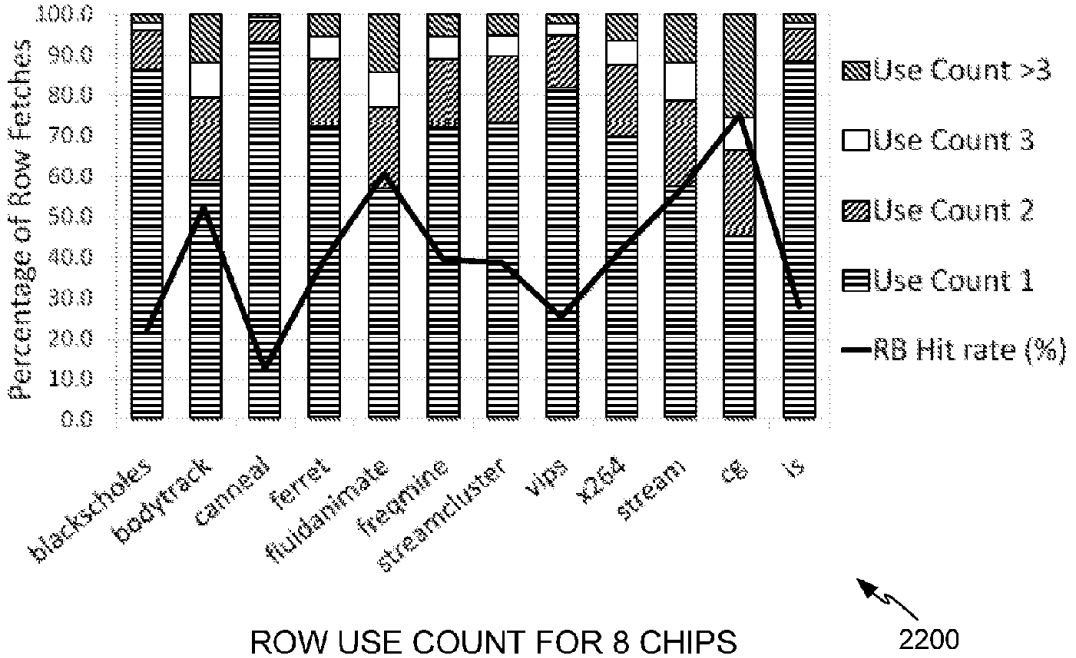
ENERGY CONSUMPTION COMPARISONS

FIG. 20



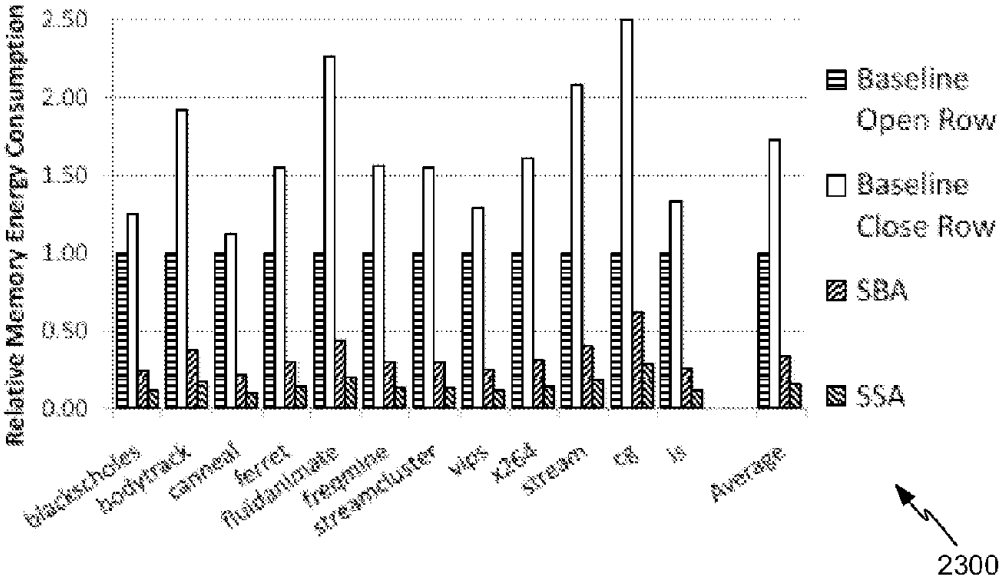
ROW BUFFER HIT RATES

FIG. 21



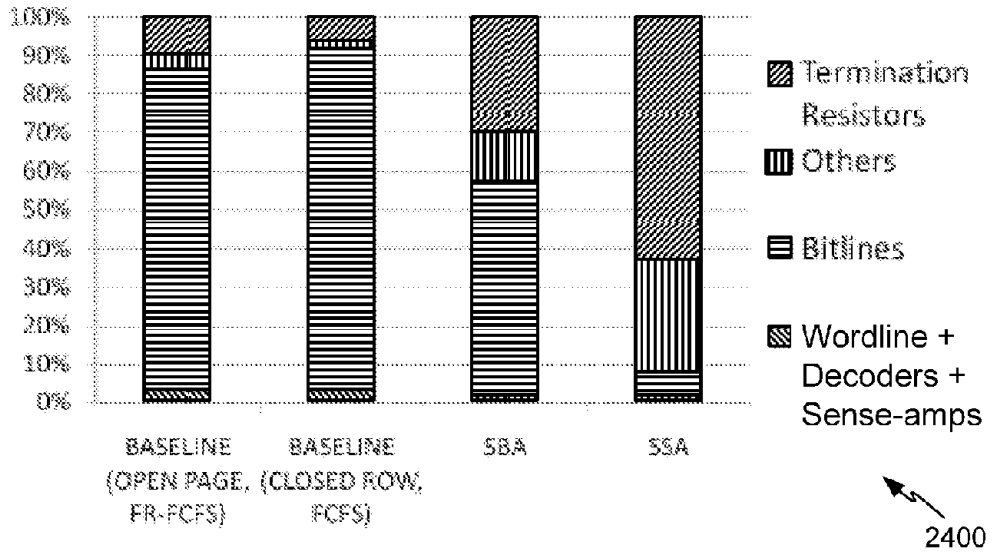
ROW USE COUNT FOR 8 CHIPS

FIG. 22



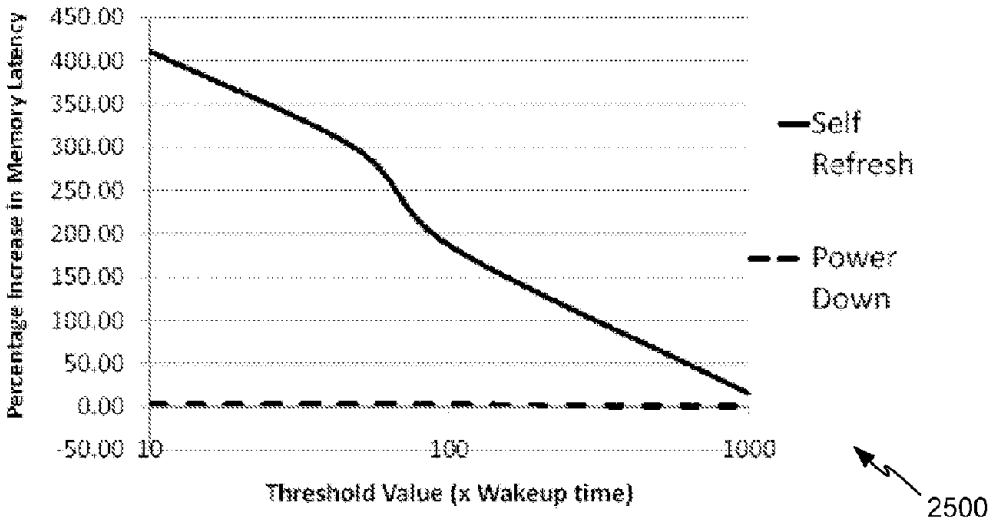
NORMALIZED TOTAL DRAM ENERGY CONSUMPTION

FIG. 23



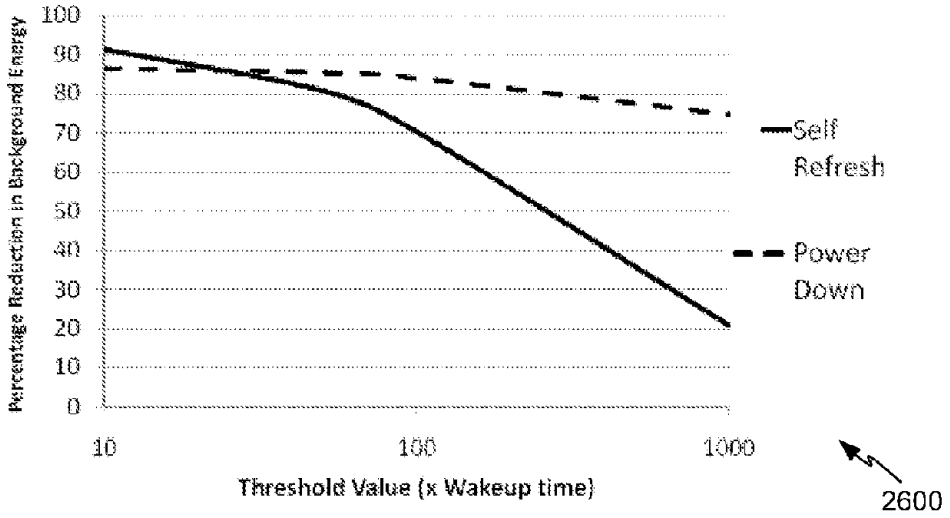
CONTRIBUTORS TO TOTAL DRAM ENERGY CONSUMPTION

FIG. 24



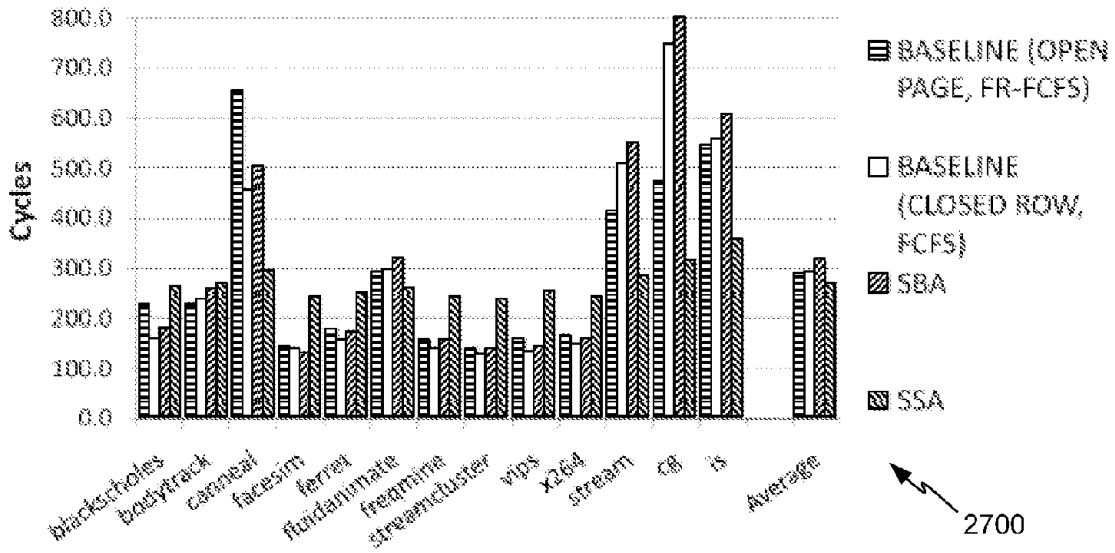
MEMORY LATENCY IMPACT OF USING LOW-POWER STATES

FIG. 25



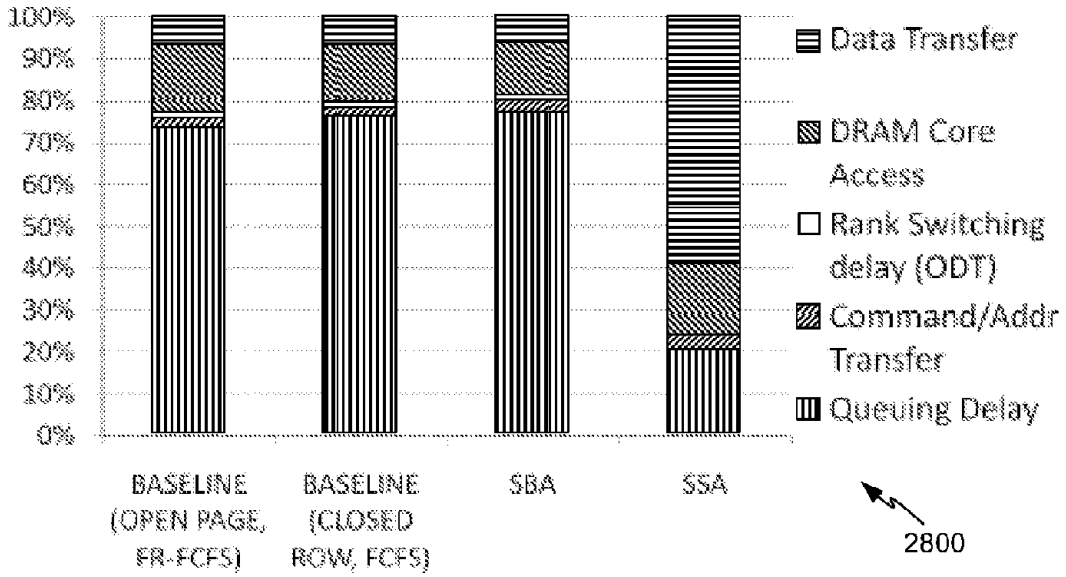
ENERGY REDUCTION USING LOW-POWER STATES

FIG. 26



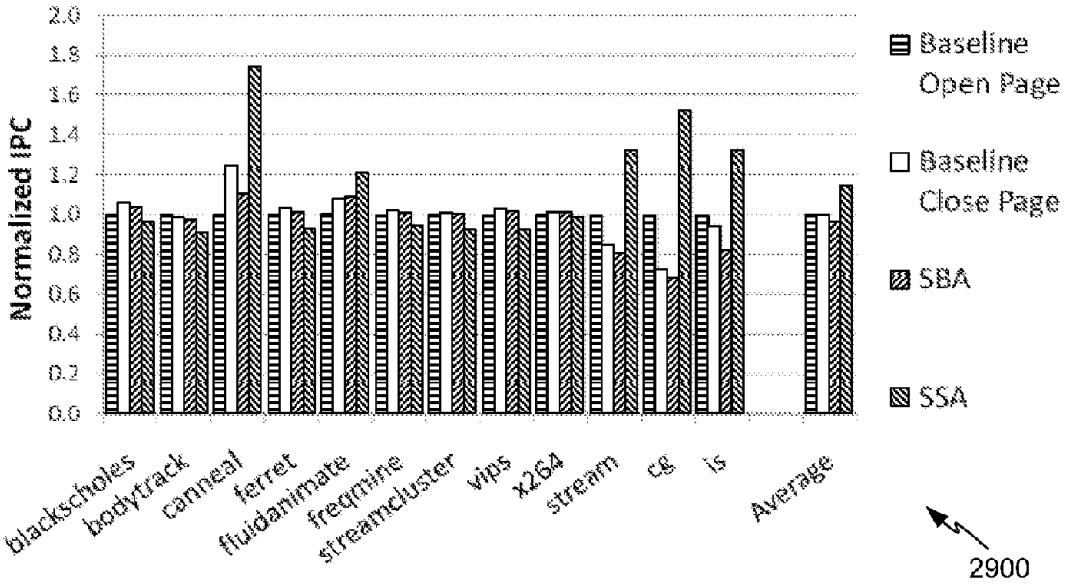
AVERAGE MAIN MEMORY LATENCY

FIG. 27



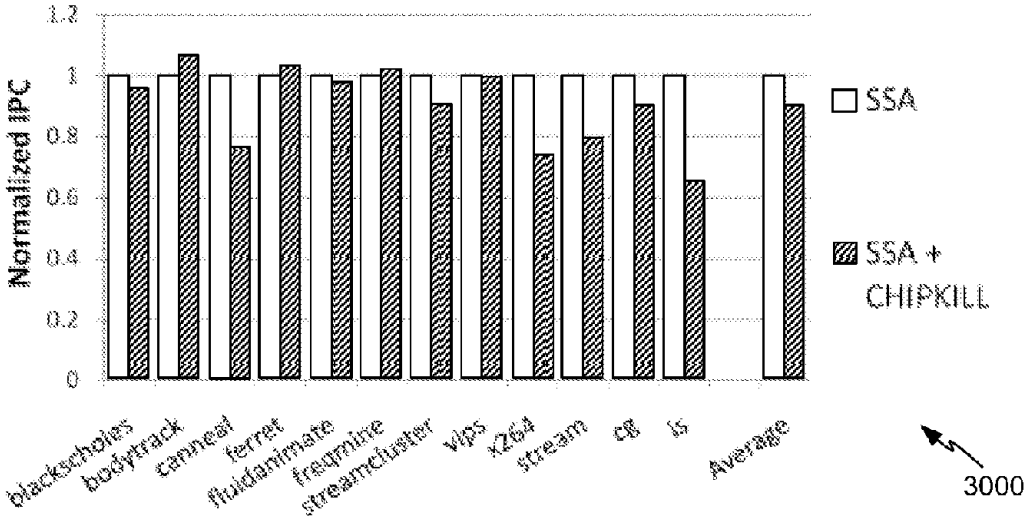
CONTRIBUTORS TO TOTAL MEMORY LATENCY

FIG. 28



NORMALIZED IPCs FOR DIFFERENT MEMORY STRUCTURES

FIG. 29



PERFORMANCE OF CHIPKILL SSA V. NON-CHIPKILL SSA

FIG. 30

Electronic Acknowledgement Receipt

EFS ID:	12849461
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	20995
Filer:	Bruce S. Itchkawitz/Daniela Lopez
Filer Authorized By:	Bruce S. Itchkawitz
Attorney Docket Number:	NETL.071A
Receipt Date:	23-MAY-2012
Filing Date:	03-NOV-2011
Time Stamp:	18:43:44
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		IDS_NETL071A.pdf	223418 bd80bd488bbeb7622d10b00b8c4794e26b306915	yes	5

	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Transmittal Letter		1	1	
	Information Disclosure Statement (IDS) Form (SB08)		2	5	
Warnings:					
Information:					
2	Foreign Reference	WO_2010_138480.pdf	1321721 3d713b680c62c0ca334c83e3629d4dccd3a a5d1f	no	30
Warnings:					
Information:					
3	Foreign Reference	WO_2011_049710.pdf	1264571 cJ842ed5fc0ed5fdecfc41841bf62815fb582 68f	no	28
Warnings:					
Information:					
4	Foreign Reference	WO_2011_094437.pdf	2972850 e00d7680a136b7176ce7452d2bd6d565ce e78e3d	no	80
Warnings:					
Information:					
5	Non Patent Literature	AHMED_ModelingofPeak_to_peakSwitchingNoiseAlongVerticalChainOfPowerDistributionTSVpairs.pdf	724646 2f412441ccc104065fb3936509bda24da16 0bd41	no	6
Warnings:					
Information:					
6	Non Patent Literature	BLACK_DieStacking3DMicroarchitecture.pdf	1448588 9eda9025ad8eac4252dfb4c52a6491f8da9. f301	no	11
Warnings:					
Information:					
7	Non Patent Literature	DANESHTALAB_CMIT_A_novelCluster_basedTopologyFor3DStacked.pdf	675937 f714738e8d5a1cc5ba14d86a18bbd2b1279 47c09	no	5
Warnings:					
Information:					
8	Non Patent Literature	FUNAYA_CachePartitioningStrategiesFor3DStackedVectorProcessors.pdf	765609 b1d084e2cc0d9008836b91d92d6c41bb36 4d45c1	no	6
Warnings:					
Information:					

9	Non Patent Literature	GHOSH_SmartRefreshAnEnhancedMemoryControllerDesignforReducingEnergy.pdf	1741753 99c81d38e64764cea5c6425b8bb86e963d3613b5	no	12
Warnings:					
Information:					
10	Non Patent Literature	KANG_8Gb_3D_DDR3DRAM_UsingThroughSiliconViaTechnology.pdf	1308333 5015946627f83e1fceefbf0a6a1ced82a9361005	no	9
Warnings:					
Information:					
11	Non Patent Literature	KANG_SignalIntegrityAndReliabilityOfANewMultiStackPackageUsingPressure.pdf	520065 2f7898d2f12658865afa6c9293809f0011fd6dce	no	4
Warnings:					
Information:					
12	Non Patent Literature	KAWANO_A_3D_PackagingTechnologyForHighDensityStackedDRAM.pdf	263692 03d6aa5b5103de18192f5033d2680d3dd19cdb2c	no	2
Warnings:					
Information:					
13	Non Patent Literature	KIM_A_QuantitativeAnalysisOfPerformanceBenefitsOf3D_DieStatckingOnMobile.pdf	734749 165a58f28561be71e1efea0c4b41bb7d0a336dbd	no	6
Warnings:					
Information:					
14	Non Patent Literature	Kurita3_D_PackagingTechWithHighlyParrallelMemoryIEICE.pdf	4270905 bc77a899514e0fa9e04609e2342d77c1577f90c2	no	12
Warnings:					
Information:					
15	Non Patent Literature	KURITA_VerticalIntegrationOfStackedDRAMandHighSpeedLogic.pdf	1224098 7390b09b5b4dc865773ce25a262c18dac962046e	no	9
Warnings:					
Information:					
16	Non Patent Literature	POST_TAGGED_3D_STACKING_chipdesignmag_10132011.pdf	2372024 75b693a9f2041ea9472f556df56ecd4ed8e1f13c	no	21
Warnings:					
Information:					
17	Non Patent Literature	RUSSELL_Gill_IntelMicronHybridMemoryCube_09192011.pdf	670368 a86af3a029657ac2ea85fb626684da60cb2917f3	no	8
Warnings:					
Information:					

18	Non Patent Literature	LOH_3D_StackedMemoryArchitecturesForMulti_coreProcessors.pdf	1726634 3a9c9d630b8890eb0c046ea1db613caefdf669f0d	no	12
Warnings:					
Information:					
19	Non Patent Literature	VAL_The3D_InterconnectionAppIsForMassMemories.pdf	1174874 b08fea281ffbc128c5502deac861eca0e2ee46cc	no	12
Warnings:					
Information:					
20	Non Patent Literature	WEIS_DesignSpaceExplorationFor3D_StackedDRAMs_Preceedings_Design.pdf	832128 9876265c234c510d616a515679e0869da5fbb9c1	no	6
Warnings:					
Information:					
21	Non Patent Literature	ZHANG_A_customized_Design_of_RAM_1119222010.pdf	369805 75c5f7f961d0e9c76e534da61508eb156f7cffc8	no	4
Warnings:					
Information:					
		Total Files Size (in bytes):	26606768		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE STATEMENT

Applicant	:	Hyun Lee
App. No	:	13/288,850
Filed	:	November 3, 2011
For	:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
Examiner	:	Amir Zarabian
Art Unit	:	2827
Conf No.	:	2466

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Submitted herewith in the above-identified application is an Information Disclosure Statement listing references for consideration. Copies of any listed foreign and non-patent literature references are being submitted.

Timing of Disclosure

This Information Disclosure Statement is being filed before the receipt of a first Office Action on the merits, and presumably no fee is required. If a first Office Action on the merits was mailed before the mailing date of this Statement, the Commissioner is authorized to charge the fee set forth in 37 CFR 1.17(p) to Deposit Account No. 11-1410.

Respectfully submitted,
KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: 5/23/12

By: _____

Bruce S. Itchkawitz, Ph.D.
Registration No. 47,677
Attorney of Record
Customer No. 20995
(949) 760-0404

13027393

PTO/SB/81 (01-09)
Approved for use through 11/30/2011. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

POWER OF ATTORNEY OR REVOCATION OF POWER OF ATTORNEY WITH A NEW POWER OF ATTORNEY AND CHANGE OF CORRESPONDENCE ADDRESS

Application Number	13/288850
Filing Date	3-Nov-2011
First Named Inventor	Lee, Hyun
Title	ARCHITECTURE FOR MEMORY MODULE WITH PACI
Art Unit	2811
Examiner Name	Unknown
Attorney Docket Number	NETL.071A

I hereby revoke all previous powers of attorney given in the above-identified application.

☐ A Power of Attorney is submitted herewith.

OR

☒ I hereby appoint Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s) to prosecute the application identified above, and to transact all business in the United States Patent and Trademark Office connected therewith:

79141

OR

☐ I hereby appoint Practitioner(s) named below as my/our attorney(s) or agent(s) to prosecute the application identified above, and to transact all business in the United States Patent and Trademark Office connected therewith:

Practitioner(s) Name	Registration Number

Please recognize or change the correspondence address for the above-identified application to:

☒ The address associated with the above-mentioned Customer Number.

OR

☐ The address associated with Customer Number:

OR

☐ Firm or Individual Name					
Address					
City		State		Zip	
Country					
Telephone		Email			

I am the:

☐ Applicant/Inventor.

OR

☒ Assignee of record of the entire interest. See 37 CFR 3.71.
Statement under 37 CFR 3.73(b) (Form PTO/SB/96) submitted herewith or filed on _____.

SIGNATURE of Applicant or Assignee of Record

Signature		Date	September 20, 2012
Name	Gail Sasaki	Telephone	949-435-0025
Title and Company	VP and CFO		

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

☐ *Total of _____ forms are submitted.

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/96 (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

STATEMENT UNDER 37 CFR 3.73(b)

Applicant/Patent Owner: Netlist, Inc.

Application No./Patent No.: 13/288,850 Filed/Issue Date: 3-Nov-2011

Titled: ARCHITECTURE FOR MEMORY MODULE WITH PACKAGES OF THREE-DIMENSIONAL STACKED (3DS) MEMORY CHIPS

Netlist, Inc. , a Corporation

(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.

states that it is:

1. ☒ the assignee of the entire right, title, and interest in;

2. ☐ an assignee of less than the entire right, title, and interest in (The extent (by percentage) of its ownership interest is _____ %); or

3. ☐ the assignee of an undivided interest in the entirety of (a complete assignment from one of the joint inventors was made)

the patent application/patent identified above, by virtue of either:

A. ☒ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel 027520, Frame 0280, or for which a copy therefore is attached.

OR

B. ☐ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

2. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

3. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet(s).

☒ As required by 37 CFR 3.73(b)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (*i.e.*, a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

Signature	SEPTEMBER 20, 2012 Date
Gail Sasaki Printed or Typed Name	VP and CFO Title

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **Send to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Electronic Acknowledgement Receipt

EFS ID:	13837358
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	20995
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	25-SEP-2012
Filing Date:	03-NOV-2011
Time Stamp:	20:34:10
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Power of Attorney	POA071A.pdf	350283 2e3afdedc811944cbd461f3f35172240ebdbe799	no	2

Warnings:

The page size in the PDF is too large. The pages should be 8.5 x 11 or A4. If this PDF is submitted, the pages will be resized upon entry into the Image File Wrapper and may affect subsequent processing	
Information:	
Total Files Size (in bytes):	350283

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
13/288,850	11/03/2011	Hyun Lee	NETL.071A

CONFIRMATION NO. 2466

79141
The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

POA ACCEPTANCE LETTER



Date Mailed: 10/03/2012

NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 09/25/2012.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

/tnnguyen/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

PTO/SB/08a/b(07-09)
Approved for use through 7/31/2012. OMB 0651-0031
US Patent & Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	13/288,850
Filing Date	03-Nov-2011
First Named Inventor	Hyun Lee
Art Unit	2827
Examiner Name	Amir Zarabian
Attorney Docket Number	NETL.071A

US PATENT DOCUMENTS

Examiner Initials *	Cite No.[1]	Document Number / Number-Kind Code [2] (if known)	Publication Date (MM-DD-YYYY)	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	1	20080025137	01-31-2008	Natarajan et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials *	Cite No.[1]	Foreign Patent Document	Publication Date (MM-DD-YYYY)	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]
	2	EP 1 816 570 A2	08-08-2007	Rambus, Inc.		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	3	International Search Report and Written Opinion, PCT/US2011/059209, January 31, 2013	

EXAMINER **DATE CONSIDERED**

* EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patents Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST. 3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document and WIPO Standard ST. 16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
Christensen, Michael R.
KNOBBE, MARTENS, OLSON & BEAR LLP
2040 Main Street, 14th Floor
Irvine, CA 92614
ETATS-UNIS D'AMERIQUE

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference NETL071WO	Date of mailing *(day/month/year)* 30 January 2013 (30-01-2013)
International application No. PCT/US2011/059209	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
Applicant NETLIST, INC.	International filing date *(day/month/year)* 3 November 2011 (03-11-2011)

1. [X] The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the International Search Report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Fascimile No.: (41-22) 338.82.70

   For more detailed instructions, see *PCT Applicant's Guide*, International Phase, paragraphs 9.004 - 9.011.

2. [ ] The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. [ ] **With regard to any protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   [ ] the protest together with the decision thereon has been transmitted to the International Bureau together with any request to forward the texts of both the protest and the decision thereon to the designated Offices.

   [ ] no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. Following the expiration of 30 months from the priority date, these comments will also be made available to the public.

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau before completion of the technical preparations for international publication (Rules 90*bis*.1 and 90*bis*.3).

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   For details about the applicable time limits, Office by Office, see www.wipo.int/pct/en/texts/time_limits.html and the *PCT Applicant's Guide*, National Chapters.

Name and mailing address of the International Searching Authority	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040 Fax: (+31-70) 340-3016	PITARD, Jacqueline Tel: +49 (0)89 2399-2562

Form PCT/ISA/220 (July 2010)

Received Orange County Docketing

FEB 05 2013

Knobbe Martens Olson & Bear LLP

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
NETL071WO		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/US2011/059209	03/11/2011	03/11/2010
Applicant		
NETLIST, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language**, the international search was carried out on the basis of:
      [X] the international application in the language in which it was filed
      [ ] a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. [ ] This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).
   c. [ ] With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. [ ] **Certain claims were found unsearchable** (See Box No. II)

3. [X] **Unity of invention is lacking** (see Box No III)

4. With regard to the **title,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the **drawings,**
   a. the figure of the **drawings** to be published with the abstract is Figure No. 2
      [X] as suggested by the applicant
      [ ] as selected by this Authority, because the applicant failed to suggest a figure
      [ ] as selected by this Authority, because this figure better characterizes the invention
   b. [ ] none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/059209

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☒ No protest accompanied the payment of additional search fees.

Petitioners
Ex. 1002, p. 237

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/059209

A. CLASSIFICATION OF SUBJECT MATTER
INV. G11C5/06 G11C7/10 G11C7/18
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G11C

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, IBM-TDB, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2008/025137 A1 (RAJAN SURESH NATARAJAN [US] ET AL) 31 January 2008 (2008-01-31) paragraph [0051]; figures 2b,2c	1-28
A	US 2006/233012 A1 (SEKIGUCHI TOMONORI [JP] ET AL) 19 October 2006 (2006-10-19) paragraph [0138]; figure 16	1-28
X	EP 1 816 570 A2 (RAMBUS INC [US]) 8 August 2007 (2007-08-08) paragraph [0019] - paragraph [0022]; figures 2,3 paragraph [0109] - paragraph [0112]; figure 15	29-34

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 January 2013	30/01/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Lindquist, Jim

4

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2011/059209

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008025137 A1	31-01-2008	US 2008025137 A1	31-01-2008
		US 2009285031 A1	19-11-2009
US 2006233012 A1	19-10-2006	CN 1841551 A	04-10-2006
		JP 4309368 B2	05-08-2009
		JP 2006277870 A	12-10-2006
		US 2006233012 A1	19-10-2006
EP 1816570 A2	08-08-2007	DE 1022642 T1	11-01-2001
		DE 00100018 T1	29-12-2005
		DE 02000378 T1	29-12-2005
		DE 06125946 T1	22-11-2007
		DE 06125954 T1	22-11-2007
		DE 06125958 T1	22-11-2007
		DE 69132121 D1	25-05-2000
		DE 69132121 T2	21-09-2000
		DE 69132501 D1	08-02-2001
		DE 69132501 T2	23-08-2001
		DE 69132721 D1	11-10-2001
		DE 69132721 T2	29-05-2002
		DE 69133500 T2	27-07-2006
		DE 69133550 T2	11-01-2007
		DE 69133565 T2	05-07-2007
		DE 69133572 T2	14-02-2008
		EP 0525068 A1	03-02-1993
		EP 1022642 A1	26-07-2000
		EP 1197830 A2	17-04-2002
		EP 1640847 A2	29-03-2006
		EP 1816569 A2	08-08-2007
		EP 1816570 A2	08-08-2007
		EP 1830241 A1	05-09-2007
		IL 96808 A	31-03-1996
		JP 3404383 B2	06-05-2003
		JP 3414393 B2	09-06-2003
		JP 3550143 B2	04-08-2004
		JP H05507374 A	21-10-1993
		JP 2001273765 A	05-10-2001
		JP 2003203008 A	18-07-2003
		US 5319755 A	07-06-1994
		US 5408129 A	18-04-1995
		US 5473575 A	05-12-1995
		US 5499385 A	12-03-1996
		US 5513327 A	30-04-1996
		US 5606717 A	25-02-1997
		US 5638334 A	10-06-1997
		US 5657481 A	12-08-1997
		US 5809263 A	15-09-1998
		US 5841580 A	24-11-1998
		US 5841715 A	24-11-1998
		US 5915105 A	22-06-1999
		US 5928343 A	27-07-1999
		US 5954804 A	21-09-1999
		US 5983320 A	09-11-1999
		US 6032214 A	29-02-2000
		US 6032215 A	29-02-2000
		US 6034918 A	07-03-2000
		US 6035365 A	07-03-2000
		US 6038195 A	14-03-2000

Petitioners
Ex. 1002, p. 239

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2011/059209

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 6044426 A	28-03-2000
		US 6049846 A	11-04-2000
		US 6067592 A	23-05-2000
		US 6070222 A	30-05-2000
		US 6085284 A	04-07-2000
		US 6101152 A	08-08-2000
		US 6128696 A	03-10-2000
		US 6182184 B1	30-01-2001
		US 6185644 B1	06-02-2001
		US 6260097 B1	10-07-2001
		US 6266285 B1	24-07-2001
		US 6304937 B1	16-10-2001
		US 6314051 B1	06-11-2001
		US 6415339 B1	02-07-2002
		US 6598171 B1	22-07-2003
		US 2001030904 A1	18-10-2001
		US 2002001253 A1	03-01-2002
		US 2002004867 A1	10-01-2002
		US 2002015351 A1	07-02-2002
		US 2002016876 A1	07-02-2002
		US 2002046314 A1	18-04-2002
		US 2002091890 A1	11-07-2002
		US 2002141281 A1	03-10-2002
		US 2004114454 A1	17-06-2004
		US 2005030802 A1	10-02-2005
		US 2005033903 A1	10-02-2005
		US 2005141332 A1	30-06-2005
		WO 9116680 A1	31-10-1991

Petitioners
Ex. 1002, p. 240

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-28

   To arrange die die interconnects (e.g. data paths) to interconnect the control die to a plurality of array dies, to minimize loads.

   ---

2. claims: 29-34

   Modes to operate the data path control circuit, to either transport data bits or command/address information, thus maximizing the use of the data path.

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:

see form PCT/ISA/220

PCT

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing
(day/month/year) see form PCT/ISA/210 (second sheet)

Applicant's or agent's file reference
see form PCT/ISA/220

FOR FURTHER ACTION
See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2011/059209	03.11.2011	03.11.2010

International Patent Classification (IPC) or both national classification and IPC
INV. G11C5/06 G11C7/10 G11C7/18

Applicant
NETLIST, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☒ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☒ Box No. VII Certain defects in the international application
- ☒ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will usually be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notifed the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

Name and mailing address of the ISA:	Date of completion of this opinion	Authorized Officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Fax: +49 89 2399 - 4465	see form PCT/ISA/210	Lindquist, Jim Telephone No. +49 89 2399-2727

Form PCT/ISA/237 (Cover Sheet) (July 2009)

Box No. I Basis of the opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed

   ☐ a translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1 (b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing filed or furnished:

   a. (means)

      ☐ on paper

      ☐ in electronic form

   b. (time)

      ☐ in the international application as filed

      ☐ together with the international application in electronic form

      ☐ subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

Box No. IV Lack of unity of invention

1. ☒ In response to the invitation (Form PCT/ISA/206) to pay additional fees, the applicant has, within the applicable time limit:

   ☒ paid additional fees

   ☐ paid additional fees under protest and, where applicable, the protest fee

   ☐ paid additional fees under protest but the applicable protest fee was not paid

   ☐ not paid additional fees

2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose not to invite the applicant to pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rule 13.1, 13.2 and 13.3 is

   ☐ complied with

   ☒ not complied with for the following reasons:

   see separate sheet

4. Consequently, this report has been established in respect of the following parts of the international application:

   ☐ all parts.

   ☒ the parts relating to claims Nos. 1-34

Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	
	No:	Claims	1-34
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-34
Industrial applicability (IA)	Yes:	Claims	1-34
	No:	Claims	

2. Citations and explanations

see separate sheet

Box No. VII Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Box No. VIII Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

V. Citations and Explanations

1 The following documents are cited in this Written Opinion:

D1 US 2008/025137 A1 (RAJAN SURESH NATARAJAN [US] ET AL) 31 January 2008 (2008-01-31)

D2 EP 1 816 570 A2 (RAMBUS INC [US]) 8 August 2007 (2007-08-08)

2 The subject matter of claim 1 lacks novelty so that the criteria of Articles 33(1) and (2) PCT are not met.

In comparison to claim 1, D1 (fig.2B, para.0051) discloses a memory package, comprising:

a plurality of array dies (206A-D) having data ports;

at least a first die interconnect (Data) and a second die interconnect (Data), the first die interconnect (Data) in electrical communication with at least one data port of a first array die (206A) of the plurality of array dies and at least one data port of a second array die (206C) of the plurality of array dies and not in electrical communication with the data ports of at least a third array die (206B,206D) of the plurality of array dies, the second die interconnect (Data) in electrical communication with at least one data port of the third array die (206B) and not in electrical communication with the data ports of the first array die (206A) and the data ports of the second array die (206C); and

a control die (202) comprising at least a first data conduit (right Data connection to 202) configured to transmit a data signal to the first die interconnect (right Data) and to not transmit the data signal to the second die interconnect (left Data), and at least a second data conduit (left Data connection to 202) configured to transmit the data signal to the second die interconnect (left Data) and to not transmit the data signal to the first die interconnect (right Data).

3 Independent claims 11 and 20 contains substantially the same features as claim 1 only differently worded. The subject-matter of these claims therefore lacks novelty for the same reasons as those set out here above.

4 The subject-matter of claim 29 lacks novelty so that Articles 33(1) and (2) PCT are not met.

In comparison to claim 29, D2 (para.0109-0112, fig.2,3,15) discloses a memory module comprising:

a register device (19) configured to receive command/address signals (20) from a memory control hub (19) and to generate data path control signals (18); and

a plurality of DRAM packages (15,16,17), each DRAM package comprising:

a control die (para.0111, in part. p.16,l.53) comprising a plurality of command/address buffers and a data path control circuit configured to control command/address time slots and data bus time slots (para.0020), the control die configured to receive data signals from the register device, and command/address signals from the register device (para.0021); and

a plurality of DDR DRAM dies (150-157, para.0109) operatively coupled to the control die to receive the data signals from the control die,

wherein the memory module is selectively configurable into at least two operational modes comprising:

a first operational mode (para.0111) in which the register device generates the data path control signals, and the control die uses the data path control signals to operate the data path control circuit (para.0111); and

a second operational mode (0020) in which the control die operate the data path control circuit to provide the command/address signals to the plurality of DDR DRAM dies without decoding the command/address signals (para.0111).

5 The features of the dependent claims are either known from D1 or D2, or derivable therefrom in an obvious manner.

VII. Certain Defects

5.1 The independent claims are not drafted in two-part form delimited against D1, contrary to Rule 6.3.(b) PCT.

5.2 Reference signs are not inserted in the claims contrary to Rule 6.2.(b) PCT.

5.3 Contrary to the requirements of Rule 5.1(a)(ii) PCT, documents D1 and D2 are not identified in the description and the relevant background art disclosed therein is not discussed.

5.4 With regard to the phrase "hereby incorporated by reference" throughout the description, according to Rule 4.14 PCT, if there is a wish to treat the international application as a continuation-in-part of another application, this application must be indicated in the PCT request of the international application.

5.5 The description also contains references to documents not retrievable by the public, e.g. application numbers. These references must either be replaced by retrievable references or they must be removed.

5.6 And finally, the reference to the spirit of the present invention in page 37 is not allowed, since it wrongly makes the impression that the extent of protection is expanded beyond the contents of the claims (see Chapter III, 4.3a of the PCT Preliminary Examination Guidelines).

VIII. Certain Observations

6 The set of claims contains too many independent claims. The various definitions of the invention given in the independent claims 1, 11, 20 and 29 are such that the claims as a whole are not clear and concise, contrary to Article 6 PCT, since it is not clear for which matter protection is sought. To meet the criteria of this article the claims may only include one independent claim in each category, with dependent claims as appropriate (Rule 6.4(a)-(c) PCT).

7 Dependent claims 30-32 are not clear, because the meaning of the term "rank multiplication" is not clear. This term must be clarified to make any positive contribution to the patentability of the previous subject-matter.

Possible steps after receipt of the international search report (ISR) and written opinion of the International Searching Authority (WO-ISA)

General information	For all international applications filed on or after 01/01/2004 the competent ISA will establish an ISR. It is accompanied by the WO-ISA. Unlike the former written opinion of the IPEA (Rule 66.2 PCT), the WO-ISA is not meant to be responded to, but to be taken into consideration for further procedural steps. This document explains about the possibilities.
Amending claims under Art. 19 PCT	Within 2 months after the date of mailing of the ISR and the WO-ISA the applicant may file amended claims under Art. 19 PCT directly with the International Bureau of WIPO. The PCT reform of 2004 did not change this procedure. For further information please see Rule 46 PCT as well as form PCT/ISA/220 and the corresponding Notes to form PCT/ISA/220.
Filing a demand for international preliminary examination	In principle, the WO-ISA will be considered as the written opinion of the IPEA. This should, in many cases, make it unnecessary to file a demand for international preliminary examination. If the applicant nevertheless wishes to file a demand this must be done before expiry of 3 months after the date of mailing of the ISR/WO-ISA or 22 months after priority date, whichever expires later (Rule 54bis PCT). Amendments under Art. 34 PCT can be filed with the IPEA as before, normally at the same time as filing the demand (Rule 66.1 (b) PCT). If a demand for international preliminary examination is filed and no comments/amendments have been received the WO-ISA will be transformed by the IPEA into an IPRP (International Preliminary Report on Patentability) which would merely reflect the content of the WO-ISA. The demand can still be withdrawn (Art. 37 PCT).
Filing informal comments	After receipt of the ISR/WO-ISA the applicant may file informal comments on the WO-ISA directly with the International Bureau of WIPO. These will be communicated to the designated Offices together with the IPRP (International Preliminary Report on Patentability) at 30 months from the priority date. Please also refer to the next box.
End of the international phase	At the end of the international phase the International Bureau of WIPO will transform the WO-ISA or, if a demand was filed, the written opinion of the IPEA into the IPRP, which will then be transmitted together with possible informal comments to the designated Offices. The IPRP replaces the former IPER (international preliminary examination report).
Relevant PCT Rules and more information	Rule 43 PCT, Rule 43bis PCT, Rule 44 PCT, Rule 44bis PCT, PCT Newsletter 12/2003, OJ 11/2003, OJ 12/2003

US 20080025137A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2008/0025137 A1**

Rajan et al. (43) **Pub. Date: Jan. 31, 2008**

(54) **SYSTEM AND METHOD FOR SIMULATING AN ASPECT OF A MEMORY CIRCUIT**

(75) Inventors: **Suresh Natarajan Rajan**, San Jose, CA (US); **Keith R. Schakel**, San Jose, CA (US); **Michael John Sebastian Smith**, Palo Alto, CA (US); **David T. Wang**, San Jose, CA (US); **Frederick Daniel Weber**, San Jose, CA (US)

Correspondence Address:
ZILKA-KOTAB, PC
P.O. BOX 721120
SAN JOSE, CA 95172-1120

(73) Assignee: **METARAM, INC.**

(21) Appl. No.: **11/461,427**

(22) Filed: **Jul. 31, 2006**

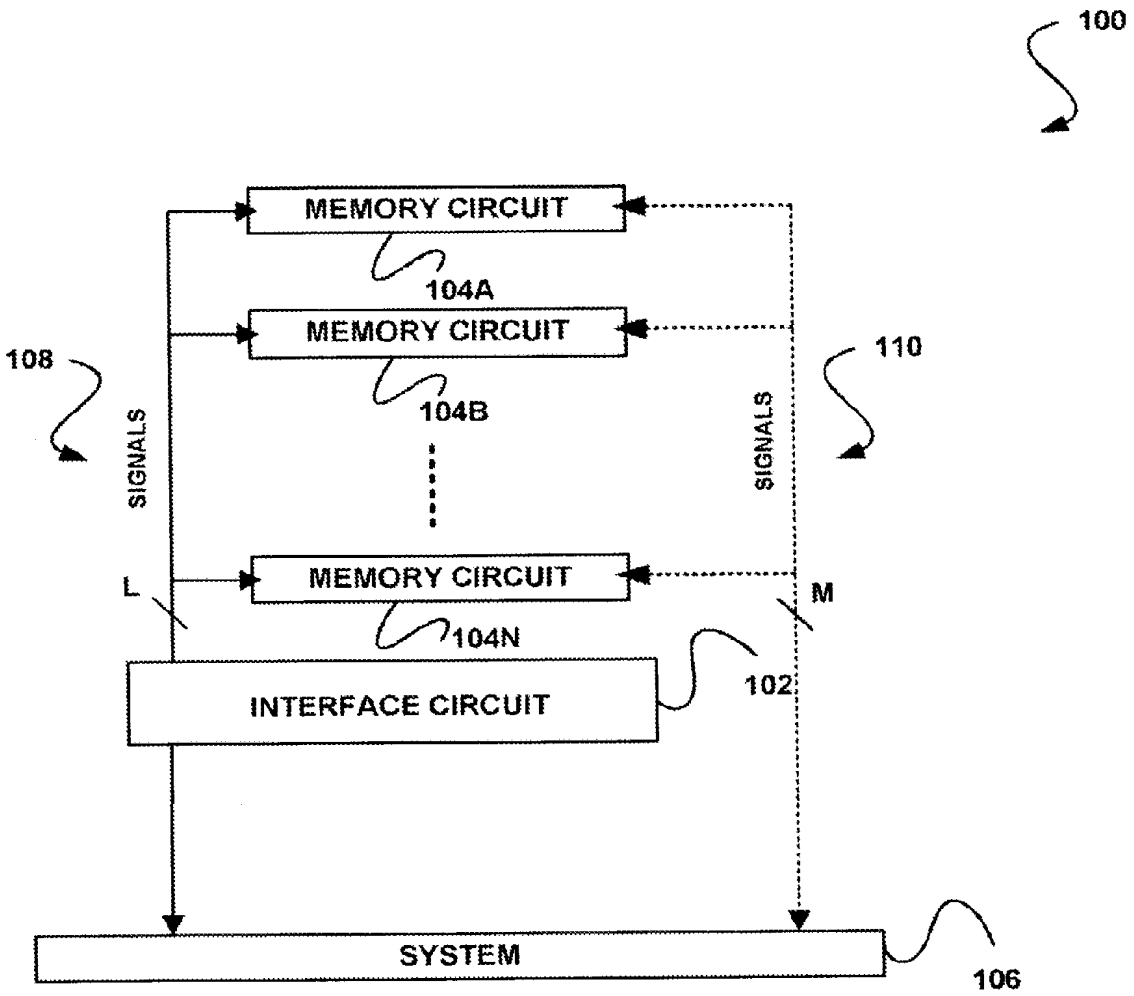
Publication Classification

(51) **Int. Cl.**
G11C 8/00 (2006.01)

(52) **U.S. Cl.** **365/239**

(57) **ABSTRACT**

A system and method are provided for simulating an aspect of a memory circuit. Included is an interface circuit that is in communication with a plurality of memory circuits and a system. Such interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of memory circuits. In accordance with various embodiments, such aspect may include a signal, a capacity, a timing, and/or a logical interface.



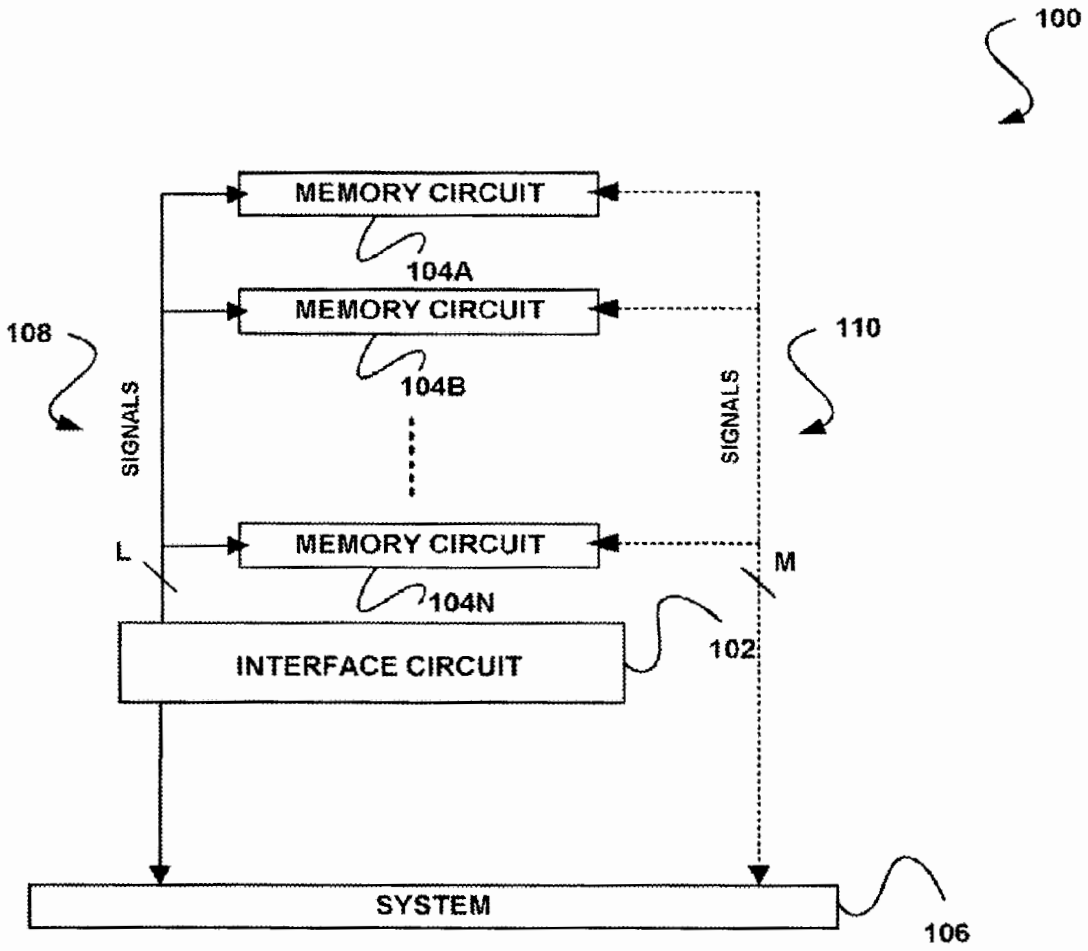

100
MEMORY CIRCUIT
104A
MEMORY CIRCUIT
104B
MEMORY CIRCUIT
104N
108
SIGNALS
110
SIGNALS
L
M
INTERFACE CIRCUIT
102
SYSTEM
106

FIGURE 1

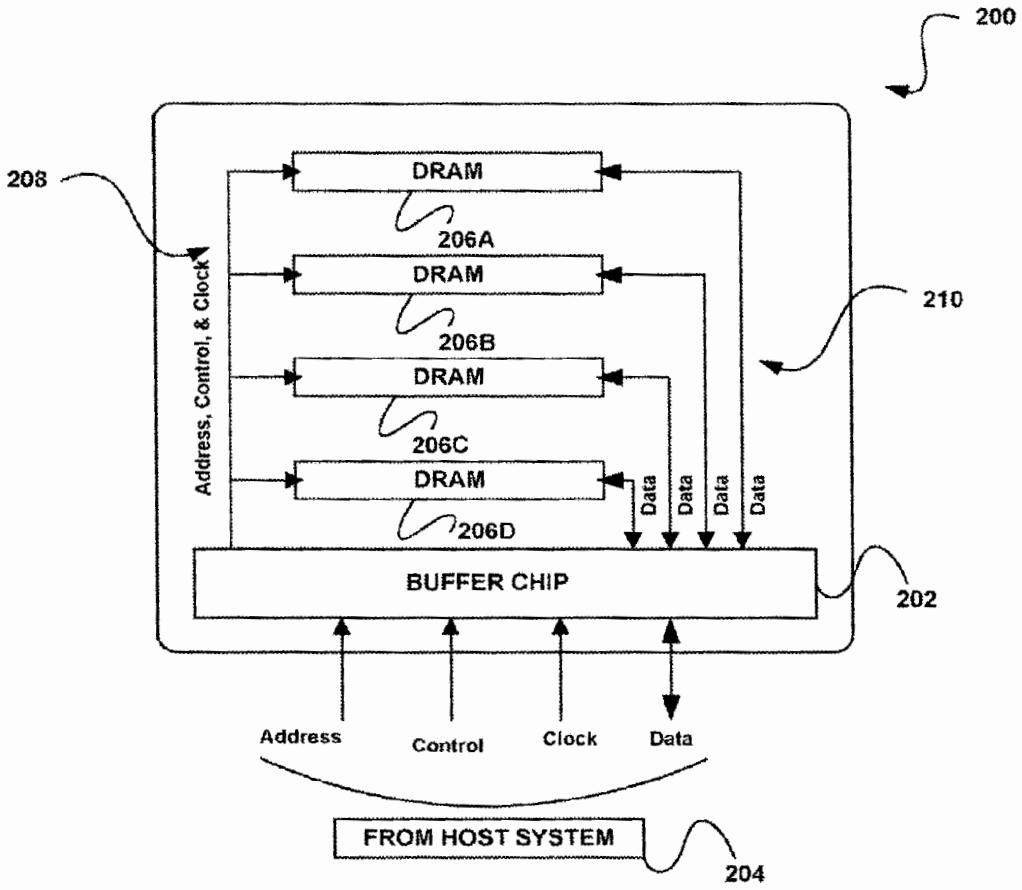

200
208
DRAM
206A
DRAM
206B
DRAM
206C
DRAM
206D
Address, Control, & Clock
Data
Data
Data
Data
210
BUFFER CHIP
202
Address
Control
Clock
Data
FROM HOST SYSTEM
204

FIGURE 2A

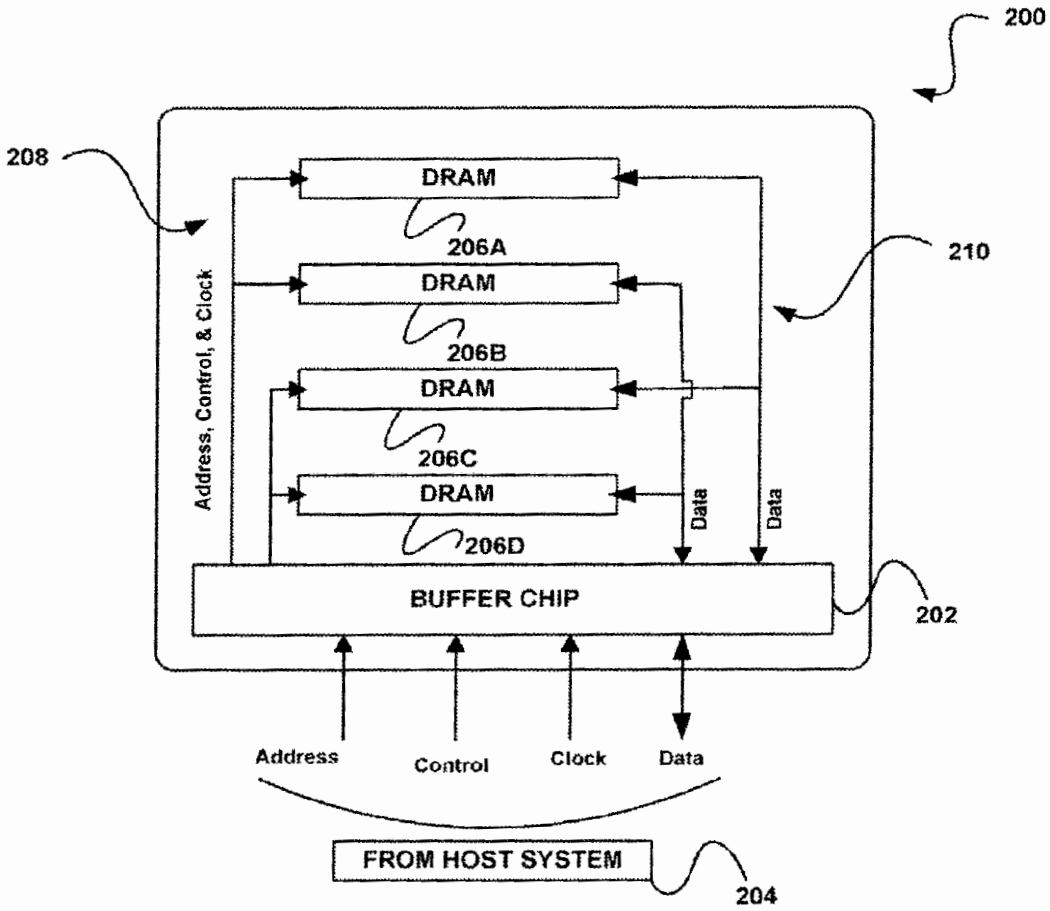

200
208
DRAM
206A
DRAM
206B
DRAM
206C
DRAM
206D
210
Address, Control, & Clock
Data
Data
BUFFER CHIP
202
Address
Control
Clock
Data
FROM HOST SYSTEM
204

FIGURE 2B

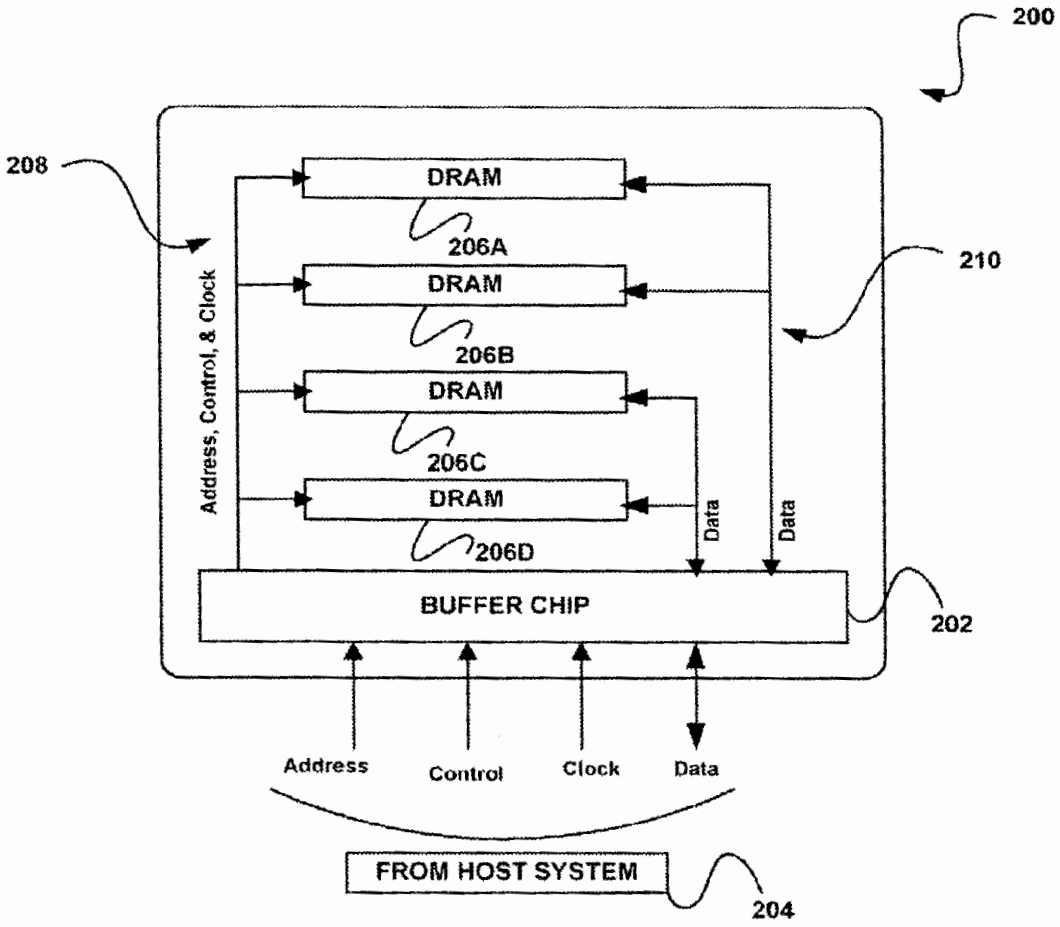

200
208
DRAM
206A
DRAM
206B
DRAM
206C
DRAM
206D
210
Address, Control, & Clock
Data
Data
BUFFER CHIP
202
Address
Control
Clock
Data
FROM HOST SYSTEM
204

FIGURE 2C

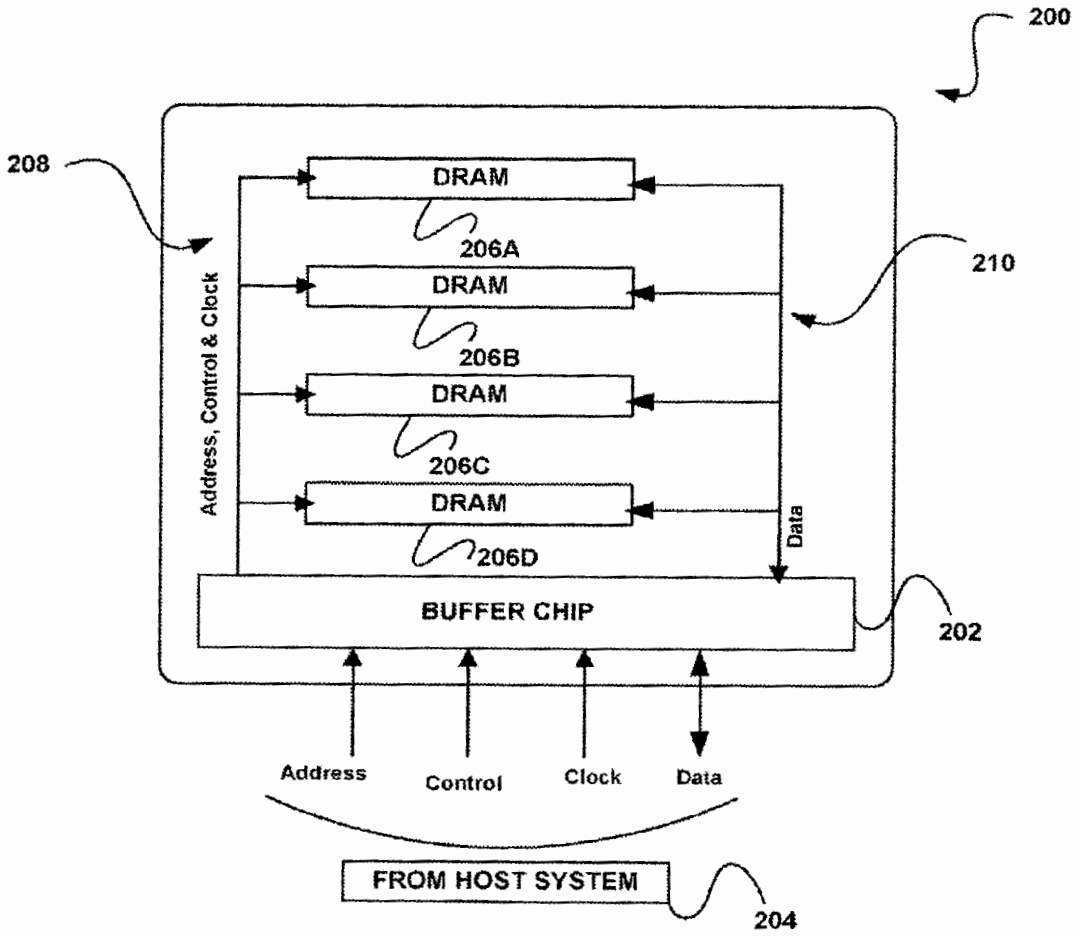

200
208
DRAM
206A
DRAM
206B
DRAM
206C
DRAM
206D
210
Address, Control & Clock
Data
BUFFER CHIP
202
Address
Control
Clock
Data
FROM HOST SYSTEM
204

FIGURE 2D

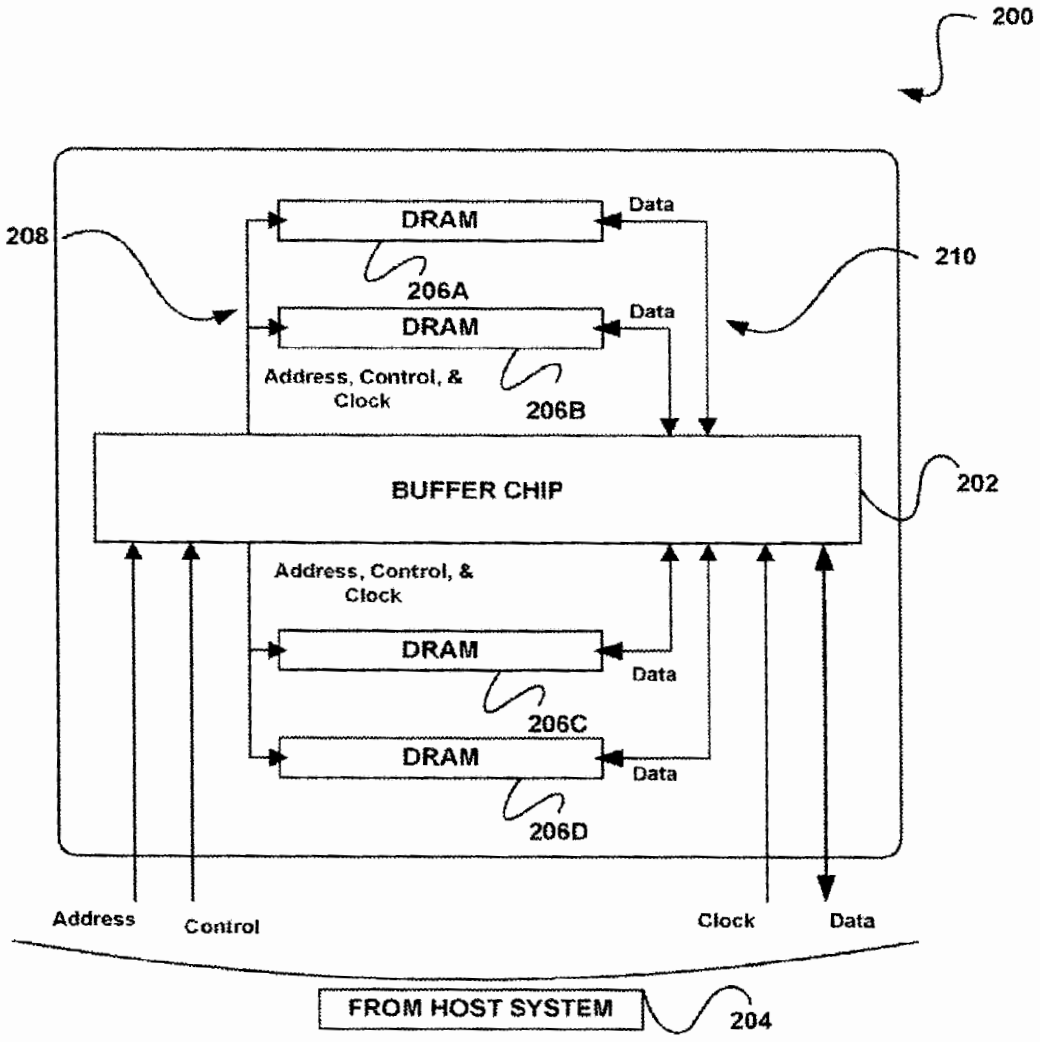

200
208
210
DRAM
Data
206A
DRAM
Data
Address, Control, &
Clock
206B
BUFFER CHIP
202
Address, Control, &
Clock
DRAM
Data
206C
DRAM
Data
206D
Address
Control
Clock
Data
FROM HOST SYSTEM
204

FIGURE 2E

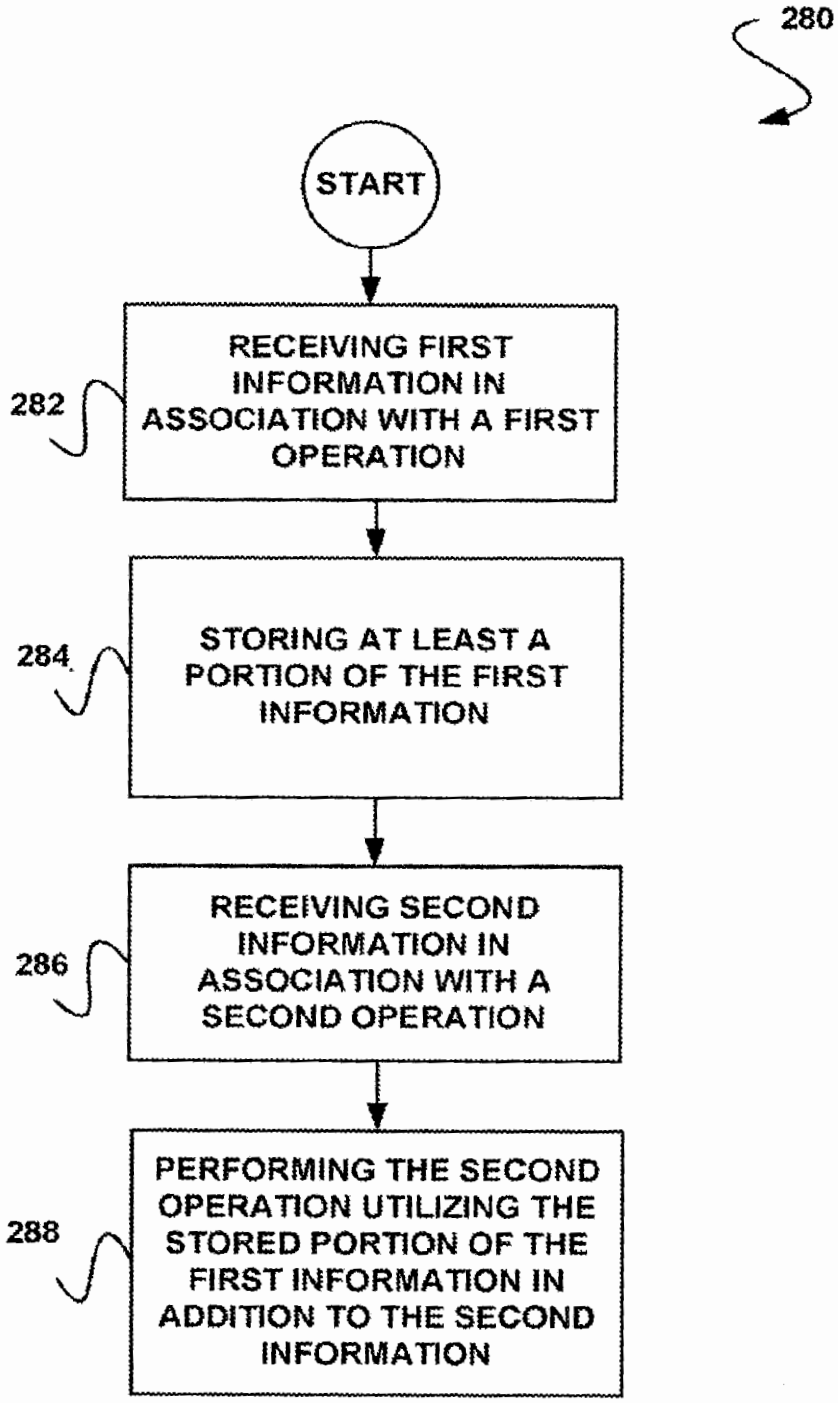

280
START
RECEIVING FIRST INFORMATION IN ASSOCIATION WITH A FIRST OPERATION
282
STORING AT LEAST A PORTION OF THE FIRST INFORMATION
284
RECEIVING SECOND INFORMATION IN ASSOCIATION WITH A SECOND OPERATION
286
PERFORMING THE SECOND OPERATION UTILIZING THE STORED PORTION OF THE FIRST INFORMATION IN ADDITION TO THE SECOND INFORMATION
288

FIGURE 2F

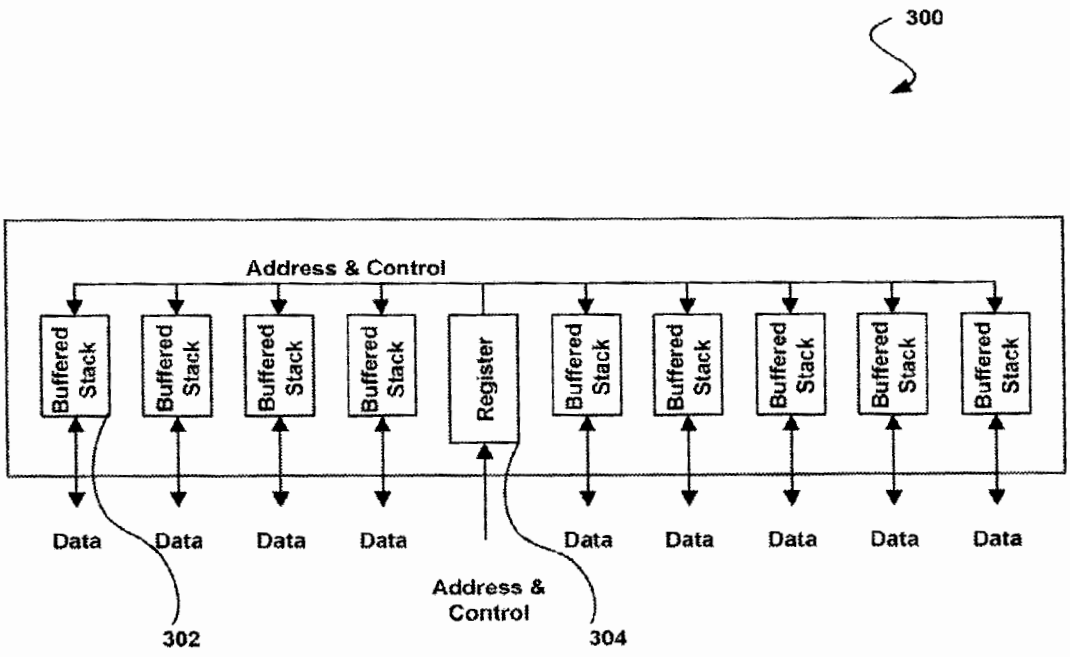

300
Address & Control
Buffered Stack
Buffered Stack
Buffered Stack
Buffered Stack
Register
Buffered Stack
Buffered Stack
Buffered Stack
Buffered Stack
Buffered Stack
Data
Data
Data
Data
Data
Data
Data
Data
Data
Address & Control
302
304

FIGURE 3

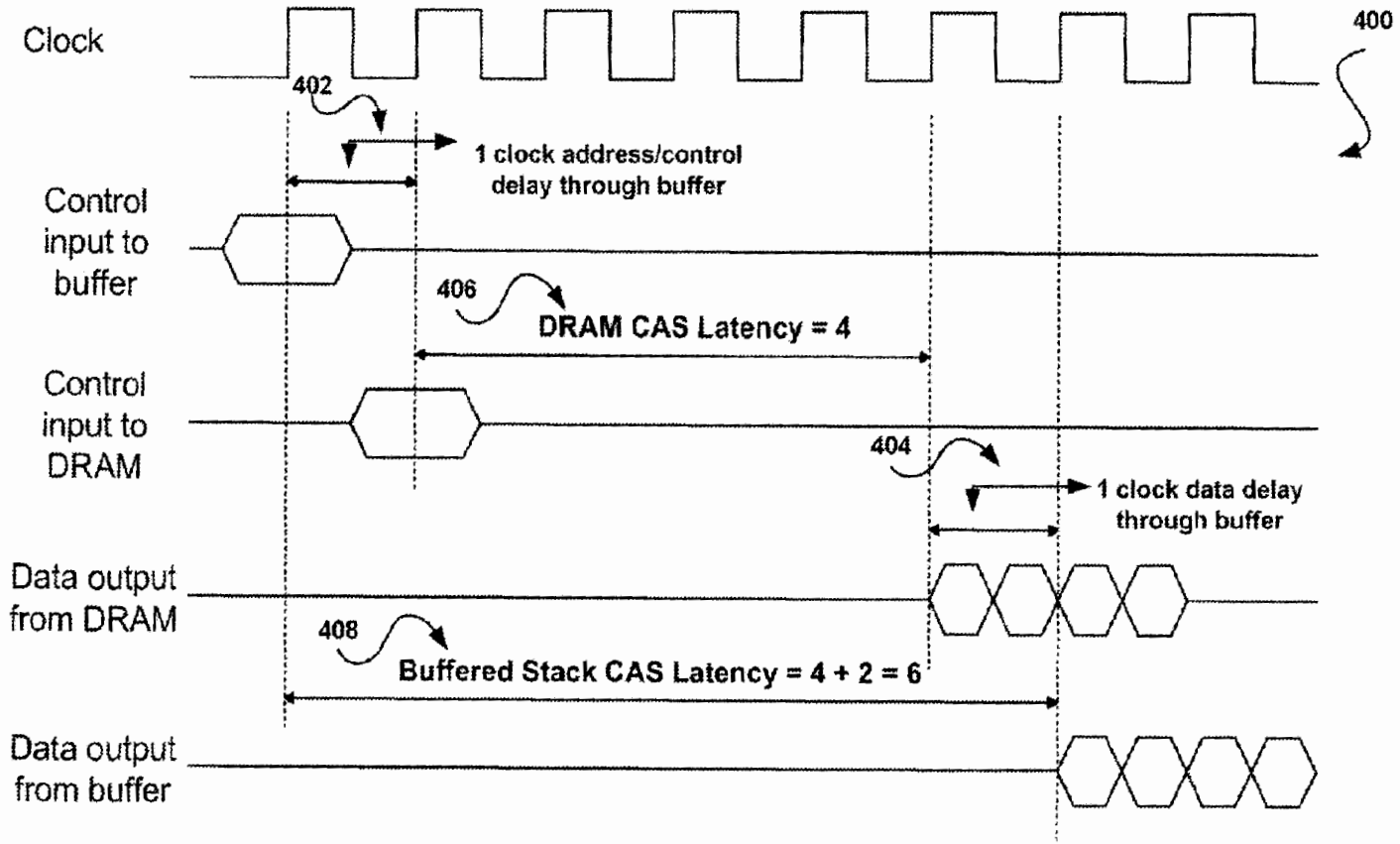

Clock
400
402
1 clock address/control delay through buffer
Control input to buffer
406
DRAM CAS Latency = 4
Control input to DRAM
404
1 clock data delay through buffer
Data output from DRAM
408
Buffered Stack CAS Latency = 4 + 2 = 6
Data output from buffer

FIGURE 4

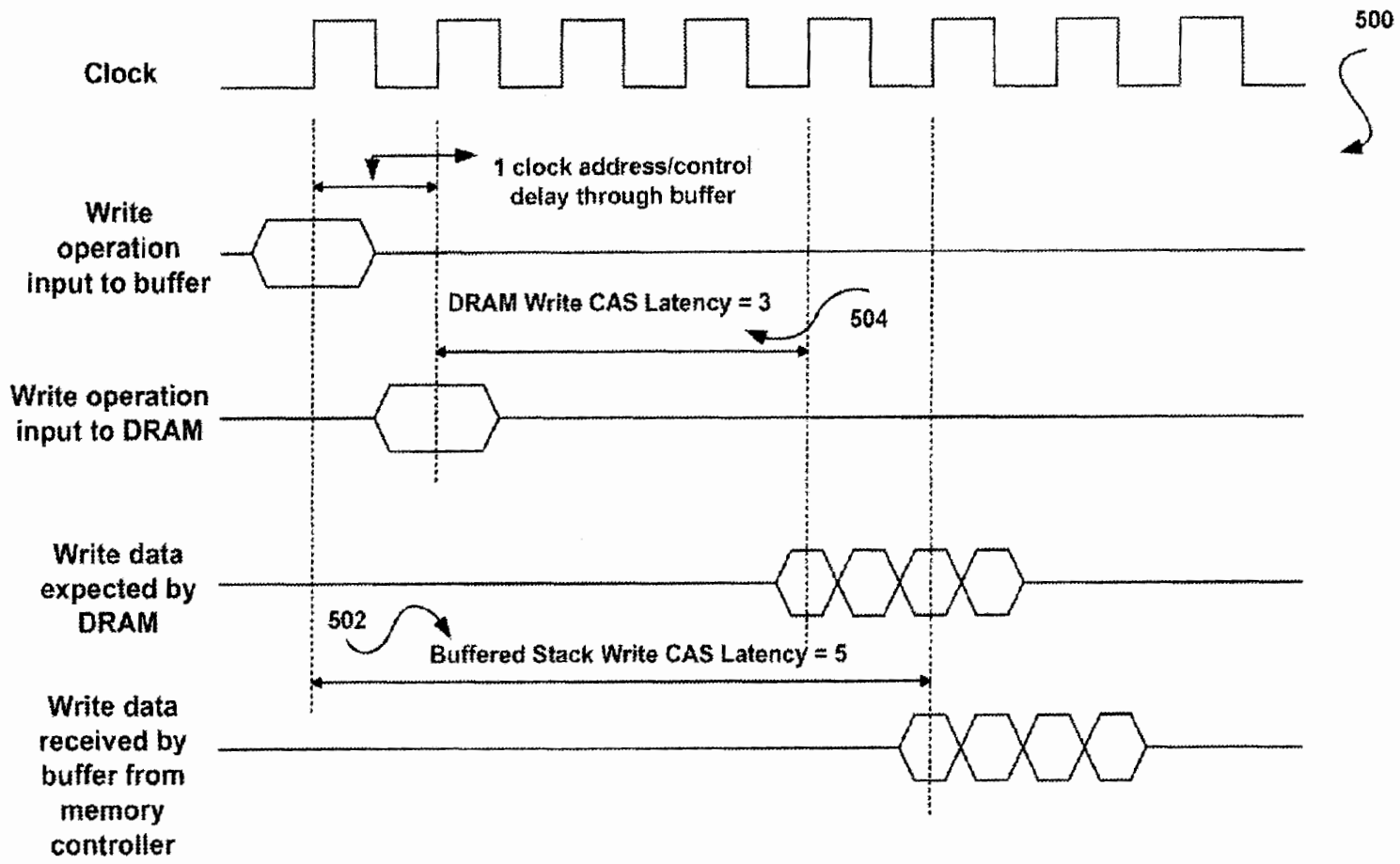

Clock
500
1 clock address/control delay through buffer
Write operation input to buffer
DRAM Write CAS Latency = 3
504
Write operation input to DRAM
Write data expected by DRAM
502
Buffered Stack Write CAS Latency = 5
Write data received by buffer from memory controller

FIGURE 5

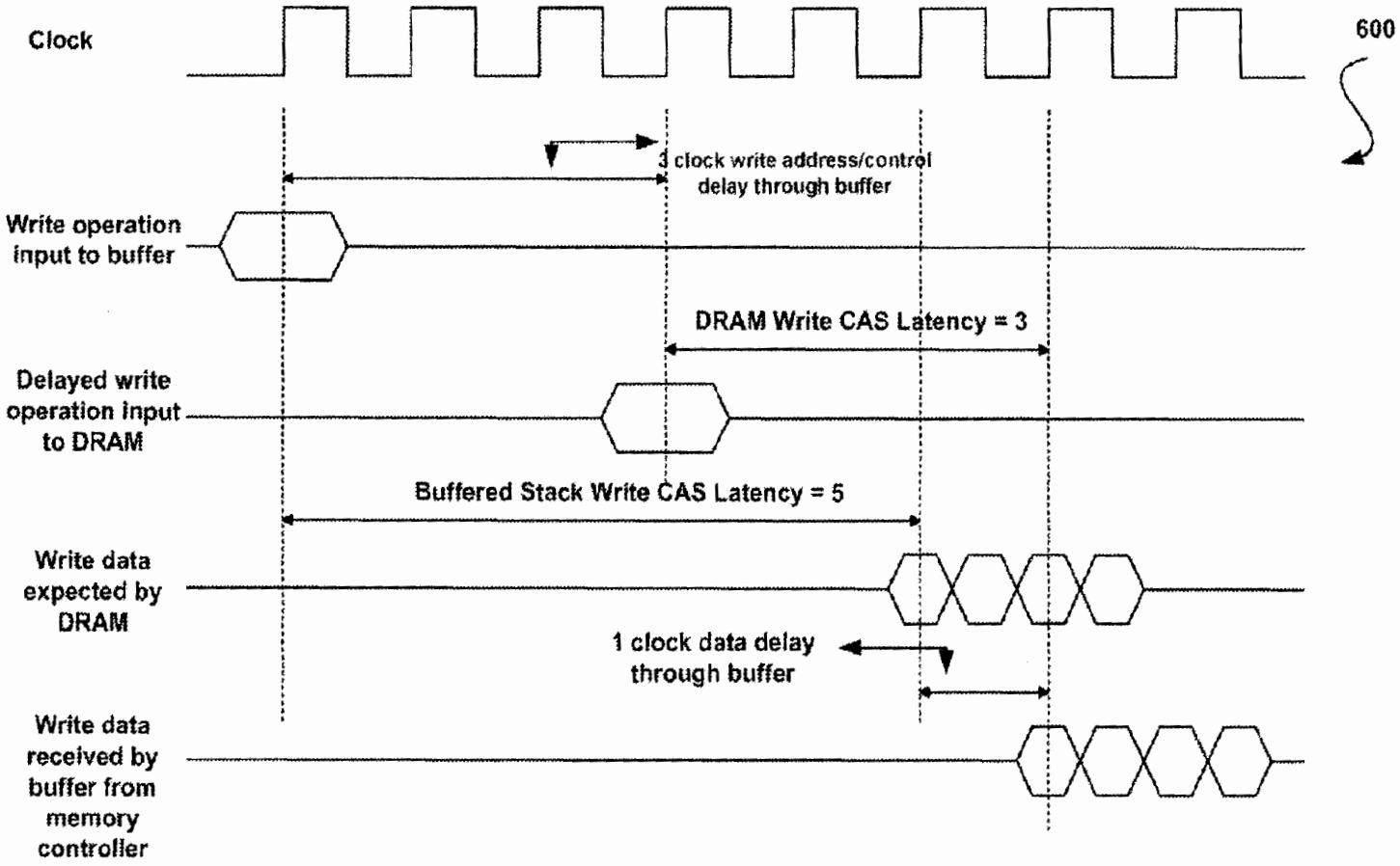

Clock
600
3 clock write address/control delay through buffer
Write operation input to buffer
DRAM Write CAS Latency = 3
Delayed write operation input to DRAM
Buffered Stack Write CAS Latency = 5
Write data expected by DRAM
1 clock data delay through buffer
Write data received by buffer from memory controller

FIGURE 6

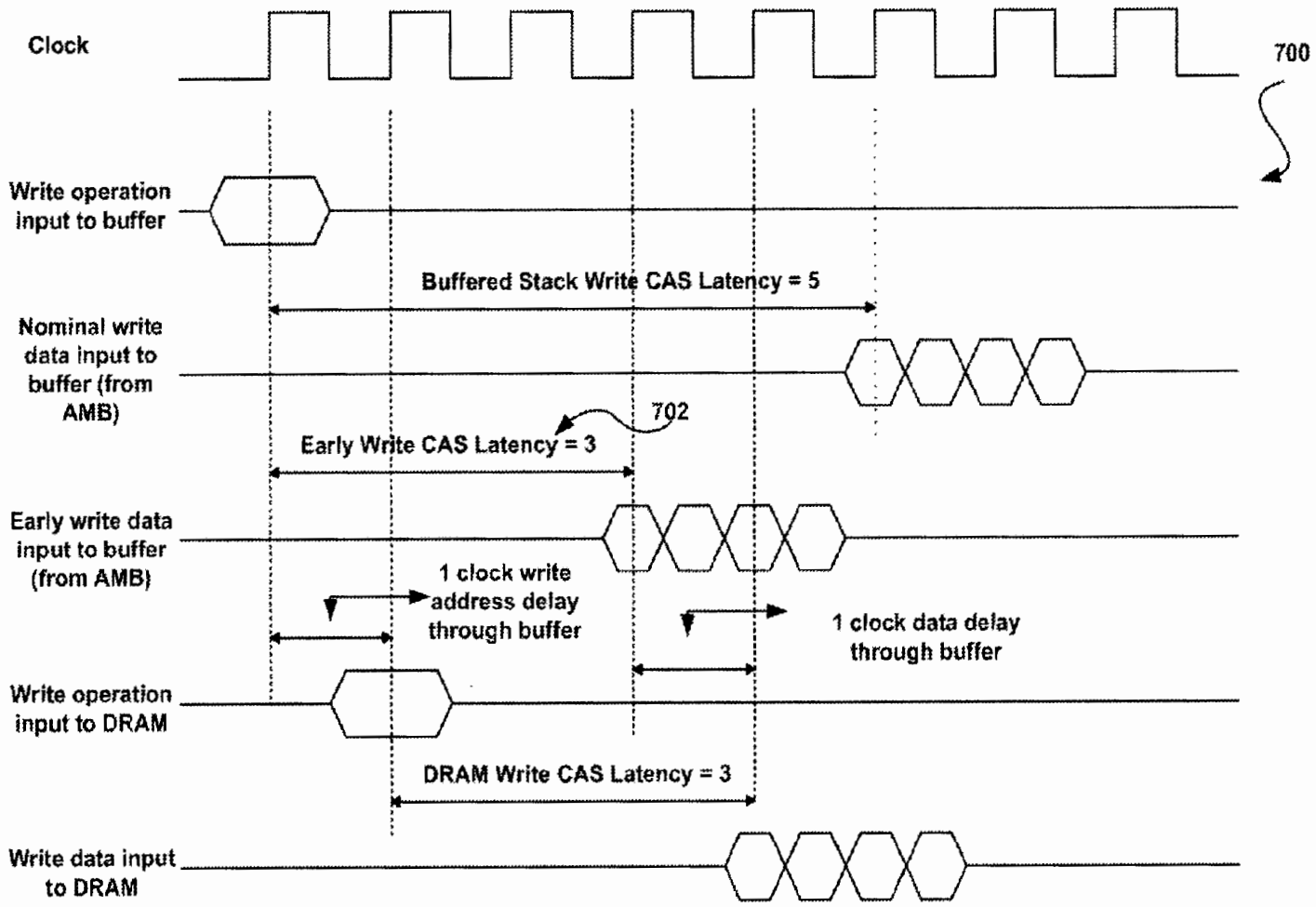

Clock
700
Write operation input to buffer
Buffered Stack Write CAS Latency = 5
Nominal write data input to buffer (from AMB)
702
Early Write CAS Latency = 3
Early write data input to buffer (from AMB)
1 clock write address delay through buffer
1 clock data delay through buffer
Write operation input to DRAM
DRAM Write CAS Latency = 3
Write data input to DRAM

FIGURE 7

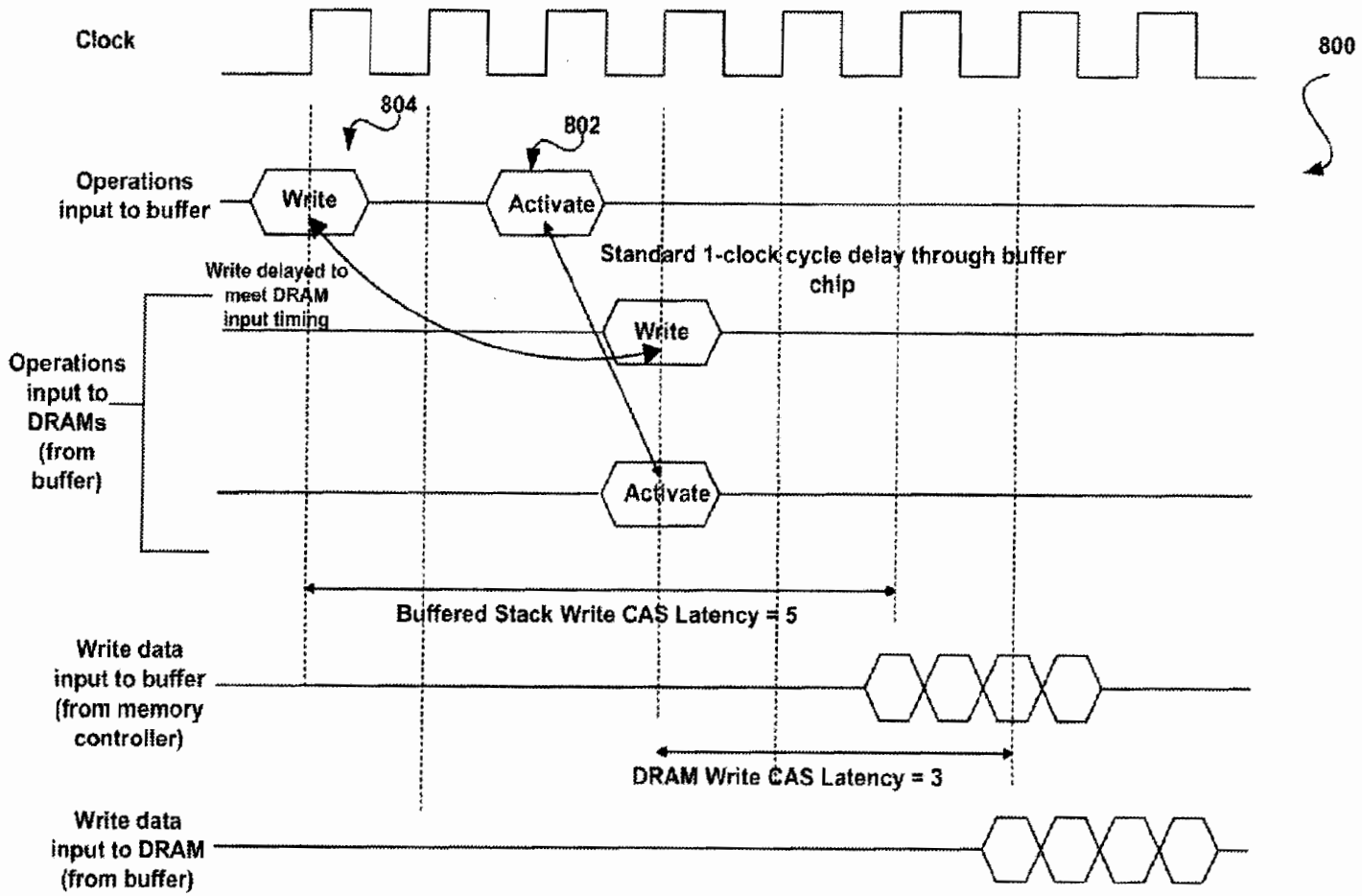

Clock
800
804
802
Operations input to buffer
Write
Activate
Standard 1-clock cycle delay through buffer chip
Write delayed to meet DRAM input timing
Write
Operations input to DRAMs (from buffer)
Activate
Buffered Stack Write CAS Latency = 5
Write data input to buffer (from memory controller)
DRAM Write CAS Latency = 3
Write data input to DRAM (from buffer)

FIGURE 8

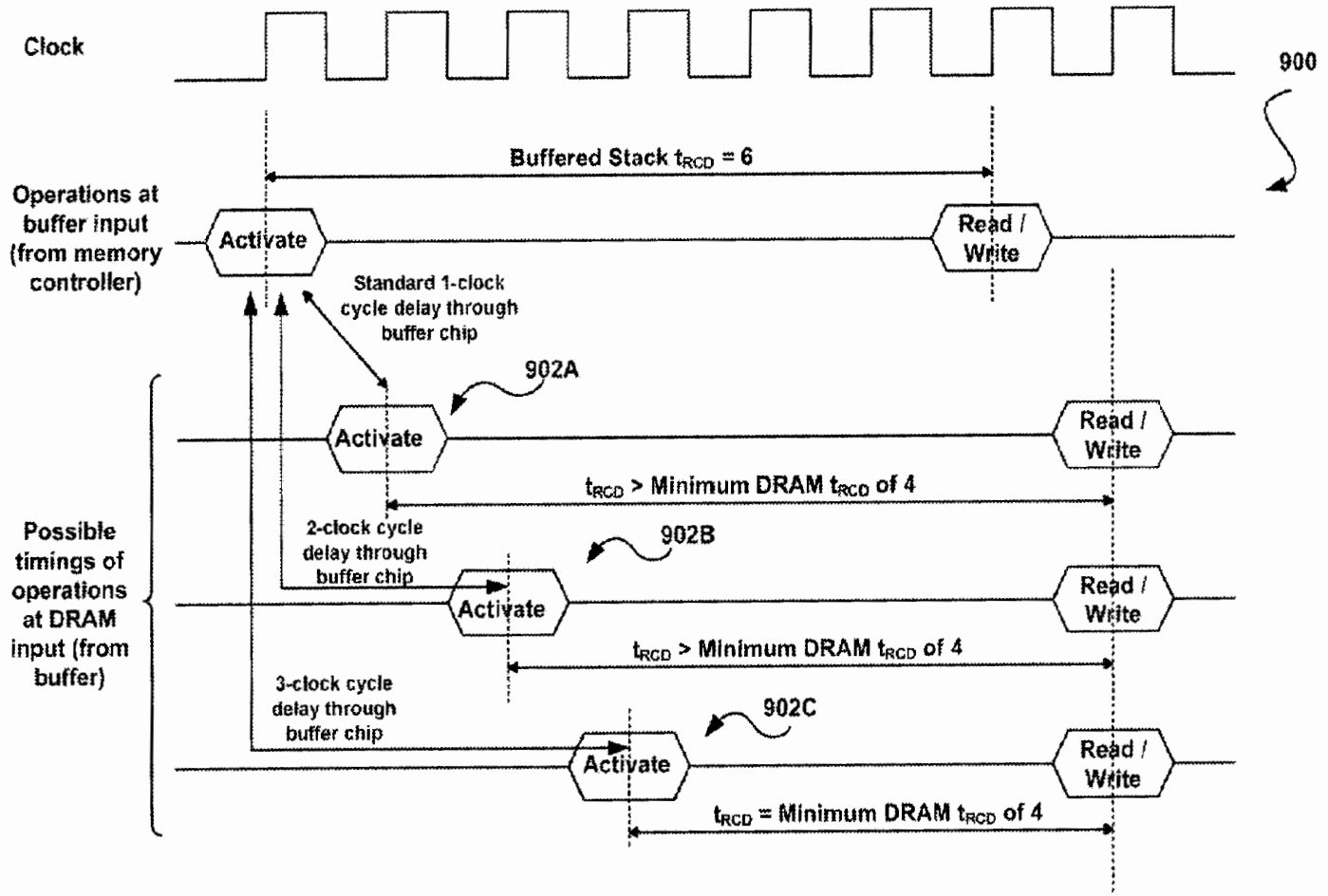

Clock
900
Buffered Stack t_{RCD} = 6
Operations at buffer input (from memory controller)
Activate
Read / Write
Standard 1-clock cycle delay through buffer chip
902A
Activate
Read / Write
t_{RCD} > Minimum DRAM t_{RCD} of 4
Possible timings of operations at DRAM input (from buffer)
2-clock cycle delay through buffer chip
902B
Activate
Read / Write
t_{RCD} > Minimum DRAM t_{RCD} of 4
3-clock cycle delay through buffer chip
902C
Activate
Read / Write
t_{RCD} = Minimum DRAM t_{RCD} of 4

FIGURE 9A

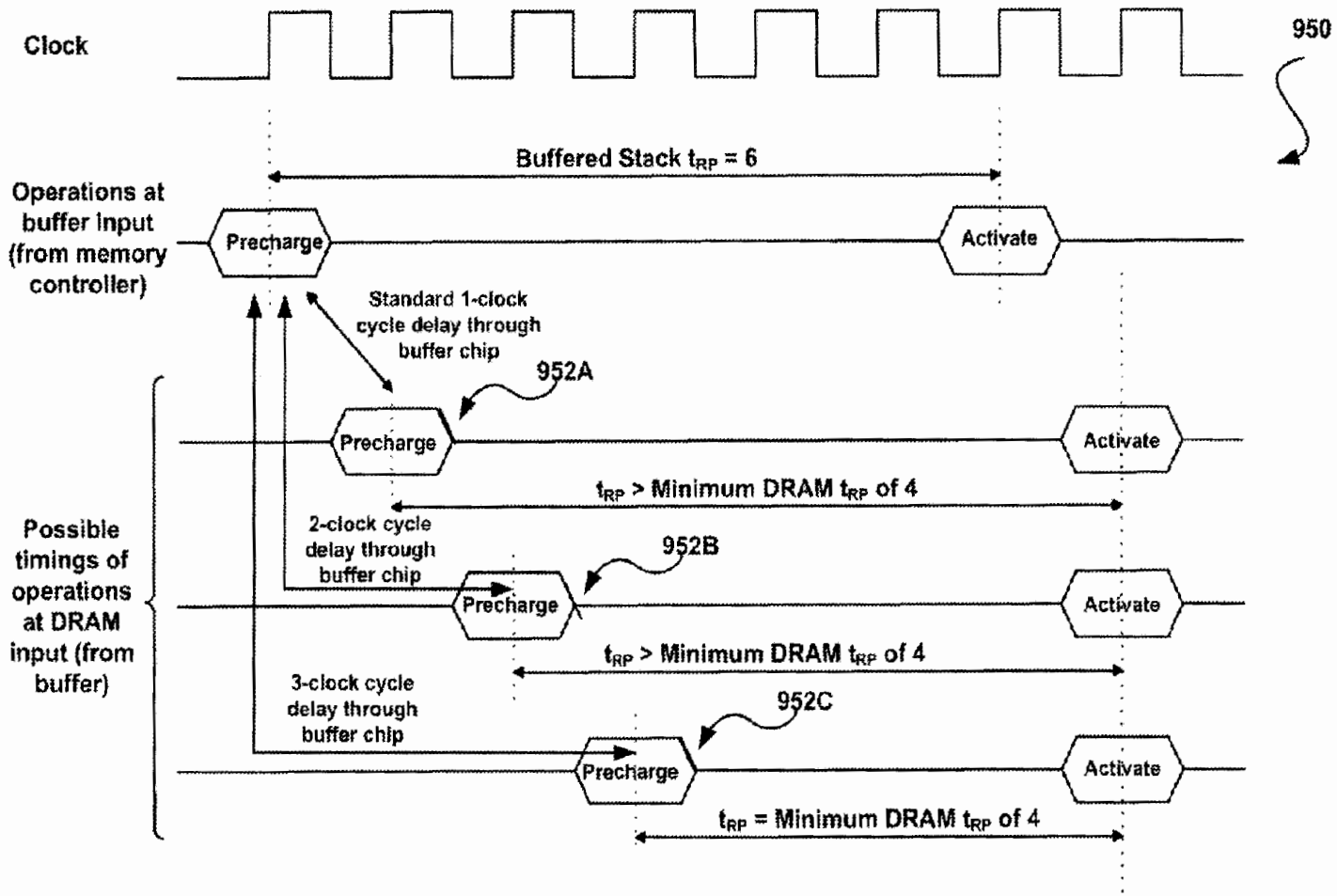

Clock
950
Buffered Stack t_{RP} = 6
Operations at buffer input (from memory controller)
Precharge
Activate
Standard 1-clock cycle delay through buffer chip
952A
Precharge
Activate
t_{RP} > Minimum DRAM t_{RP} of 4
Possible timings of operations at DRAM input (from buffer)
2-clock cycle delay through buffer chip
952B
Precharge
Activate
t_{RP} > Minimum DRAM t_{RP} of 4
3-clock cycle delay through buffer chip
952C
Precharge
Activate
t_{RP} = Minimum DRAM t_{RP} of 4

FIGURE 9B

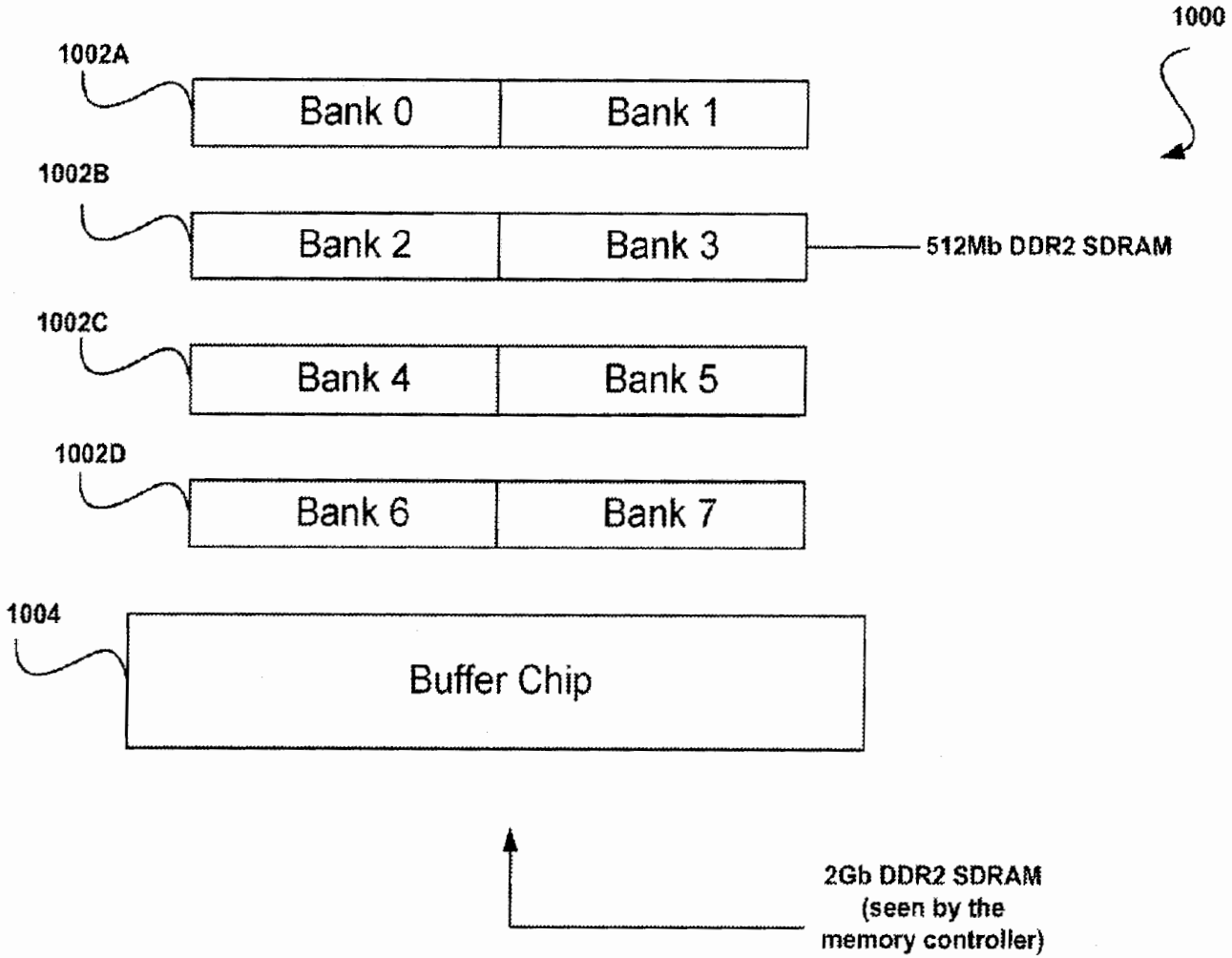

1000
1002A
Bank 0
Bank 1
1002B
Bank 2
Bank 3
512Mb DDR2 SDRAM
1002C
Bank 4
Bank 5
1002D
Bank 6
Bank 7
1004
Buffer Chip
2Gb DDR2 SDRAM
(seen by the
memory controller)

FIGURE 10

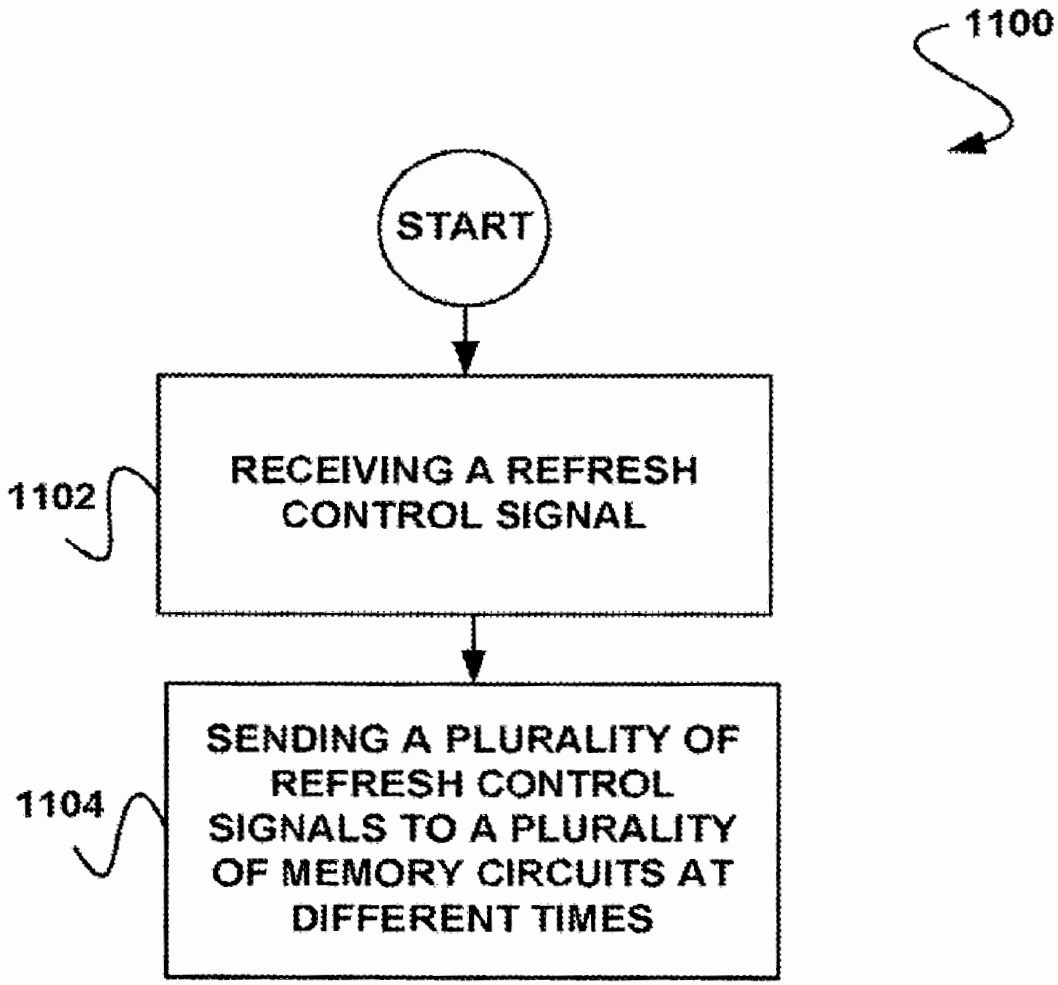

1100
START
1102
RECEIVING A REFRESH CONTROL SIGNAL
1104
SENDING A PLURALITY OF REFRESH CONTROL SIGNALS TO A PLURALITY OF MEMORY CIRCUITS AT DIFFERENT TIMES

FIGURE 11

SYSTEM AND METHOD FOR SIMULATING AN ASPECT OF A MEMORY CIRCUIT

FIELD OF THE INVENTION

[0001] The present invention relates to memory, and more particularly to multiple-memory circuit systems.

BACKGROUND

[0002] The memory capacity requirements of computers in general, and servers in particular, are increasing rapidly due to various trends such as 64-bit processors and operating systems, multi-core processors, virtualization, etc. However, other industry trends such as higher memory bus speeds and small form factor machines, etc. are reducing the number of memory module slots in such systems. Thus, a need exists in the industry for large capacity memory circuits to be used in such systems.

[0003] However, there is also an exponential relationship between a capacity of monolithic memory circuits and a price associated therewith. As a result, large capacity memory modules may be cost prohibitive. To this end, the use of multiple smaller capacity memory circuits is a cost-effective approach to increasing such memory capacity.

SUMMARY

[0004] A system and method are provided for simulating an aspect of a memory circuit. Included is an interface circuit that is in communication with a plurality of memory circuits and a system. Such interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of memory circuits. In accordance with various embodiments, such aspect may include a signal, a capacity, a timing, and/or a logical interface.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] FIG. **1** illustrates a multiple memory circuit framework, in accordance with one embodiment.

[0006] FIGS. **2A-2E** show various configurations of a buffered stack of dynamic random access memory (DRAM) circuits with a buffer chip, in accordance with various embodiments.

[0007] FIG. **2F** illustrates a method for storing at least a portion of information received in association with a first operation for use in performing a second operation, in accordance with still another embodiment.

[0008] FIG. **3** shows a high capacity dual in-line memory module (DIMM) using buffered stacks, in accordance with still yet another embodiment.

[0009] FIG. **4** shows a timing design of a buffer chip that makes a buffered stack of DRAM circuits mimic longer column address strobe (CAS) latency DRAM to a memory controller, in accordance with another embodiment.

[0010] FIG. **5** shows the write data timing expected by DRAM in a buffered stack, in accordance with yet another embodiment.

[0011] FIG. **6** shows write control signals delayed by a buffer chip, in accordance with still yet another embodiment.

[0012] FIG. **7** shows early write data from an advanced memory buffer (AMB), in accordance with another embodiment.

[0013] FIG. **8** shows address bus conflicts caused by delayed write operations, in accordance with yet another embodiment.

[0014] FIGS. **9A-B** show variable delays of operations through a buffer chip, in accordance with another embodiment.

[0015] FIG. **10** shows a buffered stack of four 512 Mb DRAM circuits mapped to a single 2 Gb DRAM circuit, in accordance with yet another embodiment.

[0016] FIG. **11** illustrates a method for refreshing a plurality of memory circuits, in accordance with still yet another embodiment.

DETAILED DESCRIPTION

[0017] FIG. **1** illustrates a multiple memory circuit framework **100**, in accordance with one embodiment. As shown, included are an interface circuit **102**, a plurality of memory circuits **104A**, **104B**, **104N**, and a system **106**. In the context of the present description, such memory circuits **104A**, **104B**, **104N** may include any circuit capable of serving as memory.

[0018] For example, in various embodiments, one or more of the memory circuits **104A**, **104B**, **104N** may include a monolithic memory circuit. For instance, such monolithic memory circuit may take the form of dynamic random access memory (DRAM). Such DRAM may take any form including, but not limited to synchronous (SDRAM), double data rate synchronous (DDR DRAM, DDR2 DRAM, DDR3 DRAM, etc.), quad data rate (QDR DRAM), direct RAMBUS (DRDRAM), fast page mode (FPM DRAM), video (VDRAM), extended data out (EDO DRAM), burst EDO (BEDO DRAM), multibank (MDRAM), synchronous graphics (SGRAM), and/or any other type of DRAM. Of course, one or more of the memory circuits **104A**, **104B**, **104N** may include other types of memory such as magnetic random access memory (MRAM), intelligent random access memory (IRAM), distributed network architecture (DNA) memory, window random access memory (WRAM), flash memory (e.g. NAND, NOR, or others, etc.), pseudostatic random access memory (PSRAM), wetware memory, and/or any other type of memory circuit that meets the above definition.

[0019] In additional embodiments, the memory circuits **104A**, **104B**, **104N** may be symmetrical or asymmetrical. For example, in one embodiment, the memory circuits **104A**, **104B**, **104N** may be of the same type, brand, and/or size, etc. Of course, in other embodiments, one or more of the memory circuits **104A**, **104B**, **104N** may be of a first type, brand, and/or size; while one or more other memory circuits **104A**, **104B**, **104N** may be of a second type, brand, and/or size, etc. Just by way of example, one or more memory circuits **104A**, **104B**, **104N** may be of a DRAM type, while one or more other memory circuits **104A**, **104B**, **104N** may be of a flash type. While three or more memory circuits **104A**, **104B**, **104N** are shown in FIG. **1** in accordance with one embodiment, it should be noted that any plurality of memory circuits **104A**, **104B**, **104N** may be employed.

[0020] Strictly as an option, the memory circuits **104A**, **104B**, **104N** may or may not be positioned on at least one dual in-line memory module (DIMM) (not shown). In various embodiments, the DIMM may include a registered DIMM (R-DIMM), a small outline-DIMM (SO-DIMM), a fully buffered-DIMM (FB-DIMM), an un-buffered DIMM,

etc. Of course, in other embodiments, the memory circuits **104**A, **104**B, **104**N may or may not be positioned on any desired entity for packaging purposes.

[0021] Further in the context of the present description, the system **106** may include any system capable of requesting and/or initiating a process that results in an access of the memory circuits **104**A, **104**B, **104**N. As an option, the system **106** may accomplish this utilizing a memory controller (not shown), or any other desired mechanism. In one embodiment, such system **106** may include a host system in the form of a desktop computer, lap-top computer, server, workstation, a personal digital assistant (PDA) device, a mobile phone device, a television, a peripheral device (e.g. printer, etc.). Of course, such examples are set forth for illustrative purposes only, as any system meeting the above definition may be employed in the context of the present framework **100**.

[0022] Turning now to the interface circuit **102**, such interface circuit **102** may include any circuit capable of indirectly or directly communicating with the memory circuits **104**A, **104**B, **104**N and the system **106**. In various optional embodiments, the interface circuit **102** may include one or more interface circuits, a buffer chip, etc. Embodiments involving such a buffer chip will be set forth hereinafter during reference to subsequent figures. In still other embodiments, the interface circuit **102** may or may not be manufactured in monolithic form.

[0023] While the memory circuits **104**A, **104**B, **104**N, interface circuit **102**, and system **106** are shown to be separate parts, it is contemplated that any of such parts (or portions thereof) may or may not be integrated in any desired manner. In various embodiments, such optional integration may involve simply packaging such parts together (e.g. stacking the parts, etc.) and/or integrating them monolithically. Just by way of example, in various optional embodiments, one or more portions (or all, for that matter) of the interface circuit **102** may or may not be packaged with one or more of the memory circuits **104**A, **104**B, **104**N (or all, for that matter). Different optional embodiments which may be implemented in accordance with the present multiple memory circuit framework **100** will be set forth hereinafter during reference to FIGS. **2**A-**2**E, and **3** et al.

[0024] In use, the interface circuit **102** may be capable of various functionality, in the context of different embodiments. More illustrative information will now be set forth regarding such optional functionality which may or may not be implemented in the context of such interface circuit **102**, per the desires of the user. It should be strongly noted that the following information is set forth for illustrative purposes and should not be construed as limiting in any manner. For example, any of the following features may be optionally incorporated with or without the exclusion of other features described.

[0025] For instance, in one optional embodiment, the interface circuit **102** interfaces a plurality of signals **108** that are communicated between the memory circuits **104**A, **104**B, **104**N and the system **106**. As shown, such signals may, for example, include address/control/clock signals, etc. In one aspect of the present embodiment, the interfaced signals **108** may represent all of the signals that are communicated between the memory circuits **104**A, **104**B, **104**N and the system **106**. In other aspects, at least a portion of signals **110** may travel directly between the memory circuits **104**A, **104**B, **104**N and the system **106** or component thereof [e.g. register, advanced memory buffer (AMB), memory controller, or any other component thereof, where the term component is defined hereinbelow]. In various embodiments, the number of the signals **108** (vs. a number of the signals **110**, etc.) may vary such that the signals **108** are a majority or more (L>M), etc.

[0026] In yet another embodiment, the interface circuit **102** may be operable to interface a first number of memory circuits **104**A, **104**B, **104**N and the system **106** for simulating at least one memory circuit of a second number. In the context of the present description, the simulation may refer to any simulating, emulating, disguising, transforming, converting, and/or the like that results in at least one aspect (e.g. a number in this embodiment, etc.) of the memory circuits **104**A, **104**B, **104**N appearing different to the system **106**. In different embodiments, the simulation may be electrical in nature, logical in nature, protocol in nature, and/or performed in any other desired manner. For instance, in the context of electrical simulation, a number of pins, wires, signals, etc. may be simulated, while, in the context of logical simulation, a particular function may be simulated. In the context of protocol, a particular protocol (e.g. DDR3, etc.) may be simulated.

[0027] In still additional aspects of the present embodiment, the second number may be more or less than the first number. Still yet, in the latter case, the second number may be one, such that a single memory circuit is simulated. Different optional embodiments which may employ various aspects of the present embodiment will be set forth hereinafter during reference to FIGS. **2**A-**2**E, and **3** et al.

[0028] In still yet another embodiment, the interface circuit **102** may be operable to interface the memory circuits **104**A, **104**B, **104**N and the system **106** for simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of the memory circuits **104**A, **104**B, **104**N. In accordance with various aspects of such embodiment, such aspect may include a signal, a capacity, a timing, a logical interface, etc. Of course, such examples of aspects are set forth for illustrative purposes only and thus should not be construed as limiting, since any aspect associated with one ore more of the memory circuits **104**A, **104**B, **104**N may be simulated differently in the foregoing manner.

[0029] In the case of the signal, such signal may refer to a control signal (e.g. an address signal; a signal associated with an activate operation, precharge operation, write operation, read operation, a mode register write operation, a mode register read operation, a refresh operation; etc.), a data signal, a logical or physical signal, or any other signal for that matter. For instance, a number of the aforementioned signals may be simulated to appear as fewer or more signals, or even simulated to correspond to a different type. In still other embodiments, multiple signals may be combined to simulated another signal. Even still, a length of time in which a signal is asserted may be simulated to be different.

[0030] In the case of protocol, such may, in one exemplary embodiment, refer to a particular standard protocol. For example, a number of memory circuits **104**A, **104**B, **104**N that obey a standard protocol (e.g. DDR2, etc.) may be used to simulate one or more memory circuits that obey a different protocol (e.g. DDR3, etc.). Also, a number of memory circuits **104**A, **104**B, **104**N that obey a version of protocol (e.g. DDR2 with 3-3-3 latency timing, etc.) may be

used to simulate one or more memory circuits that obey a different version of the same protocol (e.g. DDR2 with 5-5-5 latency timing, etc.).

[0031] In the case of capacity, such may refer to a memory capacity (which may or may not be a function of a number of the memory circuits **104**A, **104**B, **104**N; see previous embodiment). For example, the interface circuit **102** may be operable for simulating at least one memory circuit with a first memory capacity that is greater than (or less than) a second memory capacity of at least one of the memory circuits **104**A, **104**B, **104**N.

[0032] In the case where the aspect is timing-related, the timing may possibly relate to a latency (e.g. time delay, etc.). In one aspect of the present embodiment, such latency may include a column address strobe (CAS) latency, which refers to a latency associated with accessing a column of data. Still yet, the latency may include a row address to column address latency (tRCD), which refers to a latency required between the row address strobe (RAS) and CAS. Even still, the latency may include a row precharge latency (tRP), which refers a latency required to terminate access to an open row, and open access to a next row. Further, the latency may include an activate to precharge latency (tRAS), which refers to a latency required to access a certain row of data between an activate operation and a precharge operation. In any case, the interface circuit **102** may be operable for simulating at least one memory circuit with a first latency that is longer (or shorter) than a second latency of at least one of the memory circuits **104**A, **104**B, **104**N. Different optional embodiments which employ various features of the present embodiment will be set forth hereinafter during reference to FIGS. **2**A-**2**E, and **3** et al.

[0033] In still another embodiment, a component may be operable to receive a signal from the system **106** and communicate the signal to at least one of the memory circuits **104**A, **104**B, **104**N after a delay. Again, the signal may refer to a control signal (e.g. an address signal; a signal associated with an activate operation, precharge operation, write operation, read operation; etc.), a data signal, a logical or physical signal, or any other signal for that matter. In various embodiments, such delay may be fixed or variable (e.g. a function of the current signal, the previous signal, etc.). In still other embodiments, the component may be operable to receive a signal from at least one of the memory circuits **104**A, **104**B, **104**N and communicate the signal to the system **106** after a delay.

[0034] As an option, the delay may include a cumulative delay associated with any one or more of the aforementioned signals. Even still, the delay may result in a time shift of the signal forward and/or back in time (with respect to other signals). Of course, such forward and backward time shift may or may not be equal in magnitude. In one embodiment, this time shifting may be accomplished by utilizing a plurality of delay functions which each apply a different delay to a different signal. In still additional embodiments, the aforementioned shifting may be coordinated among multiple signals such that different signals are subject to shifts with different relative directions/magnitudes, in an organized fashion.

[0035] Further, it should be noted that the aforementioned component may, but need not necessarily take the form of the interface circuit **102** of FIG. **1**. For example, the component may include a register, an AMB, a component positioned on at least one DIMM, a memory controller, etc. Such register may, in various embodiments, include a Joint Electron Device Engineering Council (JEDEC) register, a JEDEC register including one or more functions set forth herein, a register with forwarding, storing, and/or buffering capabilities, etc. Different optional embodiments which employ various features of the present embodiment will be set forth hereinafter during reference to FIGS. **4**-**7**, and **9**A-B et al.

[0036] In a power-saving embodiment, at least one of a plurality of memory circuits **104**A, **104**B, **104**N may be identified that is not currently being accessed by the system **106**. In one embodiment, such identification may involve determining whether a page [i.e. any portion of any memory (s), etc.] is being accessed in at least one of the plurality of memory circuits **104**A, **104**B, **104**N. Of course, any other technique may be used that results in the identification of at least one of the memory circuits **104**A, **104**B, **104**N that is not being accessed.

[0037] In response to the identification of the at least one memory circuit **104**A, **104**B, **104**N, a power saving operation is initiated in association with the at least one memory circuit **104**A, **104**B, **104**N. In one optional embodiment, such power saving operation may involve a power down operation and, in particular, a precharge power down operation. Of course, however, it should be noted that any operation that results in at least some power savings may be employed in the context of the present embodiment.

[0038] Similar to one or more of the previous embodiments, the present functionality or a portion thereof may be carried out utilizing any desired component. For example, such component may, but need not necessarily take the form of the interface circuit **102** of FIG. **1**. In other embodiments, the component may include a register, an AMB, a component positioned on at least one DIMM, a memory controller, etc. One optional embodiment which employs various features of the present embodiment will be set forth hereinafter during reference to FIG. **10**.

[0039] In still yet another embodiment, a plurality of the aforementioned components may serve, in combination, to interface the memory circuits **104**A, **104**B, **104**N and the system **106**. In various embodiments, two, three, four, or more components may accomplish this. Also, the different components may be relatively configured in any desired manner. For example, the components may be configured in parallel, serially, or a combination thereof. In addition, any number of the components may be allocated to any number of the memory circuits **104**A, **104**B, **104**N.

[0040] Further, in the present embodiment, each of the plurality of components may be the same or different. Still yet, the components may share the same or similar interface tasks and/or perform different interface tasks. Such interface tasks may include, but are not limited to simulating one or more aspects of a memory circuit, performing a power savings/refresh operation, carrying out any one or more of the various functionalities set forth herein, and/or any other task relevant to the aforementioned interfacing. One optional embodiment which employs various features of the present embodiment will be set forth hereinafter during reference to FIG. **3**.

[0041] Additional illustrative information will now be set forth regarding various optional embodiments in which the foregoing techniques may or may not be implemented, per the desires of the user. For example, an embodiment is set forth for storing at least a portion of information received in

association with a first operation for use in performing a second operation. See FIG. **2**F. Further, a technique is provided for refreshing a plurality of memory circuits, in accordance with still yet another embodiment. See FIG. **11**.

[0042] It should again be strongly noted that the following information is set forth for illustrative purposes and should not be construed as limiting in any manner. Any of the following features may be optionally incorporated with or without the exclusion of other features described.

[0043] FIGS. 2A-2E show various configurations of a buffered stack of DRAM circuits **206**A-D with a buffer chip **202**, in accordance with various embodiments. As an option, the various configurations to be described in the following embodiments may be implemented in the context of the architecture and/or environment of FIG. **1**. Of course, however, they may also be carried out in any other desired environment (e.g. using other memory types, etc.). It should also be noted that the aforementioned definitions may apply during the present description.

[0044] As shown in each of such figures, the buffer chip **202** is placed electrically between an electronic host system **204** and a stack of DRAM circuits **206**A-D. In the context of the present description, a stack may refer to any collection of memory circuits. Further, the buffer chip **202** may include any device capable of buffering a stack of circuits (e.g. DRAM circuits **206**A-D, etc.). Specifically, the buffer chip **202** may be capable of buffering the stack of DRAM circuits **206**A-D to electrically and/or logically resemble at least one larger capacity DRAM circuit to the host system **204**. In this way, the stack of DRAM circuits **206**A-D may appear as a smaller quantity of larger capacity DRAM circuits to the host system **204**.

[0045] For example, the stack of DRAM circuits **206**A-D may include eight 512 Mb DRAM circuits. Thus, the buffer chip **202** may buffer the stack of eight 512 Mb DRAM circuits to resemble a single 4 Gb DRAM circuit to a memory controller (not shown) of the associated host system **204**. In another example, the buffer chip **202** may buffer the stack of eight 512 Mb DRAM circuits to resemble two 2 Gb DRAM circuits to a memory controller of an associated host system **204**.

[0046] Further, the stack of DRAM circuits **206**A-D may include any number of DRAM circuits. Just by way of example, a buffer chip **202** may be connected to 2, 4, 8 or more DRAM circuits **206**A-D. Also, the DRAM circuits **206**A-D may be arranged in a single stack, as shown in FIGS. **2**A-2D.

[0047] The DRAM circuits **206**A-D may be arranged on a single side of the buffer chip **202**, as shown in FIGS. **2**A-2D. Of course, however, the DRAM circuits **206**A-D may be located on both sides of the buffer chip **202** shown in FIG. 2E. Thus, for example, a buffer chip **202** may be connected to 16 DRAM circuits with 8 DRAM circuits on either side of the buffer chip **202**, where the 8 DRAM circuits on each side of the buffer chip **202** are arranged in two stacks of four DRAM circuits.

[0048] The buffer chip **202** may optionally be a part of the stack of DRAM circuits **206**A-D. Of course, however, the buffer chip **202** may also be separate from the stack of DRAM circuits **206**A-D. In addition, the buffer chip **202** may be physically located anywhere in the stack of DRAM circuits **206**A-D, where such buffer chip **202** electrically sits between the electronic host system **204** and the stack of DRAM circuits **206**A-D.

[0049] In one embodiment, a memory bus (not shown) may connect to the buffer chip **202**, and the buffer chip **202** may connect to each of the DRAM circuits **206**A-D in the stack. As shown in FIGS. 2A-D, the buffer chip **202** may be located at the bottom of the stack of DRAM circuits **206**A-D (e.g. the bottom-most device in the stack). As another option, and as shown in FIG. **2**E, the buffer chip **202** may be located in the middle of the stack of DRAM circuits **206**A-D. As still yet another option, the buffer chip **202** may be located at the top of the stack of DRAM circuits **206**A-D (e.g. the top-most device in the stack). Of course, however, the buffer chip **202** may be located anywhere between the two extremities of the stack of DRAM circuits **206**A-D.

[0050] The electrical connections between the buffer chip **202** and the stack of DRAM circuits **206**A-D may be configured in any desired manner. In one optional embodiment; address, control (e.g. command, etc.), and clock signals may be common to all DRAM circuits **206**A-D in the stack (e.g. using one common bus). As another option, there may be multiple address, control and clock busses. As yet another option, there may be individual address, control and clock busses to each DRAM circuit **206**A-D. Similarly, data signals may be wired as one common bus, several busses or as an individual bus to each DRAM circuit **206**A-D. Of course, it should be noted that any combinations of such configurations may also be utilized.

[0051] For example, as shown in FIG. **2**A, the stack of DRAM circuits **206**A-D may have one common address, control and clock bus **208** with individual data busses **210**. In another example, as shown in FIG. 2B, the stack of DRAM circuits **206**A-D may have two address, control and clock busses **208** along with two data busses **210**. In still yet another example, as shown in FIG. **2**C, the stack of DRAM circuits **206**A-D may have one address, control and clock bus **208** together with two data busses **210**. In addition, as shown in FIG. **2**D, the stack of DRAM circuits **206**A-D may have one common address, control and clock bus **208** and one common data bus **210**. It should be noted that any other permutations and combinations of such address, control, clock and data buses may be utilized.

[0052] These configurations may therefore allow for the host system **204** to only be in contact with a load of the buffer chip **202** on the memory bus. In this way, any electrical loading problems (e.g. bad signal integrity, improper signal timing, etc.) associated with the stacked DRAM circuits **206**A-D may (but not necessarily) be prevented, in the context of various optional embodiments.

[0053] FIG. **2**F illustrates a method **280** for storing at least a portion of information received in association with a first operation for use in performing a second operation, in accordance with still yet another embodiment. As an option, the method **280** may be implemented in the context of the architecture and/or environment of any one or more of FIGS. **1**-2E. For example, the method **280** may be carried out by the interface circuit **102** of FIG. **1**. Of course, however, the method **280** may be carried out in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0054] In operation **282**, first information is received in association with a first operation to be performed on at least one of a plurality of memory circuits (e.g. see the memory circuits **104**A, **104**B, **104**N of FIG. **1**, etc.). In various embodiments, such first information may or may not be received coincidently with the first operation, as long as it is

associated in some capacity. Further, the first operation may, in one embodiment, include a row operation. In such embodiment, the first information may include address information (e.g. a set of address bits, etc.).

[0055] For reasons that will soon become apparent, at least a portion of the first information is stored. Note operation **284**. Still yet, in operation **286**, second information is received in association with a second operation. Similar to the first information, the second information may or may not be received coincidently with the second operation, and may include address information. Such second operation, however, may, in one embodiment, include a column operation.

[0056] To this end, the second operation may be performed utilizing the stored portion of the first information in addition to the second information. See operation **288**. More illustrative information will now be set forth regarding various optional features with which the foregoing method **280** may or may not be implemented, per the desires of the user. Specifically, an example will be set for illustrating the manner in which the method **280** may be employed for accommodating a buffer chip that is simulating at least one aspect of a plurality of memory circuits.

[0057] In particular, the present example of the method **280** of FIG. 2F will be set forth in the context of the various components (e.g. buffer chip **202**, etc.) shown in the embodiments of FIGS. **2A-2E**. It should be noted that, since the buffered stack of DRAM circuits **206**A-D may appear to the memory controller of the host system **204** as one or more larger capacity DRAM circuits, the buffer chip **202** may receive more address bits from the memory controller than are required by the DRAM circuits **206**A-D in the stack. These extra address bits may be decoded by the buffer chip **202** to individually select the DRAM circuits **206**A-D in the stack, utilizing separate chip select signals to each of the DRAM circuits **206**A-D in the stack.

[0058] For example, a stack of four ×4 1 Gb DRAM circuits **206**A-D behind a buffer chip **202** may appear as a single ×4 4 Gb DRAM circuit to the memory controller. Thus, the memory controller may provide sixteen row address bits and three bank address bits during a row (e.g. activate) operation, and provide eleven column address bits and three bank address bits during a column (e.g. read or write) operation. However, the individual DRAM circuits **206**A-D in the stack may require only fourteen row address bits and three bank address bits for a row operation, and eleven column address bits and three bank address bits during a column operation.

[0059] As a result, during a row operation in the above example, the buffer chip **202** may receive two address bits more than are needed by each DRAM circuit **206**A-D in the stack. The buffer chip **202** may therefore use the two extra address bits from the memory controller to select one of the four DRAM circuits **206**A-D in the stack. In addition, the buffer chip **202** may receive the same number of address bits from the memory controller during a column operation as are needed by each DRAM circuit **206**A-D in the stack.

[0060] Thus, in order to select the correct DRAM circuit **206**A-D in the stack during a column operation, the buffer chip **202** may be designed to store the two extra address bits provided during a row operation and use the two stored address bits to select the correct DRAM circuit **206**A-D during the column operation. The mapping between a system address (e.g. address from the memory controller, including the chip select signal(s)) and a device address (e.g. the address, including the chip select signals, presented to the DRAM circuits **206**A-D in the stack) may be performed by the buffer chip **202** in various manners.

[0061] In one embodiment, a lower order system row address and bank address bits may be mapped directly to the device row address and bank address inputs. In addition, the most significant row address bit(s) and, optionally, the most significant bank address bit(s), may be decoded to generate the chip select signals for the DRAM circuits **206**A-D in the stack during a row operation. The address buts used to generate the chip select signals during the row operation may also be stored in an internal lookup table by the buffer chip **22** for one or more clock cycles. During a column operation, the system column address and bank address bits may be mapped directly to the device column address and bank address inputs, while the stored address bits may be decoded to generate the chip select signals.

[0062] For example, addresses may be mapped between four 512 Mb DRAM circuits **206**A-D that simulate a single 2 Gb DRAM circuits utilizing the buffer chip **202**. There may be 15 row address bits from the system **204**, such that row address bits **0** through **13** are mapped directly to the DRAM circuits **206**A-D. There may also be 3 bank address bits from the system **204**, such that bank address bits **0** through **1** are mapped directly to the DRAM circuits **206**A-D.

[0063] During a row operation, the bank address bit **2** and the row address bit **14** may be decoded to generate the 4 chip select signals for each of the four DRAM circuits **206**A-D. Row address bit **14** may be stored during the row operation using the bank address as the index. In addition, during the column operation, the stored row address bit **14** may again used with bank address bit **2** to form the four DRAM chip select signals.

[0064] As another example, addresses may be mapped between four 1 Gb DRAM circuits **206**A-D that simulate a single 4 Gb DRAM circuits utilizing the buffer chip **202**. There may be 16 row address bits from the system **204**, such that row address bits **0** through **14** are mapped directly to the DRAM circuits **206**A-D. There may also be 3 bank address bits from the system **204**, such that bank address bits **0** through **3** are mapped directly to the DRAM circuits **206**A-D.

[0065] During a row operation, row address bits **14** and **15** may be decoded to generate the 4 chip select signals for each of the four DRAM circuits **206**A-D. Row address bits **14** and **15** may also be stored during the row operation using the bank address as the index. During the column operation, the stored row address bits **14** and **15** may again be used to form the four DRAM chip select signals.

[0066] In various embodiments, this mapping technique may optionally be used to ensure that there are no unnecessary combinational logic circuits in the critical timing path between the address input pins and address output pins of the buffer chip **202**. Such combinational logic circuits may instead be used to generate the individual chip select signals. This may therefore allow the capacitive loading on the address outputs of the buffer chip **202** to be much higher than the loading on the individual chip select signal outputs of the buffer chip **202**.

[0067] In another embodiment, the address mapping may be performed by the buffer chip **202** using some of the bank address signals from the memory controller to generate the individual chip select signals. The buffer chip **202** may store

the higher order row address bits during a row operation using the bank address as the index, and then may use the stored address bits as part of the DRAM circuit bank address during a column operation. This address mapping technique may require an optional lookup table to be positioned in the critical timing path between the address inputs from the memory controller and the address outputs, to the DRAM circuits **206**A-D in the stack.

[0068] For example, addresses may be mapped between four 512 Mb DRAM circuits **206**A-D that simulate a single 2 Gb DRAM utilizing the buffer chip **202**. There may be 15 row address bits from the system **204**, where row address bits **0** through **13** are mapped directly to the DRAM circuits **206**A-D. There may also be 3 bank address bits from the system **204**, such that bank address bit **0** is used as a DRAM circuit bank address bit for the DRAM circuits **206**A-D.

[0069] In addition, row address bit **14** may be used as an additional DRAM circuit bank address bit. During a row operation, the bank address bits **1** and **2** from the system may be decoded to generate the 4 chip select signals for each of the four DRAM circuits **206**A-D. Further, row address bit **14** may be stored during the row operation. During the column operation, the stored row address bit **14** may again be used along with the bank address bit **0** from the system to form the DRAM circuit bank address.

[0070] In both of the above described address mapping techniques, the column address from the memory controller may be mapped directly as the column address to the DRAM circuits **206**A-D in the stack. Specifically, this direct mapping may be performed since each of the DRAM circuits **206**A-D in the stack, even if of the same width but different capacities (e.g. from 512 Mb to 4 Gb), may have the same page sizes. In an optional embodiment, address A[**10**] may be used by the memory controller to enable or disable auto-precharge during a column operation. Therefore, the buffer chip **202** may forward A[**10**] from the memory controller to the DRAM circuits **206**A-D in the stack without any modifications during a column operation.

[0071] In various embodiments, it may be desirable to determine whether the simulated DRAM circuit behaves according to a desired DRAM standard or other design specification. A behavior of many DRAM circuits is specified by the JEDEC standards and it may be desirable, in some embodiments, to exactly simulate a particular JEDEC standard DRAM. The JEDEC standard defines control signals that a DRAM circuit must accept and the behavior of the DRAM circuit as a result of such control signals. For example, the JEDEC specification for a DDR2 DRAM is known as JESD79-2B.

[0072] If it is desired, for example, to determine whether a JEDEC standard is met, the following algorithm may be used. Such algorithm checks, using a set of software verification tools for formal verification of logic, that protocol behavior of the simulated DRAM circuit is the same as a desired standard or other design specification. This formal verification is quite feasible because the DRAM protocol described in a DRAM standard is typically limited to a few control signals (e.g. approximately 15 control signals in the case of the JEDEC DDR2 specification, for example).

[0073] Examples of the aforementioned software verification tools include MAGELLAN supplied by SYNOPSYS, or other software verification tools, such as INCISIVE supplied by CADENCE, verification tools supplied by JASPER, VERIX supplied by REAL INTENT, 0-IN supplied by MENTOR CORPORATION, and others. These software verification tools use written assertions that correspond to the rules established by the DRAM protocol and specification. These written assertions are further included in the code that forms the logic description for the buffer chip. By writing assertions that correspond to the desired behavior of the simulated DRAM circuit, a proof may be constructed that determines whether the desired design requirements are met. In this way, one may test various embodiments for compliance with a standard, multiple standards, or other design specification.

[0074] For instance, an assertion may be written that no two DRAM control signals are allowed to be issued to an address, control and clock bus at the same time. Although one may know which of the various buffer chip/DRAM stack configurations and address mappings that have been described herein are suitable, the aforementioned algorithm may allow a designer to prove that the simulated DRAM circuit exactly meets the required standard or other design specification. If, for example, an address mapping that uses a common bus for data and a common bus for address results in a control and clock bus that does not meet a required specification, alternative designs for buffer chips with other bus arrangements or alternative designs for the interconnect between the buffer chips may be used and tested for compliance with the desired standard or other design specification.

[0075] FIG. **3** shows a high capacity DIMM **300** using buffered stacks of DRAM circuits **302**, in accordance with still yet another embodiment. As an option, the high capacity DIMM **300** may be implemented in the context of the architecture and environment of FIGS. **1** and/or **2**A-F. Of course, however, the high capacity DIMM **300** may be used in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0076] As shown, a high capacity DIMM **300** may be created utilizing buffered stacks of DRAM circuits **302**. Thus, a DIMM **300** may utilize a plurality of buffered stacks of DRAM circuits **302** instead of individual DRAM circuits, thus increasing the capacity of the DIMM. In addition, the DIMM **300** may include a register **304** for address and operation control of each of the buffered stacks of DRAM circuits **302**. It should be noted that any desired number of buffered stacks of DRAM circuits **302** may be utilized in conjunction with the DIMM **300**. Therefore, the configuration of the DIMM **300**, as shown, should not be construed as limiting in any way.

[0077] In an additional unillustrated embodiment, the register **304** may be substituted with an AMB (not shown), in the context of an FB-DIMM.

[0078] FIG. **4** shows a timing design **400** of a buffer chip that makes a buffered stack of DRAM circuits mimic longer CAS latency DRAM to a memory controller, in accordance with another embodiment. As an option, the design of the buffer chip may be implemented in the context of the architecture and environment of FIGS. **1-3**. Of course, however, the design of the buffer chip may be used in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0079] In use, any delay through a buffer chip (e.g. see the buffer chip **202** of FIGS. **2**A-F, etc.) may be made transparent to a memory controller of a host system (e.g. see the

host system **204** of FIGS. **2**A-E, etc.) utilizing the buffer chip. In particular, the buffer chip may buffer a stack of DRAM circuits such that the buffered stack of DRAM circuits appears as at least one larger capacity DRAM circuit with higher CAS latency.

[0080] Such delay may be a result of the buffer chip being located electrically between the memory bus of the host system and the stacked DRAM circuits, since most or all of the signals that connect the memory bus to the DRAM circuits pass through the buffer chip. A finite amount of time may therefore be needed for these signals to traverse through the buffer chip. With the exception of register chips and advanced memory buffers (AMB), industry standard protocols for memory [e.g. (DDR SDRAM), DDR2 SDRAM, etc.] may not comprehend the buffer chip that sits between the memory bus and the DRAM. Industry standard protocols for memory [e.g. (DDR SDRAM), DDR2 SDRAM, etc.] narrowly define the properties of chips that sit between host and memory circuits. Such industry standard protocols define the properties of a register chip and AMB but not the properties of the buffer chip **202**, etc. Thus, the signal delay through the buffer chip may violate the specifications of industry standard protocols.

[0081] In one embodiment, the buffer chip may provide a one-half clock cycle delay between the buffer chip receiving address and control signals from the memory controller (or optionally from a register chip, an AMB, etc.) and the address and control signals being valid at the inputs of the stacked DRAM circuits. Similarly, the data signals may also have a one-half clock cycle delay in traversing the buffer chip, either from the memory controller to the DRAM circuits or from the DRAM circuits to the memory controller. Of course, the one-half clock cycle delay set forth above is set forth for illustrative purposes only and thus should not be construed as limiting in any manner whatsoever. For example, other embodiments are contemplated where a one clock cycle delay, a multiple clock cycle delay (or fraction thereof), and/or any other delay amount is incorporated, for that matter. As mentioned earlier, in other embodiments, the aforementioned delay may be coordinated among multiple signals such that different signals are subject to time-shifting with different relative directions/magnitudes, in an organized fashion.

[0082] As shown in FIG. **4**, the cumulative delay through the buffer chip (e.g. the sum of a first delay **402** of the address and control signals through the buffer chip and a second delay **404** of the data signals through the buffer chip) is j clock cycles. Thus, the buffer chip may make the buffered stack appear to the memory controller as one or more larger DRAM circuits with a CAS latency **408** of i+j clocks, where i is the native CAS latency of the DRAM circuits.

[0083] In one example, if the DRAM circuits in the stack have a native CAS latency of 4 and the address and control signals along with the data signals experience a one-half clock cycle delay though the buffer chip, then the buffer chip may make the buffered stack appear to the memory controller as one or more larger DRAM circuits with a CAS latency of 5 (i.e. 4+1). In another example, if the address and control signals along with the data signals experience a 1 clock cycle delay through the buffer chip, then the buffer chip may make the buffered stack appear as one or more larger DRAM circuits with a CAS latency of 6 (i.e. 4+2).

[0084] FIG. **5** shows the write data timing **500** expected by a DRAM circuit in a buffered stack, in accordance with yet another embodiment. As an option, the write data timing **500** may be implemented in the context of the architecture and environment of FIGS. **1-4**. Of course, however, the write data timing **500** may be carried out in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0085] Designing a buffer chip (e.g. see the buffer chip **202** of FIGS. **2**A-E, etc.) so that a buffered stack appears as at least one larger capacity DRAM circuit with higher CAS latency may, in some embodiments, create a problem with the timing of write operations. For example, with respect to a buffered stack of DDR2 SDRAM circuits with a CAS latency of 4 that appear as a single larger DDR2 SDRAM with a CAS latency of 6 to the memory controller, the DDR2 SDRAM protocol may specify that the write CAS latency is one less than the read CAS latency. Therefore, since the buffered stack appears as a DDR2 SDRAM with a read CAS latency of 6, the memory controller may use a write CAS latency of 5 (see **502**) when scheduling a write operation to the buffered stack.

[0086] However, since the native read CAS latency of the DRAM circuits is 4, the DRAM circuits may require a write CAS latency of 3 (see **504**). As a result, the write data from the memory controller may arrive at the buffer chip later than when the DRAM circuits require the data. Thus, the buffer chip may delay such write operations to alleviate any of such timing problems. Such delay in write operations will be described in more detail with respect to FIG. **6** below.

[0087] FIG. **6** shows write operations **600** delayed by a buffer chip, in accordance with still yet another embodiment. As an option, the write operations **600** may be implemented in the context of the architecture and environment of FIGS. **1-5**. Of course, however, the write operations **600** may be used in any desired environment. Again, it should also be noted that the aforementioned definitions may apply during the present description.

[0088] In order to be compliant with the protocol utilized by the DRAM circuits in the stack, a buffer chip (e.g. see the buffer chip **202** of FIGS. **2**A-E, etc.) may provide an additional delay, over and beyond the delay of the address and control signals through the buffer chip, between receiving the write operation and address from the memory controller (and/or optionally from a register and/or AMB, etc.), and sending it to the DRAM circuits in the stack. The additional delay may be equal to j clocks, where j is the cumulative delay of the address and control signals through the buffer chip and the delay of the data signals through the buffer chip. As another option, the write address and operation may be delayed by a register chip on a DIMM, by a AMB, or by the memory controller.

[0089] FIG. **7** shows early write data **700** from an AMB, in accordance with another embodiment. As an option, the early write data **700** may be implemented in the context of the architecture and environment of FIGS. **1-5**. Of course, however, the early write data **700** may be used in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0090] As shown, an AMB on an FB-DIMM may be designed to send write data earlier to buffered stacks instead of delaying the write address and operation, as described in reference to FIG. **6**. Specifically, an early write latency **702** may be utilized to send the write data to the buffered stack.

Thus, correct timing of the write operation at the inputs of the DRAM circuits in the stack may be ensured.

[0091] For example, a buffer chip (e.g. see the buffer chip **202** of FIGS. **2**A-E, etc.) may have a cumulative latency of 2, in which case, the AMB may send the write data 2 clock cycles earlier to the buffered stack. It should be noted that this scheme may not be possible in the case of registered DIMMs since the memory controller sends the write data directly to the buffered stacks. As an option, a memory controller may be designed to send write data earlier so that write operations have the correct timing at the input of the DRAM circuits in the stack without requiring the buffer chip to delay the write address and operation.

[0092] FIG. **8** shows address bus conflicts **800** caused by delayed write operations, in accordance with yet another embodiment. As mentioned earlier, the delaying of the write addresses and operations may be performed by a buffer chip, or optionally a register, AMB, etc., in a manner that is completely transparent to the memory controller of a host system. However, since the memory controller is unaware of this delay, it may schedule subsequent operations, such as for example activate or precharge operations, which may collide with the delayed writes on the address bus from the buffer chip to the DRAM circuits in the stack. As shown, an activate operation **802** may interfere with a write operation **804** that has been delayed. Thus, a delay of activate operations may be employed, as will be described in further detail with respect to FIG. **9**.

[0093] FIGS. **9**A-B show variable delays **900** and **950** of operations through a buffer chip, in accordance with another embodiment. As an option, the variable delays **900** and **950** may be implemented in the context of the architecture and environment of FIGS. **1-8**. Of course, however, the variable delays **900** and **950** may be carried out in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0094] In order to prevent conflicts on an address bus between the buffer chip and its associated stack(s), either the write operation or the precharge/activate operation may be delayed. As shown, a buffer chip (e.g. see the buffer chip **202** of FIGS. **2**A-E, etc.) may delay the precharge/activate operations **952**A-C/**902**A-C. In particular, the buffer chip may make the buffered stack appear as one or more larger capacity DRAM circuits that have longer tRCD (RAS to CAS delay) and tRP (i.e. precharge time) parameters.

[0095] For example, if the cumulative latency through a buffer chip is 2 clock cycles while the native read CAS latency of the DRAM circuits is 4 clock cycles, then in order to hide the delay of the address/control signals and the data signals through the buffer chip, the buffered stack may appear as one or more larger capacity DRAM circuits with a read CAS latency of 6 clock cycles to the memory controller. In addition, if the tRCD and tRP of the DRAM circuits is 4 clock cycles each, the buffered stack may appear as one or more larger capacity DRAM circuits with tRCD of 6 clock cycles and tRP of 6 clock cycles in order to allow a buffer chip (e.g., see the buffer chip **202** of FIGS. **2**A-E, etc.) to delay the activate and precharge operations in a manner that is transparent to the memory controller. Specifically, a buffered stack that uses **4-4-4** DRAM circuits (i.e. CAS latency=4, tRCD=4, tRP=4) may appear as one or at least one larger capacity DRAM circuits with 6-6-6 timing (i.e. CAS latency=6, tRCD=6, tRP=6).

[0096] Since the buffered stack appears to the memory controller as having a tRCD of 6 clock cycles, the memory controller may schedule a column operation to a bank 6 clock cycles after an activate (e.g. row) operation to the same bank. However, the DRAM circuits in the stack may actually have a tRCD of 4 clock cycles. Thus, the buffer chip may have the ability to delay the activate operation by up to 2 clock cycles in order to avoid any conflicts on the address bus between the buffer chip and the DRAM circuits in the stack while still ensuring correct read and write timing on the channel between the memory controller and the buffered stack.

[0097] As shown, the buffer chip may issue the activate operation to the DRAM circuits one, two, or three clock cycles after it receives the activate operation from the memory controller, register, or AMB. The actual delay of the activate operation through the buffer chip may depend on the presence or absence of other DRAM operations that may conflict with the activate operation, and may optionally change from one activate operation to another.

[0098] Similarly, since the buffered stack may appear to the memory controller as at least one larger capacity DRAM circuit with a tRP of 6 clock cycles, the memory controller may schedule a subsequent activate (e.g. row) operation to a bank a minimum of 6 clock cycles after issuing a precharge operation to that bank. However, since the DRAM circuits in the stack actually have a tRP of 4 clock cycles, the buffer chip may have the ability to delay issuing the precharge operation to the DRAM circuits in the stack by up to 2 clock cycles in order to avoid any conflicts on the address bus between the buffer chip and the DRAM circuits in the stack. In addition, even if thee are no conflicts on the address bus, the buffer chip may still delay issuing a precharge operation in order to satisfy the tRAS requirement of the DRAM circuits.

[0099] In particular, if the activate operation to a bank was delayed to avoid an address bus conflict, then the precharge operation to the same bank may be delayed by the buffer chip to satisfy the tRAS requirement of the DRAM circuits. The buffer chip may issue the precharge operation to the DRAM circuits one, two, or three clock cycles after it receives the precharge operation from the memory controller, register, or AMB. The actual delay of the precharge operation through the buffer chip may depend on the presence or absence of address bus conflicts or tRAS violations, and may change from one precharge operation to another.

[0100] FIG. **10** shows a buffered stack **1000** of four 512 Mb DRAM circuits mapped to a single 2 Gb DRAM circuit, in accordance with yet another embodiment. As an option, the buffered stack **1000** may be implemented in the context of the architecture and environment of FIGS. **1-9**. Of course, however, the buffered stack **1000** may be carried out in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0101] The multiple DRAM circuits **1002**A-D buffered in the stack by the buffer chip **1004** may appear as at least one larger capacity DRAM circuit to the memory controller. However, the combined power dissipation of such DRAM circuits **1002**A-D may be much higher than the power dissipation of a monolithic DRAM of the same capacity. For example, the buffered stack may consist of four 512 Mb DDR2 SDRAM circuits that appear to the memory controller as a single 2 Gb DDR2 SDRAM circuit.

[0102] The power dissipation of all four DRAM circuits **1002**A-D in the stack may be much higher than the power dissipation of a monolithic 2 Gb DDR2 SDRAM. As a result, a DIMM containing multiple buffered stacks may dissipate much more power than a standard DIMM built using monolithic DRAM circuits. This increased power dissipation may limit the widespread adoption of DIMMs that use buffered stacks.

[0103] Thus, a power management technique that reduces the power dissipation of DIMMs that contain buffered stacks of DRAM circuits may be utilized. Specifically, the DRAM circuits **1002**A-D may be opportunistically placed in a precharge power down mode using the clock enable (CKE) pin of the DRAM circuits **1002**A-D. For example, a single rank registered DIMM (R-DIMM) may contain a plurality of buffered stacks of DRAM circuits **1002**A-D, where each stack consists of four ×4 512 Mb DDR2 SDRAM circuits **1002**A-D and appears as a single ×4 2 Gb DDR2 SDRAM circuit to the memory controller. A 2 Gb DDR2 SDRAM may generally have eight banks as specified by JEDEC. Therefore, the buffer chip **1004** may map each 512 Mb DRAM circuit in the stack to two banks of the equivalent 2 Gb DRAM, as shown.

[0104] The memory controller of the host system may open and close pages in the banks of the DRAM circuits **1002**A-D based on the memory requests it receives from the rest of the system. In various embodiments, no more than one page may be able to be open in a bank at any given time. For example, with respect to FIG. **10**, since each DRAM circuit **1002**A-D in the stack is mapped to two banks of the equivalent larger DRAM, at any given time a DRAM circuit **1002**A-D may have two open pages, one open page, or no open pages. When a DRAM circuit **1002**A-D has no open pages, the power management scheme may place that DRAM circuit **1002**A-D in the precharge power down mode by de-asserting its CKE input.

[0105] The CKE inputs of the DRAM circuits **1002**A-D in a stack may be controlled by the buffer chip **1004**, by a chip on an R-DIMM, by an AMB on a FB-DIMM, or by the memory controller in order to implement the power management scheme described hereinabove. In one embodiment, this power management scheme may be particularly efficient when the memory controller implements a closed page policy.

[0106] Another optional power management scheme may include mapping a plurality of DRAM circuits to a single bank of the larger capacity DRAM seen by the memory controller. For example, a buffered stack of sixteen ×4 256 Mb DDR2 SDRAM circuits may appear to the memory controller as a single ×4 4 Gb DDR2 SDRAM circuit. Since a 4 Gb DDR2 SDRAM circuit is specified by JEDEC to have eight banks, each bank of the 4 Gb DDR2 SDRAM circuit may be 512 Mb. Thus, two of the 256 Mb DDR2 SDRAM circuits may be mapped by the buffer chip **1004** to a single bank of the equivalent 4 Gb DDR2 SDRAM circuit seen by the memory controller.

[0107] In this way, bank **0** of the 4 Gb DDR2 SDRAM circuit may be mapped by the buffer chip to two 256 Mb DDR2 SDRAM circuits (e.g. DRAM A and DRAM B) in the stack. However, since only one page can be open in a bank at any given time, only one of DRAM A or DRAM B may be in the active state at any given time. If the memory controller opens a page in DRAM A, then DRAM B may be placed in the precharge power down mode by de-asserting its CKE input. As another option, if the memory controller opens a page in DRAM B, DRAM A may be placed in the precharge power down mode by de-asserting its CKE input. This technique may ensure that if p DRAM circuits are mapped to a bank of the larger capacity DRAM circuit seen by the memory controller, then p-1 of the p DRAM circuits may continuously (e.g. always, etc.) be subjected to a power saving operation. The power saving operation may, for example, comprise operating in precharge power down mode except when refresh is required. Of course, power-savings may also occur in other embodiments without such continuity.

[0108] FIG. **11** illustrates a method **1100** for refreshing a plurality of memory circuits, in accordance with still yet another embodiment. As an option, the method **1100** may be implemented in the context of the architecture and environment of any one or more of FIGS. **1-10**. For example, the method **1100** may be carried out by the interface circuit **102** of FIG. **1**. Of course, however, the method **1100** may be carried out in any desired environment. It should also be noted that the aforementioned definitions may apply during the present description.

[0109] As shown, a refresh control signal is received in operation **1102**. In one optional embodiment, such refresh control signal may, for example, be received from a memory controller, where such memory controller intends to refresh a simulated memory circuit(s).

[0110] In response to the receipt of such refresh control signal, a plurality of refresh control signals are sent to a plurality of the memory circuits (e.g. see the memory circuits **104**A, **104**B, **104**N of FIG. **1**, etc.), at different times. See operation **1104**. Such refresh control signals may or may not each include the refresh control signal of operation **1102** or an instantiation/copy thereof. Of course, in other embodiments, the refresh control signals may each include refresh control signals that are different in at least one aspect (e.g. format, content, etc.).

[0111] During use of still additional embodiments, at least one first refresh control signal may be sent to a first subset (e.g. of one or more) of the memory circuits at a first time and at least one second refresh control signal may be sent to a second subset (e.g. of one or more) of the memory circuits at a second time. Thus, in some embodiments, a single refresh control signal may be sent to a plurality of the memory circuits (e.g. a group of memory circuits, etc.). Further, a plurality of the refresh control signals may be sent to a plurality of the memory circuits. To this end, refresh control signals may be sent individually or to groups of memory circuits, as desired.

[0112] Thus, in still yet additional embodiments, the refresh control signals may be sent after a delay in accordance with a particular timing. In one embodiment, for example, the timing in which the refresh control signals are sent to the memory circuits may be selected to minimize a current draw. This may be accomplished in various embodiments by staggering a plurality of refresh control signals. In still other embodiments, the timing in which the refresh control signals are sent to the memory circuits may be selected to comply with a tRFC parameter associated with each of the memory circuits.

[0113] To this end, in the context of an example involving a plurality of DRAM circuits (e.g. see the embodiments of FIGS. **1-2**F, etc.), DRAM circuits of any desired size may receive periodic refresh operations to maintain the integrity

of data therein. A memory controller may initiate refresh operations by issuing refresh control signals to the DRAM circuits with sufficient frequency to prevent any loss of data in the DRAM circuits. After a refresh control signal is issued to a DRAM circuit, a minimum time (e.g. denoted by tRFC) may be required to elapse before another control signal may be issued to that DRAM circuit. The tRFC parameter may therefore increase as the size of the DRAM circuit increases.

[0114] When the buffer chip receives a refresh control signal from the memory controller, it may refresh the smaller DRAM circuits within the span of time specified by the tRFC associated with the emulated DRAM circuit. Since the tRFC of the emulated DRAM circuits is larger than that of the smaller DRAM circuits, it may not be necessary to issue refresh control signals to all of the smaller DRAM circuits simultaneously. Refresh control signals may be issued separately to individual DRAM circuits or may be issued to groups of DRAM circuits, provided that the tRFC requirement of the smaller DRAM circuits is satisfied by the time the tRFC of the emulated DRAM circuits has elapsed. In use, the refreshes may be spaced to minimize the peak current draw of the combination buffer chip and DRAM circuit set during a refresh operation.

[0115] While various embodiments have been described above, it should be understood that they have been presented by way of example only, and not limitation. For example, any of the network elements may employ any of the desired functionality set forth hereinabove. Thus, the breadth and scope of a preferred embodiment should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with the following claims and their equivalents.

What is claimed is:

1. A sub-system, comprising:

an interface circuit in communication with a plurality of memory circuits and a system, the interface circuit operable to interface the memory circuits and the system for simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of memory circuits;

wherein the at least one aspect is selected from the group consisting of a signal, a capacity, a timing, and a logical interface.

2. The sub-system of claim **1**, wherein the at least one aspect includes the signal.

3. The sub-system of claim **2**, wherein the signal is selected from the group consisting of an address signal, a control signal, and a data signal.

4. The sub-system of claim **1**, wherein the at least one aspect includes the capacity.

5. The sub-system of claim **4**, wherein the capacity includes a memory capacity.

6. The sub-system of claim **5**, wherein memory capacity relates to a number of the memory circuits.

7. The sub-system of claim **5**, wherein the interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with a first memory capacity that is greater than a second memory capacity of at least one of the plurality of memory circuits.

8. The sub-system of claim **1**, wherein the at least one aspect includes the timing.

9. The sub-system of claim **8**, wherein the timing relates to a latency.

10. The sub-system of claim **9**, wherein the latency includes a column address strobe (CAS) latency.

11. The sub-system of claim **10**, wherein the interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with a first CAS latency that is longer than a second CAS latency of at least one of the plurality of memory circuits.

12. The sub-system of claim **9**, wherein the latency includes a row address strobe (RAS) to CAS latency (tRCD).

13. The sub-system of claim **12**, wherein the interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with a first tRCD that is longer than a second tRCD of at least one of the plurality of memory circuits.

14. The sub-system of claim **9**, wherein the latency includes a row precharge latency (tRP).

15. The sub-system of claim **14**, wherein the interface circuit is operable to interface the memory circuits and the system for simulating at least one memory circuit with a first tRP that is longer than a second tRP of at least one of the plurality of memory circuits.

16. The sub-system of claim **9**, wherein the latency includes an activate to precharge latency (tRAS).

17. The sub-system of claim **1**, wherein the at least one aspect includes the logical interface.

18. The sub-system of claim **1**, wherein the memory circuits each include dynamic random access memory (DRAM).

19. A method, comprising:

interfacing a plurality of memory circuits and a system; and

simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of memory circuits;

wherein the at least one aspect is selected from the group consisting of a signal, a capacity, a timing, and a logical interface.

20. A sub-system, comprising:

an interface circuit in communication with a plurality of memory circuits and a system, the interface circuit operable to interface the memory circuits and the system for simulating at least one memory circuit with at least one aspect that is different from at least one aspect of at least one of the plurality of memory circuits;

wherein the interface circuit interfaces a majority of address or control signals of the memory circuits.

* * * * *

US 20060233012A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2006/0233012 A1**

Sekiguchi et al. (43) **Pub. Date: Oct. 19, 2006**

(54) **SEMICONDUCTOR STORAGE DEVICE HAVING A PLURALITY OF STACKED MEMORY CHIPS**

(75) Inventors: **Tomonori Sekiguchi**, Tokyo (JP); **Hideki Osaka**, Tokyo (JP); **Tatemi Ido**, Tokyo (JP); **Osamu Nagashima**, Tokyo (JP); **Mitsuaki Katagiri**, Tokyo (JP); **Ichiro Anjo**, Tokyo (JP)

Correspondence Address:
MCDERMOTT WILL & EMERY LLP
600 13TH STREET, N.W.
WASHINGTON, DC 20005-3096 (US)

(73) Assignees: **ELPIDA MEMORY, INC.**; **HITACHI, LTD.**; **INTELLECTUAL PROPERTY GROUP**

(21) Appl. No.: **11/392,805**

(22) Filed: **Mar. 30, 2006**

(30) **Foreign Application Priority Data**

Mar. 30, 2005 (JP) 2005-098027

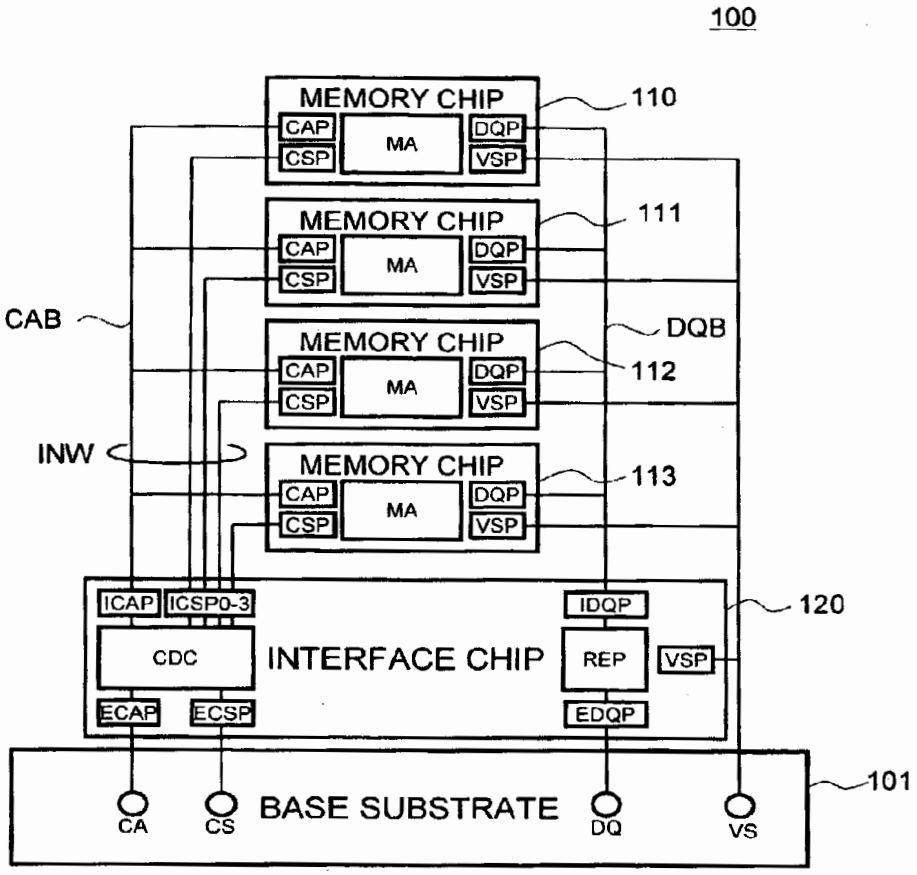
Publication Classification

(51) **Int. Cl.**
G11C 5/02 (2006.01)

(52) **U.S. Cl.** **365/51**

(57) **ABSTRACT**

A semiconductor storage employs a base substrate (**101**) having a command/address external terminal group (CA), a data input/output external terminal group (DQ), and a single chip select external terminal (CS), and also comprises a plurality of memory chips (**110**) to (**113**) mounted on a base substrate (**101**), each of which can individually carry out read and write operations. The terminals (CA), (DQ), and (CS) are connected to an interface chip (**120**). The interface chip (**120**) has a chip select signal generation circuit that can individually activate a plurality of memory chips (**110**) to (**113**) on the basis of an address signal fed by way of the terminal (CA) and on the basis of a chip select signal fed by way of the terminal (CS).



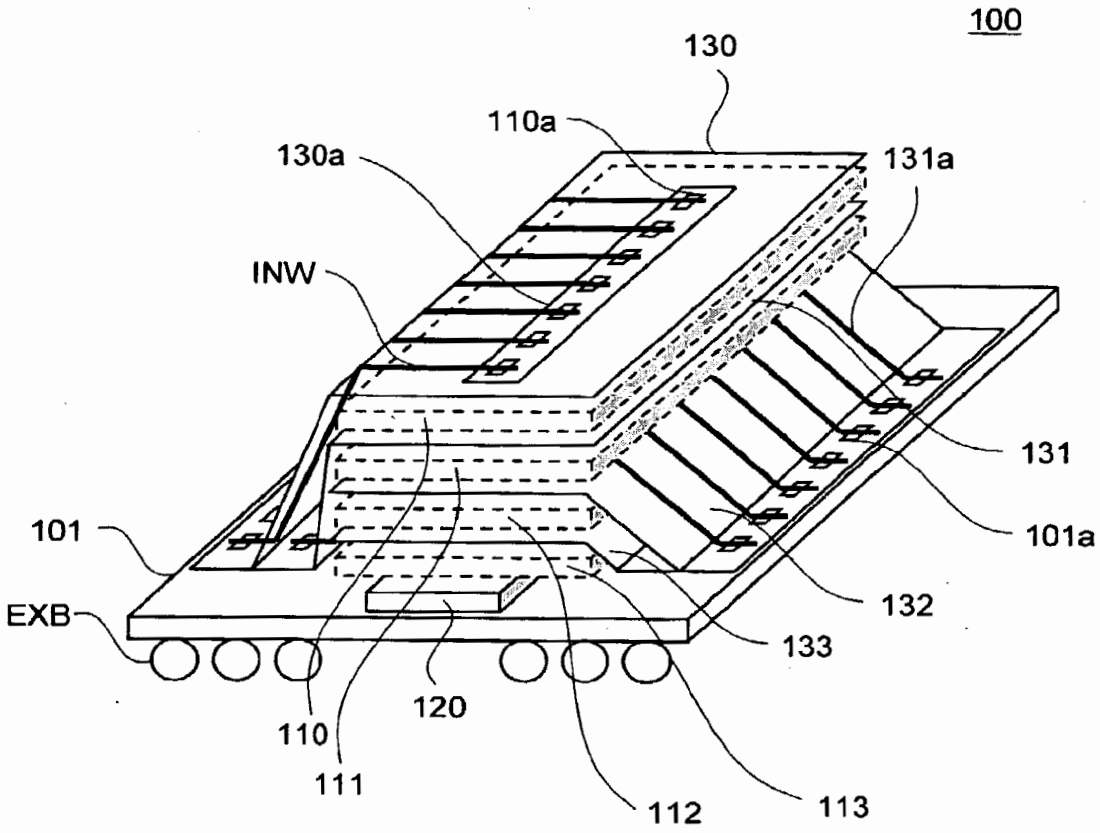

100
130
110a
130a
131a
INW
131
101a
101
132
EXB
133
120
110
111
112
113

FIG. 1

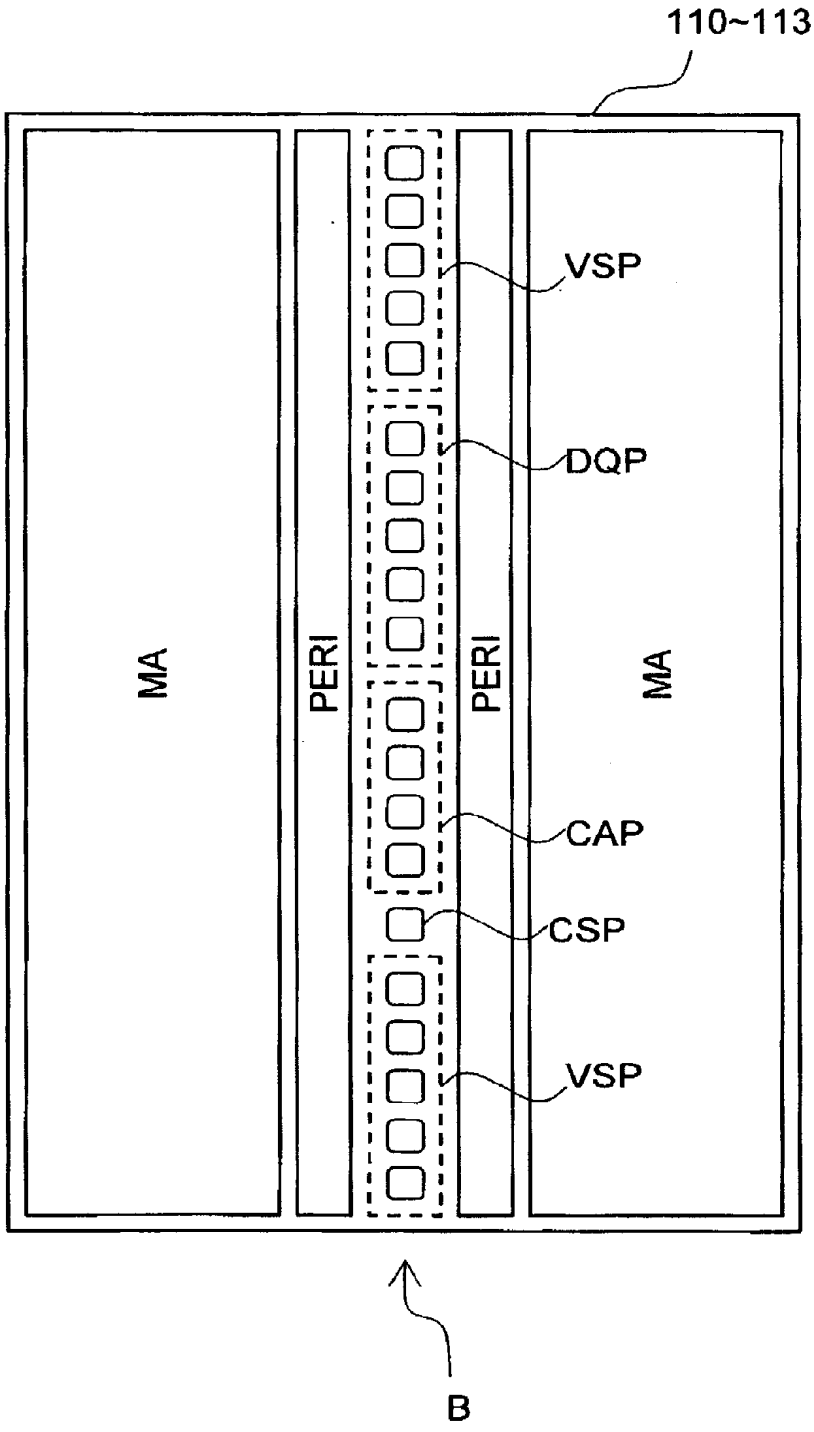


FIG.2

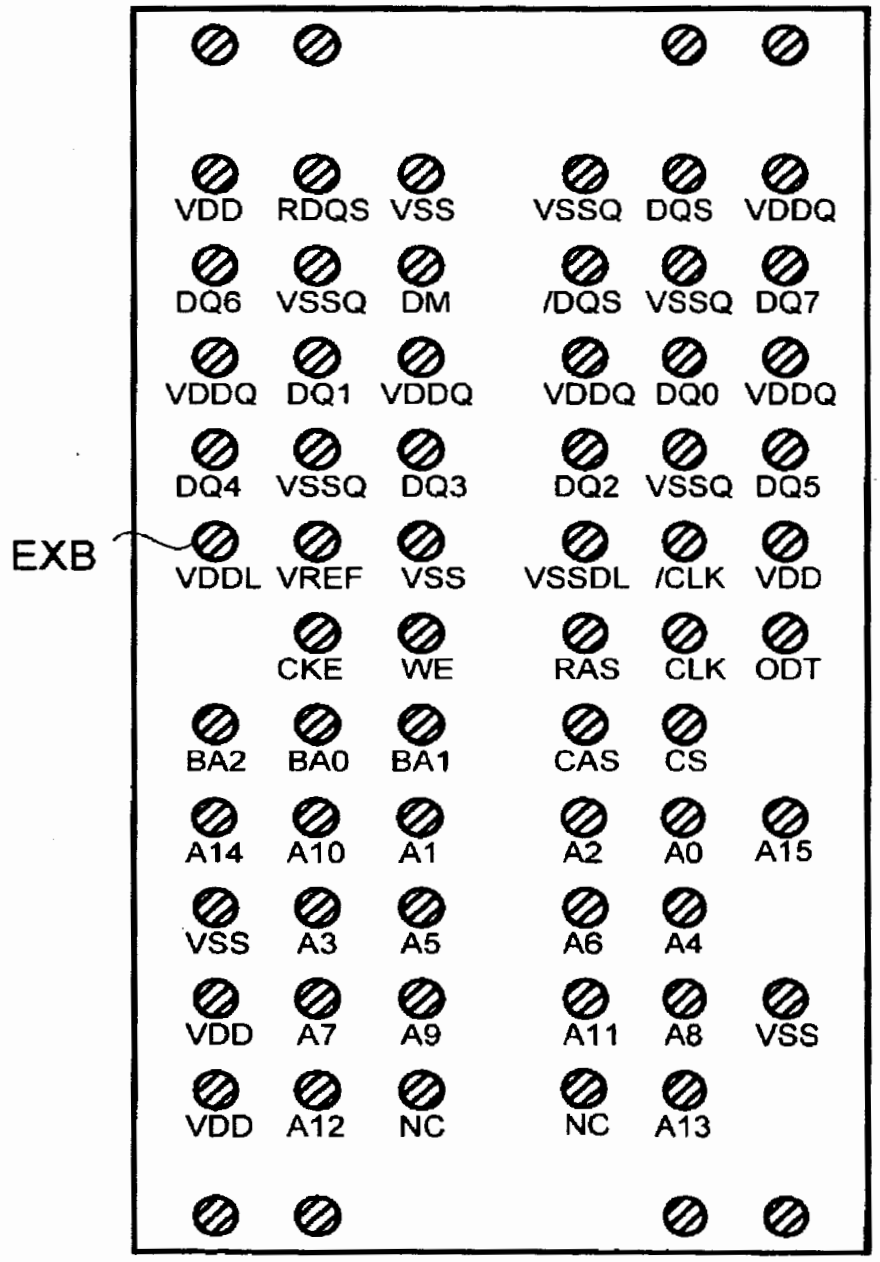

EXB
VDD RDQS VSS VSSQ DQS VDDQ
DQ6 VSSQ DM /DQS VSSQ DQ7
VDDQ DQ1 VDDQ VDDQ DQ0 VDDQ
DQ4 VSSQ DQ3 DQ2 VSSQ DQ5
VDDL VREF VSS VSSDL /CLK VDD
CKE WE RAS CLK ODT
BA2 BA0 BA1 CAS CS
A14 A10 A1 A2 A0 A15
VSS A3 A5 A6 A4
VDD A7 A9 A11 A8 VSS
VDD A12 NC NC A13

FIG.3

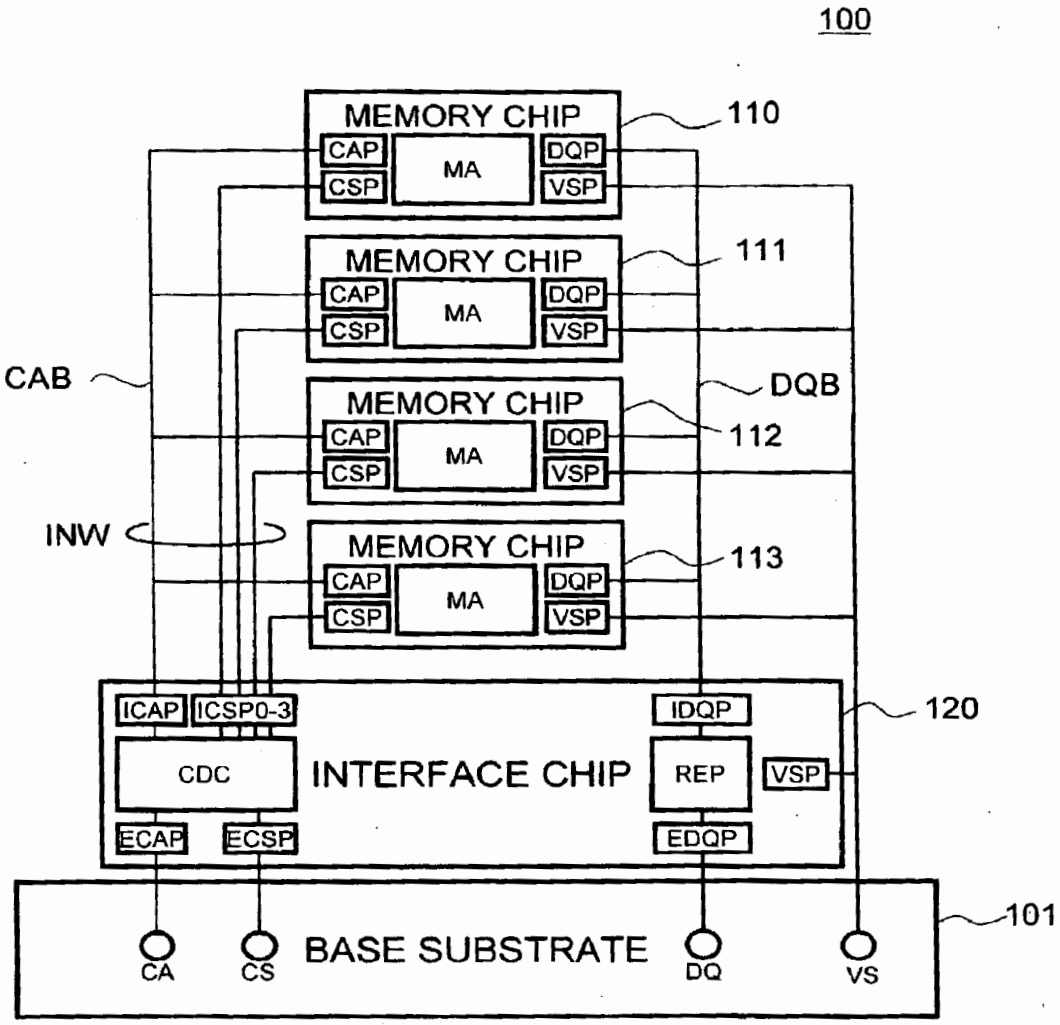

100
MEMORY CHIP
CAP
CSP
MA
DQP
VSP
110
111
112
113
CAB
DQB
INW
ICAP
ICSP0-3
IDQP
120
CDC
INTERFACE CHIP
REP
VSP
ECAP
ECSP
EDQP
101
CA
CS
BASE SUBSTRATE
DQ
VS

FIG.4

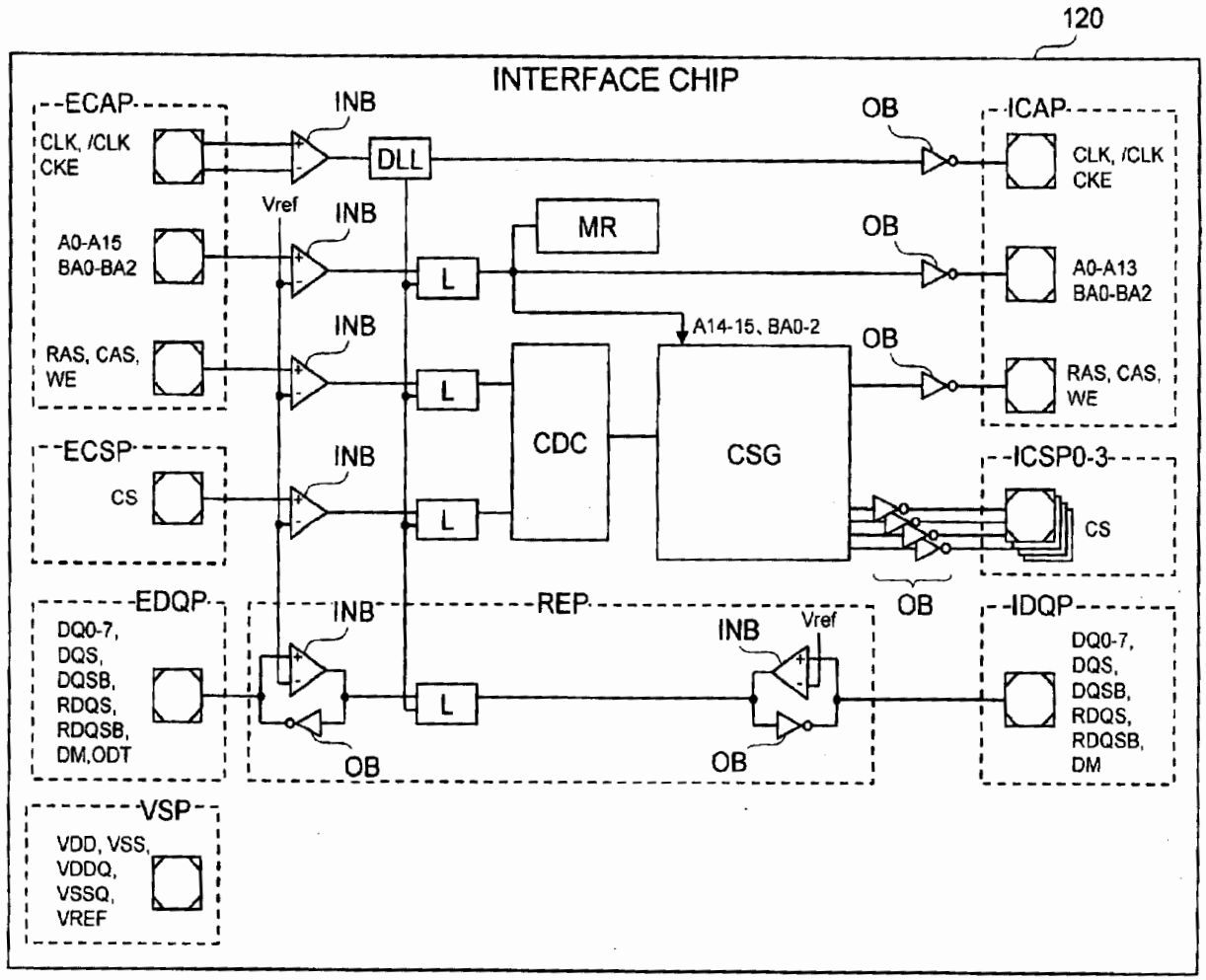

120
INTERFACE CHIP
ECAP
CLK, /CLK
CKE
A0-A15
BA0-BA2
RAS, CAS,
WE
ECSP
CS
EDQP
DQ0-7,
DQS,
DQSB,
RDQS,
RDQSB,
DM,ODT
VSP
VDD, VSS,
VDDQ,
VSSQ,
VREF
INB
Vref
DLL
L
MR
CDC
CSG
A14-15, BA0-2
OB
REP
ICAP
CLK, /CLK
CKE
A0-A13
BA0-BA2
RAS, CAS,
WE
ICSP0-3
CS
IDQP
DQ0-7,
DQS,
DQSB,
RDQS,
RDQSB,
DM

FIG.5

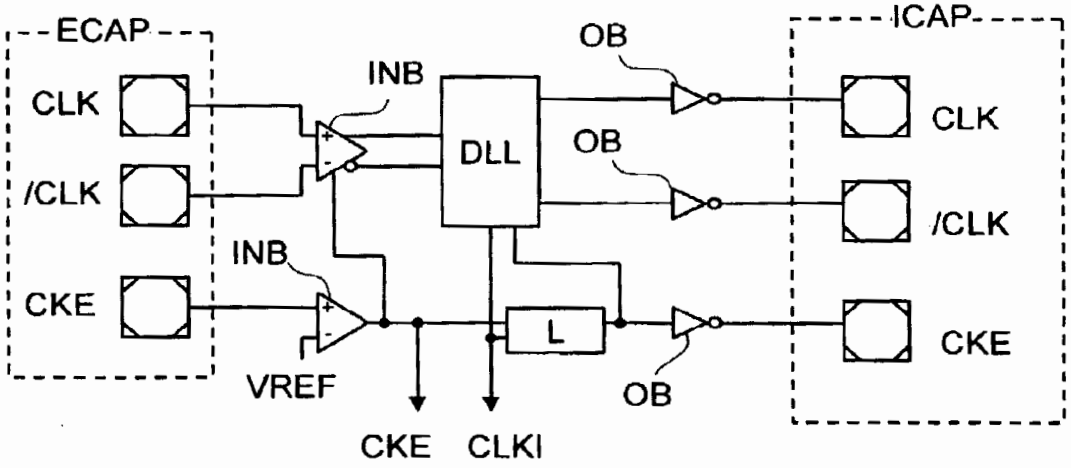

ECAP
CLK
/CLK
CKE
INB
INB
VREF
DLL
L
OB
OB
OB
CKE
CLKI
ICAP
CLK
/CLK
CKE

FIG.6

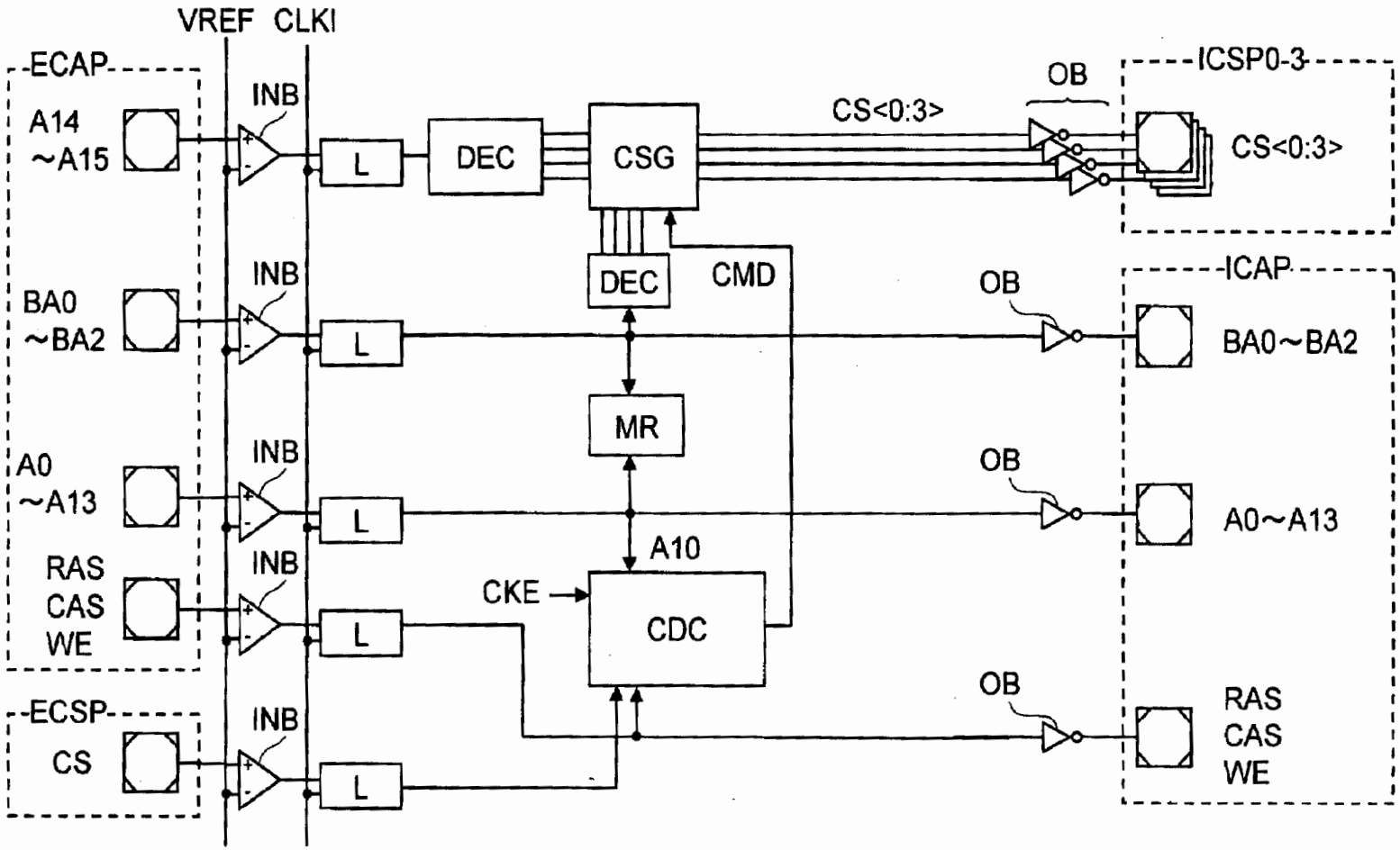

VREF
CLKI
ECAP
A14
~A15
BA0
~BA2
A0
~A13
RAS
CAS
WE
ECSP
CS
INB
L
DEC
CSG
CS<0:3>
OB
ICSP0-3
CS<0:3>
DEC
CMD
ICAP
BA0~BA2
MR
A0~A13
A10
CKE
CDC
RAS
CAS
WE

FIG.7

	COMMAND	SYMBOL	CS PUBLICATION METHOD
INITIALIZATION	Mode register set	MRS	Simultaneously activate all CS
	Extended mode register set	EMRS	
ROW SYSTEM	Auto refresh	REF	Sequentially activate with phase difference
	Self refresh entry	SELF	
	Self refresh exit	SELFX	Simultaneously activate all CS
	Single bank precharge	PRE	Activate CS stored in the active bank register
	Precharge all banks	PALL	
	Bank activate	ACT	Select CS by decoding A14, 15 and store in active bank register
COLUMN SYSTEM	Write	WRIT	Activate CS stored in the active bank register
	Write with auto precharge	WRITA	
	Read	READ	
	Read with auto precharge	READA	
OTHER	No operation	NOP	Simultaneously activate all CS
	Device deselect	DESL	
	Power down mode entry	PDEN	
	Power down mode exit	PDEX	

FIG.8

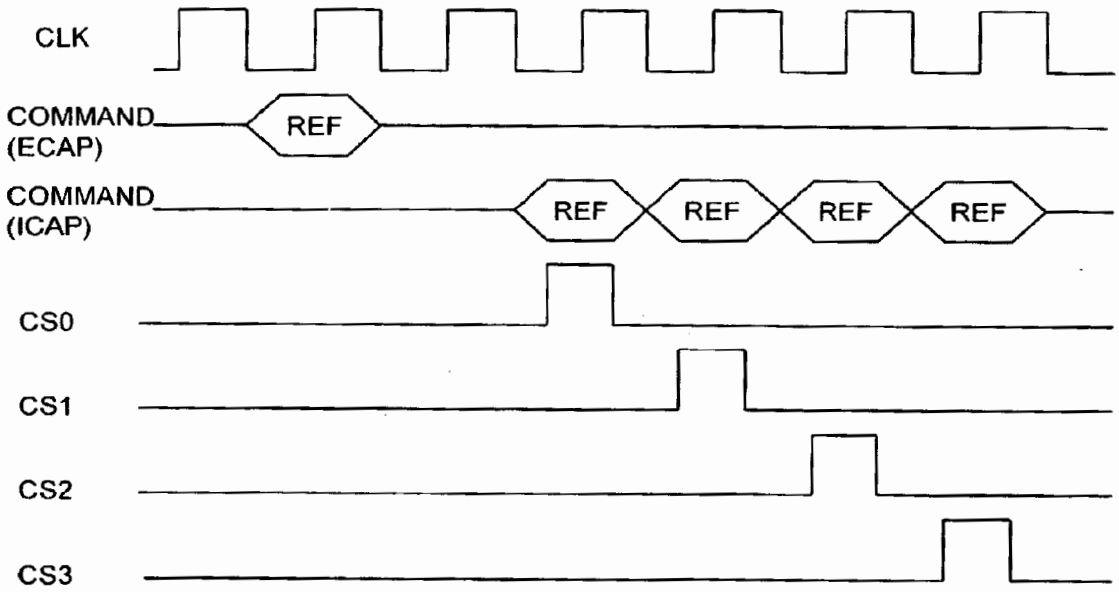

CLK
COMMAND (ECAP)
REF
COMMAND (ICAP)
REF
REF
REF
REF
CS0
CS1
CS2
CS3

FIG.9

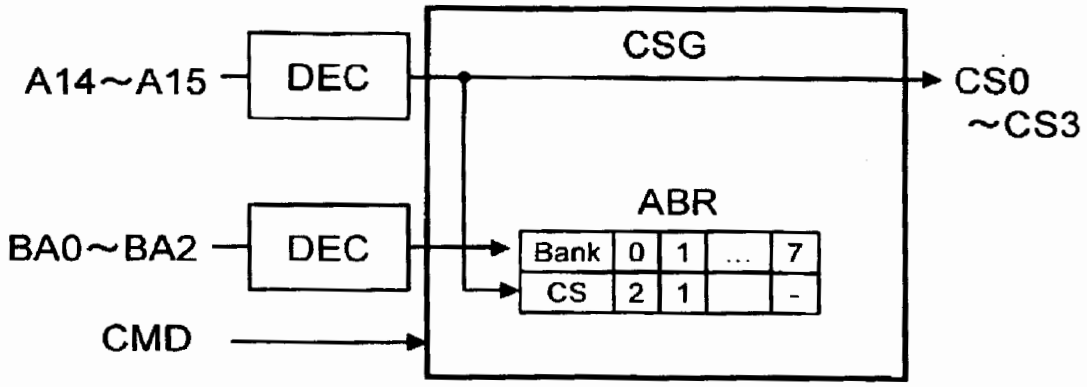


FIG.10A

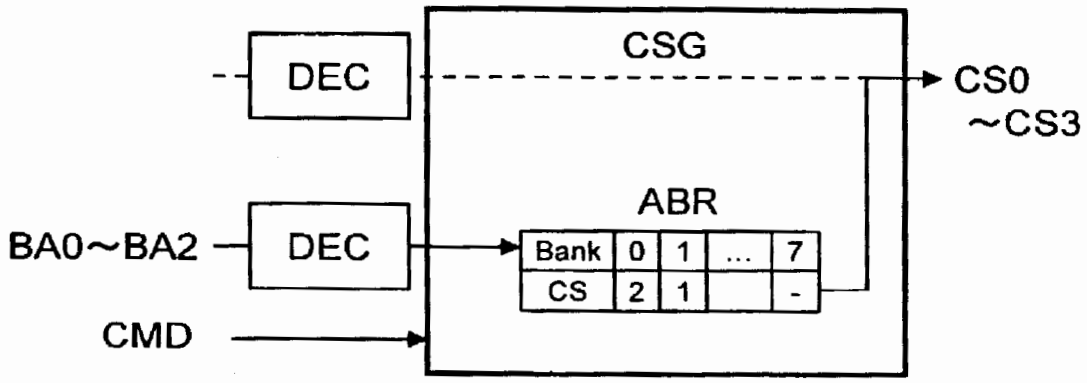


FIG.10B

FIG.11

IDQP
REP
EDQP
DM
DQ0 ~DQ7
/DQS
DQS
ODT
OB
OBV
INB
VREF
DQSI
DSG
L
RT
CLKI
ROV
VREF
RTV

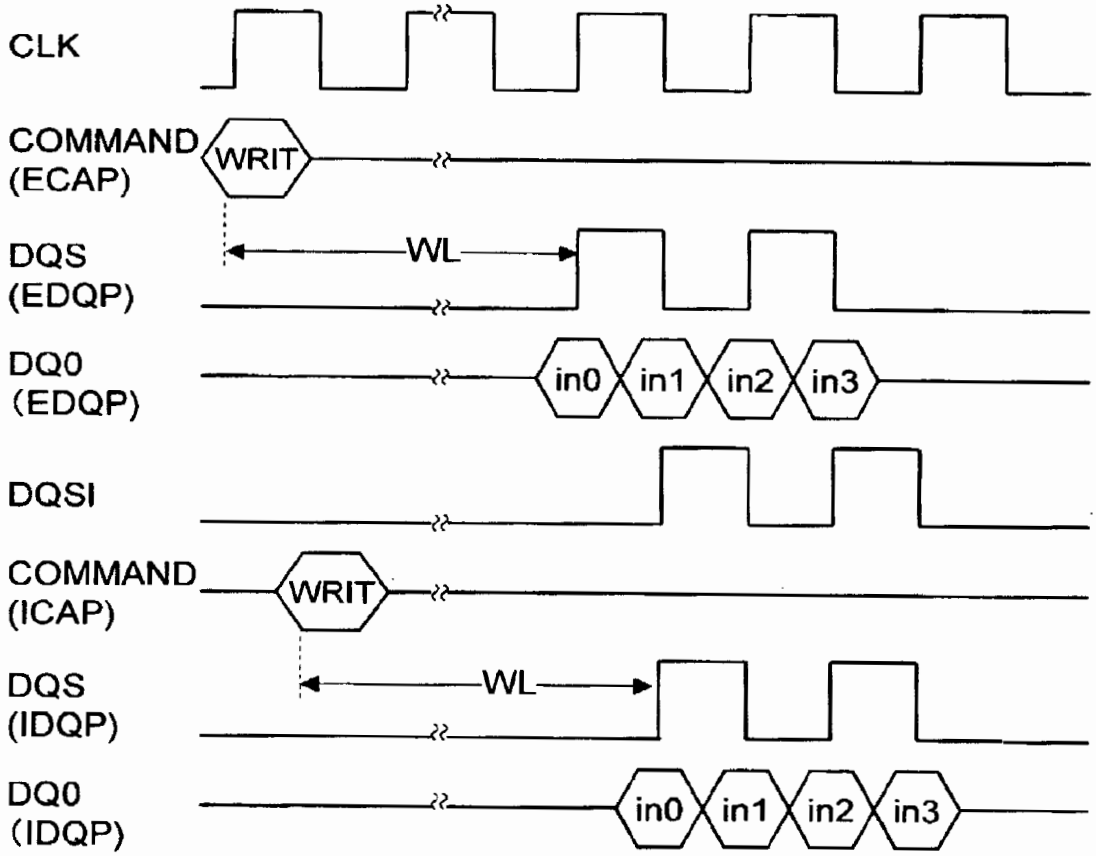

CLK
COMMAND (ECAP)
WRIT
DQS (EDQP)
WL
DQ0 (EDQP)
in0
in1
in2
in3
DQSI
COMMAND (ICAP)
WRIT
DQS (IDQP)
WL
DQ0 (IDQP)
in0
in1
in2
in3

FIG.12

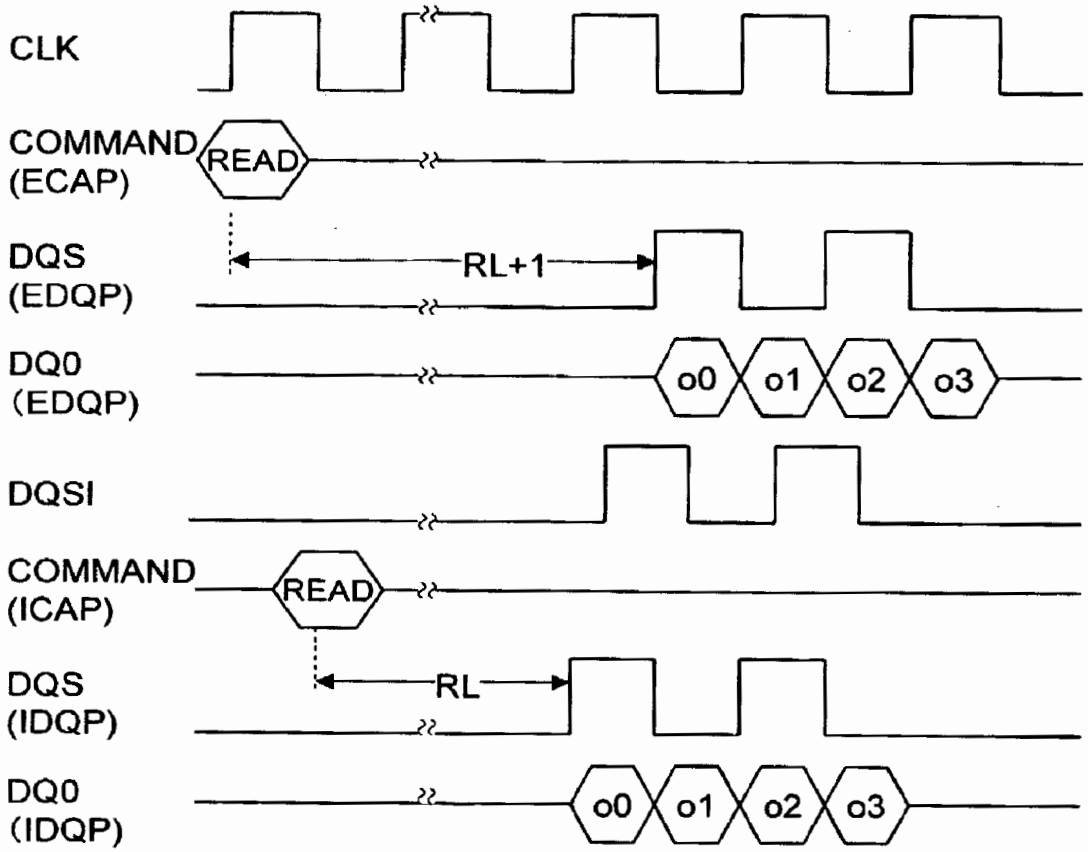

CLK
COMMAND
(ECAP)
READ
DQS
(EDQP)
RL+1
DQ0
(EDQP)
o0
o1
o2
o3
DQSI
COMMAND
(ICAP)
READ
DQS
(IDQP)
RL
DQ0
(IDQP)
o0
o1
o2
o3

FIG.13

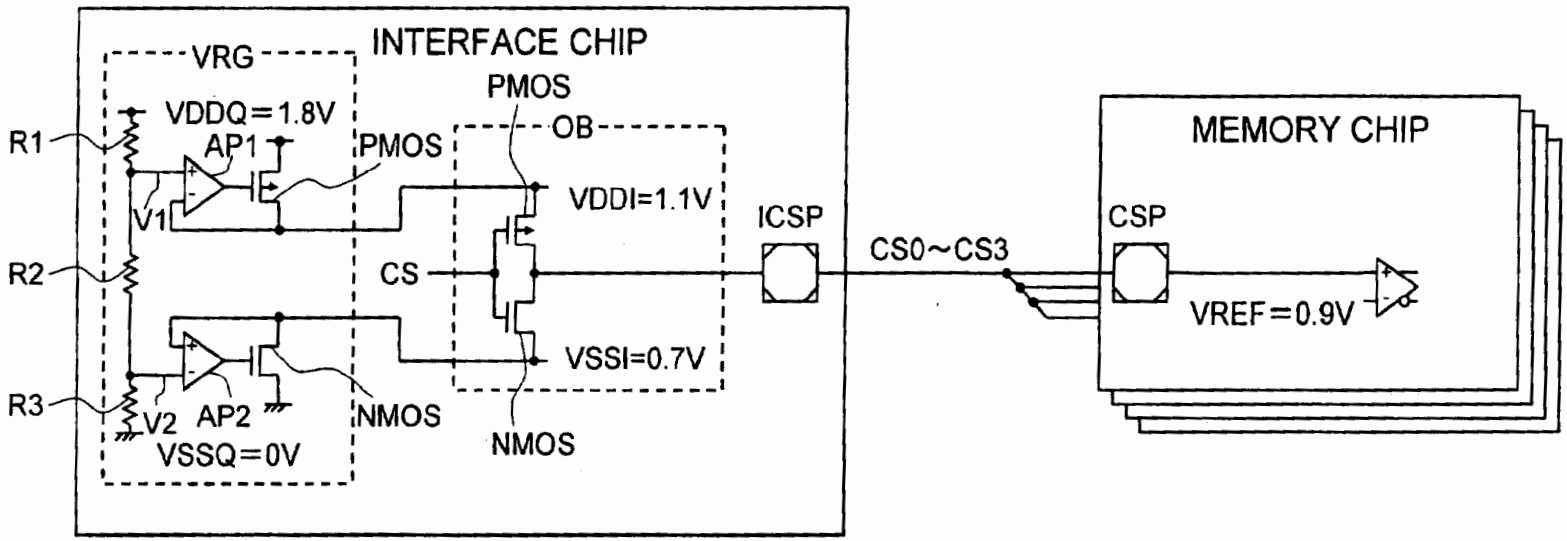

INTERFACE CHIP
VRG
VDDQ=1.8V
AP1
PMOS
PMOS
OB
R1
V1
VDDI=1.1V
ICSP
CS0~CS3
MEMORY CHIP
CSP
R2
CS
VREF=0.9V
VSSI=0.7V
R3
V2
AP2
NMOS
NMOS
VSSQ=0V

FIG.14

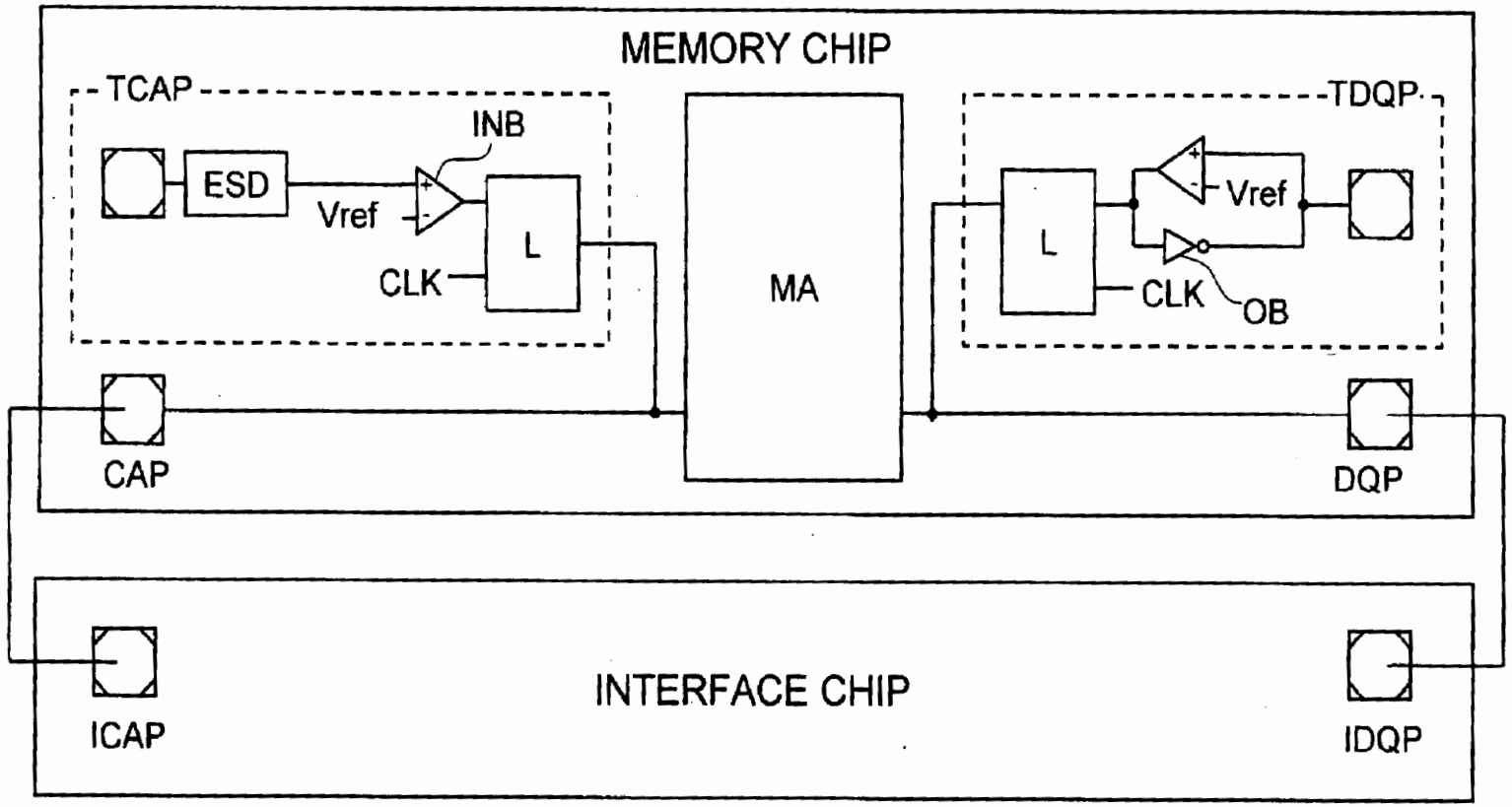

MEMORY CHIP
TCAP
INB
ESD
Vref
L
CLK
MA
TDQP
L
Vref
CLK
OB
CAP
DQP
INTERFACE CHIP
ICAP
IDQP

FIG.15

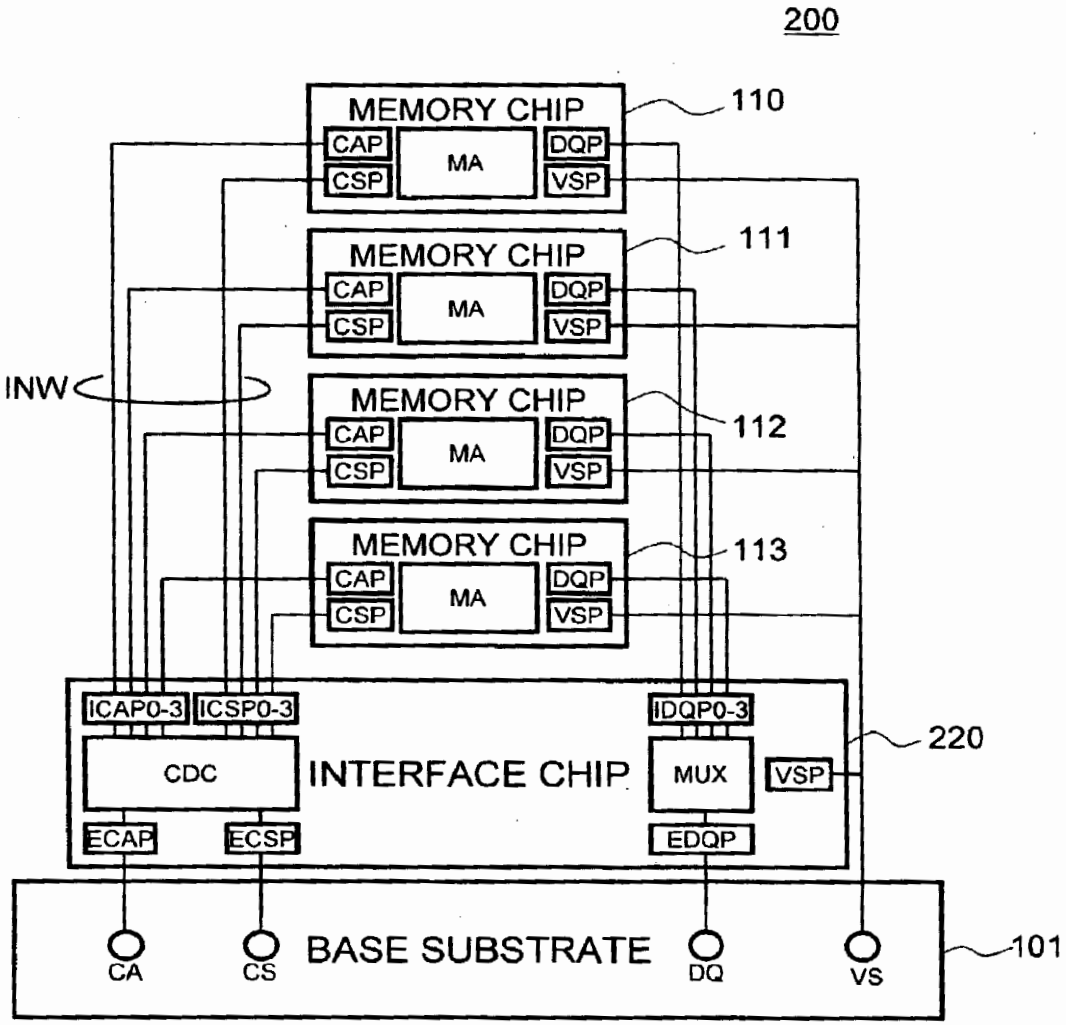

200
MEMORY CHIP
CAP
CSP
MA
DQP
VSP
110
111
112
113
INW
ICAP0-3
ICSP0-3
IDQP0-3
CDC
INTERFACE CHIP
MUX
VSP
220
ECAP
ECSP
EDQP
BASE SUBSTRATE
CA
CS
DQ
VS
101

FIG.16

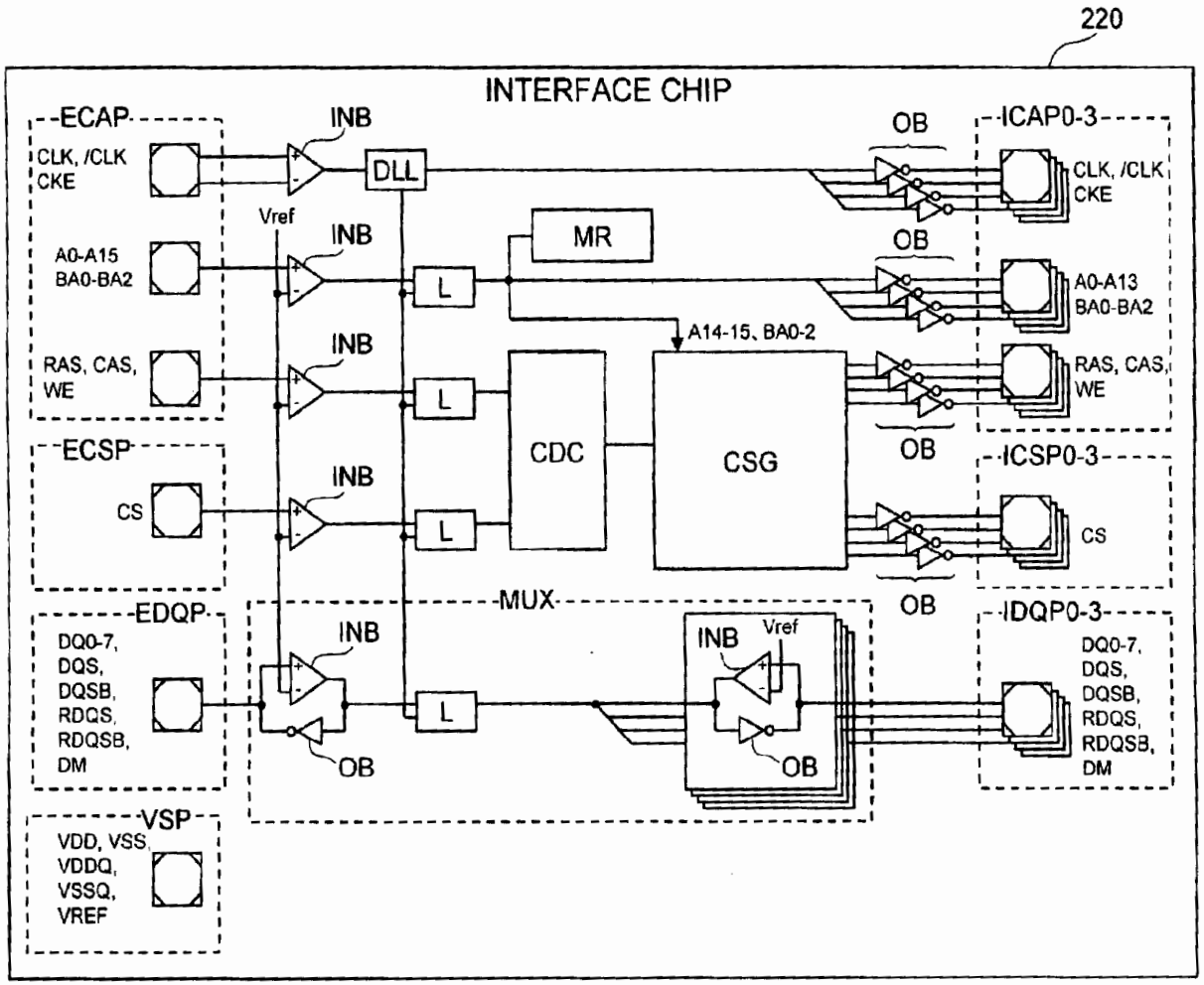

220
INTERFACE CHIP
ECAP
CLK, /CLK
CKE
A0-A15
BA0-BA2
RAS, CAS,
WE
ECSP
CS
EDQP
DQ0-7,
DQS,
DQSB,
RDQS,
RDQSB,
DM
VSP
VDD, VSS,
VDDQ,
VSSQ,
VREF
INB
Vref
DLL
MR
L
CDC
CSG
A14-15, BA0-2
OB
MUX
ICAP0-3
CLK, /CLK
CKE
A0-A13
BA0-BA2
RAS, CAS,
WE
ICSP0-3
CS
IDQP0-3
DQ0-7,
DQS,
DQSB,
RDQS,
RDQSB,
DM

FIG.17

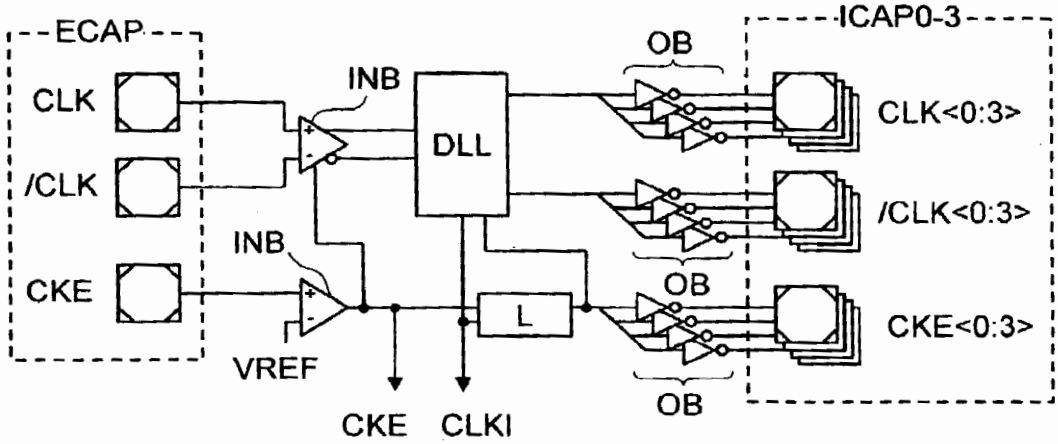

ECAP
CLK
/CLK
CKE
INB
INB
VREF
DLL
L
CKE
CLKI
OB
OB
OB
ICAP0-3
CLK<0:3>
/CLK<0:3>
CKE<0:3>

FIG.18

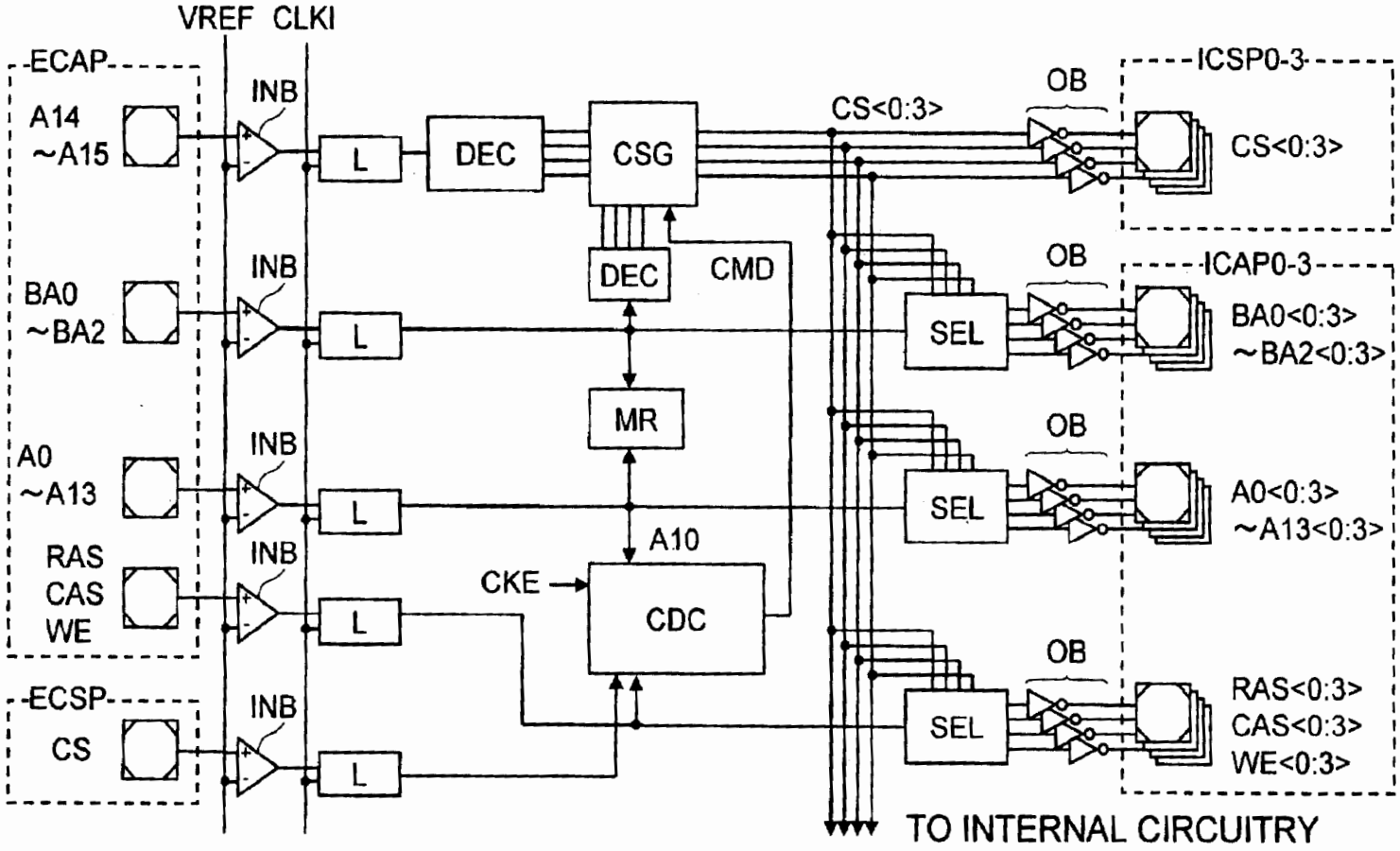

VREF
CLKI
ECAP
A14
~A15
BA0
~BA2
A0
~A13
RAS
CAS
WE
ECSP
CS
INB
L
DEC
CSG
CS<0:3>
OB
ICSP0-3
CS<0:3>
DEC
CMD
SEL
ICAP0-3
BA0<0:3>
~BA2<0:3>
MR
A10
CKE
CDC
A0<0:3>
~A13<0:3>
RAS<0:3>
CAS<0:3>
WE<0:3>
TO INTERNAL CIRCUITRY

FIG.19

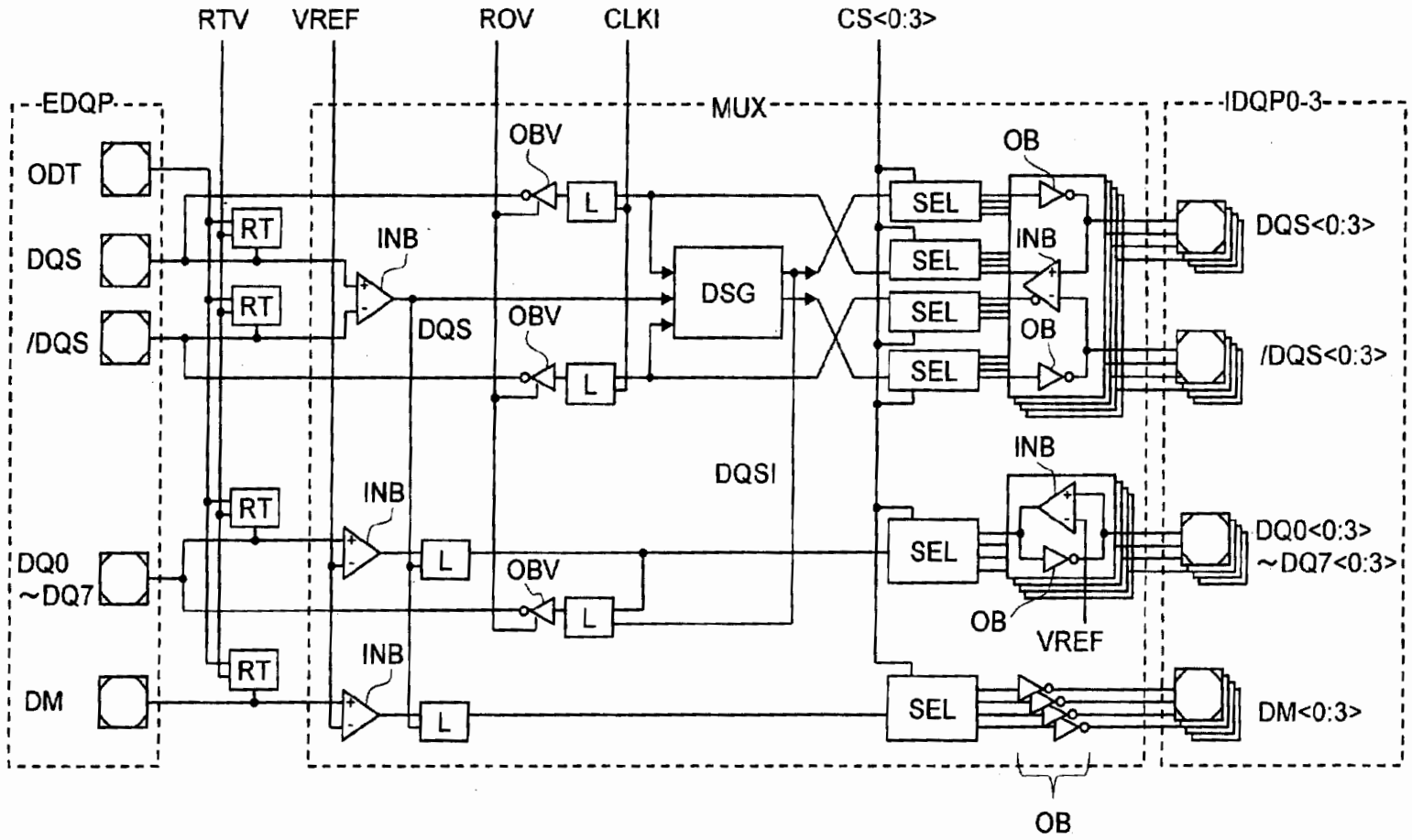

RTV
VREF
ROV
CLKI
CS<0:3>
EDQP
MUX
IDQP0-3
ODT
DQS
/DQS
DQ0
~DQ7
DM
RT
INB
OBV
L
DSG
DQS
DQSI
SEL
OB
VREF
DQS<0:3>
/DQS<0:3>
DQ0<0:3>
~DQ7<0:3>
DM<0:3>

FIG.20

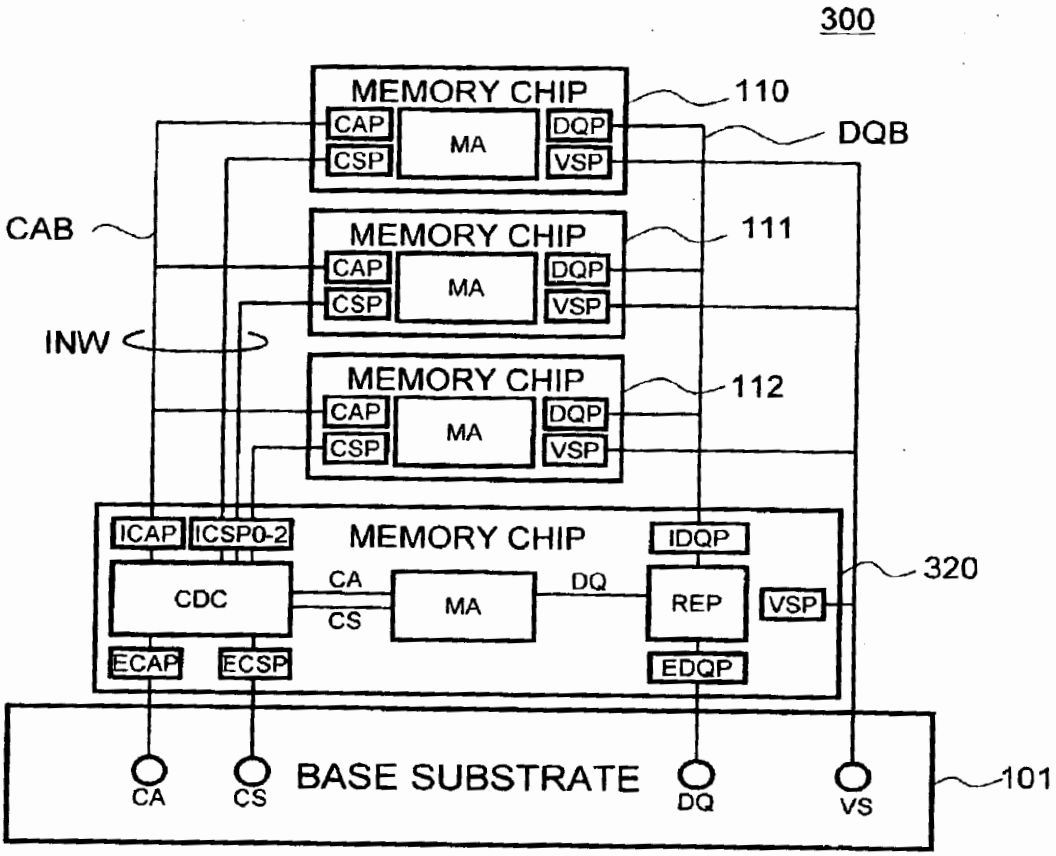

300
MEMORY CHIP
110
CAP
MA
DQP
CSP
VSP
DQB
CAB
MEMORY CHIP
111
INW
MEMORY CHIP
112
ICAP
ICSP0-2
MEMORY CHIP
IDQP
320
CDC
CA
CS
MA
DQ
REP
VSP
ECAP
ECSP
EDQP
BASE SUBSTRATE
CA
CS
DQ
VS
101

FIG.21

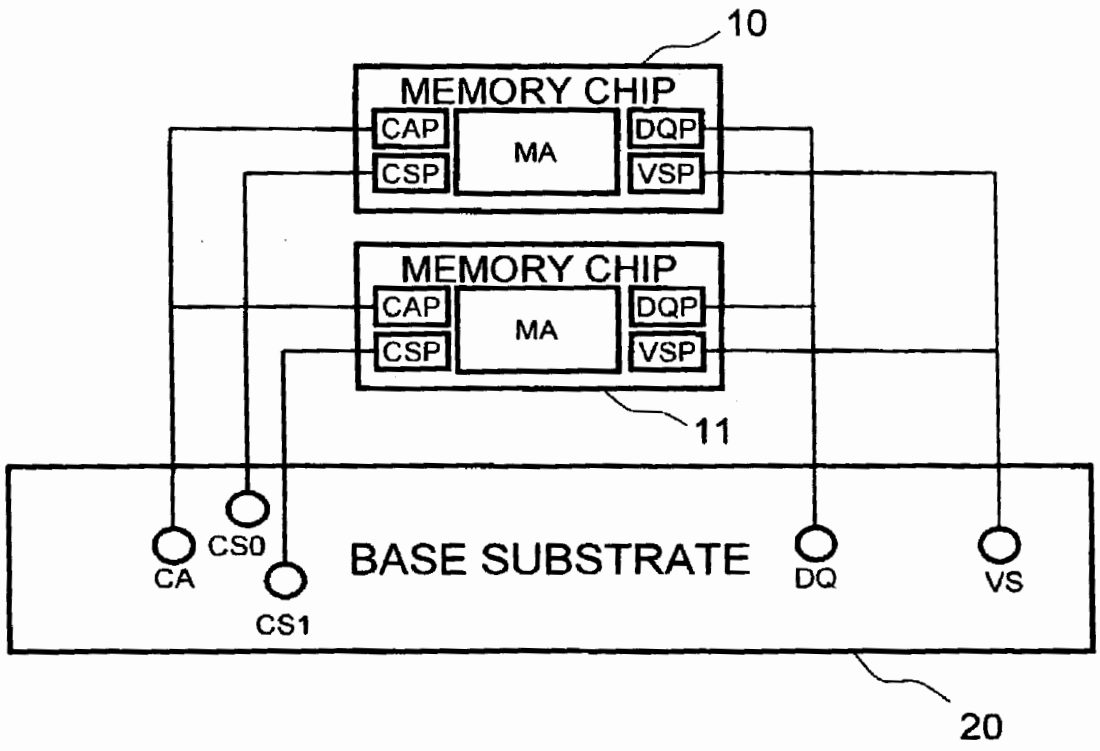


FIG.22

PRIOR ART

SEMICONDUCTOR STORAGE DEVICE HAVING A PLURALITY OF STACKED MEMORY CHIPS

TECHNICAL FIELD

[0001] The present invention relates to a semiconductor storage device, and more specifically relates to a semiconductor storage device that has a plurality of stacked memory chips.

BACKGROUND OF THE INVENTION

[0002] Semiconductor storage devices typified by DRAM (Dynamic Random Access Memory) are widely used as the main memory of servers and other computers, but with the increase in hardware speed and software complexity, the storage capacity required in main memory continues to increase. For this reason, it is difficult to meet the storage capacity needs required in main memory by using only a single memory chip, and a module referred as DIMM (Dual Inline Memory Module) in which a plurality of memory chips are mounted on a module substrate is ordinarily used.

[0003] However, the number of memory chips that can be mounted is low and the resulting storage capacity is limited when the memory chip is disposed in a flat configuration on a module substrate. Methods have been proposed in which a plurality of memory chips are three-dimensionally stacked and packaged, and the memory chips are mounted on a module substrate. See Japanese Patent Application Laid-Open Nos. H2-290048, 2001-110978, 2001-273755, and 2004-327474.

[0004] **FIG. 22** is a schematic diagram of a conventional semiconductor storage device that is three-dimensionally stacked and packaged. The semiconductor storage device shown in **FIG. 22** has a structure in which two memory chips **10** and **11** are stacked on the base substrate **20**, and the base substrate **20** is mounted on a module substrate (not shown). Each of the memory chips **10** and **11** has a memory array MA, a command/address pad CAP to which commands and addresses are supplied, a chip select signal pad CSP to which a chip select signal is supplied, a data input/output pad DQP for inputting and outputting data, and a power pad VSP for supplying power.

[0005] The command/address pad CAP, data input/output pad DQP, and power pad VSP each have a shared connection with the memory chip **10** and memory chip **11**, and are brought together in the command/address external terminal group CA, data input/output external terminal group DQ, and power supply external terminal group VS, respectively, on the base substrate **20**. On the contrary, the chip select signal pad CSP does not share a connection, and the chip select signal pad CSP of the memory chip **10** is connected to the chip select signal external terminal CS0, and the chip select signal pad CSP of the memory chip **11** is connected to the chip select signal external terminal CS1. The purpose of this configuration is to allow the memory chips. **10** and **11** to be selectively activated.

[0006] Another method has recently been proposed whereby the core unit in which the memory cell is formed and the interface unit in which the peripheral circuit for the memory cell is formed are each placed in a separate chip, and higher storage capacity and higher speeds are made possible by stacking these chips. This means that semiconductor storage devices that are conventionally composed of a single chip are divided into a plurality of chips. This method therefore allows the capacity of the chip (core chip) in which the core unit is formed to be increased, and makes it possible to obtain a very large storage capacity by furthermore stacking the core chips.

[0007] However, since the semiconductor storage device shown in **FIG. 22** is configured so that the command/address external terminal group CA and data input/output external terminal group DQ are connected to the memory chips **10** and **11**, there is a problem in that the parasitic capacitance of the command/address external terminal group CA and the data input/output external terminal group DQ is considerable in comparison with ordinary semiconductor storage devices in which only a single chip has been packaged.

[0008] Since such parasitic capacitance is the cause of signal waveform disturbance, there is a possibility that an adequate operating margin cannot be assured when the interface speed is very high, that is, when the clock speed exceeds 1 GHz, for example. This problem makes it difficult to make the capacity sufficiently large by using the method shown in **FIG. 22** because the problem is made more conspicuous as the number of memory chips to be stacked increases.

[0009] Also, since the semiconductor storage device shown in **FIG. 22** is different from an ordinary semiconductor storage device in which only a single chip has been packaged in the sense that two chip select signal external terminals are provided (CS0, CS1) compatibility with ordinary semiconductor storage devices is therefore lost. For this reason, a special-purpose memory controller must be designed in order to use the semiconductor storage device shown in **FIG. 22**, and the overall cost of the system is liable to increase.

[0010] Using the method described in Japanese Patent Application Laid-Open No. 2004-327474 appears to be able to solve the above-described problems, but this configuration is fundamentally different from the type of semiconductor storage device in which memory chips that are capable of independently carrying out read and write operations are stacked, and is rather a configuration in which core chips that are not capable of independently carrying out read and write operations are stacked, as shown in **FIG. 22**.

[0011] Therefore, an ordinary memory chip naturally cannot be used and a special-purpose chip must be newly designed. Also, with the configuration described in Japanese Patent Application Laid-Open No. 2004-327474, a through-electrode must be formed in the chips, and other advanced manufacturing techniques that are not sufficiently established must be used. It is therefore believed that manufacturing such chips at low cost under current conditions is difficult.

SUMMARY OF THE INVENTION

[0012] The present invention was contrived in order to solve such problems, and an object of the present invention is to reduce the parasitic capacitance of the command/address external terminal group and the data input/output external terminal group to an amount that is equal to that of a single chip component in a semiconductor storage device having a plurality of stacked memory chips.

[0013] Another object of the present invention is to assure compatibility with ordinary semiconductor storage devices in which a single chip is packaged, in a semiconductor storage device having a plurality of stacked memory chips.

[0014] Yet another object of the present invention is to implement the objects described above with low cost by using memory chips that are each individually capable of read and write operations.

[0015] The above and other objects of the present invention can be accomplished by a semiconductor storage device comprising: a base substrate having a command/address external terminal group to which command signals and address signals are supplied, a data input/output external terminal group for inputting and outputting data signals, and a single chip select external terminal; and a plurality of memory chips that are stacked on the base substrate and are each individually capable of reading and writing operations, wherein a plurality of terminals constituting said command/address external terminal group, a plurality of terminals constituting said data input/output external terminal group, and said single chip select external terminal are connected to a single chip that has an interface function; and the single chip that has the interface function further has at least a chip select signal generation circuit that can individually activate said plurality of memory chips based on said address signals supplied via said command/address external terminal group and based on said chip select signal supplied via said chip select external terminal.

[0016] In accordance with the present invention, since the command/address external terminal group and other external terminals are connected to a single chip, the parasitic capacitance of the terminals can be reduced to an amount that is about equal to that of a single chip component. Also, since the semiconductor storage device of the present invention is provided with only a single chip select external terminal, it is possible to assure compatibility with ordinary semiconductor storage devices in which a single chip is packaged, even though the semiconductor storage device is provided with a plurality of stacked memory chips.

[0017] In addition, the stacked memory chips are memory chips that are individually capable of read and write operations. It is therefore possible to use so-called general-purpose memory, and hence to manufacture these chips at relatively low cost.

[0018] In the present invention, the single chip that has an interface function may be an interface chip that is different than the plurality of memory chips, or may be any one of the plurality of memory chips. With the former, it is possible to use the exact same chip as the memory chip that is used, and with the latter, it is possible to reduce the number of stacked chips by one.

[0019] The command signals, address signals, and data signals may be transmitted among the plurality of memory chips and the single chip that has an interface function by way of wiring with a shared connection to the plurality of memory chips, or may be carried out by way of wiring that is individually connected to the plurality of memory chips. The former allows the number of wires to be reduced, and the latter allows the wiring capacitance to be reduced.

[0020] Thus, in accordance with the present invention, it is possible to reduce the parasitic capacitance of the interface unit of a semiconductor storage device having a plurality of stacked memory chips to an amount equal to that of a single chip component. The input/output frequency of the command signal, address signal, and data signal can thereby be increased to that equivalent to a single chip component. Also, since the memory chips are three-dimensionally stacked in the present invention, the mounting surface area can be made equal to that of a single chip component.

[0021] Furthermore, since only a single chip select external terminal is provided, it is possible to assure compatibility with ordinary semiconductor storage devices in which a single chip is packaged. For this reason, there is no requirement that a special-purpose memory controller or the like be designed, and cost increases for the entire system can be held in check.

[0022] Since the stacked memory chips are furthermore ordinary memory chips (general-purpose DRAM, for example) that are each individually capable of read and write operations, formation of a through-electrode or other special steps are not required, and an increase in manufacturing costs can be inhibited.

[0023] Since it is also possible to individually activate the plurality of memory chips by using a chip select signal generation circuit, the power consumption of the entire package can also be reduced.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] The above and other objects, features and advantages of this invention will become more apparent by reference to the following detailed description of the invention taken in conjunction with the accompanying drawings, wherein:

[0025] **FIG. 1** is a simplified perspective view that schematically shows a structure of the semiconductor storage device according to the first embodiment of the present invention;

[0026] **FIG. 2** is a simplified planar view of the memory chips shown in **FIG. 1**;

[0027] **FIG. 3** is a layout drawing of ball electrodes disposed on the reverse side of the base substrate shown in **FIG. 1**;

[0028] **FIG. 4** is a schematic diagram that describes the connection configuration between the memory chips and the interface chip in the first embodiment;

[0029] **FIG. 5** is a simplified circuit diagram showing the circuit configuration of the interface chip used in the first embodiment;

[0030] **FIG. 6** is a diagram showing in greater detail the circuit configuration of the clock system circuitry contained in the interface chip shown in **FIG. 5**;

[0031] **FIG. 7** is a diagram showing in greater detail the circuit configuration of the command/address system circuitry contained in the interface chip shown in **FIG. 5**;

[0032] **FIG. 8** is a list showing the commands and symbols that are output by the command decoder shown in **FIG. 7**, and the methods for issuing a chip select signal CS when the commands are executed;

[0033] **FIG. 9** is a timing chart showing the issue timing of the refresh command to each memory chip;

[0034] **FIG. 10A** is a diagram for describing the function of the active bank register in the case where an activation command is issued;

[0035] **FIG. 10B** is a diagram for describing the function of the active bank register in the case where a column system command is issued;

[0036] **FIG. 11** is a diagram showing in greater detail the circuit configuration of the principal components of the data system circuitry contained in the interface chip shown in **FIG. 5**;

[0037] **FIG. 12** is a timing chart showing the write operation of the semiconductor storage device according to the first embodiment;

[0038] **FIG. 13** is a timing chart showing the read operation of the semiconductor storage device according to the first embodiment;

[0039] **FIG. 14** is a schematic diagram that describes the method for reducing the amplitude of the inter-chip interface between the memory chips and the interface chip;

[0040] **FIG. 15** is a diagram for describing an improved method for connecting the pads and the memory array in the memory chips;

[0041] **FIG. 16** is a schematic diagram of the connection configuration between the memory chips and the interface chip of the second embodiment;

[0042] **FIG. 17** is a simplified circuit diagram showing the circuit configuration of the interface chip shown in **FIG. 16** used in the second embodiment;

[0043] **FIG. 18** is a diagram showing in greater detail the circuit configuration of the clock system circuit contained in the interface chip shown in **FIG. 17**;

[0044] **FIG. 19** is a diagram showing in greater detail the circuit configuration of the command/address system circuit contained in the interface chip shown in **FIG. 17**;

[0045] **FIG. 20** is a diagram showing in greater detail the circuit configuration of the data system circuit contained in the interface chip shown in **FIG. 17**;

[0046] **FIG. 21** is a schematic diagram that describes the connection configuration between the memory chips and the memory chip in the semiconductor storage device of the third embodiment of the present invention; and

[0047] **FIG. 22** is a schematic diagram of a conventional semiconductor storage device that is three-dimensionally stacked and packaged.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0048] Preferred embodiments of the present invention will now be explained in detail with reference to the drawings.

[0049] **FIG. 1** is a simplified perspective view that schematically shows the structure of the semiconductor storage device **100** according to the first embodiment of the present invention.

[0050] The semiconductor storage device **100** of the present embodiment is provided with a base substrate **101**, four memory chips **110** to **113** stacked on the base substrate **101**, and an interface chip **120** disposed between the base substrate **101** and memory chips **110** to **113**, as shown in **FIG. 1**. In the present embodiment, DRAM is selected as the memory chips **110** to **113**, although this selection imposes no particular limitation. The memory chips **110** to **113** are chips having mutually the same configuration, and are DRAM chips that are individually capable of read and write operations, that is to say, the chips are ordinary general-purpose DRAM chips.

[0051] **FIG. 2** is a simplified planar view of the memory chips **110** to **113**.

[0052] The surface of the memory chips **110** to **113** can be largely classified into a memory array area MA, a peripheral circuit area PERI, and a bonding pad area B, as shown in **FIG. 2**. The bonding pad area B is disposed in a single direction in the center of the memory chip so as to be disposed between the two peripheral circuit areas PERI, and this area includes power supply pads VSP, a data input/output pad DQP, a command/address pad CAP, and a chip select signal pad CSP.

[0053] The power supply pads VSP have a peripheral circuit operating power supply VDD, a peripheral circuit ground power supply VSS, an output circuit operating power supply VDDQ, an output circuit ground power supply VSSQ, an interface reference power supply VREF, and the like. The data input/output pad DQP has data input/output pads DQ0 to DQ7 (the same applies to the case of ×8 components) a data strobe signal pad DQS, a data mask signal pad DM, and the like. The command/address pad CAP has a clock pad CLK, a clock enable pad CKE, a row address strobe pad RAS, a column address strobe pad CAS, a write enable pad WE, a bank address pads BA0 to BA2, an address pads A0 to A13, and the like.

[0054] The pads disposed on the memory chips **110** to **113** are connected to the pads **10**la disposed on the base substrate **101** by way of tapes **130** to **133** on which wiring INW is formed, as shown in **FIG. 1**. More specifically, pads **110***a* disposed on the memory chip **110** and pads **101***a* disposed on the base substrate **101** are connected by tape **130** that is disposed so as to cover the memory chip **110**, and pads disposed on the memory chips **111** to **113** and the pads **101***a* disposed on the base substrate **101** are similarly connected by tapes **131** to **133** that are disposed so as to cover the memory chips **111** to **113**, respectively.

[0055] Wiring (not shown) is formed on the base substrate **101**, and the pads **10**la and the pads (not shown) on the memory chip side of interface chip **120** are connected by the wiring. Pads (not shown) on the external circuit side of the interface chip **120** are connected to ball electrodes EXB disposed on the reverse side (mounting side) of the base substrate **101** by way of other wiring (not shown) disposed on the base substrate **101**. The ball electrodes EXB are external terminals of the semiconductor storage device **100** of the present embodiment.

[0056] **FIG. 3** is a layout drawing of ball electrodes EXB disposed on the reverse side of the base substrate **101**.

[0057] Other than the addition of address external terminals A**14** and A**15**, the types of ball electrodes EXB (which

is an external terminal group of the semiconductor storage device **100**) match the types of pads disposed on the memory chips **110** to **113**, as shown in **FIG. 3**. Therefore, the chip select signal external terminal CS is always a single terminal. The ball electrodes EXB can be classified into a power supply external terminal group (VS), a data input/output external terminal group (DQ), a command/address external terminal group (CA), and a chip select signal external terminal (CS).

[0058] The peripheral circuit operating power supply VDD, peripheral circuit ground power supply VSS, output circuit operating power supply VDDQ, output circuit ground power supply VSSQ, interface reference power supply VREF, and the like constitute the power external terminal group (VS) in the same manner as the pads disposed on the memory chips **110** to **113**. The data input/output terminals DQ**0** to DQ**7**, data strobe signal terminal DQS, data mask signal terminal DM, and the like constitute the data input/output external terminal group (DQ). The clock terminal CLK, clock enable terminal CKE, row address strobe terminal RAS, column address strobe terminal CAS, write enable terminal WE, bank address terminals BA**0** to BA**2**, address terminals **A0** to **A13**, and the like constitute the command/address external terminal group (CA).

[0059] **FIG. 4** is a schematic diagram that describes the connection configuration between the memory chips **110** to **113** and the interface chip **120** in the present embodiment.

[0060] In the present embodiment, only the chip select signals are connected in a one-to-one relationship between the memory chips **110** to **113** and the interface chip **120**, as shown in **FIG. 4**. The other signals, that is to say, the command/address signals and data signals, are bus-connected by the command/address bus CAB and the data bus DQB. The power supply is also bus-connected.

[0061] As used herein, the term "one-to-one connection" refers to a connection in which the pads of the memory chips **110** to **113** and the pads of the interface chip **120** are individually connected for each memory chip, and the term "bus-connected" refers to a connection in which the pads of the memory chips **110** to **113** and the pads of the interface chip **120** share connections. Therefore, in the present embodiment, in the set of signals transmitted between the memory chips **110** to **113** and the interface chip **120**, only the chip select signal is a signal that is individually provided to the memory chips **110** to **113**.

[0062] Described more specifically, the command/address external terminal group CA and the chip select signal external terminal CS are connected to the command/address pad ECAP and the chip select signal pad ECSP, respectively, on the interface chip **120** by way of the ball electrodes EXB on the base substrate **101**. These signals are supplied to the command decoder CDC, and the output signals of the decoder are supplied to the command/address pad ICAP and chip select signal pad ICSP, respectively.

[0063] The signals (commands and addresses) that are output to the command/address pad ICAP are supplied in a shared configuration to the command/address pads CAP of the four memory chips **110** to **113** by way of the command/address bus CAB. On the other hand, four chip select signals that are output to the chip select signal pad ICSP are connected in a one-to-one relationship between the interface chip **120** and each of the memory chips **110** to **113**.

[0064] The data input/output external terminal group DQ is connected once to the data input/output pad EDQP on the interface chip **120** by way of the ball electrodes EXB on the base substrate **101**. The data input/output signals are input to a bidirectional repeater REP, and are connected to the data input/output pads DQP of the memory chips **110** to **113** by way of the data input/output pad IDQP. The data input/output pad IDQP and the data input/output pads DQP of the four memory chips, **110** to **113** are connected by a data bus DQB in the form of a 1:4 bus.

[0065] The power supply external terminal group VS is connected to the power supply pads VSP of the memory chips **110** to **113** and to the interface chip **120** directly by way of the ball electrodes EXB on the base substrate **101**.

[0066] Thus, in the semiconductor storage device **100** of the present embodiment, the connections between the ball electrodes EXB on the base substrate **101** and the memory chips **110** to **113** (excluding the power supply) are all carried out by way of the interface chip **120**. The chip connected to the data input/output external terminal group DQ and the command/address external terminal group CA is only the interface chip **120**. Thus, the parasitic capacitance of the external terminals can be reduced to an amount equal to that of a single chip component, and the frequency of the command/address and data input/output can be increased to a level equivalent to that of a single chip component.

[0067] In the present embodiment, since the command/address signals and data input/output signals are connected in the form of a bus inside a package, the number of wires inside the package can be kept low and manufacturing costs can thereby be held in check.

[0068] **FIG. 5** is a simplified circuit diagram showing the circuit configuration of the interface chip **120** used in the present embodiment.

[0069] The command/address pad ECAP on the interface chip **120** has pads for three systems: the clock system signals CLK, /CLK, and CKE; address system signals **A0** to **A-15** and BA**0** to BA**2**; and command system signals RAS, CAS, and WE, as shown in **FIG. 5**. The clock system signals CLK, /CLK, and CKE are supplied from outside of the semiconductor storage device **100** and are supplied to a clock regeneration circuit DLL by way of an input buffer INB. The output of the clock regeneration circuit DLL is supplied to terminals that output, respectively, the clock signals CLK, /CLK, and CKE of the command/address pad ICAP by way of the output buffer OB. The output of the clock regeneration circuit DLL is distributed within the interface chip **120** in order to latch other input signals in the latch circuits L.

[0070] The address system signals **A0** to **A15** and BA**0** to BA**2** are latched in the latch circuits L by way of the input buffers INB. The timing signals of the latch circuits L are supplied by the clock regeneration circuit DLL described above. The output of the latch circuits L is supplied to the terminals that output the address signals **A0** to **A13** and BA**0** to BA**2** of the command/address pad ICAP by way of the output buffer OB.

[0071] Here, the address signals **A0** to **A13** and BA**0** to BA**2** of the address signals **A0** to **A15** and BA**0** to BA**2** that are supplied from the exterior are directly fed to the command/address pad ICAP. The address signals **A14** and **A15** are not supplied to the command/address pad ICAP, but are

rather fed to the chip select signal generation circuit CSG together with the output of the mode register MR and the bank addresses BA0 to BA2. The chip select signal generation circuit CSG is a circuit that activates any or all of the four chip select signal terminals CS supplied to the memory chips **110** to **113**.

[0072] The command system signals RAS, CAS, and WE are supplied to the command decoder CDC by way of the input buffers INB and latch circuits L. The output of the command decoder CDC is supplied to the chip select signal generation circuit CSG. The exact same signals as those fed to a general-purpose DRAM can be used as the command system signals RAS, CAS, and WE.

[0073] The chip select signal pad ECSP on the interface chip **120** is a single pad as described above. Therefore, the same chip select signal CS as used in a general-purpose DRAM may be supplied from outside of the semiconductor storage device **100** without consideration for the number of (four) stacked layers of memory chips **110** to **113**. The chip select signal CS supplied to the chip select signal pad ECSP is input to the command decoder CDC by way of the input buffer INB and latch circuits L. As described above, the chip select signal generation circuit CSG activates any or all of the four chip select signals CS on the basis of the address signals A14 and A15, the bank addresses BA0 to BA2, and the output of the command decoder CDC.

[0074] The data input/output pad EDQP on the interface chip **120** is a terminal for inputting and outputting the data signals DQ0 to DQ7, data strobe signal DQS, data mask signal DM, and other signals, and is connected to the data input/output pad IDQP by way of the bidirectional repeater REP, as described above.

[0075] In this case, when the memory chips **110** to **113** to be stacked are the double data rate (DDR-800) type, a signal with speed of 400 Mbps is transferred over the command/address external terminal group CA and the command/address bus CAB inside the package. On the other hand, a signal with speed of 800 Mbps is transferred over the on the data input/output external terminal group DQ and the data bus DQB inside the package.

[0076] Described next in greater detail is the circuit configuration of the interface chip **120** used in the present embodiment.

[0077] **FIG. 6** is a diagram showing in greater detail the circuit configuration of the clock system circuitry contained in the interface chip **120**.

[0078] The complimentary clock signals CLK and /CLK supplied from the command/address pad ECAP are input to a differential amp input buffer INB, as shown in **FIG. 6**. The input buffer INB is activated by the output of the differential amp input buffer INB that receives the clock enable signal CKE and the reference voltage VREF. The output of the input buffer INB that receives the complimentary clock signals CLK and /CLK is supplied to the clock regeneration circuit DLL, and the output of the regeneration circuit is supplied to the complimentary clock signals CLK and /CLK of the command/address pad ICAP by way of the output buffers OB. The clock regeneration circuit DLL generates an internal clock signal CLKI, and the internal clock signal CLKI is supplied to the internal circuitry of the interface chip **120**.

[0079] The output of the input buffer INB that receives the clock enable signal CKE is latched by the latch circuit L (which uses internal clock signal CLKI as a timing signal), and is supplied to the clock enable terminal CKE of the command/address pad ICAP by way of the output buffer OB. The clock enable signal CKE buffered by the input buffer INB is also supplied to the internal circuitry of interface chip **120**. The clockenable signal CKE is fed to the clock regeneration circuit DLL, and when the clock enable signal CKE is activated, the clock feed to the memory chips **110** to **113** is stopped.

[0080] Thus, in this present embodiment, since a clock regeneration circuit DLL is provided inside the interface chip **120**, the phase difference between the input and output clocks is reduced and the timing margin of the latch circuits L is increased. The operating margin can therefore be adequately assured even when the operating frequency is higher. When the clock enable signal CKE is deactivated, the operation of the clock regeneration circuit DLL is stopped, and since the clock feed to the memory chips **110** to **113** and the internal circuitry of the interface chip **120** is stopped, the power consumption during standby can be reduced.

[0081] **FIG. 7** is a diagram showing in greater detail the circuit configuration of the command/address system circuitry contained in the interface chip **120**.

[0082] The command/address system signals A0 to A15, BA0 to BA2, RAS, CAS, and WE are input from outside of the semiconductor storage device **100** to input buffers INB by way of the command/address pad ECAP, as shown in **FIG. 7**.

[0083] The chip select signal CS is supplied from outside of the semiconductor storage device **100** to an input buffer INB by way of the chip select signal pad ECSP. The input buffers INB are differential amp input buffers in which the reference voltage VREF is used, and the outputs of the buffers are supplied to corresponding latch circuits L. The latch circuits L synchronize with the internal clock signal CLKI generated by the clock regeneration circuit DLL (refer to **FIG. 6**), and accept the output of the corresponding input buffer INB.

[0084] The address signals A14 and A15 of the address signals A0 to A15 are decoded in the decoder DEC, and decoder output is supplied to the chip select signal generation circuit CSG. The bank addresses BA0 to BA2 are supplied to the command/address pad ICAP and are also supplied to the chip select signal generation circuit CSG and the mode register MR. The address signals A0 to A13 are supplied to the command/address pad ICAP and are also supplied to the mode register MR. The address signal A10 of the address signals A0 to A13 is also supplied to the command decoder CDC together with the command system signals RAS, CAS, and WE, as well as the chip select signal CS and clock enable signal CKE. The command decoder CDC decodes the signals in the same manner as a general-purpose DRAM, and generates a command signal CMD on the basis of the decoding result.

[0085] The command system signals RAS, CAS, and WE are directly supplied to the command/address pad ICAP, and an activate command (Bank activate), read command (Read) write command (Write), precharge command (Precharge), refresh command (Refresh), and other commands are presented to the memory chips **110** to **113**.

[0086] The chip select signal generation circuit CSG receives a command signal CMD fed by the command decoder CDC and determines the operating mode. Any or all of the four chip select signals CS0 to CS3 are activated based on the operating mode thus determined. The chip select signals CS0 to CS3 are supplied to the corresponding chip select signal pads ICSP, respectively, by way of output buffers OB.

[0087] Thus, in the present embodiment, since the higher bits A14 and A15 of the address signals are decoded and the chip select signals CS0 to CS3 are activated based on the decoded signals, the memory chip appears from outside of the semiconductor storage device **100** as a single, large memory chip, and only one of the four memory chips **110** to **113** is activated in ordinary operation inside the semiconductor storage device **100**. For this reason, an increase in power consumption can be held in check while assuring compatibility with general-purpose DRAM.

[0088] **FIG. 8** is a list showing the commands and symbols that are output by the command decoder CDC, and the methods for issuing a chip select signal CS when the commands are executed.

[0089] Commands are input to the command decoder CDC by using the clock system signal CKE, chip select signal CS, address signal A10, and command system signals RAS, CAS, and WE fed from outside the semiconductor storage device **100**, and the command decoder CDC generates a command signal CMD by using a combination of the signals. The chip select signal generation circuit CSG that receives the command signal CMD activates the chip select signals CS0 to CS3 as described below in accordance with the type of command.

[0090] First, when a mode set command "Mode register set (MRS)" or "Extended mode register set (EMRS)" classified as an "Initialization" command has been input, all the chip select signals CS0 to CS3 are simultaneously activated.

[0091] When a refresh command "Auto refresh (REF)" or "Self refresh entry (SELF)" classified as a "Row system" command has been input, the chip select signals CS0 to CS3 are sequentially activated with a phase difference. In other words, when synchronization is made with the clock signal CLK from outside the semiconductor storage device **100** and a refresh command (REF or SELF) is input, the interface chip **120** continuously issues a refresh command (REF or SELF) to the memory chips **110** to **113** by way of the command/address pad ICAP, and sequentially activates the chip select signals CS0 to CS3, as shown in **FIG. 9**.

[0092] The same refresh command (REF or SELF) is thereby sequentially input with a prescribed phase difference to the four memory chips **110** to **113**. The phase difference is a single clock cycle as shown in the example of **FIG. 9**, but no limitation is imposed thereby, and the phase difference may be a plurality of clock cycles.

[0093] Thus, the reason that refresh commands (REF or SELF) are sequentially issued with a phase difference to the memory chips **110** to **113** is that a greater amount of electric current temporarily flows during a refresh operation than during ordinary read and write operations. In other words, when refresh commands (REF and SELF) are simultaneously issued to all the memory chips **110** to **113**, the four memory chips **110** to **113** are simultaneously refreshed. As a result, a considerable amount of electric current flows, and the power supply potential and ground potential are liable to fluctuate.

[0094] In the case that return from self refresh is instructed, that is to say, in the case that a "Self refresh exit (SELFX)" command has been issued, all of the chip select signals CS0 to CS3 are simultaneously activated. Inputting a precharge command "Single bank precharge (PRE)" or "Precharge all banks (PALL)" activates the chip select signals CS0 to CS3, which correspond to the chip numbers (#0 to #3) stored in the active bank register ABR.

[0095] When an activation command "Back activate (ACT)" has been input, any one of the chip select signals CS0 to CS3 is activated based on the address signals A14 and A15, and the relationship between the activated chip select signal (CS0 to CS3) and the selected bank number (#0 to #7) is stored in the active bank register ABR shown in **FIG. 10A**. The active bank register ABR is a circuit disposed inside the chip select signal generation circuit CSG, and this register is made up of a group of registers that hold the relationship between the chip numbers (#0 to #3) specified by the higher two bits A14 and A15 of the row address, and the bank numbers (#0 to #7) specified by the bank addresses BA0 to BA2.

[0096] The reason for providing such an active bank register ABR is that since chip selection is carried out by the higher bits A14 and A15 of the row address A0 to A15 and only the bank address BA0 to BA2 and the column address A0 to A9 are fed when a column system command is input, it cannot be determined which chip select signals CS0 to CS3 should be activated when a column system command has been input if the above-described relationship is not stored using the active bank register ABR. However, provided that the bank address is expanded to five bits (BA0 to BA4) in lieu of the row address A14 and A15, that the chip is selected by using the higher two bits (BA0 and BA1), and that the bank inside the chip selected by the lower three bits (BA2 to BA4) is selected, such a problem is solved because the relationship between the bank address and the chip to be selected is fixed.

[0097] Nevertheless, in order to maintain compatibility with common general-purpose DRAM having an 8-bank configuration, the bank address is preferably given a three-bit configuration (BA0 to BA2) and a portion (A14 and A15) of the row address is used rather than the bank address to select a chip, as in the present embodiment. It is for this reason that an active bank register ABR is provided in the present embodiment.

[0098] When a write command "Write (WRIT)" or "Write with auto precharge (WRITA)" classified as a "column system" command, or a read command "Read (READ)" or "Read with auto precharge (READA)" classified as a "column system" command has been input, the bank address BA0 to BA2 is decoded in the decoder DEC, as shown in **FIG. 10B**, and the chip select signals CS0 to CS3, which correspond to the chip numbers (#0 to #3) stored in the active bank register ABR, are activated.

[0099] When a "No operation (NOP)" command, "Device deselect (DESL)" command, "Power down mode entry (PDEN)" command, or "Power down mode exit (PDEX)" command classified as "another" command has been input, all of the chip select signals CS0 to CS3 are simultaneously activated.

[0100] Thus, in the present embodiment, since the chip number to be activated is stored in the active bank register ABR when an activation command "Bank activate (ACT)" has been input, data can be accessed from outside the semiconductor storage device **100** by inputting to the interface chip **120** only the bank address BA0 to BA2 and the column address A0 to A9 when a read command or write command is input.

[0101] The same address input method as that used for a general-purpose DRAM can therefore be used and compatibility can be assured. Since commands are issued with a phase difference to the memory chips **110** to **113** when a refresh operation is carried out, the peak electric current that flows at the instant the memory chips **110** to **113** carry out a refresh operation can be staggered and power supply noise can be reduced.

[0102] **FIG. 11** is a diagram showing in greater detail the circuit configuration of the principal components of the data system circuitry contained in the interface chip **120**. In **FIG. 11**, the RDQS and /RDQS included in the signals of the data system are omitted.

[0103] The data system signals ODT, DQS, /DQS, DQ0 to DQ7, and DM are supplied from outside the semiconductor storage device **100** by way of the data input/output pad EDQP, as shown in **FIG. 11**. The on-chip resistance activation signal ODT is supplied to terminating resistors RT connected to the terminals, and the resistance value data RTV issued by the mode register is supplied to the terminating resistors RT in a shared configuration.

[0104] The data system signals have a higher frequency than the command/address system signals, and the signal waveform is therefore more easily distorted, but signal reflection can be reduced by carrying out termination inside the chip without any terminating resistance on the substrate. The distortion of the signal waveform can thereby be inhibited, resulting in an increased operating margin. In an ordinary DRAM, a terminating resistance is used inside the memory chip, but in the semiconductor storage device **100** of the present embodiment, a terminating resistance is used inside the interface chip **120**.

[0105] The data system signals DQS, /DQS, DQ0 to DQ7, and DM are supplied to the data input/output pad IDQP by way of a bidirectional repeater REP. More specifically, the data signals DQ0 to DQ7 and DM that are input from outside the semiconductor storage device **100** during a write operation are amplified by differential input buffers INB that make a comparison with the reference voltage VREF and are latched in the latch circuits L. The data strobe signals DQS and /DQS input from outside the semiconductor storage device **100** during a write operation are amplified by differential input buffers INB. When signals are output to the memory chips **110** to **113**, an internal strobe signal DQSI that is offset in phase by 90° with respect to the data signals DQ0 to DQ7 is applied. The internal strobe signal DQSI is generated by a strobe signal generation circuit DSG.

[0106] Therefore, the data strobe signal DQS on the data input/output pad EDQP is offset in phase by 90° with respect to the data in0, in1, in2, and in3 that are continuously input during a write operation, as shown in **FIG. 12**. Accordingly, the latch circuits L synchronize with the rise and fall of the data strobe signal DQS, and the data in0, in1, in2, and in3 can be sequentially received. Also, the data in0, in1, in2, and in3 sequentially received in this manner synchronize with the data strobe signal DQS on the data input/output pad IDQP, and are output to the memory chips **110** to **113** selected the chip select signals CS0 to CS3.

[0107] More specifically, synchronization is made with the rising edge of the clock signal CLK during a write operation, as shown in **FIG. 12**; a write command (WRIT) is input to the command/address pad ECAP; and after a prescribed write latency WL has elapsed, the data signals DQ0 to DQ7 are input to the data input/output pad EDQP. Only a single data DQ0 of the data signals DQ0 to DQ7 is shown in **FIG. 12**. The data strobe signal DQS (which is offset in phase by 90°) is also input to the data input/output pad EDQP simultaneously with the data signals DQ0 to DQ7 to be written.

[0108] A write command (WRIT) is supplied to the interface chip **120**, after which synchronization is made with the subsequent falling clock signal CLK, and the same command (WRIT) is output to the memory chips **110** to **113** by way of the command/address pad ICAP. The data signals DQ0 to DQ7 and data strobe signal DQS are also output from the data input/output pad IDQP after a prescribed write latency WL has elapsed. The write latency WL is a value that is set in advance in the mode register on the memory chips **110** to **113** and on the interface chip **120**. In the case of a write operation, the write latency WL is set to the same value in the interface chip **120** and the memory chips **110** to **113**.

[0109] The data signals DQ0 to DQ7 output from the memory chips **110** to **113** during a read operation synchronize with the internal strobe signal DQSI that is generated by the strobe signal generation circuit DSG inside the interface chip **120**, are received by the latch circuits L, and are output to the outside of the semiconductor storage device **100**. Since the data strobe signal DQS fed from the memory chips **110** to **113** during a read operation match the phase of the data signals DQ0 to DQ7, as shown in **FIG. 13**, the data signals DQ0 to DQ7 cannot be directly received in synchronization with the data strobe signal DQS.

[0110] In view of the above, an internal strobe signal DQSI that is offset in phase by 90° from the data strobe signal DQS is generated by the strobe signal generation circuit DSG, and the read data is received in synchronization with the internal strobe signal. The data signals DQ0 to DQ7 thus received are synchronized with the internal clock signal CLK1 (not shown) and output to an impedance variable output buffer OBV. The impedance of the impedance variable output buffer OBV is controlled by the impedance setting value ROV issued from the mode register MR.

[0111] More specifically, the read command (READ) is input to the command/address pad ECAP of the interface chip **120** at the rising edge of the clock signal CLK during a read operation, as shown in **FIG. 13**. The read command (READ) is thereafter output to the memory chips **110** to **113** at the subsequent falling edge of the clock signal CLK by way of the command/address pad ICAP. After a prescribed read latency RL has elapsed, data signals DQ0 to DQ7 and DQS are output from the memory chips **110** to **113**.

[0112] The interface chip **120** receives the data signals by way of the data input/output pad IDQP, synchronizes the signals with the subsequent clock signal CLK, and outputs

the signals to the outside of the semiconductor storage device **100** by way of the data input/output pad EDQP. Since the data strobe signal DQS matches the phase of the data signals DQ0 to DQ7 during a read operation, the signal is directly matched with the clock and is output from the data input/output pad EDQP.

[0113] As viewed from outside of the semiconductor storage device **100**, the overall read latency is a value in which a single clock cycle has been added (which is the latency of the interface chip **120** itself) to the read latency RL of the memory chips **110** to **113**, that is to say, a value that is equal to RL+1, as shown in **FIG. 13**. Therefore, the value obtained by adding the latency of the interface chip **120** as such to the value of the mode register of the memory chips **110** to **113** is set as the read latency RL in the mode register MR on the interface chip **120**.

[0114] Thus, the present embodiment involves receiving data that uses an internal strobe signal DQSI generated by the strobe signal generation circuit DSG, allowing an adequate operating margin to be assured even when the operating frequency is high. Since the value of the terminating resistance and the impedance of the output buffer canbe controlled, distortion of the data transfer waveform is controlled, and, as a result, the operating margin can be expanded.

[0115] Next, the signal amplitude between the memory chips **110** to **113** and the interface chip **120** is described.

[0116] **FIG. 14** is a schematic diagram that describes the method for reducing the amplitude of the inter-chip interface between the memory chips **110** to **113** and the interface chip **120**. As an example, the case is shown in which the amplitude of the chip select signals CS0 to CS3 is reduced, but the same method may be adopted for the other signals as well.

[0117] In the present example, in the output buffer OB contained in the interface chip **120**, an inverter with a CMOS structure is used as the output buffer for providing output to the memory chips **110** to **113**, as shown in **FIG. 14**. An internal power supply voltage VDDI on the high level side of the inverter is kept at a lower level than the externally supplied power supply VDDQ, and the internal power supply voltage VSSI on the low level side is kept at a higher level than the externally supplied power supply VSSQ. If, for example, the VDDQ is 1.8 V, the VDDI may be set to about 1.1 V. Also, if the VSSQ is 0 V, the VSSI may be set to about 0.7 V.

[0118] When a signal (the chip select signal CS in **FIG. 14**) to be output is input to the input terminal of an inverter in which such internal power supply voltages VDDI and VSSI are used, the signal output has reduced amplitude. In this case, the threshold value of the NMOS and PMOS that constitute the inverter is preferably set to be lower than that of an ordinary device in order to increase speed. When set in this fashion, the amplitude of the data signals can be reduced without providing a terminating resistance in the wiring.

[0119] The internal power supply voltages VDDI and VSSI can be generated by a step-down circuit VRG. The step-down circuit VRG is composed of resistors R1 to R3 that produce reference voltages V1 and V2; differential amps AP1 and AP2 in which the reference voltages V1 and V2 are supplied to a non-inverted input terminal (+) and an inverted input terminal (−) respectively; and a PMOS and NMOS in which the outputs of the differential amps AP1 and AP2 are fed to the gates, respectively. The drains of the PMOS and NMOS are connected to the inverted input terminal (−) of the differential amp AP1 and the non-inverted input terminal (+) of the differential amp AP2, respectively; and the electric potentials are the internal power supply voltages VDDI and VSSI, respectively.

[0120] If such a low-amplitude output buffer is used, the charge and discharge power can be considerably reduced by charging and discharging the parasitic capacitance of the wiring INW and the bonding pad when a command/address or data is transferred.

[0121] Another feature of in the present embodiment is that since four memory chips **110** to **113** are mounted in a single package, there may be a problem in that the temperature inside the package may increase due to the power consumption of the memory chips **110** to **113**, but if the power consumption due to signal transfer between the chips is reduced by using a low-amplitude output buffer, the temperature increase inside the package can be inhibited, and, as a result, the data retention time of the memory chips **110** to **113** can be extended.

[0122] The low-amplitude output buffer may be used not only on the interface chip **120** side, but also on the memory chips **110** to **113** side.

[0123] As described above, the semiconductor storage device **100** of the present embodiment has a configuration in which a plurality of memory chips **110** to **113** that are capable of independently carrying out read and write operations are stacked, and the plurality of memory chips **110** to **113** are externally connected by way of an interface chip **120**. For this reason, the chip connected to the command/address external terminal group CA and the data input/output external terminal group DQ is a single chip (interface chip **120**), and parasitic capacitance of the external terminals can be reduced to an amount that is about equal to that of a single chip component.

[0124] The frequency of the command/address and data input/output can be increased thereby to an amount that is equal to about that of a single chip component. Additionally, since the memory chips **110** to **113** are three-dimensionally stacked, the mounting surface area can be reduced to an amount equal to a single chip component.

[0125] As viewed from external circuitry, the interface chip **120** can be handled as an ordinary general-purpose DRAM, making it possible to assure compatibility with ordinary semiconductor storage devices that are packaged in only a single chip. For this reason, there is no need to design a special-purpose memory controller or the like, and an increase in the cost of the system overall can be held in check.

[0126] Since the stacked memory chips **110** to **113** are ordinary memory chips (general-purpose DRAM) that are each independently capable of read and write operations, a special step for forming a through-electrode or carrying out other processes is not required. For this reason, it is also possible to hold down increases in manufacturing costs.

[0127] In the present embodiment, the command/address signals and data input/output signals are connected in the

form of a bus inside the package, making it possible to keep the number of wires inside the package low and thereby to hold manufacturing costs in check. Since any one of the chip select signals CS0 to CS3 is activated by decoding the higher bits A14 and A15 of the row address during read and write operations, the interface chip **120** can selectively activate the memory chips **110** to **113**, and the overall power consumption of the stacked package can be reduced.

[0128] In the embodiment described above, common, general-use DRAM is used for the memory chips **110** to **113**, but the data transfer speed between the memory chips **110** to **113** and the interface chip **120** can be further increased by separating the pads for normal use and the pads for testing as the pads of the memory chips **110** to **113**. This method is described in **FIG. 15**.

[0129] **FIG. 15** is a diagram for describing an improved method for connecting the pads and the memory array MA in the memory chips **110** to **113**. For the sake of simplicity, only the command/address pad CAP and the data input/output pad DQP are depicted, but the chip select signal pad CSP may also be handled in the same manner as the command/address pad CAP.

[0130] In the present example, in addition to being provided with the command/address pad CAP and the data input/output pad DQP for normal use, the memory chips **110** to **113** are also provided with a test command/address pad TCAP and a test data input/output pad TDQP, as shown in **FIG. 15**. Although not depicted, a test pad that corresponds to the chip select signal pad CSP is also provided.

[0131] The test command/address pad TCAP has an electrostatic protection element ESD, an input buffer INB, and a latch circuit L. The test data input/output pad TDQP has an output buffer OB, an input buffer INB, and a latch circuit L. In contrast to this configuration, the command/address pad CAP and data input/output pad DQP for normal use are not connected to such circuitry and are essentially directly connected to the memory array MA, as shown in **FIG. 15**.

[0132] The test command/address pad TCAP and the test data input/output pad TDQP are pads that are used when the memory chips **110** to **113** are tested. More specifically, since the parasitic capacitance of the wiring is considerable, a tester used for testing the memory chips **110** to **113** requires an output buffer with a relatively high driving force, an input buffer with high sensitivity, and an electrostatic protection element.

[0133] In contrast, the chip connections are made with very short distance wires in terms of practical use, and since the parasitic capacitance is very low in comparison with the tester, the driving force of the output buffer OB may be fairly low, and there is no requirement that an input buffer INB with such a high sensitivity be used. Also, the necessity to provide an electrostatic protection element ESD is also low. Considering this point, in the present example, the electrostatic protection element ESD, the input buffer INB, and the like are bypassed, and the pads and memory array MA are essentially directly connected.

[0134] Testing can thereby be correctly carried out during the testing procedure, the data transfer speed between the memory chips **110** to **113** and the interface chip **120** can be increased during actual use, and power consumption can be reduced.

[0135] A second preferred embodiment of the present invention is described next.

[0136] The semiconductor storage device **200** of the second embodiment of the present invention is exactly the same as the semiconductor storage device **100** of the first embodiment described above in terms of the external appearance and the memory chips **110** to **113** that are used, but are different in terms of the interface chip that is used and the manner in which the memory chips **110** to **113** and the interface chip are connected. For this reason, a duplicative description of the portions that are the same as the first embodiment is omitted.

[0137] **FIG. 16** is a schematic diagram of the connection configuration between the memory chips **110** to **113** and the interface chip **220** of the present embodiment.

[0138] In the present embodiment, all of the signals are connected in a one-to-one relationship between the memory chips **110** to **113** and the interface chip **220**, as shown in **FIG. 16**. In other words, not only are the chip select signals connected in a one-to-one relationship, but the command/address signals and data signals are also connected in such a relationship. The power supply is bus connected. Therefore, in the present embodiment, all of the signals transferred between the memory chips **110** to **113** and the interface chip **220** are individual signals for each memory chips **110** to **113**.

[0139] Described more specifically, the command/address external terminal group CA and chip select signal external terminal CS are connected, respectively, to the command/address pad ECAP and the chip select signal pad ECSP on the interface chip **220** by way of the ball electrodes EXB on the base substrate **101**. These signals are input to the command decoder CDC, and the four outputs of the decoder are output, respectively, to the command/address pads ICAP0 to ICAP3 and chip select signal pads ICSP0 to ICSP3, which are individually assigned to the memory chips **110** to **113**.

[0140] The signals (command and address) output to the command/address pads ICAP0 to ICAP3 are individually fed to the command/address pad CAP of the four memory chips **110** to **113**. The four chip select signals that are output to the chip select signal pads ICSP0 to ICSP3 are also connected in a one-to-one relationship between the interface chip **220** and memory chips **110** to **113**, respectively.

[0141] The data input/output external terminal group DQ is connected to the data input/output pad EDQP on the interface chip **220** by way of the ball electrodes EXB on the base substrate **101**. The data input/output signals are input to a bidirectional multiplexer MUX, and are connected to the data input/output pad DQP of the memory chips **110** to **113**, respectively, by way of the data input/output pads IDQP0 to IDQP3 that are individually assigned to the memory chips **110** to **113**.

[0142] To be written to the memory chips **110** to **113**, the data is input to the interface chip **220** from outside the semiconductor storage device **100** by way of the data input/output external terminal group DQ, and write data is sent to only one of the memory chips **110** to **113** selected by way of the bidirectional multiplexer MUX. When data is read from the memory chips **110** to **113**, the read data from the selected chip of the memory chips **110** to **113** is input to

the interface chip **220** and is output to the data input/output external terminal group DQ by way of the bidirectional multiplexer.

[0143] The power supply external terminal group VS is connected directly to the power supply pads VSP of the memory chips **110** to **113** and the interface chip **220** by way of the ball electrodes EXB on the base substrate **101**.

[0144] Thus, in the present embodiment, since the command/address signals and data input/output signals are connected in a one-to-one relationship inside the package, only the input/output circuits of two transmitting and receiving channels are connected to the wiring INW inside the package. For this reason, the parasitic capacitance of the wiring INW can be reduced, and the charging and discharging electric current of the parasitic capacitance (which is a problem when high speed signals are transmitted) can be reduced.

[0145] **FIG. 17** is a simplified circuit diagram showing the circuit configuration of the interface chip **220** used in the present embodiment.

[0146] The interface chip **220** used in the present embodiment, as shown **FIG. 17**, is different from the interface chip **120** shown in **FIG. 5** in that command/address pads ICAP are individually provided to the memory chips **110** to **113** (ICAP**0** to ICAP**3**), output buffers OB are individually provided in a corresponding relationship, data input/output pads IDQP are individually provided to the memory chips **110** to **113** (IDQP0 to IDQP**3**), and bidirectional repeaters are substituted for the bidirectional multiplexer MUX in a corresponding relationship. Other than the above, the configuration is the same as the interface chip **120** shown in **FIG. 5**.

[0147] **FIG. 18** is a diagram showing in greater detail the circuit configuration of the clock system circuit contained in the interface chip **220**.

[0148] The clock system circuit contained in the interface chip **220**, as shown in **FIG. 18**, is different than the clock system circuit of the interface chip **120** shown in **FIG. 6** in that the pads for feeding the complimentary clock signals CLK and /CLK and the clock enable signal CKE are individually provided to the memory chips **110** to **113** (ICAPO to ICAP**3**), and output buffers OB are independently provided in a corresponding relationship. Other than the above, the configuration is the same as the clock system circuit shown in **FIG. 6**.

[0149] **FIG. 19** is a diagram showing in greater detail the circuit configuration of the command/address system circuit contained in the interface chip **220**.

[0150] The command/address system circuit contained in the interface chip **220**, as shown in **FIG. 19**, is different than the command/address system circuit of the interface chip **120** shown in **FIG. 7** in that command/address pads ICAP are individually provided to the memory chips **110** to **113** (ICAP**0** to ICAP**3**), output buffers OB are individually provided in a corresponding relationship, and selectors SEL for selectively activating the output buffers are also provided. Chip select signals CS**0** to CS**3** are fed to the selectors SEL, and the selectors SEL feed a signal only to the output buffers OB that correspond to the activated chip select signals CS**0** to CS**3**. Therefore, commands and addresses are not fed to the unselected memory chips **110** to **113**.

[0151] **FIG. 20** is a diagram showing in greater detail the circuit configuration of the data system circuit contained in the interface chip **220**.

[0152] The data system circuit contained in the interface chip **220**, as shown in **FIG. 20**, is different than the data system circuit of the interface chip **120** shown in **FIG. 11** in that data input/output pads IDQP are individually provided to the memory chips **110** to **113** (IDQPO to IDQP**3**), input buffers INB and output buffers OB are individually provided in a corresponding relationship, and selectors SEL that selectively activate the output buffers are also provided.

[0153] In the same manner as the command/address system circuit, chip select signals CS**0** to CS**3** are fed to the selectors SEL, the selectors SEL feed a signal only to the output buffers OB that correspond to the activated chip select signals CS**0** to CS**3**, and only the output of the input buffers INB corresponding to the activated chip select signals CS**0** to CS**3** are allowed to pass. Therefore, data is not fed to the unselected memory chips **110** to **113**.

[0154] The semiconductor storage device **200** of the present embodiment can achieve substantially the same effects as the semiconductor storage device **100** of the first embodiment described above, and since the command/address signals and data input/output signals are connected in a one-to-one relationship inside the package, the parasitic capacitance of the wiring INW inside the package can be reduced. The charging and discharging electric current of the parasitic capacitance (which is a problem when high speed signals are transmitted) can thereby be reduced.

[0155] In the present embodiment, both the command/address signals and data input/output signals are connected in a one-to-one relationship, but since the speed of the command/address signals is half the speed of the data input/output signals in a DDR-type DRAM, power consumption does not always become a large problem for command/address signals. Considering this point, the command/address signals are bus-connected in the same manner as in the first embodiment, and the data input/output signals may be connected in a one-to-one relationship. According to this configuration, it is possible to reduce power consumption while keeping the number of wires inside the chip low.

[0156] A third preferred embodiment of the present invention is described next.

[0157] **FIG. 21** is a schematic diagram that describes the connection configuration between the memory chips **110** to **112** and the memory chip **320** in the semiconductor storage device **300** of the third embodiment of the present invention.

[0158] In the present embodiment, the special-purpose interface chip is dispensed with. Instead, an interface chip function is provided to the memory chip **320** that is most proximate to the base substrate **101**, as shown in **FIG. 21**. Other than the above, the configuration is the same as the semiconductor storage device **100** of the first embodiment described above. Signals other than the chip select signals CS**0** to CS**2** are bus-connected using the command/address bus CAB and the data bus DQB between the memory chips **110** to **112** and the memory chip **320**. According to the present invention, the number of chips can be reduced by

one, and the thickness of the entire package can therefore be reduced. Also, since the number of wires and the load of the internal wiring INW can be reduced, power consumption can be further lowered.

[0159] In the present embodiment, both the command/address signals and the data input/output signals are bus-connected in the same manner as the semiconductor storage device **100** of the first embodiment, although it is also possible to connect both the command/address signals and the data input/output signals in a one-to-one relationship as in the semiconductor storage device **200** of the second embodiment. Another option is to bus-connect the command/address signals as in the first embodiment, and to connect the data input/output signals in a one-to-one relationship.

[0160] The present invention is in no way limited to the aforementioned embodiments, but rather various modifications are possible within the scope of the invention as recited in the claims, and naturally these modifications are included within the scope of the invention.

[0161] In the embodiments described above, four memory chips were used as an example, but the number of memory chips that may be used is not limited to four, and any number of memory chips may be used as long as this number is two or more.

[0162] In the embodiments described above, DRAM was used as the memory chip, but the present invention is not limited to this type of chip alone, and other types of memory chips may also be used as long as the memory chips are each independently capable of read operations and write operations.

What is claimed is:

1. A semiconductor storage device comprising:

a base substrate having a command/address external terminal group to which command signals and address signals are supplied, a data input/output external terminal group for inputting and outputting data signals, and a single chip select external terminal; and

a plurality of memory chips that are stacked on the base substrate and are each individually capable of reading and writing operations, wherein

a plurality of terminals constituting said command/address external terminal group, a plurality of terminals constituting said data input/output external terminal group, and said single chip select external terminal are connected to a single chip that has an interface function; and

the single chip that has the interface function further has at least a chip select signal generation circuit that can individually activate said plurality of memory chips based on said address signals supplied via said command/address external terminal group and based on said chip select signal supplied via said chip select external terminal.

2. The semiconductor storage device as claimed in claim 1, wherein said single chip that has the interface function is an interface chip that is different from said plurality of memory chips.

3. The semiconductor storage device as claimed in claim 1, wherein said single chip that has the interface function is any one chip from among said plurality of memory chips.

4. The semiconductor storage device as claimed in claim 1, wherein at least one of the signals selected from said command signals, said address signals, and said data signals are transmitted between said plurality of memory chips and said single chip that has the interface function by way of wiring with a shared connection to said plurality of memory chips.

5. The semiconductor storage device as claimed in claim 1, wherein at least one of the signals selected from said command signals, said address signals, and said data signals are transmitted between said plurality of memory chips and said single chip that has the interface function by way of wiring that is individually connected to the plurality of memory chips.

6. The semiconductor storage device as claimed in claim 1, wherein said chip select signal generation circuit activates one of said plurality of memory chips based on a portion of said address signals supplied via said command/address external terminal group which is different from a bank address when an activation command is issued.

7. The semiconductor storage device as claimed in claim 6, wherein said chip select signal generation circuit has an active bank register for storing the relationship between the selected bank and the memory chip that is to be activated, and activates one of said plurality of memory chips by referencing said active bank register when a read command or a write command is issued together with said bank address via said command/address external terminal group.

8. The semiconductor storage device as claimed in claim 1, wherein said chip select signal generation circuit sequentially activates said plurality of memory chips when a refresh command is issued.

9. The semiconductor storage device as claimed in claim 1, wherein said single chip that has the interface function further has a clock regeneration circuit, and a clock regenerated by said clock regeneration circuit is supplied to said plurality of memory chips.

10. The semiconductor storage device as claimed in claim 1, wherein said single chip that has the interface function further has a strobe signal generation circuit, and said chip that has the interface function outputs a data signal that is read from said memory chip via said data input/output external terminal group in synchronism with an output of said strobe signal generation circuit.

11. The semiconductor storage device as claimed in claim 1, wherein said single chip that has the interface function has a terminating resistance that is connected to said data input/output external terminal group.

12. The semiconductor storage device as claimed in claim 1, wherein each of said plurality ofmemory chips has a bonding pad area, and said plurality of memory chips are connected to said base substrate by way of tape on which wiring is formed.

13. The semiconductor storage device as claimed in claim 1, wherein a first internal power supply voltage that is lower than an externally supplied high-level-side power supply voltage is applied as a high-level-side power supply voltage, and a second internal power supply voltage that is higher than the externally supplied low-level-side power supply voltage is applied as a low-level-side power supply voltage to an output buffer that is designed to provide output to said

plurality of memory chips and is selected from the output buffers included in said single chip that has the interface function.

14. The semiconductor storage device as claimed in claim 1, wherein said plurality of memory chips have a test pad in addition to an ordinary pad designed for actual use, said test pad is connected to a memory array at least via a buffer circuit, and said ordinary pad bypasses said buffer circuit and connects to said memory array.

15. A semiconductor storage device comprising:

a base substrate having a command/address external terminal group to which command signals and address signals are supplied and a single chip select external terminal to which a chip select signal is supplied;

an interface chip that is mounted on the base substrate and is connected to the command/address external terminal group and the chip select external terminal; and

a plurality of memory chips that are stacked on the base substrate and are each individually capable of reading and writing operations,

wherein said interface chip can individually activate said plurality of memory chips based on said address signals and said chip select signal.

16. The semiconductor storage device as claimed in claim 15, wherein at least one of the signals selected from said command signals and said address signals are transmitted between said plurality of memory chips and said interface chip via wiring with a shared connection to said plurality of memory chips.

17. The semiconductor storage device as claimed in claim 15, wherein at least one of the signals selected from said command signals and said address signals are transmitted between said plurality of memory chips and said interface chip via wiring that is individually connected to the plurality of memory chips.

18. A semiconductor storage device comprising:

a base substrate having a command/address external terminal group to which command signals and address signals are supplied and a single chip select external terminal to which a chip select signal is supplied; and

a plurality of memory chips that are stacked on the base substrate and are each individually capable of reading and writing operations,

wherein one of said plurality of memory chips is connected to the command/address external terminal group and the chip select external terminal and can individually activate said plurality of memory chips based on said address signals and said chip select signal.

19. The semiconductor storage device as claimed in claim 18, wherein at least one of the signals selected from said command signals and said address signals are transmitted between said one of said plurality of memory chips and a remaining memory chips via wiring with a shared connection to said remaining memory chips.

20. The semiconductor storage device as claimed in claim 18, wherein at least one of the signals selected from said command signals and said address signals are transmitted between said one of said plurality of memory chips and a remaining memory chips via wiring that is individually connected to the remaining memory chips.

* * * * *

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 816 570 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
08.08.2007 Bulletin 2007/32

(51) Int Cl.:
G06F 13/16 *(2006.01)*

(21) Application number: **06125954.5**

(22) Date of filing: **16.04.1991**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK YU

(30) Priority: **18.04.1990 US 510898**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
00100018.1 / 1 022 641
91908374.1 / 0 525 068

(71) Applicant: **RAMBUS INC.**
Los Altos,
California 94022 (US)

(72) Inventors:
- **Farmwald, Michael**
  Portola Valley, CA 94028 (US)
- **Horowitz, Mark**
  Palo Alto, CA 94306 (US)

(74) Representative: **Eisenführ, Speiser & Partner**
Patentanwälte Rechtsanwälte
Postfach 10 60 78
28060 Bremen (DE)

Remarks:
This application was filed on 12 - 12 - 2006 as a divisional application to the application mentioned under INID code 62.

(54) Integrated circuit I/O using a high performance bus interface

(57) The present invention includes a semiconductor device, the semiconductor device comprising connection means adapted to connect the semiconductor device to an external bus which is a part of a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to the external bus, wherein the external bus includes a plurality of bus lines for carrying substantially all address, data, control and device select information needed by the semiconductor device for communication with substantially every other semiconductor device connected to the external bus, and wherein the external bus has substantially fewer bus lines than the number of bits in a single address, wherein each said bus line is a terminated transmission line, and wherein the external bus carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to the semiconductor device, a bus line driver capable of producing a low-voltage-swing signal on one of said terminated transmission lines, wherein the semiconductor device is a semiconductor transceiver device to interface multiple units to the bus, wherein the multiple units are DRAMs.

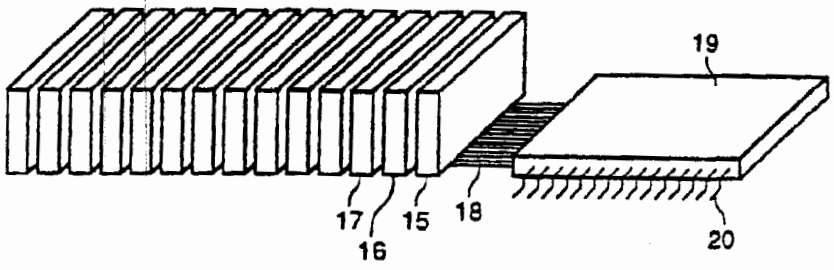


FIG. 3

EP 1 816 570 A2

Printed by Jouve, 75001 PARIS (FR)

Description

FIELD OF THE INVENTION

[0001] An integrated circuit bus interface for computer and video systems is described which allows high speed transfer of blocks of data, particularly to and from memory devices, with reduced power consumption and increased system reliability. A new method of physically implementing the bus architecture is also described.

BACKGROUND OF THE INVENTION

[0002] Semiconductor computer memories have traditionally been designed and structured to use one memory device for each bit, or small group of bits, of any individual computer word, where the word size is governed by the choice of computer. Typical word sizes range from 4 to 64 bits. Each memory device typically is connected in parallel to a series of address lines and connected to one of a series of data lines. When the computer seeks to read from or write to a specific memory location, an address is put on the address lines and some or all of the memory devices are activated using a separate device select line for each needed device. One or more devices may be connected to each data line but typically only a small number of data lines are connected to a single memory device. Thus data line 0 is connected to device(s) 0, data line 1 is connected to device(s) 1, and so on. Data is thus accessed or provided in parallel for each memory read or write operation. For the system to operate properly, every single memory bit in every memory device must operate dependably and correctly.

[0003] To understand the concept of the present invention, it is helpful to review the architecture of conventional memory devices. Internal to nearly all types of memory devices (including the most widely used Dynamic Random Access Memory (DRAM), Static RAM (SRAM) and Read Only Memory (ROM) devices), a large number of bits are accessed in parallel each time the system carries out a memory access cycle. However, only a small percentage of accessed bits which are available internally each time the memory device is cycled ever make it across the device boundary to the external world.

[0004] Referring to Fig. 1, all modern DRAM, SRAM and ROM designs have internal architectures with row (word) lines 5 and column (bit) lines 6 to allow the memory cells to tile a two dimensional area 1. One bit of data is stored at the intersection of each word and bit line. When a particular word line is enabled, all of the corresponding data bits are transferred onto the bit lines. Some prior art DRAMs take advantage of this organization to reduce the number of pins needed to transmit the address. The address of a given memory cell is split into two addresses, row and column, each of which can be multiplexed over a bus only half as wide as the memory cell address of the prior art would have required.

COMPARISON WITH PRIOR ART

[0005] Prior art memory systems have attempted to solve the problem of high speed access to memory with limited success. U.S. Patent No. 3,821,715 (Hoff et. al.), was issued to Intel Corporation for the earliest 4-bit micro-processor. That patent describes a bus connecting a single central processing unit (CPU) with multiple RAMs and ROMs. That bus multiplexes addresses and data over a 4-bit wide bus and uses point-to-point control signals to select particular RAMs or ROMs. The access time is fixed and only a single processing element is permitted. There is no block-mode type of operation, and most important, not all of the interface signals between the devices are bused (the ROM and RAM control lines and the RAM select lines are point-to-point).

[0006] In U.S. Patent No. 4,315,308 (Jackson), a bus connecting a single CPU to a bus interface unit is described. The invention uses multiplexed address, data, and control information over a single 16-bit wide bus. Block-mode operations are defined, with the length of the block sent as part of the control sequence. In addition, variable access-time operations using a "stretch" cycle signal are provided. There are no multiple processing elements and no capability for multiple outstanding requests, and again, not all of the interface signals are bused.

[0007] In U.S. Patent No. 4,449,207 (Kung, et. al.), a DRAM is described which multiplexes address and data on an internal bus. The external interface to this DRAM is conventional, with separate control, address and data connections.

[0008] In U.S. Patent Nos. 4,764,846 and 4,706,166 (Go), a 3-D package arrangement of stacked die with connections along a single edge is described. Such packages are difficult to use because of the point-to-point wiring required to interconnect conventional memory devices with processing elements. Both patents describe complex schemes for solving these problems. No attempt is made to solve the problem by changing the interface.

[0009] In U.S. Patent No. 3,969,706 (Proebsting, et. al.), the current state-of-the-art DRAM interface is described. The address is two-way multiplexed, and there are separate pins for data and control (RAS, CAS, WE, CS). The number of pins grows with the size of the DRAM, and many of the connections must be made point-to-point in a memory system using such DRAMs.

[0010] There are many backplane buses described in the prior art, but not in the combination described or having the

Petitioners
Ex. 1002, p. 315

features of this invention. Many backplane buses multiplex addresses and data on a single bus (e.g., the NU bus). ELXSI and others have implemented split-transaction buses (U.S. Patent No. 4,595,923 and 4,481,625 (Roberts)). ELXSI has also implemented a relatively low-voltage-swing current-mode ECL driver (approximately 1 V swing). Address-space registers are implemented on most backplane buses, as is some form of block mode operation.

[0011] Nearly all modern backplane buses implement some type of arbitration scheme, but the arbitration scheme used in this invention differs from each of these. U.S. Patent Nos. 4,837,682 (Culler), 4,818,985 (Ikeda), 4,779,089 (Theus) and 4,745,548 (Blahut) describe prior art schemes. All involve either log N extra signals, (Theus, Blahut), where N is the number of potential bus requestors, or additional delay to get control of the bus (Ikeda, Culler). None of the buses described in patents or other literature use only bused connections. All contain some point-to-point connections on the backplane. None of the other aspects of this invention such as power reduction by fetching each data block from a single device or compact and low-cost 3-D packaging even apply to backplane buses.

[0012] The clocking scheme used in this invention has not been used before and in fact would be difficult to implement in backplane buses due to the signal degradation caused by connector stubs. U.S. Patent No. 4,247,817 (Heller) describes a clocking scheme using two clock lines, but relies on ramp-shaped clock signals in contrast to the normal rise-time signals used in the present invention.

[0013] In U.S. Patent No. 4,646,279 (Voss), a video RAM is described which implements a parallel-load, serial-out shift register on the output of a DRAM. This generally allows greatly improved bandwidth (and has been extended to 2, 4 and greater width shift-out paths.) The rest of the interfaces to the DRAM (RAS, CAS, multiplexed address, etc.) remain the same as for conventional DRAMS.

[0014] One object of the present invention is to use a new bus interface built into semiconductor devices to support high-speed access to large blocks of data from a single memory device by an external user of the data, such as a microprocessor, in an efficient and cost-effective manner.

[0015] Another object of this invention is to provide a clocking scheme to permit high speed clock signals to be sent along the bus with minimal clock skew between devices.

[0016] Another object of this invention is to allow mapping out defective memory devices or portions of memory devices.

[0017] Another object of this invention is to provide a method for distinguishing otherwise identical devices by assigning a unique identifier to each device.

[0018] Yet another object of this invention is to provide a method for transferring address, data and control information over a relatively narrow bus and to provide a method of bus arbitration when multiple devices seek to use the bus simultaneously.

[0019] Another object of this invention is to provide a method of distributing a high-speed memory cache within the DRAM chips of a memory system which is much more effective than previous cache methods.

[0020] Another object of this invention is to provide devices, especially DRAMs, suitable for use with the bus architecture of the invention.

SUMMARY OF INVENTION

[0021] The present invention includes a memory subsystem comprising at least two semiconductor devices, including at least one memory device, connected in parallel to a bus, where the bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices, where the control information includes device-select information and the bus has substantially fewer bus lines than the number of bits in a single address, and the bus carries device-select information without the need for separate device-select lines connected directly to individual devices.

[0022] Referring to Fig. 2, a standard DRAM 13, 14, ROM (or SRAM) 12, microprocessor CPU 11, I/O device, disk controller or other special purpose device such as a high speed switch is modified to use a wholly bus-based interface rather than the prior art combination of point-to-point and bus-based wiring used with conventional versions of these devices. The new bus includes clock signals, power and multiplexed address, data and control signals. In a preferred implementation, 8 bus data lines and an AddressValid bus line carry address, data and control information for memory addresses up to 40 bits wide. Persons skilled in the art will recognize that 16 bus data lines or other numbers of bus data lines can be used to implement the teaching of this invention. The new bus is used to connect elements such as memory, peripheral, switch and processing units.

[0023] In the system of this invention, DRAMs and other devices receive address and control information over the bus and transmit or receive requested data over the same bus. Each memory device contains only a single bus interface with no other signal pins. Other devices that may be included in the system can connect to the bus and other non-bus lines, such as input/output lines. The bus supports large data block transfers and split transactions to allow a user to achieve high bus utilization. This ability to rapidly read or write a large block of data to one single device at a time is an important advantage of this invention.

[0024] The DRAMs that connect to this bus differ from conventional DRAMs in a number of ways. Registers are

Petitioners
Ex. 1002, p. 316

provided which may store control information, device identification, device-type and other information appropriate for the chip such as the address range for each independent portion of the device. New bus interface circuits must be added and the internals of prior art DRAM devices need to be modified so they can provide and accept data to and from the bus at the peak data rate of the bus. This requires changes to the column access circuitry in the DRAM, with only a minimal increase in die size. A circuit is provided to generate a low skew internal device clock for devices on the bus, and other circuits provide for demultiplexing input and multiplexing output signals.

[0025] High bus bandwidth is achieved by running the bus at a very high clock rate (hundreds of MHz). This high clock rate is made possible by the constrained environment of the bus. The bus lines are controlled-impedance, doubly-terminated lines. For a data rate of 500 MHz, the maximum bus propagation time is less than 1 ns (the physical bus length is about 10 cm). In addition, because of the packaging used, the pitch of the pins can be very close to the pitch of the pads. The loading on the bus resulting from the individual devices is very small. In a preferred implementation, this generally allows stub capacitances of 1-2 pF and inductances of 0.5 - 2 nH. Each device 15, 16, 17, shown in Figure 3, only has pins on one side and these pins connect directly to the bus 18. A transceiver device 19 can be included to interface multiple units to a higher order bus through pins 20.

[0026] A primary result of the architecture of this invention is to increase the bandwidth of DRAM access. The invention also reduces manufacturing and production costs, power consumption, and increases packing density and system reliability.

BRIEF DESCRIPTION OF THE DRAWINGS

[0027]

Figure 1 is a diagram which illustrates the basic 2-D organization of memory devices.
Figure 2 is a schematic block diagram which illustrates the parallel connection of all bus lines and the serial Reset line to each device in the system.
Figure 3 is a perspective view of a system of the invention which illustrates the 3-D packaging of semiconductor devices on the primary bus.
Figure 4 shows the format of a request packet.
Figure 5 shows the format of a retry response from a slave.
Figure 6 shows the bus cycles after a request packet collision occurs on the bus and how arbitration is handled.
Figure 7 shows the timing whereby signals from two devices can overlap temporarily and drive the bus at the same time.
Figure 8 shows the connection and timing between bus clocks and devices on the bus.
Figure 9 is a perspective view showing how transceivers can be used to connect a number of bus units to a transceiver bus. Figure 10 is a block and schematic diagram of input/output circuitry used to connect devices to the bus.
Figure 11 is a schematic diagram of a clocked sense-amplifier used as a bus input receiver.
Figure 12 is a block diagram showing how the internal device clock is generated from two bus clock signals using a set of adjustable delay lines.
Figure 13 is a timing diagram showing the relationship of signals in the block diagram of Figure 12.
Figure 14 is timing diagram of a preferred means of implementing the reset procedure of this invention.
Figure 15 is a diagram illustrating the general organization of a 4 Mbit DRAM divided into 8 subarrays.

DETAILED DESCRIPTION

[0028] The present invention is designed to provide a high speed, multiplexed bus for communication between processing devices and memory devices and to provide devices adapted for use in the bus system. The invention can also be used to connect processing devices and other devices, such as I/O interfaces or disk controllers, with or without memory devices on the bus. The bus consists of a relatively small number of lines connected in parallel to each device on the bus. The bus carries substantially all address, data and control information needed by devices for communication with other devices on the bus. In many systems using the present invention, the bus carries almost every signal between every device in the entire system. There is no need for separate device-select lines since device-select information for each device on the bus is carried over the bus. There is no need for separate address and data lines because address and data information can be sent over the same lines. Using the organization described herein, very large addresses (40 bits in the preferred implementation) and large data blocks (1024 bytes) can be sent over a small number of bus lines (8 plus one control line in the preferred implementation).

[0029] Virtually all of the signals needed by a computer system can be sent over the bus. Persons skilled in the art recognize that certain devices, such as CPUs, may be connected to other signal lines and possibly to independent buses, for example a bus to an independent cache memory, in addition to the bus of this invention. Certain devices, for

Petitioners
Ex. 1002, p. 317

example cross-point switches, could be connected to multiple, independent buses of this invention. In the preferred implementation, memory devices are provided that have no connections other than the bus connections described herein and CPUs are provided that use the bus of this invention as the principal, if not exclusive, connection to memory and to other devices on the bus.

[0030] All modern DRAM, SRAM and ROM designs have internal architectures with row (word) and column (bit) lines to efficiently tile a 2-D area. Referring to Fig. 1, one bit of data is stored at the intersection of each word line 5 and bit line 6. When a particular word line is enabled, all of the corresponding data bits are transferred onto the bit lines. This data, about 4000 bits at a time in a 4 MBit DRAM, is then loaded into column sense amplifiers 3 and held for use by the I/O circuits.

[0031] In the invention presented here, the data from the sense amplifiers is enabled 32 bits at a time onto an internal device bus running at approximately 125 MHz. This internal device bus moves the data to the periphery of the devices where the data is multiplexed into an 8-bit wide external bus interface, running at approximately 500 MHz.

[0032] The bus architecture of this invention connects master or bus controller devices, such as CPUs, Direct Memory Access devices (DMAs) or Floating Point Units (FPUs), and slave devices, such as DRAM, SRAM or ROM memory devices. A slave device responds to control signals; a master sends control signals. Persons skilled in the art realize that some devices may behave as both master and slave at various times, depending on the mode of operation and the state of the system. For example, a memory device will typically have only slave functions, while a DMA controller, disk controller or CPU may include both slave and master functions. Many other semiconductor devices, including I/O devices, disk controllers, or other special purpose devices such as high speed switches can be modified for use with the bus of this invention.

[0033] Each semiconductor device contains a set of internal registers, preferably including a device identification (device ID) register, a device-type descriptor register, control registers and other registers containing other information relevant to that type of device. In a preferred implementation, semiconductor devices connected to the bus contain registers which specify the memory addresses contained within that device and access-time registers which store a set of one or more delay times at which the device can or should be available to send or receive data.

[0034] Most of these registers can be modified and preferably are set as part of an initialization sequence that occurs when the system is powered up or reset. During the initialization sequence each device on the bus is assigned a unique device ID number, which is stored in the device ID register. A bus master can then use these device ID numbers to access and set appropriate registers in other devices, including access-time registers, control registers, and memory registers, to configure the system. Each slave may have one or several access-time registers (four in a preferred embodiment). In a preferred embodiment, one access-time register in each slave is permanently or semi-permanently programmed with a fixed value to facilitate certain control functions. A preferred implementation of an initialization sequence is described below in more detail.

[0035] All information sent between master devices and slave devices is sent over the external bus, which, for example, may be 8 bits wide. This is accomplished by defining a protocol whereby a master device, such as a microprocessor, seizes exclusive control of the external bus (i.e., becomes the bus master) and initiates a bus transaction by sending a request packet (a sequence of bytes comprising address and control information) to one or more slave devices on the bus. An address can consist of 16 to 40 or more bits according to the teachings of this invention. Each slave on the bus must decode the request packet to see if that slave needs to respond to the packet. The slave that the packet is directed to must then begin any internal processes needed to carry out the requested bus transaction at the requested time. The requesting master may also need to transact certain internal processes before the bus transaction begins. After a specified access time the slave(s) respond by returning one or more bytes (8 bits) of data or by storing information made available from the bus. More than one access time can be provided to allow different types of responses to occur at different times.

[0036] A request packet and the corresponding bus access are separated by a selected number of bus cycles, allowing the bus to be used in the intervening bus cycles by the same or other masters for additional requests or brief bus accesses. Thus multiple, independent accesses are permitted, allowing maximum utilization of the bus for transfer of short blocks of data. Transfers of long blocks of data use the bus efficiently even without overlap because the overhead due to bus address, control and access times is small compared to the total time to request and transfer the block.

Device Address Mapping

[0037] Another unique aspect of this invention is that each memory device is a complete, independent memory sub-system with all the functionality of a prior art memory board in a conventional backplane-bus computer system. Individual memory devices may contain a single memory section or may be subdivided into more than one discrete memory section. Memory devices preferably include memory address registers for each discrete memory section. A failed memory device (or even a subsection of a device) can be "mapped out" with only the loss of a small fraction of the memory, maintaining essentially full system capability. Mapping out bad devices can be accomplished in two ways, both compatible with this

Petitioners
Ex. 1002, p. 318

invention.

[0038] The preferred method uses address registers in each memory device (or independent discrete portion thereof) to store information which defines the range of bus addresses to which this memory device will respond. This is similar to prior art schemes used in memory boards in conventional backplane bus systems. The address registers can include a single pointer, usually pointing to a block of known size, a pointer and a fixed or variable block size value or two pointers, one pointing to the beginning and one to the end (or to the "top" and "bottom") of each memory block. By appropriate settings of the address registers, a series of functional memory devices or discrete memory sections can be made to respond to a contiguous range of addresses, giving the system access to a contiguous block of good memory, limited primarily by the number of good devices connected to the bus. A block of memory in a first memory device or memory section can be assigned a certain range of addresses, then a block of memory in a next memory device or memory section can be assigned addresses starting with an address one higher (or lower, depending on the memory structure) than the last address of the previous block.

[0039] Preferred devices for use in this invention include device-type register information specifying the type of chip, including how much memory is available in what configuration on that device. A master can perform an appropriate memory test, such as reading and writing each memory cell in one or more selected orders, to test proper functioning of each accessible discrete portion of memory (based in part on information like device ID number and device-type) and write address values (up to 40 bits in the preferred embodiment, 10^{12} bytes), preferably contiguous, into device address-space registers. Non-functional or impaired memory sections can be assigned a special address value which the system can interpret to avoid using that memory.

[0040] The second approach puts the burden of avoiding the bad devices on the system master or masters. CPUs and DMA controllers typically have some sort of translation look-aside buffers (TLBs) which map virtual to physical (bus) addresses. With relatively simple software, the TLBs can be programmed to use only working memory (data structures describing functional memories are easily generated). For masters which don't contain TLBs (for example, a video display generator), a small, simple RAM can be used to map a contiguous range of addresses onto the addresses of the functional memory devices.

[0041] Either scheme works and permits a system to have a significant percentage of non-functional devices and still continue to operate with the memory which remains. This means that systems built with this invention will have much improved reliability over existing systems, including the ability to build systems with almost no field failures.

Bus

[0042] The preferred bus architecture of this invention comprises 11 signals: BusData[0:7]; AddrValid; Clk1 and Clk2; plus an input reference level and power and ground lines connected in parallel to each device. Signals are driven onto the bus during conventional bus cycles. The notation "Signal[i:j]" refers to a specific range of signals or lines, for example, BusData[0:7] means BusData0, BusData1, ..., BusData7. The bus lines for BusData[0:7] signals form a byte-wide, multiplexed data/address/control bus. AddrValid is used to indicate when the bus is holding a valid address request, and instructs a slave to decode the bus data as an address and, if the address is included on that slave, to handle the pending request. The two clocks together provide a synchronized, high speed clock for all the devices on the bus. In addition to the bused signals, there is one other line (ResetIn, ResetOut) connecting each device in series for use during initialization to assign every device in the system a unique device ID number (described below in detail).

[0043] To facilitate the extremely high data rate of this external bus relative to the gate delays of the internal logic, the bus cycles are grouped into pairs of even/odd cycles. Note that all devices connected to a bus should preferably use the same even/odd labeling of bus cycles and preferably should begin operations on even cycles. This is enforced by the clocking scheme.

Protocol and Bus Operation

[0044] The bus uses a relatively simple, synchronous, split-transaction, block-oriented protocol for bus transactions. One of the goals of the system is to keep the intelligence concentrated in the masters, thus keeping the slaves as simple as possible (since there are typically many more slaves than masters). To reduce the complexity of the slaves, a slave should preferably respond to a request in a specified time, sufficient to allow the slave to begin or possibly complete a device-internal phase including any internal actions that must precede the subsequent bus access phase. The time for this bus access phase is known to all devices on the bus - each master being responsible for making sure that the bus will be free when the bus access begins. Thus the slaves never worry about arbitrating for the bus. This approach eliminates arbitration in single master systems, and also makes the slave-bus interface simpler.

[0045] In a preferred implementation of the invention, to initiate a bus transfer over the bus, a master sends out a request packet, a contiguous series of bytes containing address and control information. It is preferable to use a request packet containing an even number of bytes and also preferable to start each packet on an even bus cycle.

Petitioners
Ex. 1002, p. 319

[0046] The device-select function is handled using the bus data lines. AddrValid is driven, which instructs all slaves to decode the request packet address, determine whether they contain the requested address, and if they do, provide the data back to the master (in the case of a read request) or accept data from the master (in the case of a write request) in a data block transfer. A master can also select a specific device by transmitting a device ID number in a request packet. In a preferred implementation, a special device ID number is chosen to indicate that the packet should be interpreted by all devices on the bus. This allows a master to broadcast a message, for example to set a selected control register of all devices with the same value.

[0047] The data block transfer occurs later at a time specified in the request packet control information, preferably beginning on an even cycle. A device begins a data block transfer almost immediately with a device-internal phase as the device initiates certain functions, such as setting up memory addressing, before the bus access phase begins. The time after which a data block is driven onto the bus lines is selected from values stored in slave access-time registers. The timing of data for reads and writes is preferably the same; the only difference is which device drives the bus. For reads, the slave drives the bus and the master latches the values from the bus. For writes the master drives the bus and the selected slave latches the values from the bus.

[0048] In a preferred implementation of this invention shown in Figure 4, a request packet 22 contains 6 bytes of data -- 4.5 address bytes and 1.5 control bytes. Each request packet uses all nine bits of the multiplexed data/address lines (AddrValid 23 + BusData[0:7] 24) for all six bytes of the request packet. Setting 23 AddrValid = 1 in an otherwise unused even cycle indicates the start of an request packet (control information). In a valid request packet, AddrValid 27 must be 0 in the last byte. Asserting this signal in the last byte invalidates the request packet. This is used for the collision detection and arbitration logic (described below). Bytes 25-26 contain the first 35 address bits, Address[0:35]. The last byte contains AddrValid 27 (the invalidation switch) and 28, the remaining address bits, Address[36:39], and BlockSize [0:3] (control information).

[0049] The first byte contains two 4 bit fields containing control information, AccessType[0:3], an op code (operation code) which, for example, specifies the type of access, and Master[0:3], a position reserved for the master sending the packet to include its master ID number. Only master numbers 1 through 15 are allowed - master number 0 is reserved for special system commands. Any packet with Master[0:3] = 0 is an invalid or special packet and is treated accordingly.

[0050] The AccessType field specifies whether the requested operation is a read or write and the type of access, for example, whether it is to the control registers or other parts of the device, such as memory. In a preferred implementation, AccessType[0] is a Read/Write switch: if it is a 1, then the operation calls for a read from the slave (the slave to read the requested memory block and drive the memory contents onto the bus); if it is a 0, the operation calls for a write into the slave (the slave to read data from the bus and write it to memory). AccessType[1:3] provides up to 8 different access types for a slave. AccessType[1:2] preferably indicates the timing of the response, which is stored in an access-time register, AccessRegN. The choice of access-time register can be selected directly by having a certain op code select that register, or indirectly by having a slave respond to selected op codes with pre-selected access times (see table below). The remaining bit, AccessType[3] may be used to send additional information about the request to the slaves.

[0051] One special type of access is control register access, which involves addressing a selected register in a selected slave. In the preferred implementation of this invention, AccessType[1:3] equal to zero indicates a control register request and the address field of the packet indicates the desired control register. For example, the most significant two bytes can be the device ID number (specifying which slave is being addressed) and the least significant three bytes can specify a register address and may also represent or include data to be loaded into that control register. Control register accesses are used to initialize the access-time registers, so it is preferable to use a fixed response time which can be preprogrammed or even hard wired, for example the value in AccessReg0, preferably 8 cycles. Control register access can also be used to initialize or modify other registers, including address registers.

[0052] The method of this invention provides for access mode control specifically for the DRAMs. One such access mode determines whether the access is page mode or normal RAS access. In normal mode (in conventional DRAMS and in this invention), the DRAM column sense amps or latches have been precharged to a value intermediate between logical 0 and 1. This precharging allows access to a row in the RAM to begin as soon as the access request for either inputs (writes) or outputs (reads) is received and allows the column sense amps to sense data quickly. In page mode (both conventional and in this invention), the DRAM holds the data in the column sense amps or latches from the previous read or write operation. If a subsequent request to access data is directed td the same row, the DRAM does not need to wait for the data to be sensed (it has been sensed already) and access time for this data is much shorter than the normal access time. Page mode generally allows much faster access to data but to a smaller block of data (equal to the number of sense amps). However, if the requested data is not in the selected row, the access time is longer than the normal access time, since the request must wait for the RAM to precharge before the normal mode access can start. Two access-time registers in each DRAM preferably contain the access times to be used for normal and for page-mode accesses, respectively.

[0053] The access mode also determines whether the DRAM should precharge the sense amplifiers or should save the contents of the sense amps for a subsequent page mode access. Typical settings are "precharge after normal

Petitioners
Ex. 1002, p. 320

access" and "save after page mode access" but "precharge after page mode access" or "save after normal access" are allowed, selectable modes of operation. The DRAM can also be set to precharge the sense amps if they are not accessed for a selected period of time.

[0054] In page mode, the data stored in the DRAM sense amplifiers may be accessed within much less time than it takes to read out data in normal mode (~10-20 nS vs. 40-100 nS). This data may be kept available for long periods. However, if these sense amps (and hence bit lines) are not precharged after an access, a subsequent access to a different memory word (row) will suffer a precharge time penalty of about 40-100 nS because the sense amps must precharge before latching in a new value.

[0055] The contents of the sense amps thus may be held and used as a cache, allowing faster, repetitive access to small blocks of data. DRAM-based page-mode caches have been attempted in the prior art using conventional DRAM organizations but they are not very effective because several chips are required per computer word. Such a conventional page-mode cache contains many bits (for example, 32 chips x 4Kbits) but has very few independent storage entries. In other words, at any given point in time the sense amps hold only a few different blocks or memory "locales" (a single block of 4K words, in the example above). Simulations have shown that upwards of 100 blocks are required to achieve high hit rates (>90% of requests find the requested data already in cache memory) regardless of the size of each block. *See, for* example, Anant Agarwal, et. al., "An Analytic Cache Model," ACM Transactions on Computer Systems, Vol. 7 (2), pp. 184-215 (May 1989).

[0056] The organization of memory in the present invention allows each DRAM to hold one or more (4 for 4MBit DRAMS) separately- addressed and independent blocks of data. A personal computer or workstation with 100 such DRAMs (i.e. 400 blocks or locales) can achieve extremely high, very repeatable hit rates (98-99% on average) as compared to the lower (50-80%), widely varying hit rates using DRAMS organized in the conventional fashion. Further, because of the time penalty associated with the deferred precharge on a "miss" of the page-mode cache, the conventional DRAM-based page-mode cache generally has been found to work less well than no cache at all.

[0057] For DRAM slave access, the access types are preferably used in the following way:

AccessType[1:3]	Use	AccessTime
0	Control Register Access	Fixed, 8[AccessReg0]
1	Unused	Fixed, 8[AccessReg0]
2-3	Unused	AccessReg1
4-5	Page Mode DRAM access	AccessReg2
6-7	Normal DRAM access	AccessReg3

Persons skilled in the art will recognize that a series of available bits could be designated as switches for controlling these access modes. For example:

AccessType[2] = page mode/normal switch

AccessType[3] = precharge/save-data switch

BlockSize[0:3] specifies the size of the data block transfer. If BlockSize[0] is 0, the remaining bits are the binary representation of the block size (0-7). If BlockSize[0] is 1, then the remaining bits give the block size as a binary power of 2, from 8 to 1024. A zero-length block can be interpreted as a special command, for example, to refresh a DRAM without returning any data, or to change the DRAM from page mode to normal access mode or vice-versa.

BlockSize[0:2]	Number of Bytes in Block
0-7	0-7 respectively
8	8
9	16
10	32
11	64
12	128
13	256
14	512
15	1024

Persons skilled in the art will recognize that other block size encoding schemes or values can be used.

[0058] In most cases, a slave will respond at the selected access time by reading or writing data from or to the bus over bus lines BusData[0:7] and AddrValid will be at logical 0. In a preferred embodiment, substantially each memory

Petitioners
Ex. 1002, p. 321

access will involve only a single memory device, that is, a single block will be read from or written to a single memory device.

Retry Format

[0059] In some cases, a slave may not be able to respond correctly to a request, e.g., for a read or write. In such a situation, the slave should return an error message, sometimes called a N(o)ACK(nowledge) or retry message. The retry message can include information about the condition requiring a retry, but this increases system requirements for circuitry in both slave and masters. A simple message indicating only that an error has occurred allows for a less complex slave, and the master can take whatever action is needed to understand and correct the cause of the error.

[0060] For example, under certain conditions a slave might not be able to supply the requested data. During a page-mode access, the DRAM selected must be in page mode and the requested address must match the address of the data held in the sense amps or latches. Each DRAM can check for this match during a page-mode access. If no match is found, the DRAM begins precharging and returns a retry message to the master during the first cycle of the data block (the rest of the returned block is ignored). The master then must wait for the precharge time (which is set to accommodate the type of slave in question, stored in a special register, PreChargeReg), and then resend the request as a normal DRAM access (AccessType = 6 or 7).

[0061] In the preferred form of the present invention, a slave signals a retry by driving AddrValid true at the time the slave was supposed to begin reading or writing data. A master which expected to write to that slave must monitor AddrValid during the write and take corrective action if it detects a retry message. Figure 5 illustrates the format of a retry message 28 which is useful for read requests, consisting of 23 AddrValid=1 with Master[0:3] = 0 in the first (even) cycle. Note that AddrValid is normally 0 for data block transfers and that there is no master 0 (only 1 through 15 are allowed). All DRAMs and masters can easily recognize such a packet as an invalid request packet, and therefore a retry message. In this type of bus transaction all of the fields except for Mastery[0:3] and AddrValid 23 may be used as information fields, although in the implementation described, the contents are undefined. Persons skilled in the art recognize that another method of signifying a retry message is to add a DataInvalid line and signal to the bus. This signal could be asserted in the case of a NACK.

Bus Arbitration

[0062] In the case of a single master, there are by definition no arbitration problems. The master sends request packets and keeps track of periods when the bus will be busy in response to that packet. The master can schedule multiple requests so that the corresponding data block transfers do not overlap.

[0063] The bus architecture of this invention is also useful in configurations with multiple masters. When two or more masters are on the same bus, each master must keep track of all the pending transactions, so each master knows when it can send a request packet and access the corresponding data block transfer. Situations will arise, however, where two or more masters send a request packet at about the same time and the multiple requests must be detected, then sorted out by some sort of bus arbitration.

[0064] There are many ways for each master to keep track of when the bus is and will be busy. A simple method is for each master to maintain a bus-busy data structure, for example by maintaining two pointers, one to indicate the earliest point in the future when the bus will be busy and the other to indicate the earliest point in the future when the bus will be free, that is, the end of the latest pending data block transfer. Using this information, each master can determine whether and when there is enough time to send a request packet (as described above under Protocol) before the bus becomes busy with another data block transfer and whether the corresponding data block transfer will interfere with pending bus transactions. Thus each master must read every request packet and update its bus-busy data structure to maintain information about when the bus is and will be free.

[0065] With two or more masters on the bus, masters will occasionally transmit independent request packets during the same bus cycle. Those multiple requests will collide as each such master drives the bus simultaneously with different information, resulting in scrambled request information and neither desired data block transfer. In a preferred form of the invention, each device on the bus seeking to write a logical 1 on a BusData or AddrValid line drives that line with a current sufficient to sustain a voltage greater than or equal to the high-logic value for the system. Devices do not drive lines that should have a logical 0; those lines are simply held at a voltage corresponding to a low-logic value. Each master tests the voltage on at least some, preferably all, bus data and the AddrValid lines so the master can detect a logical '1' where the expected level is '0' on a line that it does not drive during a given bus cycle but another master does drive.

[0066] Another way to detect collisions is to select one or more bus lines for collision signalling. Each master sending a request drives that line or lines and monitors the selected lines for more than the normal drive current (or a logical value of ">1"), indicating requests by more than one master. Persons skilled in the art will recognize that this can be implemented with a protocol involving BusData and AddrValid lines or could be implemented using an additional bus line.

Petitioners
Ex. 1002, p. 322

[0067] In the preferred form of this invention, each master detects collisions by monitoring lines which it does not drive to see if another master is driving those lines. Referring to Fig. 4, the first byte of the request packet includes the number of each master attempting to use the bus (Master[0:3]). If two masters send packet requests starting at the same point in time, the master numbers will be logical "or"ed together by at least those masters, and thus one or both of the masters, by monitoring the data on the bus and comparing what it sent, can detect a collision. For instance if requests by masters number 2 (0010) and 5 (0101) collide, the bus will be driven with the value Master[0:3]=7 (0010 + 0101 = 0111). Master number 5 will detect that the signal Master[2] = 1 and master 2 will detect that Master[1] and Master[3] = 1, telling both masters that a collision has occurred. Another example is masters 2 and 11, for which the bus will be driven with the value Master[0:3]=11 (0010 + 1011 = 1011), and although master 11 can't readily detect this collision, master 2 can. When any collision is detected, each master detecting a collision drives the value of AddrValid 27 in byte 5 of the request packet 22 to 1, which is detected by all masters, including master 11 in the second example above, and forces a bus arbitration cycle, described below.

[0068] Another collision condition may arise where master A sends a request packet in cycle 0 and master B tries to send a request packet starting in cycle 2 of the first request packet, thereby overlapping the first request packet. This will occur from time to time because the bus operates at high speeds, thus the logic in a second-initiating master may not be fast enough to detect a request initiated by a first master in cycle 0 and to react fast enough by delaying its own request. Master B eventually notices that it wasn't supposed to try to send a request packet (and consequently almost surely destroyed the address that master A was trying to send), and, as in the example above of a simultaneous collision, drives a 1 on AddrValid during byte 5 of the first request packet 27 forcing an arbitration. The logic in the preferred implementation is fast enough that a master should detect a request packet by another master by cycle 3 of the first request packet, so no master is likely to attempt to send a potentially colliding request packet later than cycle 2.

[0069] Slave devices not need to detect a collision directly, but they must wait to do anything irrecoverable until the last byte (byte 5) is read to ensure that the packet is valid. A request packet with Master[0:3] equal to 0 (a retry signal) is ignored and does not cause a collision. The subsequent bytes of such a packet are ignored.

[0070] To begin arbitration after a collision, the masters wait a preselected number of cycles after the aborted request packet (4 cycles in a preferred implementation), then use the next free cycle to arbitrate for the bus (the next available even cycle in the preferred implementation). Each colliding master signals to all other colliding masters that it seeks to send a request packet, a priority is assigned to each of the colliding masters, then each master is allowed to make its request in the order of that priority.

[0071] Figure 6 illustrates one preferred way of implementing this arbitration. Each colliding master signals its intent to send a request packet by driving a single BusData line during a single bus cycle corresponding to its assigned master number (1-15 in the present example). During two-byte arbitration cycle 29, byte 0 is allocated to requests 1-7 from masters 1-7, respectively, (bit 0 is not used) and byte 1 is allocated to requests 8-15 from masters 8-15, respectively. At least one device and preferably each colliding master reads the values on the bus during the arbitration cycles to determine and store which masters desire to use the bus. Persons skilled in the art will recognize that a single byte can be allocated for arbitration requests if the system includes more bus lines than masters. More than 15 masters can be accommodated by using additional bus cycles.

[0072] A fixed priority scheme (preferably using the master numbers, selecting lowest numbers first) is then used to prioritize, then sequence the requests in a bus arbitration queue which is maintained by at least one device. These requests are queued by each master in the bus-busy data structure and no further requests are allowed until the bus arbitration queue is cleared. Persons skilled in the art will recognize that other priority schemes can be used, including assigning priority according to the physical location of each master.

System Configuration/Reset

[0073] In the bus-based system of this invention, a mechanism is provided to give each device on the bus a unique device identifier (device ID) after power-up or under other conditions as desired or needed by the system. A master can then use this device ID to access a specific device, particularly to set or modify registers of the specified device, including the control and address registers. In the preferred embodiment, one master is assigned to carry out the entire system configuration process. The master provides a series of unique device ID numbers for each unique device connected to the bus system. In the preferred embodiment, each device connected to the bus contains a special device-type register which specifies the type of device, for instance CPU, 4 MBit memory, 64 MBit memory or disk controller. The configuration master should check each device, determine the device type and set appropriate control registers, including access-time registers. The configuration master should check each memory device and set all appropriate memory address registers.

[0074] One means to set up unique device ID numbers is to have each device to select a device ID in sequence and store the value in an internal device ID register. For example, a master can pass sequential device ID numbers through shift registers in each of a series of devices, or pass a token from device to device whereby the device with the token

Petitioners
Ex. 1002, p. 323

reads in device ID information from another line or lines. In a preferred embodiment, device ID numbers are assigned to devices according to their physical relationship, for instance, their order along the bus.

[0075] In a preferred embodiment of this invention, the device ID setting is accomplished using a pair of pins on each device, ResetIn and ResetOut. These pins handle normal logic signals and are used only during device ID configuration. On each rising edge of the clock, each device copies ResetIn (an input) into a four-stage reset shift register. The output of the reset shift register is connected to ResetOut, which in turn connects to ResetIn for the next sequentially connected device. Substantially all devices on the bus are thereby daisy-chained together. A first reset signal, for example, while ResetIn at a device is a logical 1, or when a selected bit of the reset shift register goes from zero to non-zero, causes the device to hard reset, for example by clearing all internal registers and resetting all state machines. A second reset signal, for example, the falling edge of ResetIn combined with changeable values on the external bus, causes that device to latch the contents of the external bus into the internal device ID register (Device[0:7]).

[0076] To reset all devices on a bus, a master sets the ResetIn line of the first device to a "1" for long enough to ensure that all devices on the bus have been reset (4 cycles times the number of devices -- note that the maximum number of devices on the preferred bus configuration is 256 (8 bits), so that 1024 cycles is always enough time to reset all devices.) Then ResetIn is dropped to "0" and the BusData lines are driven with the first followed by successive device ID numbers, changing after every 4 clock pulses. Successive devices set those device ID numbers into the corresponding device ID register as the falling edge of ResetIn propagates through the shift registers of the daisy-chained devices. Figure 14 shows ResetIn at a first device going low while a master drives a first device ID onto the bus data lines BusData[0:3]. The first device then latches in that first device ID. After four clock cycles, the master changes BusData[0:3] to the next device ID number and ResetOut at the first device goes low, which pulls ResetIn for the next daisy-chained device low, allowing the next device to latch in the next device ID number from BusData[0:3]. In the preferred embodiment, one master is assigned device ID 0 and it is the responsibility of that master to control the ResetIn line and to drive successive device ID numbers onto the bus at the appropriate times. In the preferred embodiment, each device waits two clock cycles after ResetIn goes low before latching in a device ID number from BusData[0:3].

[0077] Persons skilled in the art recognize that longer device ID numbers could be distributed to devices by having each device read in multiple bytes from the bus and latch the values into the device ID register. Persons skilled in the art also recognize that there are alternative ways of getting device ID numbers to unique devices. For instance, a series of sequential numbers could be clocked along the ResetIn line and at a certain time each device could be instructed to latch the current reset shift register value into the device ID register.

[0078] The configuration master should choose and set an access time in each access-time register in each slave to a period sufficiently long to allow the slave to perform an actual, desired memory access. For example, for a normal DRAM access, this time must be longer than the row address strobe (RAS) access time. If this condition is not met, the slave may not deliver the correct data. The value stored in a slave access-time register is preferably one-half the number of bus cycles for which the slave device should wait before using the bus in response to a request. Thus an access time value of '1' would indicate that the slave should not access the bus until at least two cycles after the last byte of the request packet has been received. The value of AccessReg0 is preferably fixed at 8 (cycles) to facilitate access to control registers.

[0079] The bus architecture of this invention can include more than one master device. The reset or initialization sequence should also include a determination of whether there are multiple masters on the bus, and if so to assign unique master ID numbers to each. Persons skilled in the art will recognize that there are many ways of doing this. For instance, the master could poll each device to determine what type of device it is, for example, by reading a special register then, for each master device, write the next available master ID number into a special register.

ECC

[0080] Error detection and correction ("ECC") methods well known in the art can be implemented in this system. ECC information typically is calculated for a block of data at the time that block of data is first written into memory. The data block usually has an integral binary size, e.g. 256 bits, and the ECC information uses significantly fewer bits. A potential problem arises in that each binary data block in prior art schemes typically is stored with the ECC bits appended, resulting in a block size that is not an integral binary power.

[0081] In a preferred embodiment of this invention, ECC information is stored separately from the corresponding data, which can then be stored in blocks having integral binary size. ECC information and corresponding data can be stored, for example, in separate DRAM devices. Data can be read without ECC using a single request packet, but to write or read error-corrected data requires two request packets, one for the data and a second for the corresponding ECC information. ECC information may not always be stored permanently and in some situations the ECC information may be available without sending a request packet or without a bus data block transfer.

[0082] In a preferred embodiment, a standard data block size can be selected for use with ECC, and the ECC method will determine the required number of bits of information in a corresponding ECC block. RAMs containing ECC information

Petitioners
Ex. 1002, p. 324

can be programmed to store an access time that is equal to: (1) the access time of the normal RAM (containing data) plus the time to access a standard data block (for corrected data) minus the time to send a request packet (6 bytes); or (2) the access time of a normal RAM minus the time to access a standard ECC block minus the time to send a request packet. To read a data block and the corresponding ECC block, the master simply issues a request for the data immediately followed by a request for the ECC block. The ECC RAM will wait for the selected access time then drive its data onto the bus right after (in case (1) above)) the data RAM has finished driving out the data block. Persons skilled in the art will recognize that the access time described in case (2) above can be used to drive ECC data before the data is driven onto the bus lines and will recognize that writing data can be done by analogy with the method described for a read. Persons skilled in the art will also recognize the adjustments that must be made in the bus-busy structure and the request packet arbitration methods of this invention in order to accommodate these paired ECC requests.

[0083] Since this system is quite flexible, the system designer can choose the size of the data blocks and the number of ECC bits using the memory devices of this invention. Note that the data stream on the bus can be interpreted in various ways. For instance the sequence can be 2^n data bytes followed by 2^m ECC bytes (or vice versa), or the sequence can be 2^k iterations of 8 data bytes plus 1 ECC byte. Other information, such as information used by a directory-based cache coherence scheme, can also be managed this way. See, for example, Anant Agarwal, et al., "Scaleable Directory Schemes for Cache Consistency," 15th International Symposium on Computer Architecture, June 1988, pp. 280-289. Those skilled in the art will recognize alternative methods of implementing ECC schemes that are within the teachings of this invention.

Low Power 3-D Packaging

[0084] Another major advantage of this invention is that it drastically reduces the memory system power consumption. Nearly all the power consumed by a prior art DRAM is dissipated in performing row access. By using a single row access in a single RAM to supply all the bits for a block request (compared to a row-access in each of multiple RAMs in conventional memory systems) the power per bit can be made very small. Since the power dissipated by memory devices using this invention is significantly reduced, the devices potentially can be placed much closer together than with conventional designs.

[0085] The bus architecture of this invention makes possible an innovative 3-D packaging technology. By using a narrow, multiplexed (time-shared) bus, the pin count for an arbitrarily large memory device can be kept quite small - on the order of 20 pins. Moreover, this pin count can be kept constant from one generation of DRAM density to the next. The low power dissipation allows each package to be smaller, with narrower pin pitches (spacing between the IC pins). With current surface mount technology supporting pin pitches as low as 20 mils, all off-device connections can be implemented on a single edge of the memory device. Semiconductor die useful in this invention preferably have connections or pads along one edge of the die which can then be wired or otherwise connected to the package pins with wires having similar lengths. This geometry also allows for very short leads, preferably with an effective lead length of less than 4 mm. Furthermore, this invention uses only bused interconnections, i.e., each pad on each device is connected by the bus to the corresponding pad of each other device.

[0086] The use of a low pin count and an edge-connected bus permits a simple 3-D package, whereby the devices are stacked and the bus is connected along a single edge of the stack. The fact that all of the signals are bused is important for the implementation of a simple 3-D structure. Without this, the complexity of the "backplane" would be too difficult to make cost effectively with current technology. The individual devices in a stack of the present invention can be packed quite tightly because of the low power dissipated by the entire memory system, permitting the devices to be stacked bumper-to-bumper or top to bottom. Conventional plastic-injection molded small outline (SO) packages can be used with a pitch of about 2.5 mm (100 mils), but the ultimate limit would be the device die thickness, which is about an order of magnitude smaller, 0.2-0.5 mm using current wafer technology.

Bus Electrical Description

[0087] By using devices with very low power dissipation and close physical packing, the bus can be made quite short, which in turn allows for short propagation times and high data rates. The bus of a preferred embodiment of the present invention consists of a set of resistor-terminated controlled impedance transmission lines which can operate up to a data rate of 500 MHz (2 ns cycles). The characteristics of the transmission lines are strongly affected by the loading caused by the DRAMs (or other slaves) mounted on the bus. These devices add lumped capacitance to the lines which both lowers the impedance of the lines and decreases the transmission speed. In the loaded environment, the bus impedance is likely to be on the order of 25 ohms and the propagation velocity about $c/4$ (c = the speed of light) or 7.5 cm/ns. To operate at a 2 ns data rate, the transit time on the bus should preferably be kept under 1 ns, to leave 1 ns for the setup and hold time of the input receivers (described below) plus clock skew. Thus the bus lines must be kept quite short, under about 8 cm for maximum performance. Lower performance systems may have much longer lines, e.g. a 4 ns bus

Petitioners
Ex. 1002, p. 325

may have 24 cm lines (3 ns transit time, 1 ns setup and hold time).

[0088] In the preferred embodiment, the bus uses current source drivers. Each output must be able to sink 50 mA, which provides an output swing of about 500 mV or more. In the preferred embodiment of this invention, the bus is active low. The unasserted state (the high value) is preferably considered a logical zero, and the asserted value (low state) is therefore a logical 1. Those skilled in the art understand that the method of this invention can also be implemented using the opposite logical relation to voltage. The value of the unasserted state is set by the voltage on the termination resistors, and should be high enough to allow the outputs to act as current sources, while being as low as possible to reduce power dissipation. These constraints may yield a termination voltage about 2V above ground in the preferred implementation. Current source drivers cause the output voltage to be proportional to the sum of the sources driving the bus.

[0089] Referring to Fig.7, although there is no stable condition where two devices drive the bus at the same time, conditions can arise because of propagation delay on the wires where one device, A 41, can start driving its part of the bus 44 while the bus is still being driven by another device, B 42 (already asserting a logical 1 on the bus). In a system using current drivers, when B 42 is driving the bus (before time 46), the value at points 44 and 45 is logical 1. If B 42 switches off at time 46 just when A 41 switches on, the additional drive by device A 41 causes the voltage at the output 44 of A 41 to drop briefly below the normal value. The voltage returns to its normal value at time 47 when the effect of device B 42 turning off is felt. The voltage at point 45 goes to logical 0 when device B 42 turns off, then drops at time 47 when the effect of device A 41 turning on is felt. Since the logical 1 driven by current from device A 41 is propagated irrespective of the previous value on the bus, the value on the bus is guaranteed to settle after one time of flight (t_f) delay, that is, the time it takes a signal to propagate from one end of the bus to the other. If a voltage drive was used (as in ECL wired-ORing), a logical 1 on the bus (from device B 42 being previously driven) would prevent the transition put out by device A 41 being felt at the most remote part of the system, e.g., device 43, until the turnoff waveform from device B 42 reached device A 41 plus one time of flight delay, giving a worst case settling time of twice the time of flight delay.

Clocking

[0090] Clocking a high speed bus accurately without introducing error due to propagation delays can be implemented by having each device monitor two bus clock signals and then derive internally a device clock, the true system clock. The bus clock information can be sent on one or two lines to provide a mechanism for each bused device to generate an internal device clock with zero skew relative to all the other device clocks. Referring to Figure 8, in the preferred implementation, a bus clock generator 50 at one end of the bus propagates an early bus clock signal in one direction along the bus, for example on line 53 from left to right, to the far end of the bus. The same clock signal then is passed through the direct connection shown to a second line 54, and returns as a late bus clock signal along the bus from the far end to the origin, propagating from right to left. A single bus clock line can be used if it is left unterminated at the far end of the bus, allowing the early bus clock signal to reflect back along the same line as a late bus clock signal.

[0091] Figure 8b illustrates how each device 51, 52 receives each of the two bus clock signals at a different time (because of propagation delay along the wires), with constant midpoint in time between the two bus clocks along the bus. At each device 51, 52, the rising edge 55 of Clock1 53 is followed by the rising edge 56 of Clock2 54. Similarly, the falling edge 57 of Clock1 53 is followed by the falling edge 58 of Clock2 54. This waveform relationship is observed at all other devices along the bus. Devices which are closer to the clock generator have a greater separation between Clock1 and Clock2 relative to devices farther from the generator because of the longer time required for each clock pulse to traverse the bus and return along line 54, but the midpoint in time 59, 60 between corresponding rising or falling edges is fixed because, for any given device, the length of each clock line between the far end of the bus and that device is equal. Each device must sample the two bus clocks and generate its own internal device clock at the midpoint of the two.

[0092] Clock distribution problems can be further reduced by using a bus clock and device clock rate equal to the bus cycle data rate divided by two, that is, the bus clock period is twice the bus cycle period. Thus a 500 MHz bus preferably uses a 250 MHz clock rate. This reduction in frequency provides two benefits. First it makes all signals on the bus have the same worst case data rates -- data on a 500 MHz bus can only change every 2 ns. Second, clocking at half the bus cycle data rate makes the labeling of the odd and even bus cycles trivial, for example, by defining even cycles to be those when the internal device clock is 0 and odd cycles when the internal device clock is 1.

Multiple Buses

[0093] The limitation on bus length described above restricts the total number of devices that can be placed on a single bus. Using 2.5 mm spacing between devices, a single 8 cm bus will hold about 32 devices. Persons skilled in the art will recognize certain applications of the present invention wherein the overall data rate on the bus is adequate but memory or processing requirements necessitate a much larger number of devices (many more than 32). Larger systems can easily be built using the teachings of this invention by using one or more memory subsystems, designated primary

Petitioners
Ex. 1002, p. 326

bus units, each of which consists of two or more devices, typically 32 or close to the maximum allowed by bus design requirements, connected to a transceiver device.

Referring to Figure 9, each primary bus unit can be mounted on a single circuit board 66, sometimes called a memory stick. Each transceiver device 19 in turn connects to a transceiver bus 65, similar or identical in electrical and other respects to the primary bus 18 described at length above. In a preferred implementation, all masters are situated on the transceiver bus so there are no transceiver delays between masters and all memory devices are on primary bus units so that all memory accesses experience an equivalent transceiver delay, but persons skilled in the art will recognize how to implement systems which have masters on more than one bus unit and memory devices on the transceiver bus as well as on primary bus units. In general, each teaching of this invention which refers to a memory device can be practiced using a transceiver device and one or more memory devices on an attached primay bus unit. Other devices, generically referred to as peripheral devices, including disk controllers, video controllers or I/O devices can also be attached to either the transceiver bus or a primary bus unit, as desired. Persons skilled in the art will recognize how to use a single primary bus unit or multiple primary bus units as needed with a transceiver bus in certain system designs.

[0094] The transceivers are quite simple in function. They detect request packets on the transceiver bus and transmit them to their primary bus unit. If the request packet calls for a write to a device on a transceiver's primary bus unit, that transceiver keeps track of the access time and block size and forwards all data from the transceiver bus to the primary bus unit during that time. The transceivers also watch their primary bus unit, forwarding any data that occurs there to the transceiver bus. The high speed of the buses means that the transceivers will need to be pipelined, and will require an additional one or two cycle delay for data to pass through the transceiver in either direction. Access times stored in masters on the transceiver bus must be increased to account for transceiver delay but access times stored in slaves on a primary bus unit should not be modified.

[0095] Persons skilled in the art will recognize that a more sophisticated transceiver can control transmissions to and from primary bus units. An additional control line, TrncvrRW can be bused to all devices on the transceiver bus, using that line in conjunction with the AddrValid line to indicate to all devices on the transceiver bus that the information on the data lines is: 1) a request packet, 2) valid data to a slave, 3) valid data from a slave, or 4) invalid data (or idle bus). Using this extra control line obviates the need for the transceivers to keep track of when data needs to be forwarded from its primary bus to the transceiver bus - all transceivers send all data from their primary bus to the transceiver bus whenever the control signal indicates condition 2) above. In a preferred implementation of this invention, if AddrValid and TrncvrRW are both low, there is no bus activity and the transceivers should remain in an idle state. A controller sending a request packet will drive AddrValid high, indicating to all devices on the transceiver bus that a request packet is being sent which each transceiver should forward to its primary bus unit. Each controller seeking to write to a slave should drive both AddrValid and TrncvrRW high, indicating valid data for a slave is present on the data lines. Each transceiver device will then transmit all data from the transceiver bus lines to each primary bus unit. Any controller expecting to receive information from a slave should also drive the TrncvrRW line high, but not drive AddrValid, thereby indicating to each transceiver to transmit any data coming from any slave on its primary local bus to the transceiver bus. A still more sophisticated transceiver would recognize signals addressed to or coming from its primary bus unit and transmit signals only at requested times.

[0096] An example of the physical mounting of the transceivers is shown in Figure 9. One important feature of this physical arrangement is to integrate the bus of each transceiver 19 with the original bus of DRAMs or other devices 15, 16, 17 on the primary bus unit 66. The transceivers 19 have pins on two sides, and are preferably mounted flat on the primary bus unit with a first set of pins connected to primary bus 18. A second set of transceiver pins 20, preferably orthogonal to the first set of pins, are oriented to allow the transceiver 19 to be attached to the transceiver bus 65 in much the same way as the DRAMs were attached to the primary bus unit. The transceiver bus can be generally planar and in a different plane, preferably orthogonal to the plane of each primary bus unit. The transceiver bus can also be generally circular with primary bus units mounted perpendicular and tangential to the transceiver bus.

[0097] Using this two level scheme allows one to easily build a system that contains over 500 slaves (16 buses of 32 DRAMs each). Persons skilled in the art can modify the device ID scheme described above to accommodate more than 256 devices, for example by using a longer device ID or by using additional registers to hold some of the device ID. This scheme can be extended in yet a third dimension to make a second-order transceiver bus, connecting multiple transceiver buses by aligning transceiver bus units parallel to and on top of each other and busing corresponding signal lines through a suitable transceiver. Using such a second-order transceiver bus, one could connect many thousands of slave devices into what is effectively a single bus.

Device Interface

[0098] The device interface to the high-speed bus can be divided into three main parts. The first part is the electrical interface. This part includes the input receivers, bus drivers and clock generation circuitry. The second part contains the address comparison circuitry and timing registers. This part takes the input request packet and determines if the request

Petitioners
Ex. 1002, p. 327

is for this device, and if it is, starts the internal access and delivers the data to the pins at the correct time. The final part, specifically for memory devices such as DRAMs, is the DRAM column access path. This part needs to provide bandwidth into and out of the DRAM sense amps greater than the bandwidth provided by conventional DRAMs. The implementation of the electrical interface and DRAM column access path are described in more detail in the following sections. Persons skilled in the art recognize how to modify prior-art address comparison circuitry and prior-art register circuitry in order to practice the present invention.

Electrical Interface - Input/Output Circuitry

[0099] A block diagram of the preferred input/output circuit for address/data/control lines is shown in Figure 10. This circuitry is particularly well-suited for use in DRAM devices but it can be used or modified by one skilled in the art for use in other devices connected to the bus of this invention. It consists of a set of input receivers 71, 72 and output driver 76 connected to input/output line 69 and pad 75 and circuitry to use the internal clock 73 and internal clock complement 74 to drive the input interface. The clocked input receivers take advantage of the synchronous nature of the bus. To further reduce the performance requirements for device input receivers, each device pin, and thus each bus line, is connected to two clocked receivers, one to sample the even cycle inputs, the other to sample the odd cycle inputs. By thus de-multiplexing the input 70 at the pin, each clocked amplifier is given a full 2 ns cycle to amplify the bus low-voltage-swing signal into a full value CMOS logic signal. Persons skilled in the art will recognize that additional clocked input receivers can be used within the teachings of this invention. For example, four input receivers could be connected to each device pin and clocked by a modified internal device clock to transfer sequential bits from the bus to internal device circuits, allowing still higher external bus speeds or still longer settling times to amplify the bus low-voltage-swing signal into a full value CMOS logic signal.

[0100] The output drivers are quite simple, and consist of a single NMOS pulldown transistor 76. This transistor is sized so that under worst case conditions it can still sink the 50 mA required by the bus. For 0.8 micron CMOS technology, the transistor will need to be about 200 microns long. Overall bus performance can be improved by using feedback techniques to control output transistor current so that the current through the device is roughly 50 mA under all operating conditions, although this is not absolutely necessary for proper bus operation. An example of one of many methods known to persons skilled in the art for using feedback techniques to control current is described in Hans Schumacher, et al., "CMOS Subnanosecond True-ECL Output Buffer," J. Solid State Circuits, Vol. 25 (1), pp. 150-154 (Feb. 1990). Controlling this current improves performance and reduces power dissipation. This output driver which can be operated at 500 MHz, can in turn be controlled by a suitable multiplexer with two or more (preferably four) inputs connected to other internal chip circuitry, all of which can be designed according to well known prior art.

[0101] The input receivers of every slave must be able to operate during every cycle to determine whether the signal on the bus is a valid request packet. This requirement leads to a number of constraints on the input circuitry. In addition to requiring small acquisition and resolution delays, the circuits must take little or no DC power, little AC power and inject very little current back into the input or reference lines. The standard clocked DRAM sense amp shown in Figure 11 satisfies all these requirements except the need for low input currents. When this sense amp goes from sense to sample, the capacitance of the internal nodes 83 and 84 in Figure11 is discharged through the reference line 68 and input 69, respectively. This particular current is small, but the sum of such currents from all the inputs into the reference lines summed over all devices can be reasonably large.

[0102] The fact that the sign of the current depends upon on the previous received data makes matters worse. One way to solve this problem is to divide the sample period into two phases. During the first phase, the inputs are shorted to a buffered version of the reference level (which may have an offset). During the second phase, the inputs are connected to the true inputs. This scheme does not remove the input current completely, since the input must still charge nodes 83 and 84 from the reference value to the current input value, but it does reduce the total charge required by about a factor of 10 (requiring only a 0.25V change rather than a 2.5V change). Persons skilled in the art will recognize that many other methods can be used to provide a clocked amplifier that will operate on very low input currents.

[0103] One important part of the input/output circuitry generates an internal device clock based on early and late bus clocks. Controlling clock skew (the difference in clock timing between devices) is important in a system running with 2 ns cycles, thus the internal device clock is generated so the input sampler and the output driver operate as close in time as possible to midway between the two bus clocks.

[0104] A block diagram of the internal device clock generating circuit is shown in Figure 12 and the corresponding timing diagram in Figure 13. The basic idea behind this circuit is relatively simple. A DC amplifier 102 is used to convert the small-swing bus clock into a full-swing CMOS signal. This signal is then fed into a variable delay line 103. The output of delay line 103 feeds three additional delay lines: 104 having a fixed delay; 105 having the same fixed delay plus a second variable delay; and 106 having the same fixed delay plus one half of the second variable delay. The outputs 107, 108 of the delay lines 104 and 105 drive clocked input receivers 101 and 111 connected to early and late bus clock inputs 100 and 110, respectively. These input receivers 101 and 111 have the same design as the receivers described

Petitioners
Ex. 1002, p. 328

above and shown in Fig. 11. Variable delay lines 103 and 105 are adjusted via feedback lines 116, 115 so that input receivers 101 and 111 sample the bus clocks just as they transition. Delay lines 103 and 105 are adjusted so that the falling edge 120 of output 107 precedes the falling edge 121 of the early bus clock, Clock1 53, by an amount of time 128 equal to the delay in input sampler 101. Delay line 108 is adjusted in the same way so that falling edge 122 precedes the falling edge 123 of late bus clock, Clock2 54, by the delay 128 in input sampler 111.

[0105] Since the outputs 107 and 108 are synchronized with the two bus clocks and the output 73 of the last delay line 106 is midway between outputs 107 and 108, that is, output 73 follows output 107 by the same amount of time 129 that output 73 precedes output 108, output 73 provides an internal device clock midway between the bus clocks. The falling edge 124 of internal device clock 73 precedes the time of actual input sampling 125 by one sampler delay. Note that this circuit organization automatically balances the delay in substantially all device input receivers 71 and 72 (Fig. 10), since outputs 107 and 108 are adjusted so the bus clocks are sampled by input receivers 101 and 111 just as the bus clocks transition.

[0106] In the preferred embodiment, two sets of these delay lines are used, one to generate the true value of the internal device clock 73, and the other to generate the complement 74 without adding any inverter delay. The dual circuit allows generation of truly complementary clocks, with extremely small skew. The complement internal device clock is used to clock the 'even' input receivers to sample at time 127, while the true internal device clock is used to clock the 'odd' input receivers to sample at time 125. The true and complement internal device clocks are also used to select which data is driven to the output drivers. The gate delay between the internal device clock and output circuits driving the bus is slightly greater than the corresponding delay for the input circuits, which means that the new data always will be driven on the bus slightly after the old data has been sampled.

DRAM Column Access Modification

[0107] A block diagram of a conventional 4 MBit DRAM 130 is shown in Figure 15. The DRAM memory array is divided into a number of subarrays 150-157, for example, 8. Each subarray is divided into arrays 148, 149 of memory cells. Row address selection is performed by decoders 146. A column decoder 147A, 147B, including column sense amps on either side of the decoder, runs through the core of each subarray. These column sense amps can be set to precharge or latch the most-recently stored value, as described in detail above. Internal I/O lines connect each set of sense-amps, as gated by corresponding column decoders, to input and output circuitry connected ultimately to the device pins. These internal I/O lines are used to drive the data from the selected bit lines to the data pins (some of pins 131-145), or to take the data from the pins and write the selected bit lines. Such a column access path organized by prior art constraints does not have sufficient bandwidth to interface with a high speed bus. The method of this invention does not require changing the overall method used for column access, but does change implementation details. Many of these details have been implemented selectively in certain fast memory devices, but never in conjunction with the bus architecture of this invention.

[0108] Running the internal I/O lines in the conventional way at high bus cycle rates is not possible. In the preferred method, several (preferably 4) bytes are read or written during each cycle and the column access path is modified to run at a lower rate (the inverse of the number of bytes accessed per cycle, preferably 1/4 of the bus cycle rate). Three different techniques are used to provide the additional internal I/O lines required and to supply data to memory cells at this rate. First, the number of I/O bit lines in each subarray running through the column decoder 147 is increased, for example, to 16, eight for each of the two columns of column sense amps and the column decoder selects one set of columns from the "top" half 148 of subarray 150 and one set of columns from the "bottom" half 149 during each cycle, where the column decoder selects one column sense amp per I/O bit line. Second, each column I/O line is divided into two halves, carrying data independently over separate internal I/O lines from the left half 147A and right half 147B of each subarray (dividing each subarray into quadrants) and the column decoder selects sense amps from each right and left half of the subarray, doubling the number of bits available at each cycle. Thus each column decode selection turns on *n* column sense amps, where n equals four (top left and right, bottom left and right quadrants) times the number of I/O lines in the bus to each subarray quadrant (8 lines each x 4=32 lines in the preferred implementation). Finally, during each RAS cycle, two different subarrays, e.g. 157 and 153, are accessed. This doubles again the available number of I/O lines containing data. Taken together, these changes increase the internal I/O bandwidth by at least a factor of 8. Four internal buses are used to route these internal I/O lines. Increasing the number of I/O lines and then splitting them in the middle greatly reduces the capacitance of each internal I/O line which in turn reduces the column access time, increasing the column access bandwidth even further.

[0109] The multiple, gated input receivers described above allow high speed input from the device pins onto the internal I/O lines and ultimately into memory. The multiplexed output driver described above is used to keep up with the data flow available using these techniques. Control means are provided to select whether information at the device pins should be treated as an address, and therefore to be decoded, or input or output data to be driven onto or read from the internal I/O lines.

[0110] Each subarray can access 32 bits per cycle, 16 bits from the left subarray and 16 from the right subarray. With

Petitioners
Ex. 1002, p. 329

8 I/O lines per sense-amplifier column and accessing two subarrays at a time, the DRAM can provide 64 bits per cycle. This extra I/O bandwidth is not needed for reads (and is probably not used), but may be needed for writes. Availability of write bandwidth is a more difficult problem than read bandwidth because over-writing a value in a sense-amplifier may be a slow operation, depending on how the sense amplifier is connected to the bit line. The extra set of internal I/O lines provides some bandwidth margin for write operations.

[0111] Persons skilled in the art will recognize that many variations of the teachings of this invention can be practiced that still fall within the claims of this invention which follow.

[0112] It follows a list of further embodiments of the invention:

Embodiment 1. A memory subsystem comprising
two memory devices connected in parallel to a bus,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices,
said control information including device-select information,
said bus containing substantially fewer bus lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual memory devices.

Embodiment 2. The memory subsystem with the features of embodiment 1 wherein said bus contains at least 8 bus lines adapted to carry at least 16 address bits and at least 8 data bits.

Embodiment 3. The memory subsystem with the features of embodiment 1 wherein said bus also includes parallel lines for clock and power.

Embodiment 4. A system comprising
a memory subsystem with the features of embodiment 1 wherein each bus of said memory subsystem is connected to its own transceiver device,
a transceiver bus connecting said transceiver devices, and
a means for transferring information between each of said buses of said memory subsystems and said transceiver bus, whereby memory subsystems may be integrated into a larger system having more memory than an individual memory subsystem.

Embodiment 5. The system with the features of embodiment 4 having a plurality of memory subsystems.

Embodiment 6. The system with the features of embodiment 4 further comprising a master device connected to said transceiver bus.

Embodiment 7. The system with the features of embodiment 6 wherein said master device is selected from the group consisting of a central processing unit, a floating point unit and a direct memory access unit.

Embodiment 8. The system with the features of embodiment 4 further comprising a peripheral device connected to the transceiver bus, said peripheral device adapted for connection to other devices not on the bus.

Embodiment 9. The system with the features of embodiment 8 wherein said peripheral device is selected from the group consisting of an I/O interface port, a video controller and a disk controller.

Embodiment 10. The system with the features of embodiment 5 wherein said transceiver bus is in a different plane than the plane of the bus of each of said memory subsystems.

Embodiment 11. The system with the features of embodiment 5 wherein the bus of each memory subsystem lies substantially in a subsystem bus plane and said transceiver bus lies substantially in a plane orthogonal to said subsystem bus plane.

Embodiment 12. The system with the features of embodiment 4 having at least two transceiver buses, each transceiver bus having a plurality of memory subsystem buses connected through a first transceiver to said transceiver bus,
each of said transceiver buses being further connected to a second transceiver adapted to interface to a second-order transceiver bus, whereby each transceiver bus is connected through said second transceiver to form a second-order transceiver bus unit.

Embodiment 13. A semiconductor subsystem bus for interconnecting semiconductor devices comprising
a plurality of semiconductor devices connected in parallel to a bus, at least one of said semiconductor devices being a memory device or a transceiver device which in turn is connected to a memory subsystem,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including semiconductor device-select information,
said bus containing substantially fewer bus lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual semiconductor devices, and
at least one modifiable register in each of the semiconductor devices on said bus, said modifiable registers being accessible from said bus, whereby the subsystem can be configured using signals transmitted on said bus.

Embodiment 14. The semiconductor subsystem bus with the features of embodiment 13 wherein one type of modifiable

Petitioners
Ex. 1002, p. 330

register is an access-time register designed to store a time delay after which a device may take some specified action on said bus.

Embodiment 15. The semiconductor subsystem bus with the features of embodiment 13 further comprising a semiconductor device having at least two access-time registers and

one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said bus.

Embodiment 16. The semiconductor subsystem bus with the features of embodiment 13 further comprising a memory device having at least one discrete memory section and also having a modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 17. The semiconductor subsystem bus with the features of embodiment 16 wherein said memory address information comprises a pointer to said discrete memory section.

Embodiment 18. The semiconductor subsystem bus with the features of embodiment 16 wherein said discrete memory section has a top and a bottom and said memory address information comprises pointers to said top and said bottom.

Embodiment 19. The semiconductor subsystem bus with the features of embodiment 16 wherein said memory address information comprises

a pointer to said discrete memory section and

a range value indicating the size of said discrete memory section.

Embodiment 20. The semiconductor subsystem bus with the features of embodiment 16 wherein said address registers of each of said discrete memory sections of each of said memory devices connected to said bus are set to contain memory address information that is different for each discrete memory section and such that the highest memory address in each discrete memory section is one less than the lowest memory address in another discrete memory section, whereby memory may be organized into one or a small number of contiguous memory blocks.

Embodiment 21. The semiconductor subsystem bus with the features of embodiment 16 further comprising a means for testing each of said discrete memory sections of each of said memory devices for proper function, and

for each non-functional discrete memory section, a means for setting at least one address register which corresponds to said discrete memory section to indicate that said discrete memory section is non-functional,

for each functional discrete memory section, a means for setting at least one address register which corresponds to said discrete memory section to contain such corresponding address information.

Embodiment 22. The semiconductor subsystem bus with the features of embodiment 21 wherein said address registers corresponding to said discrete memory sections are set to provide one contiguous memory block within the subsystem.

Embodiment 23. The semiconductor subsystem bus with the features of embodiment 13 wherein one of said modifiable registers is a device identification register which can be modified to contain a value unique to that semiconductor device.

Embodiment 24. The semiconductor subsystem bus with the features of embodiment 23 wherein said device identification register is set to contain a unique value which is a function of the physical position of that semiconductor device either along said bus or in relationship to other semiconductor devices or said bus.

Embodiment 25. A bus subsystem comprising

two semiconductor devices connected in parallel to a bus, wherein one of said semiconductor devices is a master device, said master device including a means for initiating bus transactions,

said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said devices,

said control information including device-select information,

said bus containing substantially fewer lines than the number of bits in a single address, and

said bus carrying device-select information without the need for separate device-select lines connected directly to individual devices on said bus, whereby said master device initiates bus transactions which transfer information between said semiconductor devices on said bus.

Embodiment 26. The bus subsystem with the features of embodiment 25 wherein one of said semiconductor devices is a memory device connected to said bus, said memory device having at least one discrete memory section and also having a modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 27. The bus subsystem with the features of embodiment 26 wherein one of said semiconductor devices comprises a transceiver device connected in parallel to said bus and connected in parallel to a memory device on a bus other than said bus.

28. The bus subsystem with the features of embodiment 26 further including a means Embodiment for said master device to request said memory device to prepare for a bus transaction by sending a request packet along said bus, said memory device and said master device each having a device-internal means to prepare to begin said bus transaction during a device-internal phase and further having a bus access means to effect said bus transaction during a bus access phase, said request packet including

a sequence of bytes containing address and control information,

Petitioners
Ex. 1002, p. 331

said control information including information about the requested bus transaction and about the access time, which corresponds to a number of bus cycles, which needs to intervene before beginning said bus-access phase, and
said address information pointing to at least one memory location within one of said discrete memory sections of said memory device.

Embodiment 29. The bus subsystem with the features of embodiment 28 wherein said memory device includes a means to read said control information and initiate said device-internal means at a time so as to complete said device-internal phase within said access time and begin said bus access phase after said number of bus cycles.

Embodiment 30. The bus subsystem with the features of embodiment 28 wherein said control information comprises an op code.

Embodiment 31. The bus subsystem with the features of embodiment 30 wherein said memory device includes sense amplifiers adapted to hold a bit of information or to precharge after a selected time and a means to transfer a data block during a data block transfer either reading data from said memory device or writing data into said memory device, and wherein said op code instructs said memory device to activate a response means, said response means including a means to

initiate a data block transfer,
select the size of said data block,
select the time to initiate said data block transfer,
access a control register, including reading from or writing to said control register,
precharge said sense amplifiers after each of said data block transfers is complete,
hold a bit of information in each of said sense amplifiers after each of said data block transfers is complete, or
select normal or page-mode access.

Embodiment 32. The bus subsystem with the features of embodiment 31 wherein said data block transfer comprises a read from or a write to memory within a single memory device.

Embodiment 33. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to a specific one of said semiconductor devices on said bus by including in said request packet a device identification number unique to said semiconductor device.

Embodiment 34. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to a selected one of said discrete memory portions by including in said request packet a specific memory address.

Embodiment 35. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to substantially all semiconductor devices on said bus by including in said request packet a special device identification number which is recognized by said semiconductor devices.

Embodiment 15 36. The bus subsystem with the features of embodiment 28 wherein said control information specifies directly or indirectly the number of bus cycles for said master device and said memory device to wait before beginning said bus access phase.

Embodiment 37. The bus subsystem with the features of embodiment 36 wherein, for a data block transfer, said master device and said memory device use the same access time and same data block size regardless of whether said data block transfer is a read or write operation.

Embodiment 38. The bus subsystem with the features of embodiment 28 wherein said control information further includes a block-size value that encodes and specifies the size of the block of data to be transferred.

Embodiment 39. The bus subsystem with the features of embodiment 38 wherein said block-size value is encoded as a linear value for relatively small block sizes values and is encoded as a logarithmic value for relatively larger block sizes.

Embodiment 40. The bus subsystem with the features of embodiment 38 wherein said block-size value is encoded using four bits, and where the encoded value is

Encoded Value	Block Size (Bytes)
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	16

Petitioners
Ex. 1002, p. 332

(continued)

Encoded Value	Block Size (Bytes)
10	32
11	64
12	128
13	256
14	512
15	1024

Embodiment 41. The bus subsystem with the features of embodiment 26 wherein said memory device is a DRAM device containing
a plurality of sense amplifiers,
a means to hold said sense amplifiers in an unmodified state after a read or write operation, leaving the device in page mode,
a means to precharge said sense amplifiers and
a means for selecting whether to precharge said sense amplifiers or to hold said sense amplifiers in an unmodified state.
Embodiment 42. The bus subsystem with the features of embodiment 28 wherein said request packet comprises an even number of bytes.
Embodiment 43. The bus subsystem with the features of embodiment 28 further including a means for generating and controlling a plurality of bus cycles, during which said bus carries said address, data and control information, and wherein alternate said bus cycles are designated odd cycles and even cycles, respectively, and wherein said request packet begins only on an even cycle.
Embodiment 44. The bus subsystem with the features of embodiment 28 further including a means for generating ECC information corresponding to a block of data and a means for using said ECC information to correct errors in storing or reading said block of data, wherein said ECC information may be stored separately from said block of data.
Embodiment 45. The bus subsystem with the features of embodiment 44 further comprising at least two of said memory devices wherein said ECC information and said corresponding block of data are stored in a first and a second said memory device, respectively, and said master device includes a means to write or read said block of data with error correction by sending separate ones of said request packets for said ECC information and for said corresponding block of data.
Embodiment 46. A bus subsystem comprising
a memory device and a master device connected in parallel on a bus,
a means for said master device to send a request packet and initiate a bus transaction and
a means for said master device to keep track of current and pending bus transactions,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices,
said bus containing substantially fewer lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual devices on said bus, whereby said master device initiates bus transactions which transfer information between devices on said bus and collisions on said bus are avoided because said master device avoids initiating bus transactions which would conflict with current or pending bus transactions.
Embodiment 47. The bus subsystem with the features of embodiment 46 having at least two of said master devices and including
a collision detecting means whereby a first said master device sending a first said request packet can detect a second said master device sending one of said colliding request packets, where one of said said colliding request packet may be sent simultaneous with the initial sending of or overlapping the sending of said first request packet, and
an arbitration means whereby said first and said second master devices select a priority order in which each of said master devices will be allowed to access said bus sequentially .
Embodiment 48. The bus subsystem with the features of embodiment 47 wherein each of said master devices has a master ID number and each of said request packets includes a master ID position which is a predetermined number of bits in a predetermined position in said request packet, and wherein said collision detection means comprises
a means included in each master device for sending a request packet including said master ID number of said master device in said master ID position of said request packet and
a means to detect a collision and invoke said arbitration means if any master device detects any other master ID number in said master ID position.
Embodiment 49. The bus subsystem with the features of embodiment 47 wherein each of said master devices includes

Petitioners
Ex. 1002, p. 333

a means for sending a request packet,
a means for driving a selected bus line or lines during at least one selected bus cycle while said request packet is being sent,
a means for monitoring said selected bus line or lines to see if a said master device is sending a colliding request packet and
a means for informing all other master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 50. The bus subsystem with the features of embodiment 47 wherein each of said master devices includes
a means, when sending a request packet, to drive a selected bus line or lines with a certain current during at least one selected bus cycle,
a means for monitoring said selected bus line or lines for a greater than normal current to see if another master device is driving that line or lines,
a means for detecting said greater than normal current, and
a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 51. The bus subsystem with the features of embodiment 47 wherein said arbitration means comprises
a means for initiating an arbitration cycle,
a means for allocating a single bus line to each master device during at least one selected bus cycle relative to the start of said arbitration cycle,
a means for allocating each master device to a single bus line during one of said selected bus cycles if there are more master devices than available bus lines,
a means for each of said master devices which sent a colliding request packet to drive said bus line allocated to said master device during said selected bus cycle, and
a means in at least one of said master devices for storing information about which master devices sent a colliding request packet,
whereby said master devices can monitor selected bus lines during said arbitration cycle and identify each said master device which sent a colliding request packet.

Embodiment 52. The bus subsystem with the features of embodiment 47 wherein said arbitration means comprises
a means included in a first one of said master devices which sent colliding request packets for identifying each of said master devices which sent colliding request packets,
a means for assigning a priority to each said master device which sent a colliding request packet, and
a means for allowing each said master device which sent a colliding request packet to access the bus sequentially according to that priority.

Embodiment 53. The bus subsystem with the features of embodiment 52 wherein said priority is based on the physical location of each of said master devices.

Embodiment 54. The bus subsystem with the features of embodiment 52 wherein said priority is based on said master ID number of said master devices.

Embodiment 55. The bus subsystem with the features of embodiment 52 wherein each of said master devices includes
a means, when sending a colliding request packet, for deciding which master device can send the next request packet in what order or at what time, whereby no master device may send a new request packet until responses to each pending request packet have been completed or scheduled.

Embodiment 56. A bus subsystem comprising
a plurality of semiconductor devices connected in parallel to a bus,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including device-select information,
said bus containing substantially fewer lines than the number of bits in a single address,
said bus carrying said device-select information without the need for separate device-select lines connected directly to individual semiconductor devices,
said semiconductor devices including a reset means having an input and an output, the output of the reset means of one semiconductor device being connected to the input of the reset means of the next semiconductor device in series.

Embodiment 57. The bus subsystem with the features of embodiment 56 further including system reset means comprising
a means for generating a first and a second reset signal,
a means for passing said first reset signal to a first of said semiconductor devices and then to subsequent ones of said semiconductor devices in series and
a means for passing a second reset signal to said first semiconductor device and then to said subsequent semiconductor devices in series,
said bus subsystem including one of said semiconductor devices containing
a device identification register adapted to contain a number unique to said semiconductor device within said bus subsystem,
a device identification register setting means, and
a device reset means for resetting said semiconductor device to some desired, known reset state in response to said

Petitioners
Ex. 1002, p. 334

first reset signal and for setting said device identification register in response to said second reset signal,
whereby said bus subsystem can be reset to a known reset state with a unique device identification value in said device identification register of each of said semiconductor devices.
Embodiment 58. The bus subsystem with the features of embodiment 57 wherein said desired, known reset state is where all registers in the semiconductor device are cleared and the state machines are reset.
Embodiment 59. The bus subsystem with the features of embodiment 57 wherein said device identification register setting means comprises
a means for detecting said second reset signal,
a means for reading a device identification number from said bus lines at a specific time relative to said second reset signal and
a means for storing said device identification number in said device identification register of said semiconductor device.
Embodiment 60. The bus subsystem with the features of embodiment 57 wherein said second reset signal comprises multiple pulse sequences and wherein said device identification setting means includes
a means for interpreting said pulse sequences as a device identification number and
a means for storing said device identification number in said device identification register of said semiconductor device.
Embodiment 61. The bus subsystem with the features of embodiment 57 wherein said device reset means comprises an n-stage shift register capable of storing n-bit values, wherein said device reset means interprets a specific value in said shift register as said first reset signal and interprets a specific value in said shift register as said second reset signal.
Embodiment 62. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices is a master device, said master device including a means for generating said first and said second reset signals.
Embodiment 63. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices is a master device, said master device including
a master ID register,
a means for assigning a master ID number to said master device and
a means for storing said master ID number in said master ID register.
Embodiment 64. The bus subsystem with the features of embodiment 63 further comprising a second one of said master devices, and a means for a first one of said master devices to assign a master ID number to substantially all other said master devices, whereby said first master device assigns one of said master ID numbers to each of said master devices on said bus subsystem and each said master device stores said assigned master ID number in said master ID register.
Embodiment 65. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices includes a device-type register adapted to contain an identifier characteristic of that type of semiconductor device, and one or more modifiable registers, at least one of which is an access-time register adapted for storing access times.
Embodiment 66. The bus subsystem with the features of embodiment 65 wherein one of said semiconductor devices is a master device having
a means for selecting a semiconductor device,
a means for reading said device-type register of said selected semiconductor device,
a means for determining the device type of said selected semiconductor device,
a means for determining access-time values appropriate for said selected semiconductor device and for storing said access-time values in said access-time registers of said selected semiconductor device, and
a means for selecting and storing other values appropriate for said selected semiconductor device in corresponding registers of said selected semiconductor device,
whereby said master device can select a semiconductor device, determine what type it is, and set said access-time and other registers to contain appropriate values.
Embodiment 67. The bus subsystem with the features of embodiment 66 further comprising a memory device having at least one discrete memory section and at least one modifiable address register adapted to store memory address information which corresponds to each of said discrete memory sections, and
said master device further comprising a means for selecting and testing each of said discrete memory sections and a means for storing address information in said address registers corresponding to each of said discrete memory sections,
whereby said master device can test all said discrete memory sections and assign unique address values thereto.
Embodiment 68. A bus subsystem comprising
two semiconductor devices connected in parallel to a bus, one of said semiconductor devices being a master device,
said bus including a plurality of bus data lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including device-select information,
said bus containing substantially fewer of said bus data lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual semiconductor devices,
wherein all of said bus data lines are terminated transmission lines and all of said address, data and control information

Petitioners
Ex. 1002, p. 335

is carried on said bus data lines as a sequential series of bits in the form of low-voltage-swing signals.

Embodiment 69. The bus subsystem with the features of embodiment 68 further comprising a semiconductor device including a current-mode driver connected to drive one of said bus data lines.

Embodiment 70. The bus subsystem with the features of embodiment 69 further comprising a semiconductor device having a means to measure the voltage of said low-voltage-swing signals on a selected one of said bus data lines, whereby said semiconductor device can determine whether zero, one, or more than one of said current-mode drivers are driving said selected bus data line.

Embodiment 71. The bus subsystem with the features of embodiment 70 further comprising a semiconductor device having

a plurality of input receivers connected to one of said bus data lines, and

a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits.

Embodiment 72. The bus subsystem with the features of embodiment 70 further comprising a semiconductor device having two input receivers connected to one of said bus data lines.

Embodiment 73. A bus subsystem comprising

two semiconductor devices connected in parallel to a bus having a first and a second end, said bus including a bus clock line, said bus clock line having first and second ends corresponding to said first and second ends of said bus, respectively,

a clock generator connected to said first end of said bus clock line to generate early bus clock signals with a normal rise time, and

signal return means at said second end of said bus clock line to return said early bus clock signals to said first end of said bus as corresponding late bus clock signals,

whereby each of said early bus clock signals will propagate from said clock generator along said clock line starting from said first end to said second end of said bus and then return at a later time to said first end of said bus as a corresponding late bus clock signal, whereby each semiconductor device on said bus can detect said early bus clock signals and said corresponding late bus clock signals.

Embodiment 74. The bus subsystem with the features of embodiment 73 further comprising a first and a second said bus clock line having first and second ends at said first and said second ends of said bus, respectively, wherein said signal return means directly connects said second ends of said first and said second bus clock lines whereby each of said early bus clock signals will propagate from said clock generator at said first end of said bus along said first bus clock line to said second end of said bus and then return on said second bus clock line to said first end of said bus as one of said corresponding late bus clock signals.

Embodiment 75. The bus subsystem with the features of embodiment 73 wherein said signal return means comprises said first bus clock line without a line terminator at said second end thereof whereby each of said early bus clock signals reaching said second end of said first bus clock line will be reflected back along said first bus clock line as said corresponding late bus clock signals.

Embodiment 76. The bus subsystem with the features of embodiment 73 further comprising

a means for operating said bus in bus cycles timed to have a certain bus cycle frequency and a corresponding bus cycle period and

a means for operating said clock generator with a period of twice the bus cycle period.

Embodiment 77. The bus subsystem with the features of embodiment 76 wherein said bus cycle frequency is greater than approximately 50 MHz and less than or equal to approximately 500 MHz.

Embodiment 78. The bus subsystem with the features of embodiment 73 further including a semiconductor device having an internal device clock generating means to derive the midpoint time between said early and corresponding late bus clock signals and to generate an internal device clock synchronized to said midpoint time.

Embodiment 79. The bus subsystem with the features of embodiment 73 further including a semiconductor device having a low-skew clock generator circuit comprising

a first delay line having an input, an output and a basic delay and means for synchronizing the output of said first delay line with said early bus clock signal,

a second delay line having said basic delay plus a variable delay, said second delay line having an output and a means for synchronizing the output of said second delay line with said late bus clock signal, and

a third delay line having a third delay and a means to set said third delay midway between the delays of said first and second delay lines, said third delay line having an output which provides an internal device clock signal synchronized to a time halfway between said early and said late bus clock signals.

Embodiment 80. The bus subsystem with the features of embodiment 73 wherein said early and said late bus clock signals are low-voltage-swing signals that transition cyclically between low and high logical values, and further including a semiconductor device having a low-skew clock generator circuit comprising

a DC amplifier to convert said early and said late bus clock signals into full-swing logic signals,

a first variable delay line having a first variable delay and an input and an output, the input of said first variable delay

Petitioners
Ex. 1002, p. 336

line being connected to said DC amplifier
a first, a second and a third additional delay line, each having an input and an output, the input of each of said additional delay lines being connected to the output of said first delay line,
said first additional delay line having a fixed delay,
said second additional delay line having said fixed delay plus a second variable delay, and
said third additional delay line having said fixed delay plus one half of said second variable delay,
a first clocked input receiver connected to sample said early bus clock signal and gated by said output of said first additional delay line,
a means for adjusting said first variable delay so said first clocked input receiver samples said early bus clock signal just as said early bus clock signal transitions,
a second clocked input receiver connected to sample said late bus clock signal and gated by said output of said second additional delay line,
a means for adjusting said second variable delay so said second clocked input receiver samples said late bus clock signal just as said late bus clock signal transitions,
whereby said output of said third additional delay line is synchronized to a time halfway between said outputs of said first and said second additional delay lines, and said output of said third additional delay line provides an internal device clock signal.

Embodiment 81. The bus subsystem with the features of embodiment 80 further comprising a semiconductor device having
a first one of said low-skew clock generator circuits which generates a "true" internal device clock signal and
a second one of said low-skew clock generator circuits connected to generate a "complement" internal device clock signal synchronized with but opposite in logical value to said "true" internal device clock signal.

Embodiment 82. A DRAM device designed to be connected to an external bus having a plurality of bus lines for carrying substantially all address, data and control information needed by said DRAM device as a sequential series of bits, said control information including device-select information, said external bus containing substantially fewer said bus lines than the number of bits in a single address, and said bus carrying device-select information without the need for separate device-select lines connected directly to said DRAM device, said DRAM device comprising
an array of memory cells connected in rows and columns, each of said memory cells adapted to store one of said bits,
a row address selection means for selecting one of said rows,
a column sense amp connected to each of said columns, each of said column sense amps adapted to latch one of said bits as a binary logical value or to precharge to a selected state,
a column decoding means connected to each of said column sense amps for selecting a plurality of said column sense amps for inputting one of said bits to or outputting one of said bits from said memory cells,
an internal I/O bus having a plurality of internal I/O lines wherein each of said internal I/O lines is connected to a plurality of said column sense amps, and
a plurality of bus connection means designed to connect said internal I/O lines to said external bus,
whereby a selected bit of said sequential series of bits can be transferred from said external bus to a selected one of said memory cells or said bit contained in a selected one of said memory cells can be transferred to said external bus.

Embodiment 83. The DRAM device with the features of embodiment 82 further comprising
an output driver connected to one said bus connection means,
an output multiplexer having an output connected to said output driver and a plurality of inputs, each of said inputs being connected to one of said internal I/O lines, and
a control means to select whether said output driver can drive said external bus,
whereby a plurality of memory cells are selected using said row address selection means and said column decoding means and a plurality of bits contained in said plurality of memory cells are output through said column sense amps to said internal I/O bus to said output multiplexer to said output driver to said external bus.

Embodiment 84. The DRAM device with the features of embodiment 82 further comprising
a plurality of input receivers connected to one of said bus data lines and to said internal I/O bus,
a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits, and
a control means to select whether an input receiver can drive said internal I/O bus, whereby a bit of said sequential series of bits is input from said external bus through one of said input receivers to one of said internal I/O lines to one of said column sense amps to one of said memory cells.

Embodiment 85. The DRAM device with the features of embodiment 82 further comprising
a first and a second half-array of said memory cells wherein each said row of said array of said memory cells is subdivided into two parts,
a first and a second one of said internal I/O buses connected to said column sense amps in said first and said second half-arrays, respectively, and

Petitioners
Ex. 1002, p. 337

a column decoder means to gate selected ones of said column sense amps connected to said memory cells in a selected row of said first and said second half-arrays simultaneously.

Embodiment 86. The DRAM device with the features of embodiment 85 wherein said column decoder means selects sixteen column sense amps at a time.

Embodiment 87. The DRAM device with the features of embodiment 82 wherein said external bus operates at a certain speed and wherein said DRAM device includes four of said internal I/O buses, each of which operates at one-fourth the speed of said external bus.

Embodiment 88. The DRAM device with the features of embodiment 82 further comprising

a means for precharging one of said column sense amps to a precharged state from which a binary logical value can quickly be loaded into said column sense amp,

if said column sense amp contains a binary logical value, a means for latching the logical value currently contained in said column sense amp and

a means for instructing said DRAM device to precharge said column sense amp or latch said binary logical value in said column sense amp.

Embodiment 89. The DRAM device with the features of embodiment 88 further comprising a means for instructing said DRAM device to precharge said column sense amp without further instruction whenever said row address selection means selects a different one of said rows.

Embodiment 90. The DRAM device with the features of embodiment 88 further comprising a means for instructing said DRAM device to precharge said column sense amp without further instruction at a first or a second preselected time after latching the latest said binary logical value, said first preselected time being long enough for said DRAM to latch said binary logical value into said column sense amp and transfer said binary logical value into memory or onto one of said internal I/O lines, and said second preselected time being a variable which can be stored in said DRAM device whereby said DRAM can latch a binary logical value into said column sense amp for transferring said binary logical' value into or out of a selected said memory cell, then precharge to allow a faster subsequent read or write.

Embodiment 91. A package containing

a semiconductor die having a side, circuitry and a plurality of connecting areas positioned along or near said side, spaced at a selected pitch and connected to said circuitry,

said package comprising a plurality of bus connecting means for connecting to a plurality of external bus lines, each of said external bus lines corresponding to one of said connecting areas, each of said bus connecting means being positioned on a first side of said package,

connected to one said external bus line and to said corresponding connecting area on said semiconductor die, and spaced at a pitch substantially identical to said selected pitch of said connecting areas,

whereby each of said external bus lines can be connected to said corresponding connecting area on said semiconductor die by bus connection means positioned along a single side of said package.

Embodiment 92. The package with the features of embodiment 91 further comprising a plurality of said bus connecting means wherein each of said bus connecting means includes

a pin adapted for connection to one of said external bus lines and

a wire connecting said pin to one of said connecting areas on said semiconductor die,

said wire having an effective lead length less than about 4 millimeters and wherein the effective lead length of said wire of each of said bus connection means for said package is approximately equal.

Embodiment 93. A plurality of packages with the features of embodiment 91 wherein at least two of said semiconductor die are memory devices, each of said packages being generally flat, having a top and a bottom, and wherein

said packages are physically secured adjacent and parallel to each other in a stack,

where a first one of said packages is adjacent to a second one of said packages in said stack, said top of said first package is substantially aligned with said bottom of said second package, and

said bus connecting means of each of said packages are substantially aligned and are lying substantially in a plane.

Embodiment 94. The plurality of packages with the features of embodiment 93 further comprising a plurality of stacks wherein each of said bus connecting means can be electrically connected to corresponding said bus connecting means in each of said stacks.

Embodiment 95. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

at least one modifiable identification register accessible to said bus through said connection means, whereby data may

Petitioners
Ex. 1002, p. 338

be transmitted to said register via said bus and enable said device thereafter to be uniquely identified.

Embodiment 96. The semiconductor device with the features of embodiment 95 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 97. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and at least one modifiable register to hold device address information, said modifiable register accessible to said bus through said connection means, whereby data may be transmitted to said register via said bus which enables said device thereafter to respond to a predetermined range of addresses.

Embodiment 98. The semiconductor device with the features of embodiment 97 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 99. The semiconductor device with the features of embodiment 98 wherein said memory device has at least one discrete memory section and also has at least one modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 100. The semiconductor device with the features of embodiment 99 wherein said memory address information comprises a pointer to said discrete memory section.

Embodiment 101. The semiconductor device with the features of embodiment 100 wherein said discrete memory section has a top and a bottom and said memory address information comprises pointers to said top and said bottom.

Embodiment 102. The semiconductor device with the features of embodiment 100 wherein said memory address information comprises

a pointer to said discrete memory section and

a range value indicating the size of said discrete memory section.

Embodiment 103. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

at least one modifiable access-time register accessible to said bus through said connection means, whereby data may be transmitted to said register via said bus which establishes a predetermined amount of time that said semiconductor device thereafter must wait before using said bus in response to a request.

Embodiment 104. The semiconductor device with the features of embodiment 103 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 105. The semiconductor device with the features of embodiment 103 further comprising at least two access-time registers and one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said bus.

Embodiment 106. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, and wherein each said bus line is a terminated transmission line, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

a bus line driver capable of producing a low-voltage-swing signal on one of said terminated transmission lines.

Embodiment 107. The semiconductor device with the features of embodiment 106 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 108. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying

Petitioners
Ex. 1002, p. 339

substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said bus further including at least one bus clock line for carrying early and late bus clock signals, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

an internal device clock generating means which generates an internal device clock synchronized to a time halfway between said early and said late bus clock signals.

embodiment 109. The semiconductor device with the features of embodiment 108 wherein said bus further includes a first and a second one of said bus clock lines, said first bus clock line carries said early bus clock signal and said second bus clock line carries said late bus clock signal, said semiconductor device further comprising a means to detect said early bus clock signal on said first bus clock line and a means to detect said late bus clock signal on said second bus clock line.

Embodiment 110. The semiconductor device with the features of embodiment 109 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 111. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying as a sequential series of bits substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus,

a plurality of input receivers connected to one of said bus data lines and

a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits.

Embodiment 112. The semiconductor device with the features of embodiment 111 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 113. The semiconductor device with the features of embodiment 112 wherein two input receivers are connected to one of said bus lines.

Embodiment 114. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said bus system includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus, and

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus.

Embodiment 115. The semiconductor device with the features of embodiment 114 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 116. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus,

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and

Petitioners
Ex. 1002, p. 340

at least one modifiable identification register accessible to said system bus through said connection means, whereby data may be transmitted to said register via said system bus and which enables said device thereafter to be uniquely identified.

Embodiment 117. The semiconductor device with the features of embodiment 116 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 118. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus,

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and

at least one modifiable register to hold device address information, said modifiable register accessible to said system bus through said connection means, whereby data may be transmitted to said register via said system bus which enables said device thereafter to respond to a predetermined range of addresses.

Embodiment 119. The semiconductor device with the features of embodiment 118 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 120. The semiconductor device with the features of embodiment 119 wherein said memory device has at least one discrete memory section and also has at least one modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 121. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus,

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and

at least one modifiable access-time register accessible to said system bus through said connection means, whereby data may be transmitted to said register via said system bus which establishes a predetermined amount of time that said semiconductor device thereafter must wait before using said system bus in response to a request.

Embodiment 122. The semiconductor device with the features of embodiment 121 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 123. The semiconductor device with the features of embodiment 121 further comprising at least two access-time registers and one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said system bus.

Embodiment 124. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, wherein said address, data, control and device-select information is carried over said bus in the form of request packets and bus transactions, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus,

a means to receive said request packets over said bus,

a means to decode information in said request packets, and

a means to respond to said information in said request packets.

Embodiment 125. The semiconductor device with the features of embodiment 124 wherein said means to decode information in said request packet further comprises

Petitioners
Ex. 1002, p. 341

a means to identify and decode said control information in said request packet,
a means to identify and decode said device-select information in said request packet,
a means to identify and decode said address information in said request packet and
a means to determine whether said control information or said address information instructs said semiconductor device to begin a response.

Embodiment 126. The semiconductor device with the features of embodiment 124 wherein each of said bus transactions is carried out in response to said address and said control information in one of said request packets, and wherein said means to identify and decode information in said request packets includes a means to identify a sequence of bytes on said bus as one of said request packets containing said address and said control information, said control information including information about the type of said bus transaction being requested and the access time which needs to intervene before beginning said bus transaction over said bus and said address and said control information includes device-select information instructing one or more said semiconductor devices to respond to said address and said control information.

Embodiment 127. The semiconductor device with the features of embodiment 124 further comprising
a plurality of sense amplifiers adapted to precharge to a selected state or to latch a bit of information,
a means to hold said sense amplifiers in an unmodified state after latching one of said bits of information,
a means to precharge said sense amplifiers and
a means for selecting whether said semiconductor device should precharge said sense amplifiers or should hold said sense amplifiers in an unmodified state.

Embodiment with the features of embodiment 128. The semiconductor device 124 wherein said means to respond to said information, where said information is control information, further comprises a means to
transfer a data block during a data block transfer, further including a means to
read data from said semiconductor device and
write data into said semiconductor device, and
initiate a data block transfer,
transfer a data block of a selected size,
transfer a data block at a selected time,
access a control register, including a means to read from or write to said control register, or
select normal or page-mode access.

Embodiment 129. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes a device identification number unique to said semiconductor device.

Embodiment 130. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes a special device identification number which calls for said semiconductor device to respond.

Embodiment 131. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes an address unique to said semiconductor device.

Embodiment 132. The semiconductor device with the features of embodiment 124 further comprising a means to interpret said control information and decode the time to wait before beginning said bus transaction over said bus.

Embodiment 133. The semiconductor device with the features of embodiment 124 further comprising a means to interpret said control information and decode the size of a data block to transfer during one of said bus transactions.

Embodiment 134. The semiconductor device with the features of embodiment 124, 125, 126, 127, 128, 129, 130, 131, 132 or 133 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 135. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, wherein said address, data, control and device-select information is carried over said bus in the form of request packets and bus transactions, said semiconductor device comprising
connection means adapted to connect said semiconductor device to said bus,
a means to encode address and control information in said request packets and
a means to send said request packets over said bus.

Embodiment 136. The semiconductor device 135 further comprising a means to request a bus transaction wherein each of said bus transactions is carried out in response to said address and said control information in one of said request packets, and wherein said means to encode information in said request packets includes a means to mark a sequence

Petitioners
Ex. 1002, p. 342

of bytes on said bus as one of said request packets, said control information including information about the type of said bus transaction being requested and the access time which needs to intervene before beginning said bus transaction over said bus and said address and said control information includes device-select information instructing one or more said semiconductor devices to respond to said address and said control information.

Embodiment 137. The semiconductor device with the features of embodiment 135 wherein one or more of said plurality of semiconductor devices has a unique device identification number, said semiconductor device further comprising a means to send control information to a specific one of said plurality of semiconductor devices by including in said request packet a selected said device identification number.

Embodiment 138. The semiconductor device with the features of embodiment 135 wherein each of said plurality of semiconductor devices is adapted to respond to a special device identification number, said semiconductor device further comprising a means to send control information to each of said plurality of semiconductor devices by including in said request packet said special device identification number.

Embodiment 139. The semiconductor device with the features of embodiment 135 wherein one or more of said plurality of semiconductor devices is a memory device having a plurality of addresses, said semiconductor device further comprising a means to send control information to a specific address or range of addresses in one of said plurality of semiconductor devices by including said specific address or range of addresses in said request packet.

Embodiment 140. The semiconductor device with the features of embodiment 135 wherein at least one of said request packets is a request packet requesting a bus transaction which is followed by a corresponding one of said bus transactions, said semiconductor device further comprising a means to encode said control information to specify directly or indirectly the time between the end of said request packet requesting a bus transaction and said corresponding bus transaction over said bus.

Embodiment 141. The semiconductor device with the features of embodiment 140 wherein one type of said bus transactions is a transfer of a data block, said semiconductor device further comprising a means to encode said control information to specify the size of said data block to transfer.

Embodiment 142. The semiconductor device with the features of embodiment 140 further comprising a means to keep track of current and pending bus transactions, whereby collisions on said bus are avoided because said semiconductor device avoids initiating bus transactions which would conflict with current or pending bus transactions.

Embodiment 143. The semiconductor device with the features of embodiment 135 wherein said semiconductor device is a first master device and one of said plurality of semiconductor devices is a second master device, further comprising a collision detecting means whereby said first master device when sending a first one of said request packets can detect said second master device sending a colliding one of said request packets, where said colliding request packet may be sent simultaneous with the initial sending of or overlapping the sending of said first request packet, and

an arbitration means whereby said first and said second master devices select a priority order in which each of said master devices will be allowed to access said bus sequentially.

Embodiment 144. The semiconductor device with the features of embodiment 143 wherein said semiconductor device is a master device and at least one of said plurality of semiconductor devices is a master device, each of said master devices has a master ID number and each of said request packets includes a master ID position which is a predetermined number of bits in a predetermined position in said request packet, and wherein said collision detection means comprises a means for said semiconductor device to send its master ID number in said request packet and

a means to detect a collision and invoke said arbitration means if said semiconductor device detects any other master ID number in said master ID position.

Embodiment 145. The semiconductor device with the features of embodiment 144 wherein said system bus architecture includes a means for carrying information on said bus during bus cycles, said semiconductor device further comprising a means for driving a selected bus line or lines during at least one selected bus cycle while sending each said request packet,

a means for monitoring said selected bus line or lines to see if another said master device is sending one of said colliding request packets and

a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 146. The semiconductor device with the features of embodiment 145 further comprising

a means, when sending a request packet, for driving a selected bus line or lines with a certain current during at least one selected bus cycle,

a means for monitoring said selected bus line or lines for a greater than normal current to see if another said master device is driving that line or lines,

a means for detecting said greater than normal current, and

a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 147. The semiconductor device with the features of embodiment 143 wherein said arbitration means comprises

a means for initiating an arbitration cycle,

Petitioners
Ex. 1002, p. 343

a means for allocating a single bus line to each said master device during at least one selected bus cycle relative to the start of said arbitration cycle,
a means for allocating each said master device to a single bus line during one of said selected bus cycles if there are more master devices than available bus lines,
a means for each of said master devices which sent one of said colliding request packets to drive said bus line allocated to said master device during said selected bus cycle, and
a means in at least one of said master devices for storing information about which master devices sent one of said colliding request packets,
whereby said master devices can monitor selected bus lines during said arbitration cycle and identify each said master device which sent one of said colliding request packets.

Embodiment 148. The semiconductor device with the features of embodiment 143 wherein said arbitration means comprises
a means for identifying each of said master devices which sent one of said colliding request packets,
a means for assigning a priority to each said master device which sent one of said colliding request packets, and
a means for allowing each said master device which sent one of said colliding request packets to access the bus sequentially according to that priority.

Embodiment 149. The semiconductor device with the features of embodiment 143 wherein said priority is based on the physical location of each of said master devices.

Embodiment 150. The semiconductor device with the features of embodiment 143 wherein said priority is based on said master ID number of said master devices.

Claims

1. A semiconductor device, the semiconductor device comprising:

connection means adapted to connect the semiconductor device to an external bus which is a part of a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to the external bus, wherein the external bus includes a plurality of bus lines for carrying substantially all address, data, control and device select information needed by the semiconductor device for communication with substantially every other semiconductor device connected to the external bus, and wherein the external bus has substantially fewer bus lines than the number of bits in a single address, wherein each bus line of the plurality of bus lines is a terminated transmission line, and wherein the external bus carries device-select information for the semiconductor device without the need for a separate device-select line connected directly to the semiconductor device,
a bus line driver capable of producing a low-voltage-swing signal on one of the terminated transmission lines,

wherein the semiconductor device is a semiconductor transceiver device to interface multiple units to the bus, wherein the multiple units are DRAMs.

2. The semiconductor device of claim 1,
wherein the bus line driver is capable of producing a low-voltage-swing signal having a voltage swing of not more than 1 Volt on one of said terminated transmission lines.

3. The semiconductor device of claim 1 or 2,
wherein the bus line driver is capable of producing a low-voltage-swing signal having a voltage swing of at least 500 mV on one of said terminated transmission lines.

Petitioners
Ex. 1002, p. 344

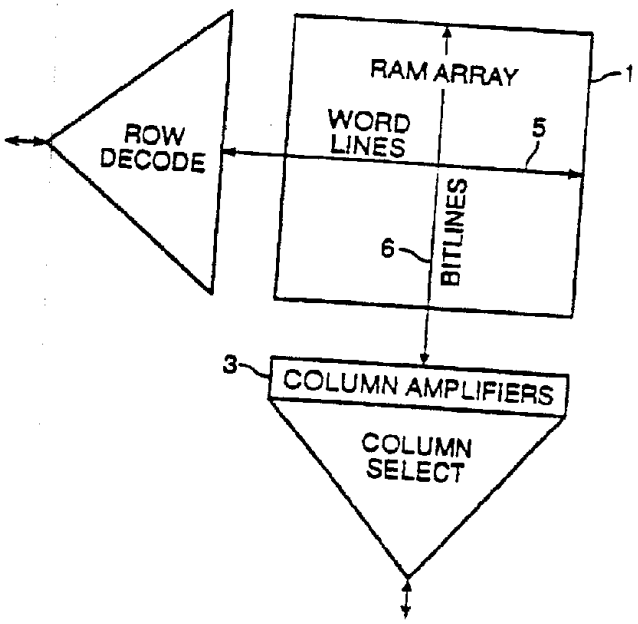


FIG. 1

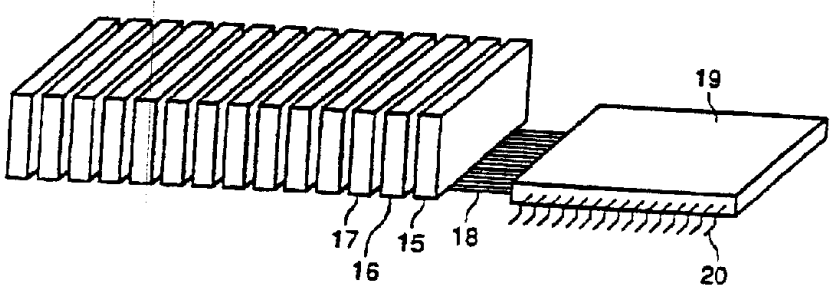


FIG. 3

Petitioners
Ex. 1002, p. 345

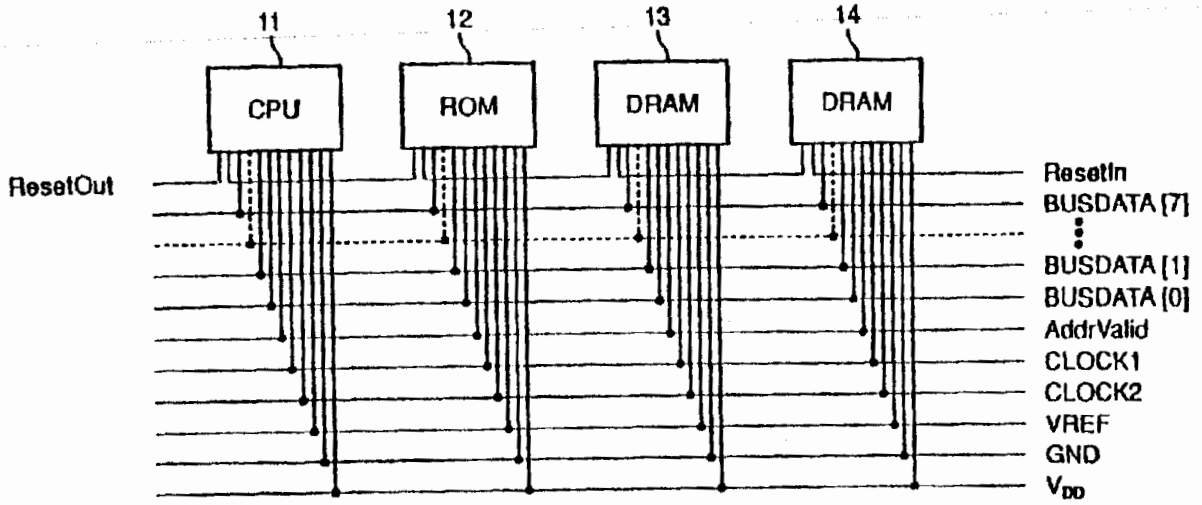


FIG. 2

Petitioners

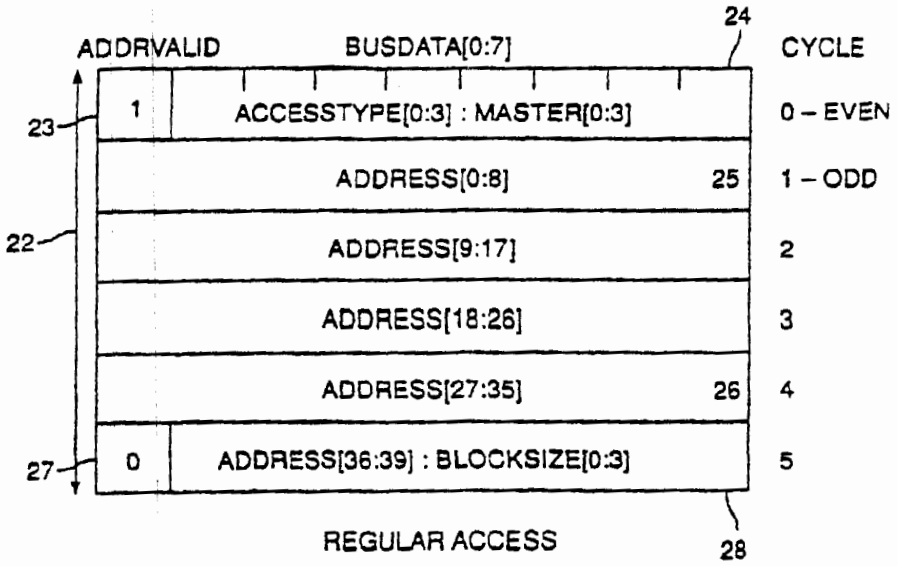


FIG. 4

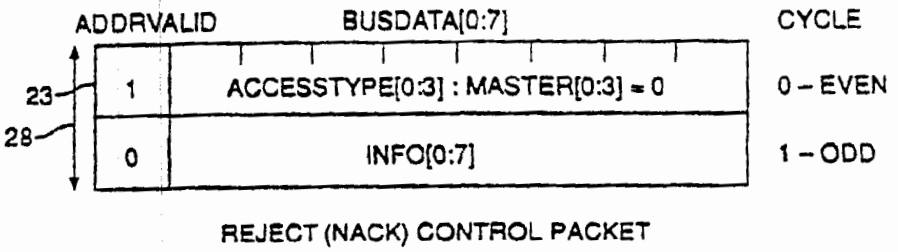


FIG. 5

Petitioners
Ex. 1002, p. 847

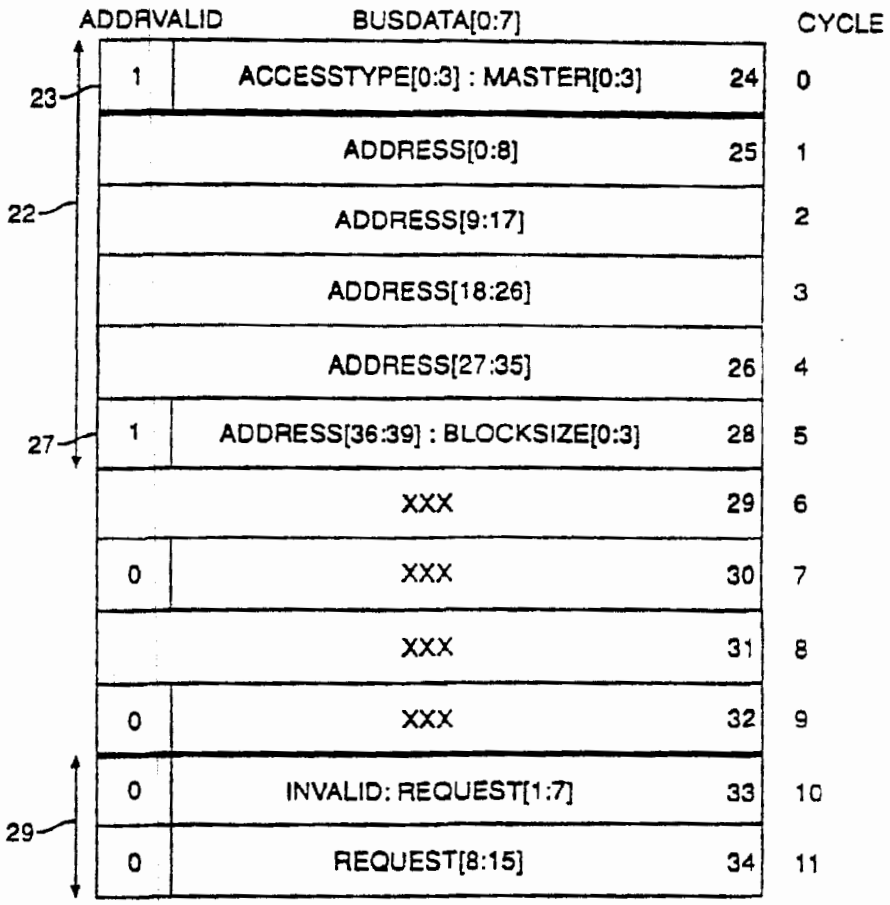


FIG. 6

Petitioners
Ex. 1002, p. 348

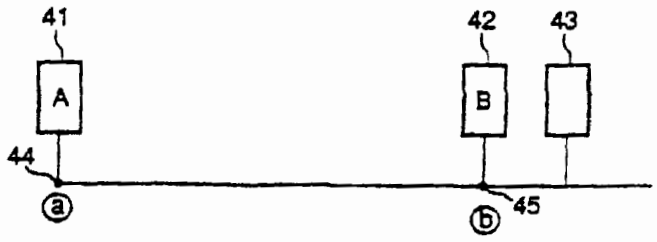


FIG. 7a

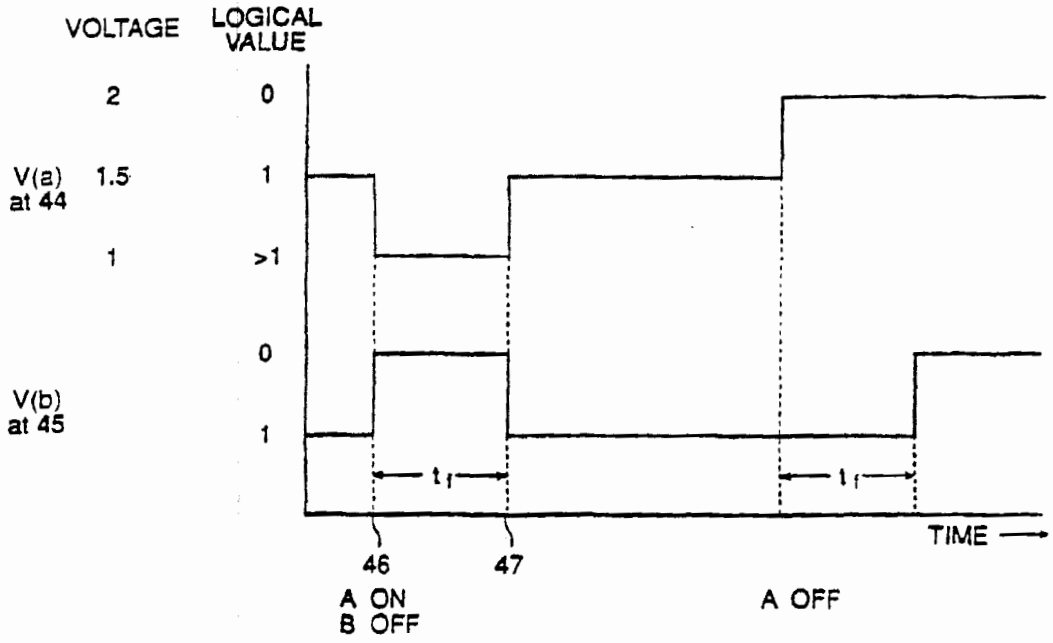


FIG. 7b

Petitioners
Ex. 1002, p. 349

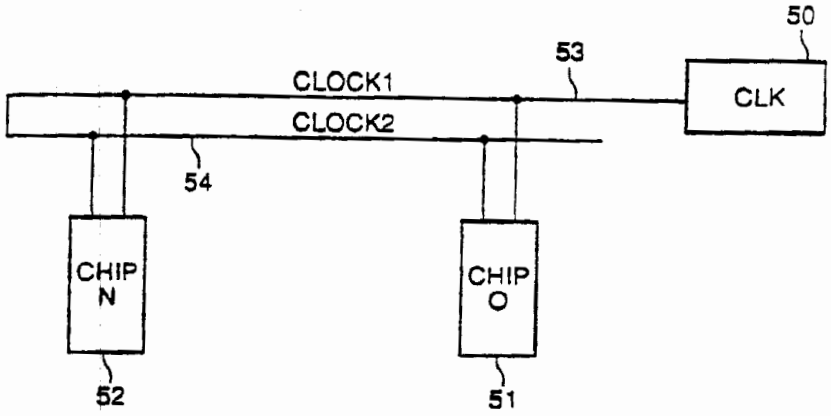


FIG. 8a

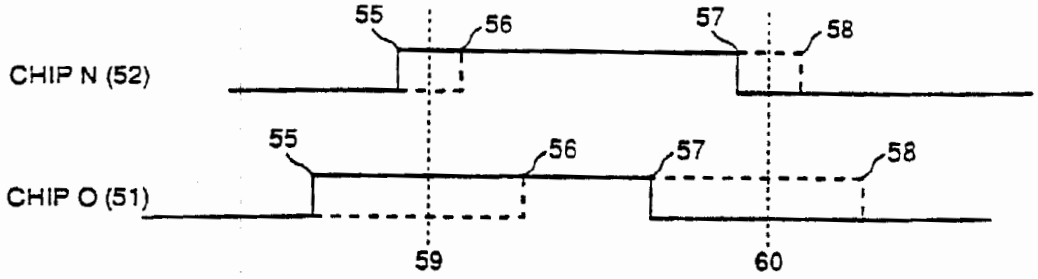


FIG. 8b

Petitioners
Ex. 1002, p. 350

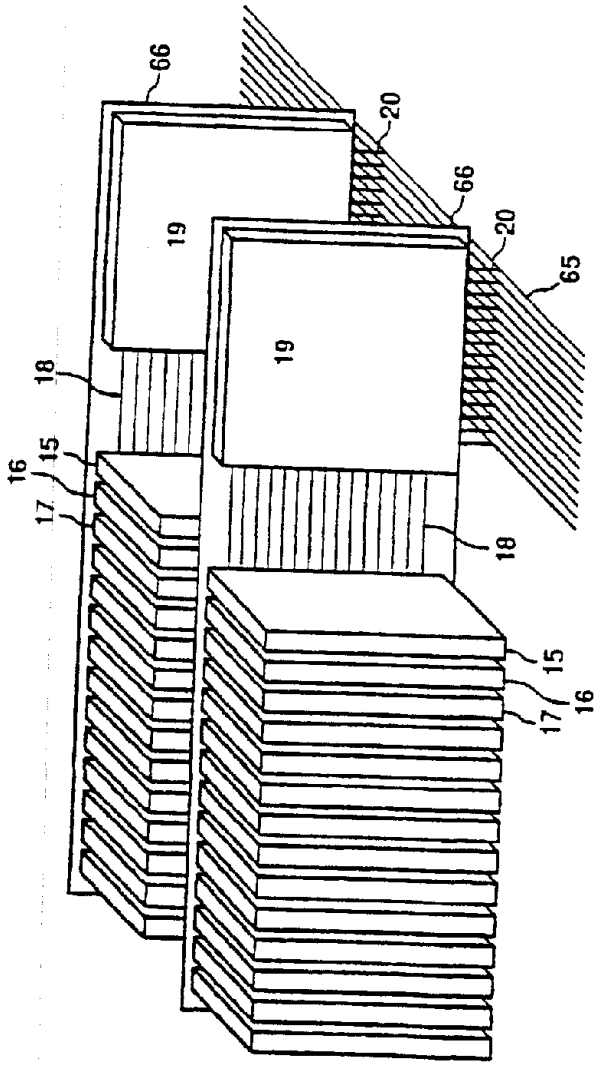


FIG. 9

Petitioners
Ex. 1002, p. 351

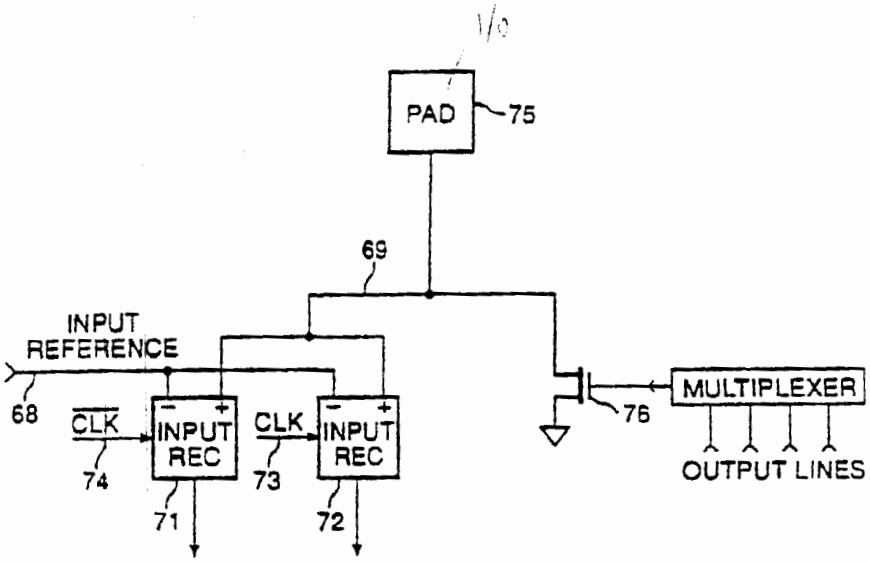


FIG. 10

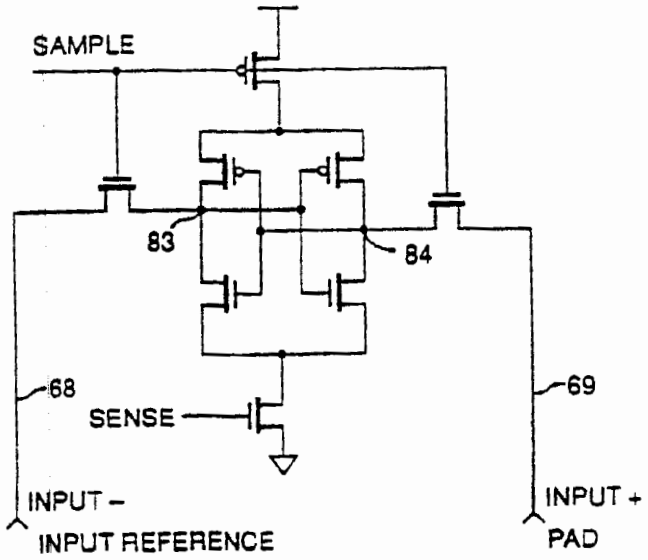


FIG. 11

Petitioners

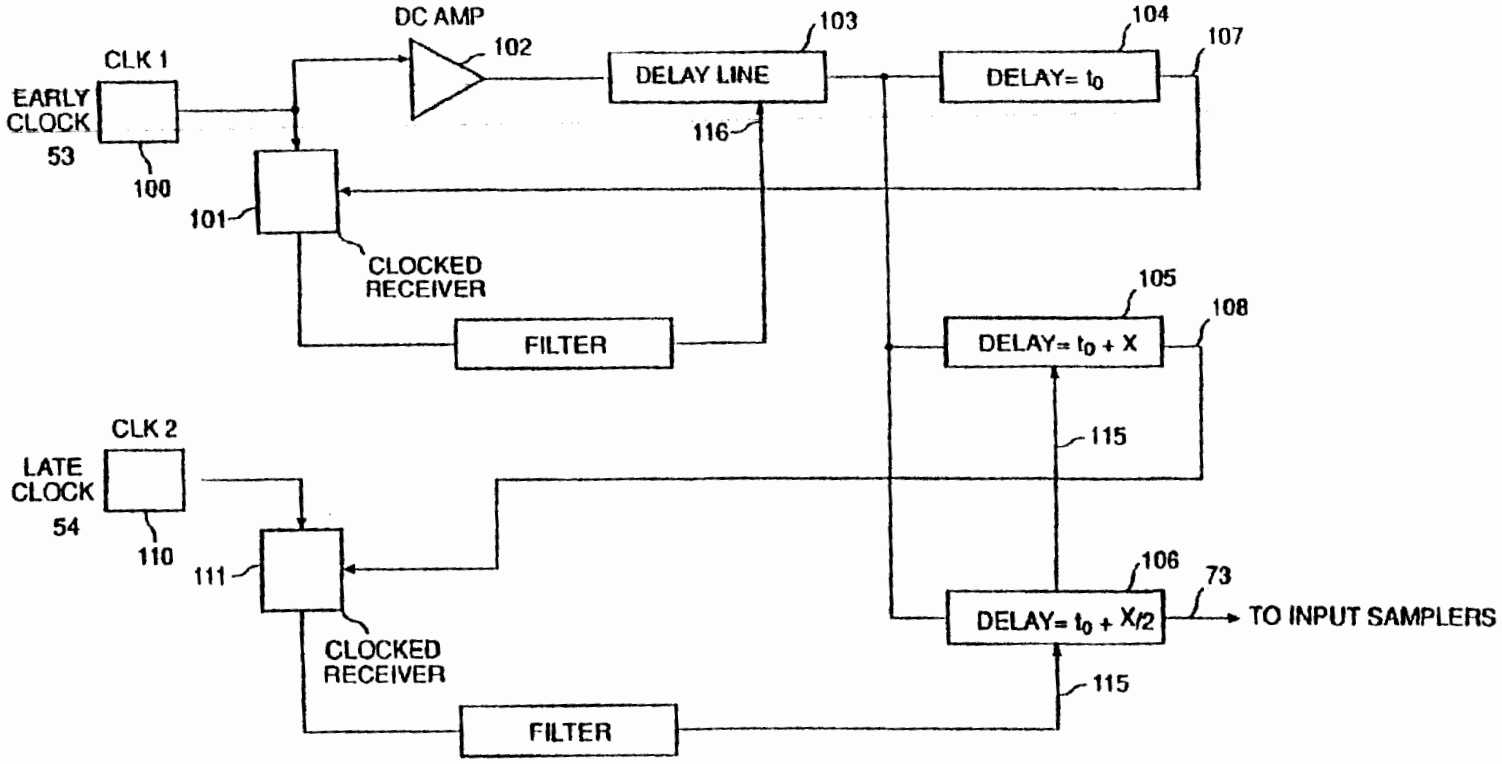


FIG. 12

Petitioners
EX. 1002, p. 353

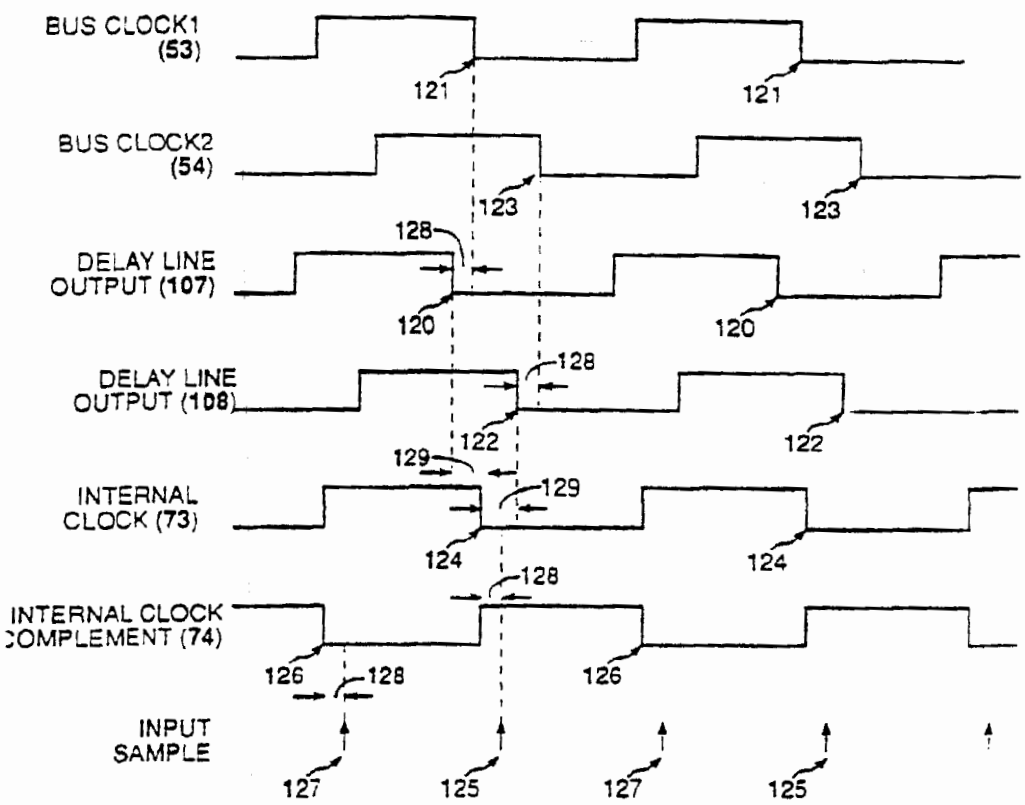


FIG. 13

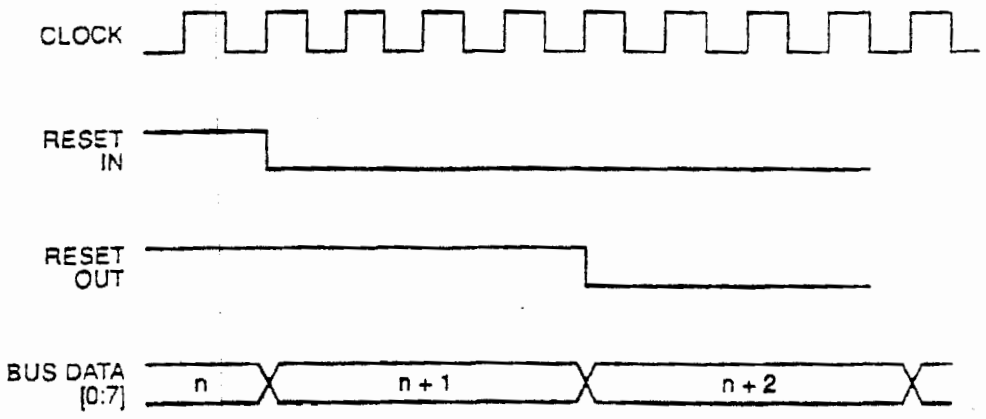


FIG. 14

Petitioners
Ex. 1002, p. 354

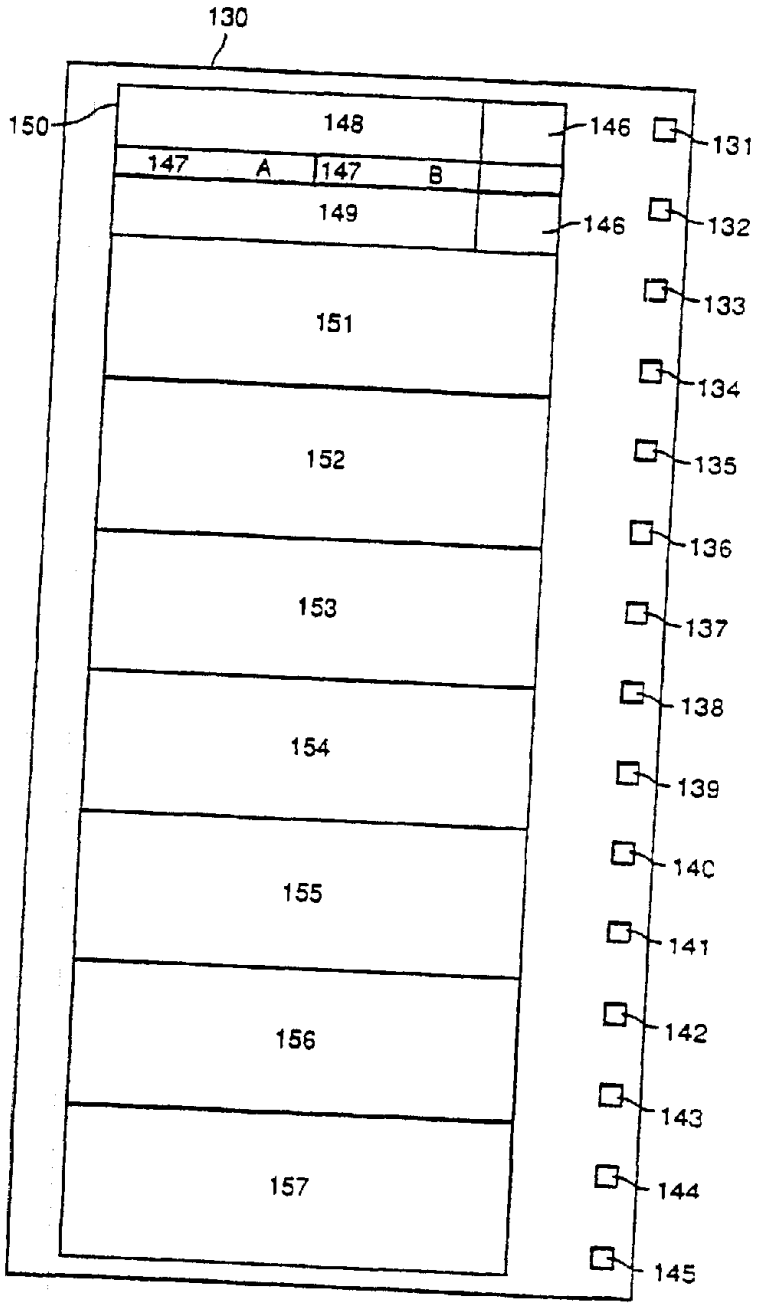


FIG. 15

Petitioners
Ex. 1002, p. 355

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 3821715 A, Hoff **[0005]**
- US 4315308 A, Jackson **[0006]**
- US 4449207 A, Kung **[0007]**
- US 4764846 A **[0008]**
- US 4706166 A, Go **[0008]**
- US 3969706 A, Proebsting **[0009]**
- US 4595923 A **[0010]**
- US 4481625 A, Roberts **[0010]**
- US 4837682 A, Culler **[0011]**
- US 4818985 A, Ikeda **[0011]**
- US 4779089 A, Theus **[0011]**
- US 4745548 A, Blahut **[0011]**
- US 4247817 A, Heller **[0012]**
- US 4646279 A, Voss **[0013]**

Non-patent literature cited in the description

- **ANANT AGARWAL.** An Analytic Cache Model. *ACM Transactions on Computer Systems,* May 1989, vol. 7 (2), 184-215 **[0055]**
- **ANANT AGARWAL et al.** Scaleable Directory Schemes for Cache Consistency. *15th International Symposium on Computer Architecture,* June 1988, 280-289 **[0083]**
- **HANS SCHUMACHER et al.** CMOS Subnanosecond True-ECL Output Buffer. *J. Solid State Circuits,* February 1990, vol. 25 (1), 150-154 **[0100]**

Petitioners
Ex. 1002, p. 356

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 816 570 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
15.08.2007 Bulletin 2007/33

(43) Date of publication A2:
08.08.2007 Bulletin 2007/32

(21) Application number: **06125954.5**

(22) Date of filing: **16.04.1991**

(51) Int Cl.:
G06F 1/04 (2006.01) *G06F 1/12* (2006.01)
G06F 13/00 (2006.01) *G06F 13/16* (2006.01)
G06F 13/36 (2006.01) *G06F 13/38* (2006.01)
G11C 7/00 (2006.01) *G11C 8/00* (2006.01)
G11C 8/04 (2006.01)

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK YU

(30) Priority: **18.04.1990 US 510898**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
00100018.1 / 1 022 641
91908374.1 / 0 525 068

(71) Applicant: **RAMBUS INC.**
Los Altos,
California 94022 (US)

(72) Inventors:
- **Farmwald, Michael**
  Portola Valley, CA 94028 (US)
- **Horowitz, Mark**
  Palo Alto, CA 94306 (US)

(74) Representative: **Eisenführ, Speiser & Partner**
Patentanwälte Rechtsanwälte
Postfach 10 60 78
28060 Bremen (DE)

(54) Integrated circuit I/O using a high performance bus interface

(57) The present invention includes a semiconductor device, the semiconductor device comprising connection means adapted to connect the semiconductor device to an external bus which is a part of a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to the external bus, wherein the external bus includes a plurality of bus lines for carrying substantially all address, data, control and device select information needed by the semiconductor device for communication with substantially every other semiconductor device connected to the external bus, and wherein the external bus has substantially fewer bus lines than the number of bits in a single address, wherein each said bus line is a terminated transmission line, and wherein the external bus carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to the semiconductor device, a bus line driver capable of producing a low-voltage-swing signal on one of said terminated transmission lines, wherein the semiconductor device is a semiconductor transceiver device to interface multiple units to the bus, wherein the multiple units are DRAMs.

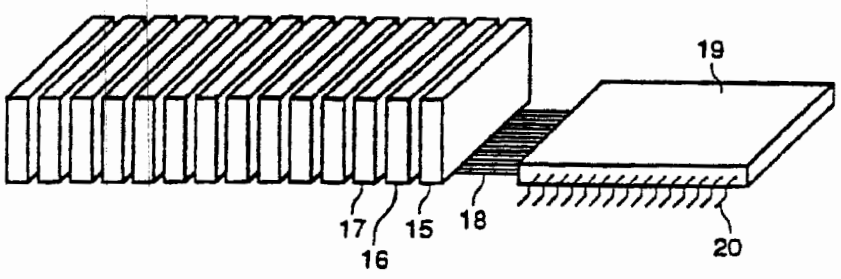


FIG. 3

EP 1 816 570 A3

Printed by Jouve, 75001 PARIS (FR)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 06 12 5954

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	GB 1 262 691 A (BURROUGHS CORP [US]) 2 February 1972 (1972-02-02) * page 2, line 56 - line 82 * * page 5, line 104 - line 119 * -----	1-3	INV. G06F1/04 G06F1/12 G06F13/00 G06F13/16
A	US 4 630 193 A (KRIS BRYAN) 16 December 1986 (1986-12-16) * column 1, line 58 - column 2, line 4 * * column 2, line 40 - column 5, line 33 * * abstract; figure 1 * -----	1-3	G06F13/36 G06F13/38 G11C7/00 G11C8/00 G11C8/04

TECHNICAL FIELDS SEARCHED (IPC)
G06F G11C

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	10 July 2007	NGUYEN XUAN HIEP, C

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Petitioners
Ex. 1002, p. 359

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 06 12 5954

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-07-2007

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
GB 1262691 A	02-02-1972	BE 741528 A	16-04-1970
		DE 1956515 A1	11-06-1970
		FR 2023120 A5	07-08-1970
		JP 52041604 B	19-10-1977
US 4630193 A	16-12-1986	EP 0077328 A1	27-04-1983
		WO 8203931 A1	11-11-1982

EPO FORM P0459

Petitioners
Ex. 1002, p. 360

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 816 570 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
08.08.2007 Bulletin 2007/32

(51) Int Cl.:
G06F 13/16 *(2006.01)*

(21) Application number: **06125954.5**

(22) Date of filing: **16.04.1991**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK YU

(30) Priority: **18.04.1990 US 510898**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
00100018.1 / 1 022 641
91908374.1 / 0 525 068

(71) Applicant: **RAMBUS INC.**
Los Altos,
California 94022 (US)

(72) Inventors:
- **Farmwald, Michael**
  Portola Valley, CA 94028 (US)
- **Horowitz, Mark**
  Palo Alto, CA 94306 (US)

(74) Representative: **Eisenführ, Speiser & Partner**
Patentanwälte Rechtsanwälte
Postfach 10 60 78
28060 Bremen (DE)

Remarks:
This application was filed on 12 - 12 - 2006 as a divisional application to the application mentioned under INID code 62.

(54) Integrated circuit I/O using a high performance bus interface

(57) The present invention includes a semiconductor device, the semiconductor device comprising connection means adapted to connect the semiconductor device to an external bus which is a part of a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to the external bus, wherein the external bus includes a plurality of bus lines for carrying substantially all address, data, control and device select information needed by the semiconductor device for communication with substantially every other semiconductor device connected to the external bus, and wherein the external bus has substantially fewer bus lines than the number of bits in a single address, wherein each said bus line is a terminated transmission line, and wherein the external bus carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to the semiconductor device, a bus line driver capable of producing a low-voltage-swing signal on one of said terminated transmission lines, wherein the semiconductor device is a semiconductor transceiver device to interface multiple units to the bus, wherein the multiple units are DRAMs.

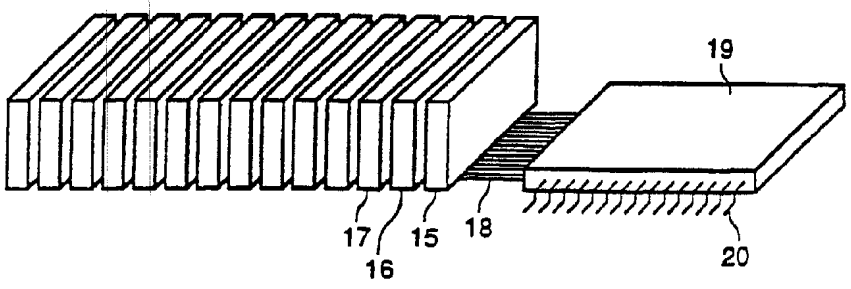


FIG. 3

EP 1 816 570 A2

Petitioners
Ex. 1002, p. 361

Description

FIELD OF THE INVENTION

[0001] An integrated circuit bus interface for computer and video systems is described which allows high speed transfer of blocks of data, particularly to and from memory devices, with reduced power consumption and increased system reliability. A new method of physically implementing the bus architecture is also described.

BACKGROUND OF THE INVENTION

[0002] Semiconductor computer memories have traditionally been designed and structured to use one memory device for each bit, or small group of bits, of any individual computer word, where the word size is governed by the choice of computer. Typical word sizes range from 4 to 64 bits. Each memory device typically is connected in parallel to a series of address lines and connected to one of a series of data lines. When the computer seeks to read from or write to a specific memory location, an address is put on the address lines and some or all of the memory devices are activated using a separate device select line for each needed device. One or more devices may be connected to each data line but typically only a small number of data lines are connected to a single memory device. Thus data line 0 is connected to device(s) 0, data line 1 is connected to device(s) 1, and so on. Data is thus accessed or provided in parallel for each memory read or write operation. For the system to operate properly, every single memory bit in every memory device must operate dependably and correctly.

[0003] To understand the concept of the present invention, it is helpful to review the architecture of conventional memory devices. Internal to nearly all types of memory devices (including the most widely used Dynamic Random Access Memory (DRAM), Static RAM (SRAM) and Read Only Memory (ROM) devices), a large number of bits are accessed in parallel each time the system carries out a memory access cycle. However, only a small percentage of accessed bits which are available internally each time the memory device is cycled ever make it across the device boundary to the external world.

[0004] Referring to Fig. 1, all modern DRAM, SRAM and ROM designs have internal architectures with row (word) lines 5 and column (bit) lines 6 to allow the memory cells to tile a two dimensional area 1. One bit of data is stored at the intersection of each word and bit line. When a particular word line is enabled, all of the corresponding data bits are transferred onto the bit lines. Some prior art DRAMs take advantage of this organization to reduce the number of pins needed to transmit the address. The address of a given memory cell is split into two addresses, row and column, each of which can be multiplexed over a bus only half as wide as the memory cell address of the prior art would have required.

COMPARISON WITH PRIOR ART

[0005] Prior art memory systems have attempted to solve the problem of high speed access to memory with limited success. U.S. Patent No. 3,821,715 (Hoff et. al.), was issued to Intel Corporation for the earliest 4-bit micro-processor. That patent describes a bus connecting a single central processing unit (CPU) with multiple RAMs and ROMs. That bus multiplexes addresses and data over a 4-bit wide bus and uses point-to-point control signals to select particular RAMs or ROMs. The access time is fixed and only a single processing element is permitted. There is no block-mode type of operation, and most important, not all of the interface signals between the devices are bused (the ROM and RAM control lines and the RAM select lines are point-to-point).

[0006] In U.S. Patent No. 4,315,308 (Jackson), a bus connecting a single CPU to a bus interface unit is described. The invention uses multiplexed address, data, and control information over a single 16-bit wide bus. Block-mode operations are defined, with the length of the block sent as part of the control sequence. In addition, variable access-time operations using a "stretch" cycle signal are provided. There are no multiple processing elements and no capability for multiple outstanding requests, and again, not all of the interface signals are bused.

[0007] In U.S. Patent No. 4,449,207 (Kung, et. al.), a DRAM is described which multiplexes address and data on an internal bus. The external interface to this DRAM is conventional, with separate control, address and data connections.

[0008] In U.S. Patent Nos. 4,764,846 and 4,706,166 (Go), a 3-D package arrangement of stacked die with connections along a single edge is described. Such packages are difficult to use because of the point-to-point wiring required to interconnect conventional memory devices with processing elements. Both patents describe complex schemes for solving these problems. No attempt is made to solve the problem by changing the interface.

[0009] In U.S. Patent No. 3,969,706 (Proebsting, et. al.), the current state-of-the-art DRAM interface is described. The address is two-way multiplexed, and there are separate pins for data and control (RAS, CAS, WE, CS). The number of pins grows with the size of the DRAM, and many of the connections must be made point-to-point in a memory system using such DRAMs.

[0010] There are many backplane buses described in the prior art, but not in the combination described or having the

Petitioners
Ex. 1002, p. 362

features of this invention. Many backplane buses multiplex addresses and data on a single bus (e.g., the NU bus). ELXSI and others have implemented split-transaction buses (U.S. Patent No. 4,595,923 and 4,481,625 (Roberts)). ELXSI has also implemented a relatively low-voltage-swing current-mode ECL driver (approximately 1 V swing). Address-space registers are implemented on most backplane buses, as is some form of block mode operation.

[0011] Nearly all modern backplane buses implement some type of arbitration scheme, but the arbitration scheme used in this invention differs from each of these. U.S. Patent Nos. 4,837,682 (Culler), 4,818,985 (Ikeda), 4,779,089 (Theus) and 4,745,548 (Blahut) describe prior art schemes. All involve either log N extra signals, (Theus, Blahut), where N is the number of potential bus requestors, or additional delay to get control of the bus (Ikeda, Culler). None of the buses described in patents or other literature use only bused connections. All contain some point-to-point connections on the backplane. None of the other aspects of this invention such as power reduction by fetching each data block from a single device or compact and low-cost 3-D packaging even apply to backplane buses.

[0012] The clocking scheme used in this invention has not been used before and in fact would be difficult to implement in backplane buses due to the signal degradation caused by connector stubs. U.S. Patent No. 4,247,817 (Heller) describes a clocking scheme using two clock lines, but relies on ramp-shaped clock signals in contrast to the normal rise-time signals used in the present invention.

[0013] In U.S. Patent No. 4,646,279 (Voss), a video RAM is described which implements a parallel-load, serial-out shift register on the output of a DRAM. This generally allows greatly improved bandwidth (and has been extended to 2, 4 and greater width shift-out paths.) The rest of the interfaces to the DRAM (RAS, CAS, multiplexed address, etc.) remain the same as for conventional DRAMS.

[0014] One object of the present invention is to use a new bus interface built into semiconductor devices to support high-speed access to large blocks of data from a single memory device by an external user of the data, such as a microprocessor, in an efficient and cost-effective manner.

[0015] Another object of this invention is to provide a clocking scheme to permit high speed clock signals to be sent along the bus with minimal clock skew between devices.

[0016] Another object of this invention is to allow mapping out defective memory devices or portions of memory devices.

[0017] Another object of this invention is to provide a method for distinguishing otherwise identical devices by assigning a unique identifier to each device.

[0018] Yet another object of this invention is to provide a method for transferring address, data and control information over a relatively narrow bus and to provide a method of bus arbitration when multiple devices seek to use the bus simultaneously.

[0019] Another object of this invention is to provide a method of distributing a high-speed memory cache within the DRAM chips of a memory system which is much more effective than previous cache methods.

[0020] Another object of this invention is to provide devices, especially DRAMs, suitable for use with the bus architecture of the invention.

SUMMARY OF INVENTION

[0021] The present invention includes a memory subsystem comprising at least two semiconductor devices, including at least one memory device, connected in parallel to a bus, where the bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices, where the control information includes device-select information and the bus has substantially fewer bus lines than the number of bits in a single address, and the bus carries device-select information without the need for separate device-select lines connected directly to individual devices.

[0022] Referring to Fig. 2, a standard DRAM 13, 14, ROM (or SRAM) 12, microprocessor CPU 11, I/O device, disk controller or other special purpose device such as a high speed switch is modified to use a wholly bus-based interface rather than the prior art combination of point-to-point and bus-based wiring used with conventional versions of these devices. The new bus includes clock signals, power and multiplexed address, data and control signals. In a preferred implementation, 8 bus data lines and an AddressValid bus line carry address, data and control information for memory addresses up to 40 bits wide. Persons skilled in the art will recognize that 16 bus data lines or other numbers of bus data lines can be used to implement the teaching of this invention. The new bus is used to connect elements such as memory, peripheral, switch and processing units.

[0023] In the system of this invention, DRAMs and other devices receive address and control information over the bus and transmit or receive requested data over the same bus. Each memory device contains only a single bus interface with no other signal pins. Other devices that may be included in the system can connect to the bus and other non-bus lines, such as input/output lines. The bus supports large data block transfers and split transactions to allow a user to achieve high bus utilization. This ability to rapidly read or write a large block of data to one single device at a time is an important advantage of this invention.

[0024] The DRAMs that connect to this bus differ from conventional DRAMs in a number of ways. Registers are

Petitioners
Ex. 1002, p. 363

provided which may store control information, device identification, device-type and other information appropriate for the chip such as the address range for each independent portion of the device. New bus interface circuits must be added and the internals of prior art DRAM devices need to be modified so they can provide and accept data to and from the bus at the peak data rate of the bus. This requires changes to the column access circuitry in the DRAM, with only a minimal increase in die size. A circuit is provided to generate a low skew internal device clock for devices on the bus, and other circuits provide for demultiplexing input and multiplexing output signals.

[0025] High bus bandwidth is achieved by running the bus at a very high clock rate (hundreds of MHz). This high clock rate is made possible by the constrained environment of the bus. The bus lines are controlled-impedance, doubly-terminated lines. For a data rate of 500 MHz, the maximum bus propagation time is less than 1 ns (the physical bus length is about 10 cm). In addition, because of the packaging used, the pitch of the pins can be very close to the pitch of the pads. The loading on the bus resulting from the individual devices is very small. In a preferred implementation, this generally allows stub capacitances of 1-2 pF and inductances of 0.5 - 2 nH. Each device 15, 16, 17, shown in Figure 3, only has pins on one side and these pins connect directly to the bus 18. A transceiver device 19 can be included to interface multiple units to a higher order bus through pins 20.

[0026] A primary result of the architecture of this invention is to increase the bandwidth of DRAM access. The invention also reduces manufacturing and production costs, power consumption, and increases packing density and system reliability.

BRIEF DESCRIPTION OF THE DRAWINGS

[0027]

Figure 1 is a diagram which illustrates the basic 2-D organization of memory devices.
Figure 2 is a schematic block diagram which illustrates the parallel connection of all bus lines and the serial Reset line to each device in the system.
Figure 3 is a perspective view of a system of the invention which illustrates the 3-D packaging of semiconductor devices on the primary bus.
Figure 4 shows the format of a request packet.
Figure 5 shows the format of a retry response from a slave.
Figure 6 shows the bus cycles after a request packet collision occurs on the bus and how arbitration is handled.
Figure 7 shows the timing whereby signals from two devices can overlap temporarily and drive the bus at the same time.
Figure 8 shows the connection and timing between bus clocks and devices on the bus.
Figure 9 is a perspective view showing how transceivers can be used to connect a number of bus units to a transceiver bus. Figure 10 is a block and schematic diagram of input/output circuitry used to connect devices to the bus.
Figure 11 is a schematic diagram of a clocked sense-amplifier used as a bus input receiver.
Figure 12 is a block diagram showing how the internal device clock is generated from two bus clock signals using a set of adjustable delay lines.
Figure 13 is a timing diagram showing the relationship of signals in the block diagram of Figure 12.
Figure 14 is timing diagram of a preferred means of implementing the reset procedure of this invention.
Figure 15 is a diagram illustrating the general organization of a 4 Mbit DRAM divided into 8 subarrays.

DETAILED DESCRIPTION

[0028] The present invention is designed to provide a high speed, multiplexed bus for communication between processing devices and memory devices and to provide devices adapted for use in the bus system. The invention can also be used to connect processing devices and other devices, such as I/O interfaces or disk controllers, with or without memory devices on the bus. The bus consists of a relatively small number of lines connected in parallel to each device on the bus. The'bus carries substantially all address, data and control information needed by devices for communication with other devices on the bus. In many systems using the present invention, the bus carries almost every signal between every device in the entire system. There is no need for separate device-select lines since device-select information for each device on the bus is carried over the bus. There is no need for separate address and data lines because address and data information can be sent over the same lines. Using the organization described herein, very large addresses (40 bits in the preferred implementation) and large data blocks (1024 bytes) can be sent over a small number of bus lines (8 plus one control line in the preferred implementation).

[0029] Virtually all of the signals needed by a computer system can be sent over the bus. Persons skilled in the art recognize that certain devices, such as CPUs, may be connected to other signal lines and possibly to independent buses, for example a bus to an independent cache memory, in addition to the bus of this invention. Certain devices, for

Petitioners
Ex. 1002, p. 364

example cross-point switches, could be connected to multiple, independent buses of this invention. In the preferred implementation, memory devices are provided that have no connections other than the bus connections described herein and CPUs are provided that use the bus of this invention as the principal, if not exclusive, connection to memory and to other devices on the bus.

[0030] All modern DRAM, SRAM and ROM designs have internal architectures with row (word) and column (bit) lines to efficiently tile a 2-D area. Referring to Fig. 1, one bit of data is stored at the intersection of each word line 5 and bit line 6. When a particular word line is enabled, all of the corresponding data bits are transferred onto the bit lines. This data, about 4000 bits at a time in a 4 MBit DRAM, is then loaded into column sense amplifiers 3 and held for use by the I/O circuits.

[0031] In the invention presented here, the data from the sense amplifiers is enabled 32 bits at a time onto an internal device bus running at approximately 125 MHz. This internal device bus moves the data to the periphery of the devices where the data is multiplexed into an 8-bit wide external bus interface, running at approximately 500 MHz.

[0032] The bus architecture of this invention connects master or bus controller devices, such as CPUs, Direct Memory Access devices (DMAs) or Floating Point Units (FPUs), and slave devices, such as DRAM, SRAM or ROM memory devices. A slave device responds to control signals; a master sends control signals. Persons skilled in the art realize that some devices may behave as both master and slave at various times, depending on the mode of operation and the state of the system. For example, a memory device will typically have only slave functions, while a DMA controller, disk controller or CPU may include both slave and master functions. Many other semiconductor devices, including I/O devices, disk controllers, or other special purpose devices such as high speed switches can be modified for use with the bus of this invention.

[0033] Each semiconductor device contains a set of internal registers, preferably including a device identification (device ID) register, a device-type descriptor register, control registers and other registers containing other information relevant to that type of device. In a preferred implementation, semiconductor devices connected to the bus contain registers which specify the memory addresses contained within that device and access-time registers which store a set of one or more delay times at which the device can or should be available to send or receive data.

[0034] Most of these registers can be modified and preferably are set as part of an initialization sequence that occurs when the system is powered up or reset. During the initialization sequence each device on the bus is assigned a unique device ID number, which is stored in the device ID register. A bus master can then use these device ID numbers to access and set appropriate registers in other devices, including access-time registers, control registers, and memory registers, to configure the system. Each slave may have one or several access-time registers (four in a preferred embodiment). In a preferred embodiment, one access-time register in each slave is permanently or semi-permanently programmed with a fixed value to facilitate certain control functions. A preferred implementation of an initialization sequence is described below in more detail.

[0035] All information sent between master devices and slave devices is sent over the external bus, which, for example, may be 8 bits wide. This is accomplished by defining a protocol whereby a master device, such as a microprocessor, seizes exclusive control of the external bus (i.e., becomes the bus master) and initiates a bus transaction by sending a request packet (a sequence of bytes comprising address and control information) to one or more slave devices on the bus. An address can consist of 16 to 40 or more bits according to the teachings of this invention. Each slave on the bus must decode the request packet to see if that slave needs to respond to the packet. The slave that the packet is directed to must then begin any internal processes needed to carry out the requested bus transaction at the requested time. The requesting master may also need to transact certain internal processes before the bus transaction begins. After a specified access time the slave(s) respond by returning one or more bytes (8 bits) of data or by storing information made available from the bus. More than one access time can be provided to allow different types of responses to occur at different times.

[0036] A request packet and the corresponding bus access are separated by a selected number of bus cycles, allowing the bus to be used in the intervening bus cycles by the same or other masters for additional requests or brief bus accesses. Thus multiple, independent accesses are permitted, allowing maximum utilization of the bus for transfer of short blocks of data. Transfers of long blocks of data use the bus efficiently even without overlap because the overhead due to bus address, control and access times is small compared to the total time to request and transfer the block.

Device Address Mapping

[0037] Another unique aspect of this invention is that each memory device is a complete, independent memory sub-system with all the functionality of a prior art memory board in a conventional backplane-bus computer system. Individual memory devices may contain a single memory section or may be subdivided into more than one discrete memory section. Memory devices preferably include memory address registers for each discrete memory section. A failed memory device (or even a subsection of a device) can be "mapped out" with only the loss of a small fraction of the memory, maintaining essentially full system capability. Mapping out bad devices can be accomplished in two ways, both compatible with this

Petitioners
Ex. 1002, p. 365

invention.

[0038] The preferred method uses address registers in each memory device (or independent discrete portion thereof) to store information which defines the range of bus addresses to which this memory device will respond. This is similar to prior art schemes used in memory boards in conventional backplane bus systems. The address registers can include a single pointer, usually pointing to a block of known size, a pointer and a fixed or variable block size value or two pointers, one pointing to the beginning and one to the end (or to the "top" and "bottom") of each memory block. By appropriate settings of the address registers, a series of functional memory devices or discrete memory sections can be made to respond to a contiguous range of addresses, giving the system access to a contiguous block of good memory, limited primarily by the number of good devices connected to the bus. A block of memory in a first memory device or memory section can be assigned a certain range of addresses, then a block of memory in a next memory device or memory section can be assigned addresses starting with an address one higher (or lower, depending on the memory structure) than the last address of the previous block.

[0039] Preferred devices for use in this invention include device-type register information specifying the type of chip, including how much memory is available in what configuration on that device. A master can perform an appropriate memory test, such as reading and writing each memory cell in one or more selected orders, to test proper functioning of each accessible discrete portion of memory (based in part on information like device ID number and device-type) and write address values (up to 40 bits in the preferred embodiment, 10^{12} bytes), preferably contiguous, into device address-space registers. Non-functional or impaired memory sections can be assigned a special address value which the system can interpret to avoid using that memory.

[0040] The second approach puts the burden of avoiding the bad devices on the system master or masters. CPUs and DMA controllers typically have some sort of translation look-aside buffers (TLBs) which map virtual to physical (bus) addresses. With relatively simple software, the TLBs can be programmed to use only working memory (data structures describing functional memories are easily generated). For masters which don't contain TLBs (for example, a video display generator), a small, simple RAM can be used to map a contiguous range of addresses onto the addresses of the functional memory devices.

[0041] Either scheme works and permits a system to have a significant percentage of non-functional devices and still continue to operate with the memory which remains. This means that systems built with this invention will have much improved reliability over existing systems, including the ability to build systems with almost no field failures.

Bus

[0042] The preferred bus architecture of this invention comprises 11 signals: BusData[0:7]; AddrValid; Clkl and Clk2; plus an input reference level and power and ground lines connected in parallel to each device. Signals are driven onto the bus during conventional bus cycles. The notation "Signal[i:j]" refers to a specific range of signals or lines, for example, BusData[0:7] means BusData0, BusData1, ..., BusData7. The bus lines for BusData[0:7] signals form a byte-wide, multiplexed data/address/control bus. AddrValid is used to indicate when the bus is holding a valid address request, and instructs a slave to decode the bus data as an address and, if the address is included on that slave, to handle the pending request. The two clocks together provide a synchronized, high speed clock for all the devices on the bus. In addition to the bused signals, there is one other line (ResetIn, ResetOut) connecting each device in series for use during initialization to assign every device in the system a unique device ID number (described below in detail).

[0043] To facilitate the extremely high data rate of this external bus relative to the gate delays of the internal logic, the bus cycles are grouped into pairs of even/odd cycles. Note that all devices connected to a bus should preferably use the same even/odd labeling of bus cycles and preferably should begin operations on even cycles. This is enforced by the clocking scheme.

Protocol and Bus Operation

[0044] The bus uses a relatively simple, synchronous, split-transaction, block-oriented protocol for bus transactions. One of the goals of the system is to keep the intelligence concentrated in the masters, thus keeping the slaves as simple as possible (since there are typically many more slaves than masters). To reduce the complexity of the slaves, a slave should preferably respond to a request in a specified time, sufficient to allow the slave to begin or possibly complete a device-internal phase including any internal actions that must precede the subsequent bus access phase. The time for this bus access phase is known to all devices on the bus - each master being responsible for making sure that the bus will be free when the bus access begins. Thus the slaves never worry about arbitrating for the bus. This approach eliminates arbitration in single master systems, and also makes the slave-bus interface simpler.

[0045] In a preferred implementation of the invention, to initiate a bus transfer over the bus, a master sends out a request packet, a contiguous series of bytes containing address and control information. It is preferable to use a request packet containing an even number of bytes and also preferable to start each packet on an even bus cycle.

Petitioners
Ex. 1002, p. 366

[0046] The device-select function is handled using the bus data lines. AddrValid is driven, which instructs all slaves to decode the request packet address, determine whether they contain the requested address, and if they do, provide the data back to the master (in the case of a read request) or accept data from the master (in the case of a write request) in a data block transfer. A master can also select a specific device by transmitting a device ID number in a request packet. In a preferred implementation, a special device ID number is chosen to indicate that the packet should be interpreted by all devices on the bus. This allows a master to broadcast a message, for example to set a selected control register of all devices with the same value.

[0047] The data block transfer occurs later at a time specified in the request packet control information, preferably beginning on an even cycle. A device begins a data block transfer almost immediately with a device-internal phase as the device initiates certain functions, such as setting up memory addressing, before the bus access phase begins. The time after which a data block is driven onto the bus lines is selected from values stored in slave access-time registers. The timing of data for reads and writes is preferably the same; the only difference is which device drives the bus. For reads, the slave drives the bus and the master latches the values from the bus. For writes the master drives the bus and the selected slave latches the values from the bus.

[0048] In a preferred implementation of this invention shown in Figure 4, a request packet 22 contains 6 bytes of data -- 4.5 address bytes and 1.5 control bytes. Each request packet uses all nine bits of the multiplexed data/address lines (AddrValid 23 + BusData[0:7] 24) for all six bytes of the request packet. Setting 23 AddrValid = 1 in an otherwise unused even cycle indicates the start of an request packet (control information). In a valid request packet, AddrValid 27 must be 0 in the last byte. Asserting this signal in the last byte invalidates the request packet. This is used for the collision detection and arbitration logic (described below). Bytes 25-26 contain the first 35 address bits, Address[0:35]. The last byte contains AddrValid 27 (the invalidation switch) and 28, the remaining address bits, Address[36:39], and BlockSize [0:3] (control information).

[0049] The first byte contains two 4 bit fields containing control information, AccessType[0:3], an op code (operation code) which, for example, specifies the type of access, and Master[0:3], a position reserved for the master sending the packet to include its master ID number. Only master numbers 1 through 15 are allowed - master number 0 is reserved for special system commands. Any packet with Master[0:3] = 0 is an invalid or special packet and is treated accordingly.

[0050] The AccessType field specifies whether the requested operation is a read or write and the type of access, for example, whether it is to the control registers or other parts of the device, such as memory. In a preferred implementation, AccessType[0] is a Read/Write switch: if it is a 1, then the operation calls for a read from the slave (the slave to read the requested memory block and drive the memory contents onto the bus); if it is a 0, the operation calls for a write into the slave (the slave to read data from the bus and write it to memory). AccessType[1:3] provides up to 8 different access types for a slave. AccessType[1:2] preferably indicates the timing of the response, which is stored in an access-time register, AccessRegN. The choice of access-time register can be selected directly by having a certain op code select that register, or indirectly by having a slave respond to selected op codes with pre-selected access times (see table below). The remaining bit, AccessType[3] may be used to send additional information about the request to the slaves.

[0051] One special type of access is control register access, which involves addressing a selected register in a selected slave. In the preferred implementation of this invention, AccessType[1:3] equal to zero indicates a control register request and the address field of the packet indicates the desired control register. For example, the most significant two bytes can be the device ID number (specifying which slave is being addressed) and the least significant three bytes can specify a register address and may also represent or include data to be loaded into that control register. Control register accesses are used to initialize the access-time registers, so it is preferable to use a fixed response time which can be preprogrammed or even hard wired, for example the value in AccessReg0, preferably 8 cycles. Control register access can also be used to initialize or modify other registers, including address registers.

[0052] The method of this invention provides for access mode control specifically for the DRAMs. One such access mode determines whether the access is page mode or normal RAS access. In normal mode (in conventional DRAMS and in this invention), the DRAM column sense amps or latches have been precharged to a value intermediate between logical 0 and 1. This precharging allows access to a row in the RAM to begin as soon as the access request for either inputs (writes) or outputs (reads) is received and allows the column sense amps to sense data quickly. In page mode (both conventional and in this invention), the DRAM holds the data in the column sense amps or latches from the previous read or write operation. If a subsequent request to access data is directed td the same row, the DRAM does not need to wait for the data to be sensed (it has been sensed already) and access time for this data is much shorter than the normal access time. Page mode generally allows much faster access to data but to a smaller block of data (equal to the number of sense amps). However, if the requested data is not in the selected row, the access time is longer than the normal access time, since the request must wait for the RAM to precharge before the normal mode access can start. Two access-time registers in each DRAM preferably contain the access times to be used for normal and for page-mode accesses, respectively.

[0053] The access mode also determines whether the DRAM should precharge the sense amplifiers or should save the contents of the sense amps for a subsequent page mode access. Typical settings are "precharge after normal

Petitioners
Ex. 1002, p. 367

access" and "save after page mode access" but "precharge after page mode access" or "save after normal access" are allowed, selectable modes of operation. The DRAM can also be set to precharge the sense amps if they are not accessed for a selected period of time.

[0054] In page mode, the data stored in the DRAM sense amplifiers may be accessed within much less time than it takes to read out data in normal mode (~10-20 nS vs. 40-100 nS). This data may be kept available for long periods. However, if these sense amps (and hence bit lines) are not precharged after an access, a subsequent access to a different memory word (row) will suffer a precharge time penalty of about 40-100 nS because the sense amps must precharge before latching in a new value.

[0055] The contents of the sense amps thus may be held and used as a cache, allowing faster, repetitive access to small blocks of data. DRAM-based page-mode caches have been attempted in the prior art using conventional DRAM organizations but they are not very effective because several chips are required per computer word. Such a conventional page-mode cache contains many bits (for example, 32 chips x 4Kbits) but has very few independent storage entries. In other words, at any given point in time the sense amps hold only a few different blocks or memory "locales" (a single block of 4K words, in the example above). Simulations have shown that upwards of 100 blocks are required to achieve high hit rates (>90% of requests find the requested data already in cache memory) regardless of the size of each block. *See, for* example, Anant Agarwal, et. al., "An Analytic Cache Model," ACM Transactions on Computer Systems, Vol. 7 (2), pp. 184-215 (May 1989).

[0056] The organization of memory in the present invention allows each DRAM to hold one or more (4 for 4MBit DRAMS) separately- addressed and independent blocks of data. A personal computer or workstation with 100 such DRAMs (i.e. 400 blocks or locales) can achieve extremely high, very repeatable hit rates (98-99% on average) as compared to the lower (50-80%), widely varying hit rates using DRAMS organized in the conventional fashion. Further, because of the time penalty associated with the deferred precharge on a "miss" of the page-mode cache, the conventional DRAM-based page-mode cache generally has been found to work less well than no cache at all.

[0057] For DRAM slave access, the access types are preferably used in the following way:

AccessType[1:3]	Use	AccessTime
0	Control Register Access	Fixed, 8[AccessReg0]
1	Unused	Fixed, 8[AccessReg0]
2-3	Unused	AccessReg1
4-5	Page Mode DRAM access	AccessReg2
6-7	Normal DRAM access	AccessReg3

Persons skilled in the art will recognize that a series of available bits could be designated as switches for controlling these access modes. For example:

AccessType[2] = page mode/normal switch

AccessType[3] = precharge/save-data switch

BlockSize[0:3] specifies the size of the data block transfer. If BlockSize[0] is 0, the remaining bits are the binary representation of the block size (0-7). If BlockSize[0] is 1, then the remaining bits give the block size as a binary power of 2, from 8 to 1024. A zero-length block can be interpreted as a special command, for example, to refresh a DRAM without returning any data, or to change the DRAM from page mode to normal access mode or vice-versa.

BlockSize[0:2]	Number of Bytes in Block
0-7	0-7 respectively
8	8
9	16
10	32
11	64
12	128
13	256
14	512
15	1024

Persons skilled in the art will recognize that other block size encoding schemes or values can be used.

[0058] In most cases, a slave will respond at the selected access time by reading or writing data from or to the bus over bus lines BusData[0:7] and AddrValid will be at logical 0. In a preferred embodiment, substantially each memory

Petitioners
Ex. 1002, p. 368

access will involve only a single memory device, that is, a single block will be read from or written to a single memory device.

Retry Format

[0059] In some cases, a slave may not be able to respond correctly to a request, e.g., for a read or write. In such a situation, the slave should return an error message, sometimes called a N(o)ACK(nowledge) or retry message. The retry message can include information about the condition requiring a retry, but this increases system requirements for circuitry in both slave and masters. A simple message indicating only that an error has occurred allows for a less complex slave, and the master can take whatever action is needed to understand and correct the cause of the error.

[0060] For example, under certain conditions a slave might not be able to supply the requested data. During a page-mode access, the DRAM selected must be in page mode and the requested address must match the address of the data held in the sense amps or latches. Each DRAM can check for this match during a page-mode access. If no match is found, the DRAM begins precharging and returns a retry message to the master during the first cycle of the data block (the rest of the returned block is ignored). The master then must wait for the precharge time (which is set to accommodate the type of slave in question, stored in a special register, PreChargeReg), and then resend the request as a normal DRAM access (AccessType = 6 or 7).

[0061] In the preferred form of the present invention, a slave signals a retry by driving AddrValid true at the time the slave was supposed to begin reading or writing data. A master which expected to write to that slave must monitor AddrValid during the write and take corrective action if it detects a retry message. Figure 5 illustrates the format of a retry message 28 which is useful for read requests, consisting of 23 AddrValid=1 with Master[0:3] = 0 in the first (even) cycle. Note that AddrValid is normally 0 for data block transfers and that there is no master 0 (only 1 through 15 are allowed). All DRAMs and masters can easily recognize such a packet as an invalid request packet, and therefore a retry message. In this type of bus transaction all of the fields except for Mastery[0:3] and AddrValid 23 may be used as information fields, although in the implementation described, the contents are undefined. Persons skilled in the art recognize that another method of signifying a retry message is to add a DataInvalid line and signal to the bus. This signal could be asserted in the case of a NACK.

Bus Arbitration

[0062] In the case of a single master, there are by definition no arbitration problems. The master sends request packets and keeps track of periods when the bus will be busy in response to that packet. The master can schedule multiple requests so that the corresponding data block transfers do not overlap.

[0063] The bus architecture of this invention is also useful in configurations with multiple masters. When two or more masters are on the same bus, each master must keep track of all the pending transactions, so each master knows when it can send a request packet and access the corresponding data block transfer. Situations will arise, however, where two or more masters send a request packet at about the same time and the multiple requests must be detected, then sorted out by some sort of bus arbitration.

[0064] There are many ways for each master to keep track of when the bus is and will be busy. A simple method is for each master to maintain a bus-busy data structure, for example by maintaining two pointers, one to indicate the earliest point in the future when the bus will be busy and the other to indicate the earliest point in the future when the bus will be free, that is, the end of the latest pending data block transfer. Using this information, each master can determine whether and when there is enough time to send a request packet (as described above under Protocol) before the bus becomes busy with another data block transfer and whether the corresponding data block transfer will interfere with pending bus transactions. Thus each master must read every request packet and update its bus-busy data structure to maintain information about when the bus is and will be free.

[0065] With two or more masters on the bus, masters will occasionally transmit independent request packets during the same bus cycle. Those multiple requests will collide as each such master drives the bus simultaneously with different information, resulting in scrambled request information and neither desired data block transfer. In a preferred form of the invention, each device on the bus seeking to write a logical 1 on a BusData or AddrValid line drives that line with a current sufficient to sustain a voltage greater than or equal to the high-logic value for the system. Devices do not drive lines that should have a logical 0; those lines are simply held at a voltage corresponding to a low-logic value. Each master tests the voltage on at least some, preferably all, bus data and the AddrValid lines so the master can detect a logical '1' where the expected level is '0' on a line that it does not drive during a given bus cycle but another master does drive.

[0066] Another way to detect collisions is to select one or more bus lines for collision signalling. Each master sending a request drives that line or lines and monitors the selected lines for more than the normal drive current (or a logical value of ">1"), indicating requests by more than one master. Persons skilled in the art will recognize that this can be implemented with a protocol involving BusData and AddrValid lines or could be implemented using an additional bus line.

Petitioners
Ex. 1002, p. 369

[0067] In the preferred form of this invention, each master detects collisions by monitoring lines which it does not drive to see if another master is driving those lines. Referring to Fig. 4, the first byte of the request packet includes the number of each master attempting to use the bus (Master[0:3]). If two masters send packet requests starting at the same point in time, the master numbers will be logical "or"ed together by at least those masters, and thus one or both of the masters, by monitoring the data on the bus and comparing what it sent, can detect a collision. For instance if requests by masters number 2 (0010) and 5 (0101) collide, the bus will be driven with the value Master[0:3]=7 (0010 + 0101 = 0111). Master number 5 will detect that the signal Master[2] = 1 and master 2 will detect that Master[1] and Master[3] = 1, telling both masters that a collision has occurred. Another example is masters 2 and 11, for which the bus will be driven with the value Master[0:3]=11 (0010 + 1011 = 1011), and although master 11 can't readily detect this collision, master 2 can. When any collision is detected, each master detecting a collision drives the value of AddrValid 27 in byte 5 of the request packet 22 to 1, which is detected by all masters, including master 11 in the second example above, and forces a bus arbitration cycle, described below.

[0068] Another collision condition may arise where master A sends a request packet in cycle 0 and master B tries to send a request packet starting in cycle 2 of the first request packet, thereby overlapping the first request packet. This will occur from time to time because the bus operates at high speeds, thus the logic in a second-initiating master may not be fast enough to detect a request initiated by a first master in cycle 0 and to react fast enough by delaying its own request. Master B eventually notices that it wasn't supposed to try to send a request packet (and consequently almost surely destroyed the address that master A was trying to send), and, as in the example above of a simultaneous collision, drives a 1 on AddrValid during byte 5 of the first request packet 27 forcing an arbitration. The logic in the preferred implementation is fast enough that a master should detect a request packet by another master by cycle 3 of the first request packet, so no master is likely to attempt to send a potentially colliding request packet later than cycle 2.

[0069] Slave devices not need to detect a collision directly, but they must wait to do anything irrecoverable until the last byte (byte 5) is read to ensure that the packet is valid. A request packet with Master[0:3] equal to 0 (a retry signal) is ignored and does not cause a collision. The subsequent bytes of such a packet are ignored.

[0070] To begin arbitration after a collision, the masters wait a preselected number of cycles after the aborted request packet (4 cycles in a preferred implementation), then use the next free cycle to arbitrate for the bus (the next available even cycle in the preferred implementation). Each colliding master signals to all other colliding masters that it seeks to send a request packet, a priority is assigned to each of the colliding masters, then each master is allowed to make its request in the order of that priority.

[0071] Figure 6 illustrates one preferred way of implementing this arbitration. Each colliding master signals its intent to send a request packet by driving a single BusData line during a single bus cycle corresponding to its assigned master number (1-15 in the present example). During two-byte arbitration cycle 29, byte 0 is allocated to requests 1-7 from masters 1-7, respectively, (bit 0 is not used) and byte 1 is allocated to requests 8-15 from masters 8-15, respectively. At least one device and preferably each colliding master reads the values on the bus during the arbitration cycles to determine and store which masters desire to use the bus. Persons skilled in the art will recognize that a single byte can be allocated for arbitration requests if the system includes more bus lines than masters. More than 15 masters can be accommodated by using additional bus cycles.

[0072] A fixed priority scheme (preferably using the master numbers, selecting lowest numbers first) is then used to prioritize, then sequence the requests in a bus arbitration queue which is maintained by at least one device. These requests are queued by each master in the bus-busy data structure and no further requests are allowed until the bus arbitration queue is cleared. Persons skilled in the art will recognize that other priority schemes can be used, including assigning priority according to the physical location of each master.

System Configuration/Reset

[0073] In the bus-based system of this invention, a mechanism is provided to give each device on the bus a unique device identifier (device ID) after power-up or under other conditions as desired or needed by the system. A master can then use this device ID to access a specific device, particularly to set or modify registers of the specified device, including the control and address registers. In the preferred embodiment, one master is assigned to carry out the entire system configuration process. The master provides a series of unique device ID numbers for each unique device connected to the bus system. In the preferred embodiment, each device connected to the bus contains a special device-type register which specifies the type of device, for instance CPU, 4 MBit memory, 64 MBit memory or disk controller. The configuration master should check each device, determine the device type and set appropriate control registers, including access-time registers. The configuration master should check each memory device and set all appropriate memory address registers.

[0074] One means to set up unique device ID numbers is to have each device to select a device ID in sequence and store the value in an internal device ID register. For example, a master can pass sequential device ID numbers through shift registers in each of a series of devices, or pass a token from device to device whereby the device with the token

Petitioners
Ex. 1002, p. 370

reads in device ID information from another line or lines. In a preferred embodiment, device ID numbers are assigned to devices according to their physical relationship, for instance, their order along the bus.

[0075] In a preferred embodiment of this invention, the device ID setting is accomplished using a pair of pins on each device, ResetIn and ResetOut. These pins handle normal logic signals and are used only during device ID configuration. On each rising edge of the clock, each device copies ResetIn (an input) into a four-stage reset shift register. The output of the reset shift register is connected to ResetOut, which in turn connects to ResetIn for the next sequentially connected device. Substantially all devices on the bus are thereby daisy-chained together. A first reset signal, for example, while ResetIn at a device is a logical 1, or when a selected bit of the reset shift register goes from zero to non-zero, causes the device to hard reset, for example by clearing all internal registers and resetting all state machines. A second reset signal, for example, the falling edge of ResetIn combined with changeable values on the external bus, causes that device to latch the contents of the external bus into the internal device ID register (Device[0:7]).

[0076] To reset all devices on a bus, a master sets the ResetIn line of the first device to a "1" for long enough to ensure that all devices on the bus have been reset (4 cycles times the number of devices -- note that the maximum number of devices on the preferred bus configuration is 256 (8 bits), so that 1024 cycles is always enough time to reset all devices.) Then ResetIn is dropped to "0" and the BusData lines are driven with the first followed by successive device ID numbers, changing after every 4 clock pulses. Successive devices set those device ID numbers into the corresponding device ID register as the falling edge of ResetIn propagates through the shift registers of the daisy-chained devices. Figure 14 shows ResetIn at a first device going low while a master drives a first device ID onto the bus data lines BusData[0:3]. The first device then latches in that first device ID. After four clock cycles, the master changes BusData[0:3] to the next device ID number and ResetOut at the first device goes low, which pulls ResetIn for the next daisy-chained device low, allowing the next device to latch in the next device ID number from BusData[0:3]. In the preferred embodiment, one master is assigned device ID 0 and it is the responsibility of that master to control the ResetIn line and to drive successive device ID numbers onto the bus at the appropriate times. In the preferred embodiment, each device waits two clock cycles after ResetIn goes low before latching in a device ID number from BusData[0:3].

[0077] Persons skilled in the art recognize that longer device ID numbers could be distributed to devices by having each device read in multiple bytes from the bus and latch the values into the device ID register. Persons skilled in the art also recognize that there are alternative ways of getting device ID numbers to unique devices. For instance, a series of sequential numbers could be clocked along the ResetIn line and at a certain time each device could be instructed to latch the current reset shift register value into the device ID register.

[0078] The configuration master should choose and set an access time in each access-time register in each slave to a period sufficiently long to allow the slave to perform an actual, desired memory access. For example, for a normal DRAM access, this time must be longer than the row address strobe (RAS) access time. If this condition is not met, the slave may not deliver the correct data. The value stored in a slave access-time register is preferably one-half the number of bus cycles for which the slave device should wait before using the bus in response to a request. Thus an access time value of '1' would indicate that the slave should not access the bus until at least two cycles after the last byte of the request packet has been received. The value of AccessReg0 is preferably fixed at 8 (cycles) to facilitate access to control registers.

[0079] The bus architecture of this invention can include more than one master device. The reset or initialization sequence should also include a determination of whether there are multiple masters on the bus, and if so to assign unique master ID numbers to each. Persons skilled in the art will recognize that there are many ways of doing this. For instance, the master could poll each device to determine what type of device it is, for example, by reading a special register then, for each master device, write the next available master ID number into a special register.

ECC

[0080] Error detection and correction ("ECC") methods well known in the art can be implemented in this system. ECC information typically is calculated for a block of data at the time that block of data is first written into memory. The data block usually has an integral binary size, e.g. 256 bits, and the ECC information uses significantly fewer bits. A potential problem arises in that each binary data block in prior art schemes typically is stored with the ECC bits appended, resulting in a block size that is not an integral binary power.

[0081] In a preferred embodiment of this invention, ECC information is stored separately from the corresponding data, which can then be stored in blocks having integral binary size. ECC information and corresponding data can be stored, for example, in separate DRAM devices. Data can be read without ECC using a single request packet, but to write or read error-corrected data requires two request packets, one for the data and a second for the corresponding ECC information. ECC information may not always be stored permanently and in some situations the ECC information may be available without sending a request packet or without a bus data block transfer.

[0082] In a preferred embodiment, a standard data block size can be selected for use with ECC, and the ECC method will determine the required number of bits of information in a corresponding ECC block. RAMs containing ECC information

Petitioners
Ex. 1002, p. 371

can be programmed to store an access time that is equal to: (1) the access time of the normal RAM (containing data) plus the time to access a standard data block (for corrected data) minus the time to send a request packet (6 bytes); or (2) the access time of a normal RAM minus the time to access a standard ECC block minus the time to send a request packet. To read a data block and the corresponding ECC block, the master simply issues a request for the data immediately followed by a request for the ECC block. The ECC RAM will wait for the selected access time then drive its data onto the bus right after (in case (1) above)) the data RAM has finished driving out the data block. Persons skilled in the art will recognize that the access time described in case (2) above can be used to drive ECC data before the data is driven onto the bus lines and will recognize that writing data can be done by analogy with the method described for a read. Persons skilled in the art will also recognize the adjustments that must be made in the bus-busy structure and the request packet arbitration methods of this invention in order to accommodate these paired ECC requests.

[0083] Since this system is quite flexible, the system designer can choose the size of the data blocks and the number of ECC bits using the memory devices of this invention. Note that the data stream on the bus can be interpreted in various ways. For instance the sequence can be 2^n data bytes followed by 2^m ECC bytes (or vice versa), or the sequence can be 2^k iterations of 8 data bytes plus 1 ECC byte. Other information, such as information used by a directory-based cache coherence scheme, can also be managed this way. See, for example, Anant Agarwal, et al., "Scaleable Directory Schemes for Cache Consistency," 15th International Symposium on Computer Architecture, June 1988, pp. 280-289. Those skilled in the art will recognize alternative methods of implementing ECC schemes that are within the teachings of this invention.

Low Power 3-D Packaging

[0084] Another major advantage of this invention is that it drastically reduces the memory system power consumption. Nearly all the power consumed by a prior art DRAM is dissipated in performing row access. By using a single row access in a single RAM to supply all the bits for a block request (compared to a row-access in each of multiple RAMs in conventional memory systems) the power per bit can be made very small. Since the power dissipated by memory devices using this invention is significantly reduced, the devices potentially can be placed much closer together than with conventional designs.

[0085] The bus architecture of this invention makes possible an innovative 3-D packaging technology. By using a narrow, multiplexed (time-shared) bus, the pin count for an arbitrarily large memory device can be kept quite small - on the order of 20 pins. Moreover, this pin count can be kept constant from one generation of DRAM density to the next. The low power dissipation allows each package to be smaller, with narrower pin pitches (spacing between the IC pins). With current surface mount technology supporting pin pitches as low as 20 mils, all off-device connections can be implemented on a single edge of the memory device. Semiconductor die useful in this invention preferably have connections or pads along one edge of the die which can then be wired or otherwise connected to the package pins with wires having similar lengths. This geometry also allows for very short leads, preferably with an effective lead length of less than 4 mm. Furthermore, this invention uses only bused interconnections, i.e., each pad on each device is connected by the bus to the corresponding pad of each other device.

[0086] The use of a low pin count and an edge-connected bus permits a simple 3-D package, whereby the devices are stacked and the bus is connected along a single edge of the stack. The fact that all of the signals are bused is important for the implementation of a simple 3-D structure. Without this, the complexity of the "backplane" would be too difficult to make cost effectively with current technology. The individual devices in a stack of the present invention can be packed quite tightly because of the low power dissipated by the entire memory system, permitting the devices to be stacked bumper-to-bumper or top to bottom. Conventional plastic-injection molded small outline (SO) packages can be used with a pitch of about 2.5 mm (100 mils), but the ultimate limit would be the device die thickness, which is about an order of magnitude smaller, 0.2-0.5 mm using current wafer technology.

Bus Electrical Description

[0087] By using devices with very low power dissipation and close physical packing, the bus can be made quite short, which in turn allows for short propagation times and high data rates. The bus of a preferred embodiment of the present invention consists of a set of resistor-terminated controlled impedance transmission lines which can operate up to a data rate of 500 MHz (2 ns cycles). The characteristics of the transmission lines are strongly affected by the loading caused by the DRAMs (or other slaves) mounted on the bus. These devices add lumped capacitance to the lines which both lowers the impedance of the lines and decreases the transmission speed. In the loaded environment, the bus impedance is likely to be on the order of 25 ohms and the propagation velocity about $c/4$ (c = the speed of light) or 7.5 cm/ns. To operate at a 2 ns data rate, the transit time on the bus should preferably be kept under 1 ns, to leave 1 ns for the setup and hold time of the input receivers (described below) plus clock skew. Thus the bus lines must be kept quite short, under about 8 cm for maximum performance. Lower performance systems may have much longer lines, e.g. a 4 ns bus

Petitioners
Ex. 1002, p. 372

may have 24 cm lines (3 ns transit time, 1 ns setup and hold time).

[0088] In the preferred embodiment, the bus uses current source drivers. Each output must be able to sink 50 mA, which provides an output swing of about 500 mV or more. In the preferred embodiment of this invention, the bus is active low. The unasserted state (the high value) is preferably considered a logical zero, and the asserted value (low state) is therefore a logical 1. Those skilled in the art understand that the method of this invention can also be implemented using the opposite logical relation to voltage. The value of the unasserted state is set by the voltage on the termination resistors, and should be high enough to allow the outputs to act as current sources, while being as low as possible to reduce power dissipation. These constraints may yield a termination voltage about 2V above ground in the preferred implementation. Current source drivers cause the output voltage to be proportional to the sum of the sources driving the bus.

[0089] Referring to Fig.7, although there is no stable condition where two devices drive the bus at the same time, conditions can arise because of propagation delay on the wires where one device, A 41, can start driving its part of the bus 44 while the bus is still being driven by another device, B 42 (already asserting a logical 1 on the bus). In a system using current drivers, when B 42 is driving the bus (before time 46), the value at points 44 and 45 is logical 1. If B 42 switches off at time 46 just when A 41 switches on, the additional drive by device A 41 causes the voltage at the output 44 of A 41 to drop briefly below the normal value. The voltage returns to its normal value at time 47 when the effect of device B 42 turning off is felt. The voltage at point 45 goes to logical 0 when device B 42 turns off, then drops at time 47 when the effect of device A 41 turning on is felt. Since the logical 1 driven by current from device A 41 is propagated irrespective of the previous value on the bus, the value on the bus is guaranteed to settle after one time of flight (t_f) delay, that is, the time it takes a signal to propagate from one end of the bus to the other. If a voltage drive was used (as in ECL wired-ORing), a logical 1 on the bus (from device B 42 being previously driven) would prevent the transition put out by device A 41 being felt at the most remote part of the system, e.g., device 43, until the turnoff waveform from device B 42 reached device A 41 plus one time of flight delay, giving a worst case settling time of twice the time of flight delay.

Clocking

[0090] Clocking a high speed bus accurately without introducing error due to propagation delays can be implemented by having each device monitor two bus clock signals and then derive internally a device clock, the true system clock. The bus clock information can be sent on one or two lines to provide a mechanism for each bused device to generate an internal device clock with zero skew relative to all the other device clocks. Referring to Figure 8, in the preferred implementation, a bus clock generator 50 at one end of the bus propagates an early bus clock signal in one direction along the bus, for example on line 53 from left to right, to the far end of the bus. The same clock signal then is passed through the direct connection shown to a second line 54, and returns as a late bus clock signal along the bus from the far end to the origin, propagating from right to left. A single bus clock line can be used if it is left unterminated at the far end of the bus, allowing the early bus clock signal to reflect back along the same line as a late bus clock signal.

[0091] Figure 8b illustrates how each device 51, 52 receives each of the two bus clock signals at a different time (because of propagation delay along the wires), with constant midpoint in time between the two bus clocks along the bus. At each device 51, 52, the rising edge 55 of Clock1 53 is followed by the rising edge 56 of Clock2 54. Similarly, the falling edge 57 of Clock1 53 is followed by the falling edge 58 of Clock2 54. This waveform relationship is observed at all other devices along the bus. Devices which are closer to the clock generator have a greater separation between Clock1 and Clock2 relative to devices farther from the generator because of the longer time required for each clock pulse to traverse the bus and return along line 54, but the midpoint in time 59, 60 between corresponding rising or falling edges is fixed because, for any given device, the length of each clock line between the far end of the bus and that device is equal. Each device must sample the two bus clocks and generate its own internal device clock at the midpoint of the two.

[0092] Clock distribution problems can be further reduced by using a bus clock and device clock rate equal to the bus cycle data rate divided by two, that is, the bus clock period is twice the bus cycle period. Thus a 500 MHz bus preferably uses a 250 MHz clock rate. This reduction in frequency provides two benefits. First it makes all signals on the bus have the same worst case data rates -- data on a 500 MHz bus can only change every 2 ns. Second, clocking at half the bus cycle data rate makes the labeling of the odd and even bus cycles trivial, for example, by defining even cycles to be those when the internal device clock is 0 and odd cycles when the internal device clock is 1.

Multiple Buses

[0093] The limitation on bus length described above restricts the total number of devices that can be placed on a single bus. Using 2.5 mm spacing between devices, a single 8 cm bus will hold about 32 devices. Persons skilled in the art will recognize certain applications of the present invention wherein the overall data rate on the bus is adequate but memory or processing requirements necessitate a much larger number of devices (many more than 32). Larger systems can easily be built using the teachings of this invention by using one or more memory subsystems, designated primary

Petitioners
Ex. 1002, p. 373

bus units, each of which consists of two or more devices, typically 32 or close to the maximum allowed by bus design requirements, connected to a transceiver device.

Referring to Figure 9, each primary bus unit can be mounted on a single circuit board 66, sometimes called a memory stick. Each transceiver device 19 in turn connects to a transceiver bus 65, similar or identical in electrical and other respects to the primary bus 18 described at length above. In a preferred implementation, all masters are situated on the transceiver bus so there are no transceiver delays between masters and all memory devices are on primary bus units so that all memory accesses experience an equivalent transceiver delay, but persons skilled in the art will recognize how to implement systems which have masters on more than one bus unit and memory devices on the transceiver bus as well as on primary bus units. In general, each teaching of this invention which refers to a memory device can be practiced using a transceiver device and one or more memory devices on an attached primay bus unit. Other devices, generically referred to as peripheral devices, including disk controllers, video controllers or I/O devices can also be attached to either the transceiver bus or a primary bus unit, as desired. Persons skilled in the art will recognize how to use a single primary bus unit or multiple primary bus units as needed with a transceiver bus in certain system designs.

[0094] The transceivers are quite simple in function. They detect request packets on the transceiver bus and transmit them to their primary bus unit. If the request packet calls for a write to a device on a transceiver's primary bus unit, that transceiver keeps track of the access time and block size and forwards all data from the transceiver bus to the primary bus unit during that time. The transceivers also watch their primary bus unit, forwarding any data that occurs there to the transceiver bus. The high speed of the buses means that the transceivers will need to be pipelined, and will require an additional one or two cycle delay for data to pass through the transceiver in either direction. Access times stored in masters on the transceiver bus must be increased to account for transceiver delay but access times stored in slaves on a primary bus unit should not be modified.

[0095] Persons skilled in the art will recognize that a more sophisticated transceiver can control transmissions to and from primary bus units. An additional control line, TrncvrRW can be bused to all devices on the transceiver bus, using that line in conjunction with the AddrValid line to indicate to all devices on the transceiver bus that the information on the data lines is: 1) a request packet, 2) valid data to a slave, 3) valid data from a slave, or 4) invalid data (or idle bus). Using this extra control line obviates the need for the transceivers to keep track of when data needs to be forwarded from its primary bus to the transceiver bus - all transceivers send all data from their primary bus to the transceiver bus whenever the control signal indicates condition 2) above. In a preferred implementation of this invention, if AddrValid and TrncvrRW are both low, there is no bus activity and the transceivers should remain in an idle state. A controller sending a request packet will drive AddrValid high, indicating to all devices on the transceiver bus that a request packet is being sent which each transceiver should forward to its primary bus unit. Each controller seeking to write to a slave should drive both AddrValid and TrncvrRW high, indicating valid data for a slave is present on the data lines. Each transceiver device will then transmit all data from the transceiver bus lines to each primary bus unit. Any controller expecting to receive information from a slave should also drive the TrncvrRW line high, but not drive AddrValid, thereby indicating to each transceiver to transmit any data coming from any slave on its primary local bus to the transceiver bus. A still more sophisticated transceiver would recognize signals addressed to or coming from its primary bus unit and transmit signals only at requested times.

[0096] An example of the physical mounting of the transceivers is shown in Figure 9. One important feature of this physical arrangement is to integrate the bus of each transceiver 19 with the original bus of DRAMs or other devices 15, 16, 17 on the primary bus unit 66. The transceivers 19 have pins on two sides, and are preferably mounted flat on the primary bus unit with a first set of pins connected to primary bus 18. A second set of transceiver pins 20, preferably orthogonal to the first set of pins, are oriented to allow the transceiver 19 to be attached to the transceiver bus 65 in much the same way as the DRAMs were attached to the primary bus unit. The transceiver bus can be generally planar and in a different plane, preferably orthogonal to the plane of each primary bus unit. The transceiver bus can also be generally circular with primary bus units mounted perpendicular and tangential to the transceiver bus.

[0097] Using this two level scheme allows one to easily build a system that contains over 500 slaves (16 buses of 32 DRAMs each). Persons skilled in the art can modify the device ID scheme described above to accommodate more than 256 devices, for example by using a longer device ID or by using additional registers to hold some of the device ID. This scheme can be extended in yet a third dimension to make a second-order transceiver bus, connecting multiple transceiver buses by aligning transceiver bus units parallel to and on top of each other and busing corresponding signal lines through a suitable transceiver. Using such a second-order transceiver bus, one could connect many thousands of slave devices into what is effectively a single bus.

Device Interface

[0098] The device interface to the high-speed bus can be divided into three main parts. The first part is the electrical interface. This part includes the input receivers, bus drivers and clock generation circuitry. The second part contains the address comparison circuitry and timing registers. This part takes the input request packet and determines if the request

Petitioners
Ex. 1002, p. 374

is for this device, and if it is, starts the internal access and delivers the data to the pins at the correct time. The final part, specifically for memory devices such as DRAMs, is the DRAM column access path. This part needs to provide bandwidth into and out of the DRAM sense amps greater than the bandwidth provided by conventional DRAMs. The implementation of the electrical interface and DRAM column access path are described in more detail in the following sections. Persons skilled in the art recognize how to modify prior-art address comparison circuitry and prior-art register circuitry in order to practice the present invention.

Electrical Interface - Input/Output Circuitry

[0099] A block diagram of the preferred input/output circuit for address/data/control lines is shown in Figure 10. This circuitry is particularly well-suited for use in DRAM devices but it can be used or modified by one skilled in the art for use in other devices connected to the bus of this invention. It consists of a set of input receivers 71, 72 and output driver 76 connected to input/output line 69 and pad 75 and circuitry to use the internal clock 73 and internal clock complement 74 to drive the input interface. The clocked input receivers take advantage of the synchronous nature of the bus. To further reduce the performance requirements for device input receivers, each device pin, and thus each bus line, is connected to two clocked receivers, one to sample the even cycle inputs, the other to sample the odd cycle inputs. By thus de-multiplexing the input 70 at the pin, each clocked amplifier is given a full 2 ns cycle to amplify the bus low-voltage-swing signal into a full value CMOS logic signal. Persons skilled in the art will recognize that additional clocked input receivers can be used within the teachings of this invention. For example, four input receivers could be connected to each device pin and clocked by a modified internal device clock to transfer sequential bits from the bus to internal device circuits, allowing still higher external bus speeds or still longer settling times to amplify the bus low-voltage-swing signal into a full value CMOS logic signal.

[0100] The output drivers are quite simple, and consist of a single NMOS pulldown transistor 76. This transistor is sized so that under worst case conditions it can still sink the 50 mA required by the bus. For 0.8 micron CMOS technology, the transistor will need to be about 200 microns long. Overall bus performance can be improved by using feedback techniques to control output transistor current so that the current through the device is roughly 50 mA under all operating conditions, although this is not absolutely necessary for proper bus operation. An example of one of many methods known to persons skilled in the art for using feedback techniques to control current is described in Hans Schumacher, et al., "CMOS Subnanosecond True-ECL Output Buffer," J. Solid State Circuits, Vol. 25 (1), pp. 150-154 (Feb. 1990). Controlling this current improves performance and reduces power dissipation. This output driver which can be operated at 500 MHz, can in turn be controlled by a suitable multiplexer with two or more (preferably four) inputs connected to other internal chip circuitry, all of which can be designed according to well known prior art.

[0101] The input receivers of every slave must be able to operate during every cycle to determine whether the signal on the bus is a valid request packet. This requirement leads to a number of constraints on the input circuitry. In addition to requiring small acquisition and resolution delays, the circuits must take little or no DC power, little AC power and inject very little current back into the input or reference lines. The standard clocked DRAM sense amp shown in Figure 11 satisfies all these requirements except the need for low input currents. When this sense amp goes from sense to sample, the capacitance of the internal nodes 83 and 84 in Figure11 is discharged through the reference line 68 and input 69, respectively. This particular current is small, but the sum of such currents from all the inputs into the reference lines summed over all devices can be reasonably large.

[0102] The fact that the sign of the current depends upon on the previous received data makes matters worse. One way to solve this problem is to divide the sample period into two phases. During the first phase, the inputs are shorted to a buffered version of the reference level (which may have an offset). During the second phase, the inputs are connected to the true inputs. This scheme does not remove the input current completely, since the input must still charge nodes 83 and 84 from the reference value to the current input value, but it does reduce the total charge required by about a factor of 10 (requiring only a 0.25V change rather than a 2.5V change). Persons skilled in the art will recognize that many other methods can be used to provide a clocked amplifier that will operate on very low input currents.

[0103] One important part of the input/output circuitry generates an internal device clock based on early and late bus clocks. Controlling clock skew (the difference in clock timing between devices) is important in a system running with 2 ns cycles, thus the internal device clock is generated so the input sampler and the output driver operate as close in time as possible to midway between the two bus clocks.

[0104] A block diagram of the internal device clock generating circuit is shown in Figure 12 and the corresponding timing diagram in Figure 13. The basic idea behind this circuit is relatively simple. A DC amplifier 102 is used to convert the small-swing bus clock into a full-swing CMOS signal. This signal is then fed into a variable delay line 103. The output of delay line 103 feeds three additional delay lines: 104 having a fixed delay; 105 having the same fixed delay plus a second variable delay; and 106 having the same fixed delay plus one half of the second variable delay. The outputs 107, 108 of the delay lines 104 and 105 drive clocked input receivers 101 and 111 connected to early and late bus clock inputs 100 and 110, respectively. These input receivers 101 and 111 have the same design as the receivers described

Petitioners
Ex. 1002, p. 375

above and shown in Fig. 11. Variable delay lines 103 and 105 are adjusted via feedback lines 116, 115 so that input receivers 101 and 111 sample the bus clocks just as they transition. Delay lines 103 and 105 are adjusted so that the falling edge 120 of output 107 precedes the falling edge 121 of the early bus clock, Clock1 53, by an amount of time 128 equal to the delay in input sampler 101. Delay line 108 is adjusted in the same way so that falling edge 122 precedes the falling edge 123 of late bus clock, Clock2 54, by the delay 128 in input sampler 111.

[0105] Since the outputs 107 and 108 are synchronized with the two bus clocks and the output 73 of the last delay line 106 is midway between outputs 107 and 108, that is, output 73 follows output 107 by the same amount of time 129 that output 73 precedes output 108, output 73 provides an internal device clock midway between the bus clocks. The falling edge 124 of internal device clock 73 precedes the time of actual input sampling 125 by one sampler delay. Note that this circuit organization automatically balances the delay in substantially all device input receivers 71 and 72 (Fig. 10), since outputs 107 and 108 are adjusted so the bus clocks are sampled by input receivers 101 and 111 just as the bus clocks transition.

[0106] In the preferred embodiment, two sets of these delay lines are used, one to generate the true value of the internal device clock 73, and the other to generate the complement 74 without adding any inverter delay. The dual circuit allows generation of truly complementary clocks, with extremely small skew. The complement internal device clock is used to clock the 'even' input receivers to sample at time 127, while the true internal device clock is used to clock the 'odd' input receivers to sample at time 125. The true and complement internal device clocks are also used to select which data is driven to the output drivers. The gate delay between the internal device clock and output circuits driving the bus is slightly greater than the corresponding delay for the input circuits, which means that the new data always will be driven on the bus slightly after the old data has been sampled.

DRAM Column Access Modification

[0107] A block diagram of a conventional 4 MBit DRAM 130 is shown in Figure 15. The DRAM memory array is divided into a number of subarrays 150-157, for example, 8. Each subarray is divided into arrays 148, 149 of memory cells. Row address selection is performed by decoders 146. A column decoder 147A, 147B, including column sense amps on either side of the decoder, runs through the core of each subarray. These column sense amps can be set to precharge or latch the most-recently stored value, as described in detail above. Internal I/O lines connect each set of sense-amps, as gated by corresponding column decoders, to input and output circuitry connected ultimately to the device pins. These internal I/O lines are used to drive the data from the selected bit lines to the data pins (some of pins 131-145), or to take the data from the pins and write the selected bit lines. Such a column access path organized by prior art constraints does not have sufficient bandwidth to interface with a high speed bus. The method of this invention does not require changing the overall method used for column access, but does change implementation details. Many of these details have been implemented selectively in certain fast memory devices, but never in conjunction with the bus architecture of this invention.

[0108] Running the internal I/O lines in the conventional way at high bus cycle rates is not possible. In the preferred method, several (preferably 4) bytes are read or written during each cycle and the column access path is modified to run at a lower rate (the inverse of the number of bytes accessed per cycle, preferably 1/4 of the bus cycle rate). Three different techniques are used to provide the additional internal I/O lines required and to supply data to memory cells at this rate. First, the number of I/O bit lines in each subarray running through the column decoder 147 is increased, for example, to 16, eight for each of the two columns of column sense amps and the column decoder selects one set of columns from the "top" half 148 of subarray 150 and one set of columns from the "bottom" half 149 during each cycle, where the column decoder selects one column sense amp per I/O bit line. Second, each column I/O line is divided into two halves, carrying data independently over separate internal I/O lines from the left half 147A and right half 147B of each subarray (dividing each subarray into quadrants) and the column decoder selects sense amps from each right and left half of the subarray, doubling the number of bits available at each cycle. Thus each column decode selection turns on *n* column sense amps, where n equals four (top left and right, bottom left and right quadrants) times the number of I/O lines in the bus to each subarray quadrant (8 lines each x 4=32 lines in the preferred implementation). Finally, during each RAS cycle, two different subarrays, e.g. 157 and 153, are accessed. This doubles again the available number of I/O lines containing data. Taken together, these changes increase the internal I/O bandwidth by at least a factor of 8. Four internal buses are used to route these internal I/O lines. Increasing the number of I/O lines and then splitting them in the middle greatly reduces the capacitance of each internal I/O line which in turn reduces the column access time, increasing the column access bandwidth even further.

[0109] The multiple, gated input receivers described above allow high speed input from the device pins onto the internal I/O lines and ultimately into memory. The multiplexed output driver described above is used to keep up with the data flow available using these techniques. Control means are provided to select whether information at the device pins should be treated as an address, and therefore to be decoded, or input or output data to be driven onto or read from the internal I/O lines.

[0110] Each subarray can access 32 bits per cycle, 16 bits from the left subarray and 16 from the right subarray. With

Petitioners
Ex. 1002, p. 376

8 I/O lines per sense-amplifier column and accessing two subarrays at a time, the DRAM can provide 64 bits per cycle. This extra I/O bandwidth is not needed for reads (and is probably not used), but may be needed for writes. Availability of write bandwidth is a more difficult problem than read bandwidth because over-writing a value in a sense-amplifier may be a slow operation, depending on how the sense amplifier is connected to the bit line. The extra set of internal I/O lines provides some bandwidth margin for write operations.

[0111] Persons skilled in the art will recognize that many variations of the teachings of this invention can be practiced that still fall within the claims of this invention which follow.

[0112] It follows a list of further embodiments of the invention:

Embodiment 1. A memory subsystem comprising
two memory devices connected in parallel to a bus,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices,
said control information including device-select information,
said bus containing substantially fewer bus lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual memory devices.

Embodiment 2. The memory subsystem with the features of embodiment 1 wherein said bus contains at least 8 bus lines adapted to carry at least 16 address bits and at least 8 data bits.

Embodiment 3. The memory subsystem with the features of embodiment 1 wherein said bus also includes parallel lines for clock and power.

Embodiment 4. A system comprising
a memory subsystem with the features of embodiment 1 wherein each bus of said memory subsystem is connected to its own transceiver device,
a transceiver bus connecting said transceiver devices, and
a means for transferring information between each of said buses of said memory subsystems and said transceiver bus, whereby memory subsystems may be integrated into a larger system having more memory than an individual memory subsystem.

Embodiment 5. The system with the features of embodiment 4 having a plurality of memory subsystems.

Embodiment 6. The system with the features of embodiment 4 further comprising a master device connected to said transceiver bus.

Embodiment 7. The system with the features of embodiment 6 wherein said master device is selected from the group consisting of a central processing unit, a floating point unit and a direct memory access unit.

Embodiment 8. The system with the features of embodiment 4 further comprising a peripheral device connected to the transceiver bus, said peripheral device adapted for connection to other devices not on the bus.

Embodiment 9. The system with the features of embodiment 8 wherein said peripheral device is selected from the group consisting of an I/O interface port, a video controller and a disk controller.

Embodiment 10. The system with the features of embodiment 5 wherein said transceiver bus is in a different plane than the plane of the bus of each of said memory subsystems.

Embodiment 11. The system with the features of embodiment 5 wherein the bus of each memory subsystem lies substantially in a subsystem bus plane and said transceiver bus lies substantially in a plane orthogonal to said subsystem bus plane.

Embodiment 12. The system with the features of embodiment 4 having at least two transceiver buses, each transceiver bus having a plurality of memory subsystem buses connected through a first transceiver to said transceiver bus,
each of said transceiver buses being further connected to a second transceiver adapted to interface to a second-order transceiver bus, whereby each transceiver bus is connected through said second transceiver to form a second-order transceiver bus unit.

Embodiment 13. A semiconductor subsystem bus for interconnecting semiconductor devices comprising
a plurality of semiconductor devices connected in parallel to a bus, at least one of said semiconductor devices being a memory device or a transceiver device which in turn is connected to a memory subsystem,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including semiconductor device-select information,
said bus containing substantially fewer bus lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual semiconductor devices, and
at least one modifiable register in each of the semiconductor devices on said bus, said modifiable registers being accessible from said bus, whereby the subsystem can be configured using signals transmitted on said bus.

Embodiment 14. The semiconductor subsystem bus with the features of embodiment 13 wherein one type of modifiable

Petitioners
Ex. 1002, p. 377

register is an access-time register designed to store a time delay after which a device may take some specified action on said bus.

Embodiment 15. The semiconductor subsystem bus with the features of embodiment 13 further comprising a semiconductor device having at least two access-time registers and

one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said bus.

Embodiment 16. The semiconductor subsystem bus with the features of embodiment 13 further comprising a memory device having at least one discrete memory section and also having a modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 17. The semiconductor subsystem bus with the features of embodiment 16 wherein said memory address information comprises a pointer to said discrete memory section.

Embodiment 18. The semiconductor subsystem bus with the features of embodiment 16 wherein said discrete memory section has a top and a bottom and said memory address information comprises pointers to said top and said bottom.

Embodiment 19. The semiconductor subsystem bus with the features of embodiment 16 wherein said memory address information comprises

a pointer to said discrete memory section and

a range value indicating the size of said discrete memory section.

Embodiment 20. The semiconductor subsystem bus with the features of embodiment 16 wherein said address registers of each of said discrete memory sections of each of said memory devices connected to said bus are set to contain memory address information that is different for each discrete memory section and such that the highest memory address in each discrete memory section is one less than the lowest memory address in another discrete memory section, whereby memory may be organized into one or a small number of contiguous memory blocks.

Embodiment 21. The semiconductor subsystem bus with the features of embodiment 16 further comprising a means for testing each of said discrete memory sections of each of said memory devices for proper function, and

for each non-functional discrete memory section, a means for setting at least one address register which corresponds to said discrete memory section to indicate that said discrete memory section is non-functional,

for each functional discrete memory section, a means for setting at least one address register which corresponds to said discrete memory section to contain such corresponding address information.

Embodiment 22. The semiconductor subsystem bus with the features of embodiment 21 wherein said address registers corresponding to said discrete memory sections are set to provide one contiguous memory block within the subsystem.

Embodiment 23. The semiconductor subsystem bus with the features of embodiment 13 wherein one of said modifiable registers is a device identification register which can be modified to contain a value unique to that semiconductor device.

Embodiment 24. The semiconductor subsystem bus with the features of embodiment 23 wherein said device identification register is set to contain a unique value which is a function of the physical position of that semiconductor device either along said bus or in relationship to other semiconductor devices or said bus.

Embodiment 25. A bus subsystem comprising

two semiconductor devices connected in parallel to a bus, wherein one of said semiconductor devices is a master device, said master device including a means for initiating bus transactions,

said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said devices,

said control information including device-select information,

said bus containing substantially fewer lines than the number of bits in a single address, and

said bus carrying device-select information without the need for separate device-select lines connected directly to individual devices on said bus, whereby said master device initiates bus transactions which transfer information between said semiconductor devices on said bus.

Embodiment 26. The bus subsystem with the features of embodiment 25 wherein one of said semiconductor devices is a memory device connected to said bus, said memory device having at least one discrete memory section and also having a modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 27. The bus subsystem with the features of embodiment 26 wherein one of said semiconductor devices comprises a transceiver device connected in parallel to said bus and connected in parallel to a memory device on a bus other than said bus.

28. The bus subsystem with the features of embodiment 26 further including a means Embodiment for said master device to request said memory device to prepare for a bus transaction by sending a request packet along said bus, said memory device and said master device each having a device-internal means to prepare to begin said bus transaction during a device-internal phase and further having a bus access means to effect said bus transaction during a bus access phase, said request packet including

a sequence of bytes containing address and control information,

Petitioners
Ex. 1002, p. 378

said control information including information about the requested bus transaction and about the access time, which corresponds to a number of bus cycles, which needs to intervene before beginning said bus-access phase, and
said address information pointing to at least one memory location within one of said discrete memory sections of said memory device.

Embodiment 29. The bus subsystem with the features of embodiment 28 wherein said memory device includes a means to read said control information and initiate said device-internal means at a time so as to complete said device-internal phase within said access time and begin said bus access phase after said number of bus cycles.

Embodiment 30. The bus subsystem with the features of embodiment 28 wherein said control information comprises an op code.

Embodiment 31. The bus subsystem with the features of embodiment 30 wherein said memory device includes sense amplifiers adapted to hold a bit of information or to precharge after a selected time and a means to transfer a data block during a data block transfer either reading data from said memory device or writing data into said memory device, and wherein said op code instructs said memory device to activate a response means, said response means including a means to

initiate a data block transfer,
select the size of said data block,
select the time to initiate said data block transfer,
access a control register, including reading from or writing to said control register,
precharge said sense amplifiers after each of said data block transfers is complete,
hold a bit of information in each of said sense amplifiers after each of said data block transfers is complete, or
select normal or page-mode access.

Embodiment 32. The bus subsystem with the features of embodiment 31 wherein said data block transfer comprises a read from or a write to memory within a single memory device.

Embodiment 33. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to a specific one of said semiconductor devices on said bus by including in said request packet a device identification number unique to said semiconductor device.

Embodiment 34. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to a selected one of said discrete memory portions by including in said request packet a specific memory address.

Embodiment 35. The bus subsystem with the features of embodiment 28 further comprising a means for said master device to send control information to substantially all semiconductor devices on said bus by including in said request packet a special device identification number which is recognized by said semiconductor devices.

Embodiment 15 36. The bus subsystem with the features of embodiment 28 wherein said control information specifies directly or indirectly the number of bus cycles for said master device and said memory device to wait before beginning said bus access phase.

Embodiment 37. The bus subsystem with the features of embodiment 36 wherein, for a data block transfer, said master device and said memory device use the same access time and same data block size regardless of whether said data block transfer is a read or write operation.

Embodiment 38. The bus subsystem with the features of embodiment 28 wherein said control information further includes a block-size value that encodes and specifies the size of the block of data to be transferred.

Embodiment 39. The bus subsystem with the features of embodiment 38 wherein said block-size value is encoded as a linear value for relatively small block sizes values and is encoded as a logarithmic value for relatively larger block sizes.

Embodiment 40. The bus subsystem with the features of embodiment 38 wherein said block-size value is encoded using four bits, and where the encoded value is

Encoded Value	Block Size (Bytes)
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	16

Petitioners
Ex. 1002, p. 379

(continued)

Encoded Value	Block Size (Bytes)
10	32
11	64
12	128
13	256
14	512
15	1024

Embodiment 41. The bus subsystem with the features of embodiment 26 wherein said memory device is a DRAM device containing
a plurality of sense amplifiers,
a means to hold said sense amplifiers in an unmodified state after a read or write operation, leaving the device in page mode,
a means to precharge said sense amplifiers and
a means for selecting whether to precharge said sense amplifiers or to hold said sense amplifiers in an unmodified state.
Embodiment 42. The bus subsystem with the features of embodiment 28 wherein said request packet comprises an even number of bytes.
Embodiment 43. The bus subsystem with the features of embodiment 28 further including a means for generating and controlling a plurality of bus cycles, during which said bus carries said address, data and control information, and wherein alternate said bus cycles are designated odd cycles and even cycles, respectively, and wherein said request packet begins only on an even cycle.
Embodiment 44. The bus subsystem with the features of embodiment 28 further including a means for generating ECC information corresponding to a block of data and a means for using said ECC information to correct errors in storing or reading said block of data, wherein said ECC information may be stored separately from said block of data.
Embodiment 45. The bus subsystem with the features of embodiment 44 further comprising at least two of said memory devices wherein said ECC information and said corresponding block of data are stored in a first and a second said memory device, respectively, and said master device includes a means to write or read said block of data with error correction by sending separate ones of said request packets for said ECC information and for said corresponding block of data.
Embodiment 46. A bus subsystem comprising
a memory device and a master device connected in parallel on a bus,
a means for said master device to send a request packet and initiate a bus transaction and
a means for said master device to keep track of current and pending bus transactions,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said memory devices,
said bus containing substantially fewer lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual devices on said bus, whereby said master device initiates bus transactions which transfer information between devices on said bus and collisions on said bus are avoided because said master device avoids initiating bus transactions which would conflict with current or pending bus transactions.
Embodiment 47. The bus subsystem with the features of embodiment 46 having at least two of said master devices and including
a collision detecting means whereby a first said master device sending a first said request packet can detect a second said master device sending one of said colliding request packets, where one of said said colliding request packet may be sent simultaneous with the initial sending of or overlapping the sending of said first request packet, and
an arbitration means whereby said first and said second master devices select a priority order in which each of said master devices will be allowed to access said bus sequentially .
Embodiment 48. The bus subsystem with the features of embodiment 47 wherein each of said master devices has a master ID number and each of said request packets includes a master ID position which is a predetermined number of bits in a predetermined position in said request packet, and wherein said collision detection means comprises
a means included in each master device for sending a request packet including said master ID number of said master device in said master ID position of said request packet and
a means to detect a collision and invoke said arbitration means if any master device detects any other master ID number in said master ID position.
Embodiment 49. The bus subsystem with the features of embodiment 47 wherein each of said master devices includes

Petitioners
Ex. 1002, p. 380

a means for sending a request packet,
a means for driving a selected bus line or lines during at least one selected bus cycle while said request packet is being sent,
a means for monitoring said selected bus line or lines to see if a said master device is sending a colliding request packet and
a means for informing all other master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 50. The bus subsystem with the features of embodiment 47 wherein each of said master devices includes
a means, when sending a request packet, to drive a selected bus line or lines with a certain current during at least one selected bus cycle,
a means for monitoring said selected bus line or lines for a greater than normal current to see if another master device is driving that line or lines,
a means for detecting said greater than normal current, and
a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 51. The bus subsystem with the features of embodiment 47 wherein said arbitration means comprises
a means for initiating an arbitration cycle,
a means for allocating a single bus line to each master device during at least one selected bus cycle relative to the start of said arbitration cycle,
a means for allocating each master device to a single bus line during one of said selected bus cycles if there are more master devices than available bus lines,
a means for each of said master devices which sent a colliding request packet to drive said bus line allocated to said master device during said selected bus cycle, and
a means in at least one of said master devices for storing information about which master devices sent a colliding request packet,
whereby said master devices can monitor selected bus lines during said arbitration cycle and identify each said master device which sent a colliding request packet.

Embodiment 52. The bus subsystem with the features of embodiment 47 wherein said arbitration means comprises
a means included in a first one of said master devices which sent colliding request packets for identifying each of said master devices which sent colliding request packets,
a means for assigning a priority to each said master device which sent a colliding request packet, and
a means for allowing each said master device which sent a colliding request packet to access the bus sequentially according to that priority.

Embodiment 53. The bus subsystem with the features of embodiment 52 wherein said priority is based on the physical location of each of said master devices.

Embodiment 54. The bus subsystem with the features of embodiment 52 wherein said priority is based on said master ID number of said master devices.

Embodiment 55. The bus subsystem with the features of embodiment 52 wherein each of said master devices includes
a means, when sending a colliding request packet, for deciding which master device can send the next request packet in what order or at what time, whereby no master device may send a new request packet until responses to each pending request packet have been completed or scheduled.

Embodiment 56. A bus subsystem comprising
a plurality of semiconductor devices connected in parallel to a bus,
said bus including a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including device-select information,
said bus containing substantially fewer lines than the number of bits in a single address,
said bus carrying said device-select information without the need for separate device-select lines connected directly to individual semiconductor devices,
said semiconductor devices including a reset means having an input and an output, the output of the reset means of one semiconductor device being connected to the input of the reset means of the next semiconductor device in series.

Embodiment 57. The bus subsystem with the features of embodiment 56 further including system reset means comprising
a means for generating a first and a second reset signal,
a means for passing said first reset signal to a first of said semiconductor devices and then to subsequent ones of said semiconductor devices in series and
a means for passing a second reset signal to said first semiconductor device and then to said subsequent semiconductor devices in series,
said bus subsystem including one of said semiconductor devices containing
a device identification register adapted to contain a number unique to said semiconductor device within said bus subsystem,
a device identification register setting means, and
a device reset means for resetting said semiconductor device to some desired, known reset state in response to said

Petitioners
Ex. 1002, p. 381

first reset signal and for setting said device identification register in response to said second reset signal,
whereby said bus subsystem can be reset to a known reset state with a unique device identification value in said device identification register of each of said semiconductor devices.

Embodiment 58. The bus subsystem with the features of embodiment 57 wherein said desired, known reset state is where all registers in the semiconductor device are cleared and the state machines are reset.

Embodiment 59. The bus subsystem with the features of embodiment 57 wherein said device identification register setting means comprises
a means for detecting said second reset signal,
a means for reading a device identification number from said bus lines at a specific time relative to said second reset signal and
a means for storing said device identification number in said device identification register of said semiconductor device.

Embodiment 60. The bus subsystem with the features of embodiment 57 wherein said second reset signal comprises multiple pulse sequences and wherein said device identification setting means includes
a means for interpreting said pulse sequences as a device identification number and
a means for storing said device identification number in said device identification register of said semiconductor device.

Embodiment 61. The bus subsystem with the features of embodiment 57 wherein said device reset means comprises an n-stage shift register capable of storing n-bit values, wherein said device reset means interprets a specific value in said shift register as said first reset signal and interprets a specific value in said shift register as said second reset signal.

Embodiment 62. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices is a master device, said master device including a means for generating said first and said second reset signals.

Embodiment 63. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices is a master device, said master device including
a master ID register,
a means for assigning a master ID number to said master device and
a means for storing said master ID number in said master ID register.

Embodiment 64. The bus subsystem with the features of embodiment 63 further comprising a second one of said master devices, and a means for a first one of said master devices to assign a master ID number to substantially all other said master devices, whereby said first master device assigns one of said master ID numbers to each of said master devices on said bus subsystem and each said master device stores said assigned master ID number in said master ID register.

Embodiment 65. The bus subsystem with the features of embodiment 57 wherein one of said semiconductor devices includes a device-type register adapted to contain an identifier characteristic of that type of semiconductor device, and one or more modifiable registers, at least one of which is an access-time register adapted for storing access times.

Embodiment 66. The bus subsystem with the features of embodiment 65 wherein one of said semiconductor devices is a master device having
a means for selecting a semiconductor device,
a means for reading said device-type register of said selected semiconductor device,
a means for determining the device type of said selected semiconductor device,
a means for determining access-time values appropriate for said selected semiconductor device and for storing said access-time values in said access-time registers of said selected semiconductor device, and
a means for selecting and storing other values appropriate for said selected semiconductor device in corresponding registers of said selected semiconductor device,
whereby said master device can select a semiconductor device, determine what type it is, and set said access-time and other registers to contain appropriate values.

Embodiment 67. The bus subsystem with the features of embodiment 66 further comprising a memory device having at least one discrete memory section and at least one modifiable address register adapted to store memory address information which corresponds to each of said discrete memory sections, and
said master device further comprising a means for selecting and testing each of said discrete memory sections and a means for storing address information in said address registers corresponding to each of said discrete memory sections,
whereby said master device can test all said discrete memory sections and assign unique address values thereto.

Embodiment 68. A bus subsystem comprising
two semiconductor devices connected in parallel to a bus, one of said semiconductor devices being a master device,
said bus including a plurality of bus data lines for carrying substantially all address, data and control information needed by said semiconductor devices,
said control information including device-select information,
said bus containing substantially fewer of said bus data lines than the number of bits in a single address, and
said bus carrying device-select information without the need for separate device-select lines connected directly to individual semiconductor devices,
wherein all of said bus data lines are terminated transmission lines and all of said address, data and control information

Petitioners
Ex. 1002, p. 382

is carried on said bus data lines as a sequential series of bits in the form of low-voltage-swing signals.

Embodiment 69. The bus subsystem with the features of embodiment 68 further comprising a semiconductor device including a current-mode driver connected to drive one of said bus data lines.

Embodiment 70. The bus subsystem with the features of embodiment 69 further comprising a semiconductor device having a means to measure the voltage of said low-voltage-swing signals on a selected one of said bus data lines, whereby said semiconductor device can determine whether zero, one, or more than one of said current-mode drivers are driving said selected bus data line.

Embodiment 71. The bus subsystem with the features of embodiment 70 further comprising a semiconductor device having

a plurality of input receivers connected to one of said bus data lines, and

a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits.

Embodiment 72. The bus subsystem with the features of embodiment 70 further comprising a semiconductor device having two input receivers connected to one of said bus data lines.

Embodiment 73. A bus subsystem comprising

two semiconductor devices connected in parallel to a bus having a first and a second end, said bus including a bus clock line, said bus clock line having first and second ends corresponding to said first and second ends of said bus, respectively,

a clock generator connected to said first end of said bus clock line to generate early bus clock signals with a normal rise time, and

signal return means at said second end of said bus clock line to return said early bus clock signals to said first end of said bus as corresponding late bus clock signals,

whereby each of said early bus clock signals will propagate from said clock generator along said clock line starting from said first end to said second end of said bus and then return at a later time to said first end of said bus as a corresponding late bus clock signal, whereby each semiconductor device on said bus can detect said early bus clock signals and said corresponding late bus clock signals.

Embodiment 74. The bus subsystem with the features of embodiment 73 further comprising a first and a second said bus clock line having first and second ends at said first and said second ends of said bus, respectively, wherein said signal return means directly connects said second ends of said first and said second bus clock lines whereby each of said early bus clock signals will propagate from said clock generator at said first end of said bus along said first bus clock line to said second end of said bus and then return on said second bus clock line to said first end of said bus as one of said corresponding late bus clock signals.

Embodiment 75. The bus subsystem with the features of embodiment 73 wherein said signal return means comprises said first bus clock line without a line terminator at said second end thereof whereby each of said early bus clock signals reaching said second end of said first bus clock line will be reflected back along said first bus clock line as said corresponding late bus clock signals.

Embodiment 76. The bus subsystem with the features of embodiment 73 further comprising

a means for operating said bus in bus cycles timed to have a certain bus cycle frequency and a corresponding bus cycle period and

a means for operating said clock generator with a period of twice the bus cycle period.

Embodiment 77. The bus subsystem with the features of embodiment 76 wherein said bus cycle frequency is greater than approximately 50 MHz and less than or equal to approximately 500 MHz.

Embodiment 78. The bus subsystem with the features of embodiment 73 further including a semiconductor device having an internal device clock generating means to derive the midpoint time between said early and corresponding late bus clock signals and to generate an internal device clock synchronized to said midpoint time.

Embodiment 79. The bus subsystem with the features of embodiment 73 further including a semiconductor device having a low-skew clock generator circuit comprising

a first delay line having an input, an output and a basic delay and means for synchronizing the output of said first delay line with said early bus clock signal,

a second delay line having said basic delay plus a variable delay, said second delay line having an output and a means for synchronizing the output of said second delay line with said late bus clock signal, and

a third delay line having a third delay and a means to set said third delay midway between the delays of said first and second delay lines, said third delay line having an output which provides an internal device clock signal synchronized to a time halfway between said early and said late bus clock signals.

Embodiment 80. The bus subsystem with the features of embodiment 73 wherein said early and said late bus clock signals are low-voltage-swing signals that transition cyclically between low and high logical values, and further including a semiconductor device having a low-skew clock generator circuit comprising

a DC amplifier to convert said early and said late bus clock signals into full-swing logic signals,

a first variable delay line having a first variable delay and an input and an output, the input of said first variable delay

Petitioners
Ex. 1002, p. 383

line being connected to said DC amplifier
a first, a second and a third additional delay line, each having an input and an output, the input of each of said additional delay lines being connected to the output of said first delay line,
said first additional delay line having a fixed delay,
said second additional delay line having said fixed delay plus a second variable delay, and
said third additional delay line having said fixed delay plus one half of said second variable delay,
a first clocked input receiver connected to sample said early bus clock signal and gated by said output of said first additional delay line,
a means for adjusting said first variable delay so said first clocked input receiver samples said early bus clock signal just as said early bus clock signal transitions,
a second clocked input receiver connected to sample said late bus clock signal and gated by said output of said second additional delay line,
a means for adjusting said second variable delay so said second clocked input receiver samples said late bus clock signal just as said late bus clock signal transitions,
whereby said output of said third additional delay line is synchronized to a time halfway between said outputs of said first and said second additional delay lines, and said output of said third additional delay line provides an internal device clock signal.

Embodiment 81. The bus subsystem with the features of embodiment 80 further comprising a semiconductor device having
a first one of said low-skew clock generator circuits which generates a "true" internal device clock signal and
a second one of said low-skew clock generator circuits connected to generate a "complement" internal device clock signal synchronized with but opposite in logical value to said "true" internal device clock signal.

Embodiment 82. A DRAM device designed to be connected to an external bus having a plurality of bus lines for carrying substantially all address, data and control information needed by said DRAM device as a sequential series of bits, said control information including device-select information, said external bus containing substantially fewer said bus lines than the number of bits in a single address, and said bus carrying device-select information without the need for separate device-select lines connected directly to said DRAM device, said DRAM device comprising
an array of memory cells connected in rows and columns, each of said memory cells adapted to store one of said bits,
a row address selection means for selecting one of said rows,
a column sense amp connected to each of said columns, each of said column sense amps adapted to latch one of said bits as a binary logical value or to precharge to a selected state,
a column decoding means connected to each of said column sense amps for selecting a plurality of said column sense amps for inputting one of said bits to or outputting one of said bits from said memory cells,
an internal I/O bus having a plurality of internal I/O lines wherein each of said internal I/O lines is connected to a plurality of said column sense amps, and
a plurality of bus connection means designed to connect said internal I/O lines to said external bus,
whereby a selected bit of said sequential series of bits can be transferred from said external bus to a selected one of said memory cells or said bit contained in a selected one of said memory cells can be transferred to said external bus.

Embodiment 83. The DRAM device with the features of embodiment 82 further comprising
an output driver connected to one said bus connection means,
an output multiplexer having an output connected to said output driver and a plurality of inputs, each of said inputs being connected to one of said internal I/O lines, and
a control means to select whether said output driver can drive said external bus,
whereby a plurality of memory cells are selected using said row address selection means and said column decoding means and a plurality of bits contained in said plurality of memory cells are output through said column sense amps to said internal I/O bus to said output multiplexer to said output driver to said external bus.

Embodiment 84. The DRAM device with the features of embodiment 82 further comprising
a plurality of input receivers connected to one of said bus data lines and to said internal I/O bus,
a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits, and
a control means to select whether an input receiver can drive said internal I/O bus, whereby a bit of said sequential series of bits is input from said external bus through one of said input receivers to one of said internal I/O lines to one of said column sense amps to one of said memory cells.

Embodiment 85. The DRAM device with the features of embodiment 82 further comprising
a first and a second half-array of said memory cells wherein each said row of said array of said memory cells is subdivided into two parts,
a first and a second one of said internal I/O buses connected to said column sense amps in said first and said second half-arrays, respectively, and

Petitioners
Ex. 1002, p. 384

a column decoder means to gate selected ones of said column sense amps connected to said memory cells in a selected row of said first and said second half-arrays simultaneously.

Embodiment 86. The DRAM device with the features of embodiment 85 wherein said column decoder means selects sixteen column sense amps at a time.

Embodiment 87. The DRAM device with the features of embodiment 82 wherein said external bus operates at a certain speed and wherein said DRAM device includes four of said internal I/O buses, each of which operates at one-fourth the speed of said external bus.

Embodiment 88. The DRAM device with the features of embodiment 82 further comprising

a means for precharging one of said column sense amps to a precharged state from which a binary logical value can quickly be loaded into said column sense amp,

if said column sense amp contains a binary logical value, a means for latching the logical value currently contained in said column sense amp and

a means for instructing said DRAM device to precharge said column sense amp or latch said binary logical value in said column sense amp.

Embodiment 89. The DRAM device with the features of embodiment 88 further comprising a means for instructing said DRAM device to precharge said column sense amp without further instruction whenever said row address selection means selects a different one of said rows.

Embodiment 90. The DRAM device with the features of embodiment 88 further comprising a means for instructing said DRAM device to precharge said column sense amp without further instruction at a first or a second preselected time after latching the latest said binary logical value, said first preselected time being long enough for said DRAM to latch said binary logical value into said column sense amp and transfer said binary logical value into memory or onto one of said internal I/O lines, and said second preselected time being a variable which can be stored in said DRAM device whereby said DRAM can latch a binary logical value into said column sense amp for transferring said binary logical' value into or out of a selected said memory cell, then precharge to allow a faster subsequent read or write.

Embodiment 91. A package containing

a semiconductor die having a side, circuitry and a plurality of connecting areas positioned along or near said side, spaced at a selected pitch and connected to said circuitry,

said package comprising a plurality of bus connecting means for connecting to a plurality of external bus lines, each of said external bus lines corresponding to one of said connecting areas, each of said bus connecting means being positioned on a first side of said package,

connected to one said external bus line and to said corresponding connecting area on said semiconductor die, and spaced at a pitch substantially identical to said selected pitch of said connecting areas,

whereby each of said external bus lines can be connected to said corresponding connecting area on said semiconductor die by bus connection means positioned along a single side of said package.

Embodiment 92. The package with the features of embodiment 91 further comprising a plurality of said bus connecting means wherein each of said bus connecting means includes

a pin adapted for connection to one of said external bus lines and

a wire connecting said pin to one of said connecting areas on said semiconductor die,

said wire having an effective lead length less than about 4 millimeters and wherein the effective lead length of said wire of each of said bus connection means for said package is approximately equal.

Embodiment 93. A plurality of packages with the features of embodiment 91 wherein at least two of said semiconductor die are memory devices, each of said packages being generally flat, having a top and a bottom, and wherein

said packages are physically secured adjacent and parallel to each other in a stack,

where a first one of said packages is adjacent to a second one of said packages in said stack, said top of said first package is substantially aligned with said bottom of said second package, and

said bus connecting means of each of said packages are substantially aligned and are lying substantially in a plane.

Embodiment 94. The plurality of packages with the features of embodiment 93 further comprising a plurality of stacks wherein each of said bus connecting means can be electrically connected to corresponding said bus connecting means in each of said stacks.

Embodiment 95. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

at least one modifiable identification register accessible to said bus through said connection means, whereby data may

Petitioners
Ex. 1002, p. 385

be transmitted to said register via said bus and enable said device thereafter to be uniquely identified.

Embodiment 96. The semiconductor device with the features of embodiment 95 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 97. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and at least one modifiable register to hold device address information, said modifiable register accessible to said bus through said connection means, whereby data may be transmitted to said register via said bus which enables said device thereafter to respond to a predetermined range of addresses.

Embodiment 98. The semiconductor device with the features of embodiment 97 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 99. The semiconductor device with the features of embodiment 98 wherein said memory device has at least one discrete memory section and also has at least one modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 100. The semiconductor device with the features of embodiment 99 wherein said memory address information comprises a pointer to said discrete memory section.

Embodiment 101. The semiconductor device with the features of embodiment 100 wherein said discrete memory section has a top and a bottom and said memory address information comprises pointers to said top and said bottom.

Embodiment 102. The semiconductor device with the features of embodiment 100 wherein said memory address information comprises

a pointer to said discrete memory section and

a range value indicating the size of said discrete memory section.

Embodiment 103. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

at least one modifiable access-time register accessible to said bus through said connection means, whereby data may be transmitted to said register via said bus which establishes a predetermined amount of time that said semiconductor device thereafter must wait before using said bus in response to a request.

Embodiment 104. The semiconductor device with the features of embodiment 103 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 105. The semiconductor device with the features of embodiment 103 further comprising at least two access-time registers and one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said bus.

Embodiment 106. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, and wherein each said bus line is a terminated transmission line, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

a bus line driver capable of producing a low-voltage-swing signal on one of said terminated transmission lines.

Embodiment 107. The semiconductor device with the features of embodiment 106 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 108. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying

Petitioners
Ex. 1002, p. 386

substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said bus further including at least one bus clock line for carrying early and late bus clock signals, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus, and

an internal device clock generating means which generates an internal device clock synchronized to a time halfway between said early and said late bus clock signals.

embodiment 109. The semiconductor device with the features of embodiment 108 wherein said bus further includes a first and a second one of said bus clock lines, said first bus clock line carries said early bus clock signal and said second bus clock line carries said late bus clock signal, said semiconductor device further comprising a means to detect said early bus clock signal on said first bus clock line and a means to detect said late bus clock signal on said second bus clock line.

Embodiment 110. The semiconductor device with the features of embodiment 109 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 111. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying as a sequential series of bits substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said bus,

a plurality of input receivers connected to one of said bus data lines and

a selection means for selecting said input receivers one by one to sense and store, one at a time, the bits of said sequential series of bits.

Embodiment 112. The semiconductor device with the features of embodiment 111 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 113. The semiconductor device with the features of embodiment 112 wherein two input receivers are connected to one of said bus lines.

Embodiment 114. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said bus system includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus, and

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus.

Embodiment 115. The semiconductor device with the features of embodiment 114 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 116. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising

connection means adapted to connect said semiconductor device to said system bus,

an internal input/output bus within said semiconductor device having more lines than said system bus,

a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and

Petitioners
Ex. 1002, p. 387

at least one modifiable identification register accessible to said system bus through said connection means,
whereby data may be transmitted to said register via said system bus and which enables said device thereafter to be uniquely identified.

Embodiment 117. The semiconductor device with the features of embodiment 116 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 118. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, said semiconductor device comprising
connection means adapted to connect said semiconductor device to said system bus,
an internal input/output bus within said semiconductor device having more lines than said system bus,
a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and
at least one modifiable register to hold device address information, said modifiable register accessible to said system bus through said connection means, whereby data may be transmitted to said register via said system bus which enables said device thereafter to respond to a predetermined range of addresses.

Embodiment 119. The semiconductor device with the features of embodiment 118 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 120. The semiconductor device with the features of embodiment 119 wherein said memory device has at least one discrete memory section and also has at least one modifiable address register adapted to store memory address information which corresponds to each said discrete memory section.

Embodiment 121. A semiconductor device capable of use in an architecture for a semiconductor system bus including a plurality of semiconductor devices connected in parallel to a bus wherein said system bus includes a plurality of bus lines for carrying substantially all address, data and control information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said system bus, and has substantially fewer bus lines than the number of bits in a single address, said semiconductor device comprising
connection means adapted to connect said semiconductor device to said system bus,
an internal input/output bus within said semiconductor device having more lines than said system bus,
a means for multiplexing the lines of said internal bus to the lines of said system bus, whereby said system bus can run at a higher speed than said internal bus, and
at least one modifiable access-time register accessible to said system bus through said connection means, whereby data may be transmitted to said register via said system bus which establishes a predetermined amount of time that said semiconductor device thereafter must wait before using said system bus in response to a request.

Embodiment 122. The semiconductor device with the features of embodiment 121 wherein said semiconductor device is a memory device which connects substantially only to said system bus and sends and receives substantially all address, data and control information over said system bus.

Embodiment 123. The semiconductor device with the features of embodiment 121 further comprising at least two access-time registers and one of said access-time registers is permanently programmed to contain a fixed value and at least one of said access-time registers can be modified by information carried on said system bus.

Embodiment 124. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, wherein said address, data, control and device-select information is carried over said bus in the form of request packets and bus transactions, said semiconductor device comprising
connection means adapted to connect said semiconductor device to said bus,
a means to receive said request packets over said bus,
a means to decode information in said request packets, and
a means to respond to said information in said request packets.

Embodiment 125. The semiconductor device with the features of embodiment 124 wherein said means to decode information in said request packet further comprises

Petitioners
Ex. 1002, p. 388

a means to identify and decode said control information in said request packet,
a means to identify and decode said device-select information in said request packet,
a means to identify and decode said address information in said request packet and
a means to determine whether said control information or said address information instructs said semiconductor device to begin a response.

Embodiment 126. The semiconductor device with the features of embodiment 124 wherein each of said bus transactions is carried out in response to said address and said control information in one of said request packets, and wherein said means to identify and decode information in said request packets includes a means to identify a sequence of bytes on said bus as one of said request packets containing said address and said control information, said control information including information about the type of said bus transaction being requested and the access time which needs to intervene before beginning said bus transaction over said bus and said address and said control information includes device-select information instructing one or more said semiconductor devices to respond to said address and said control information.

Embodiment 127. The semiconductor device with the features of embodiment 124 further comprising
a plurality of sense amplifiers adapted to precharge to a selected state or to latch a bit of information,
a means to hold said sense amplifiers in an unmodified state after latching one of said bits of information,
a means to precharge said sense amplifiers and
a means for selecting whether said semiconductor device should precharge said sense amplifiers or should hold said sense amplifiers in an unmodified state.

Embodiment with the features of embodiment 128. The semiconductor device 124 wherein said means to respond to said information, where said information is control information, further comprises a means to
transfer a data block during a data block transfer, further including a means to
read data from said semiconductor device and
write data into said semiconductor device, and
initiate a data block transfer,
transfer a data block of a selected size,
transfer a data block at a selected time,
access a control register, including a means to read from or write to said control register, or
select normal or page-mode access.

Embodiment 129. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes a device identification number unique to said semiconductor device.

Embodiment 130. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes a special device identification number which calls for said semiconductor device to respond.

Embodiment 131. The semiconductor device with the features of embodiment 124 further comprising a means to respond to said information in said request packet if said information includes an address unique to said semiconductor device.

Embodiment 132. The semiconductor device with the features of embodiment 124 further comprising a means to interpret said control information and decode the time to wait before beginning said bus transaction over said bus.

Embodiment 133. The semiconductor device with the features of embodiment 124 further comprising a means to interpret said control information and decode the size of a data block to transfer during one of said bus transactions.

Embodiment 134. The semiconductor device with the features of embodiment 124, 125, 126, 127, 128, 129, 130, 131, 132 or 133 wherein said semiconductor device is a memory device which connects substantially only to said bus and sends and receives substantially all address, data and control information over said bus.

Embodiment 135. A semiconductor device capable of use in a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to a bus wherein said bus includes a plurality of bus lines for carrying substantially all address, data, control and device-select information needed by said semiconductor device for communication with substantially every other semiconductor device connected to said bus, and has substantially fewer bus lines than the number of bits in a single address, and carries device-select information for said semiconductor device without the need for a separate device-select line connected directly to said individual semiconductor device, wherein said address, data, control and device-select information is carried over said bus in the form of request packets and bus transactions, said semiconductor device comprising
connection means adapted to connect said semiconductor device to said bus,
a means to encode address and control information in said request packets and
a means to send said request packets over said bus.

Embodiment 136. The semiconductor device 135 further comprising a means to request a bus transaction wherein each of said bus transactions is carried out in response to said address and said control information in one of said request packets, and wherein said means to encode information in said request packets includes a means to mark a sequence

Petitioners
Ex. 1002, p. 389

of bytes on said bus as one of said request packets, said control information including information about the type of said bus transaction being requested and the access time which needs to intervene before beginning said bus transaction over said bus and said address and said control information includes device-select information instructing one or more said semiconductor devices to respond to said address and said control information.

Embodiment 137. The semiconductor device with the features of embodiment 135 wherein one or more of said plurality of semiconductor devices has a unique device identification number, said semiconductor device further comprising a means to send control information to a specific one of said plurality of semiconductor devices by including in said request packet a selected said device identification number.

Embodiment 138. The semiconductor device with the features of embodiment 135 wherein each of said plurality of semiconductor devices is adapted to respond to a special device identification number, said semiconductor device further comprising a means to send control information to each of said plurality of semiconductor devices by including in said request packet said special device identification number.

Embodiment 139. The semiconductor device with the features of embodiment 135 wherein one or more of said plurality of semiconductor devices is a memory device having a plurality of addresses, said semiconductor device further comprising a means to send control information to a specific address or range of addresses in one of said plurality of semiconductor devices by including said specific address or range of addresses in said request packet.

Embodiment 140. The semiconductor device with the features of embodiment 135 wherein at least one of said request packets is a request packet requesting a bus transaction which is followed by a corresponding one of said bus transactions, said semiconductor device further comprising a means to encode said control information to specify directly or indirectly the time between the end of said request packet requesting a bus transaction and said corresponding bus transaction over said bus.

Embodiment 141. The semiconductor device with the features of embodiment 140 wherein one type of said bus transactions is a transfer of a data block, said semiconductor device further comprising a means to encode said control information to specify the size of said data block to transfer.

Embodiment 142. The semiconductor device with the features of embodiment 140 further comprising a means to keep track of current and pending bus transactions, whereby collisions on said bus are avoided because said semiconductor device avoids initiating bus transactions which would conflict with current or pending bus transactions.

Embodiment 143. The semiconductor device with the features of embodiment 135 wherein said semiconductor device is a first master device and one of said plurality of semiconductor devices is a second master device, further comprising

a collision detecting means whereby said first master device when sending a first one of said request packets can detect said second master device sending a colliding one of said request packets, where said colliding request packet may be sent simultaneous with the initial sending of or overlapping the sending of said first request packet, and

an arbitration means whereby said first and said second master devices select a priority order in which each of said master devices will be allowed to access said bus sequentially.

Embodiment 144. The semiconductor device with the features of embodiment 143 wherein said semiconductor device is a master device and at least one of said plurality of semiconductor devices is a master device, each of said master devices has a master ID number and each of said request packets includes a master ID position which is a predetermined number of bits in a predetermined position in said request packet, and wherein said collision detection means comprises

a means for said semiconductor device to send its master ID number in said request packet and

a means to detect a collision and invoke said arbitration means if said semiconductor device detects any other master ID number in said master ID position.

Embodiment 145. The semiconductor device with the features of embodiment 144 wherein said system bus architecture includes a means for carrying information on said bus during bus cycles, said semiconductor device further comprising

a means for driving a selected bus line or lines during at least one selected bus cycle while sending each said request packet,

a means for monitoring said selected bus line or lines to see if another said master device is sending one of said colliding request packets and

a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 146. The semiconductor device with the features of embodiment 145 further comprising

a means, when sending a request packet, for driving a selected bus line or lines with a certain current during at least one selected bus cycle,

a means for monitoring said selected bus line or lines for a greater than normal current to see if another said master device is driving that line or lines,

a means for detecting said greater than normal current, and

a means for informing all said master devices that a collision has occurred and for invoking said arbitration means.

Embodiment 147. The semiconductor device with the features of embodiment 143 wherein said arbitration means comprises

a means for initiating an arbitration cycle,

Petitioners
Ex. 1002, p. 390

a means for allocating a single bus line to each said master device during at least one selected bus cycle relative to the start of said arbitration cycle,
a means for allocating each said master device to a single bus line during one of said selected bus cycles if there are more master devices than available bus lines,
a means for each of said master devices which sent one of said colliding request packets to drive said bus line allocated to said master device during said selected bus cycle, and
a means in at least one of said master devices for storing information about which master devices sent one of said colliding request packets,
whereby said master devices can monitor selected bus lines during said arbitration cycle and identify each said master device which sent one of said colliding request packets.
Embodiment 148. The semiconductor device with the features of embodiment 143 wherein said arbitration means comprises
a means for identifying each of said master devices which sent one of said colliding request packets,
a means for assigning a priority to each said master device which sent one of said colliding request packets, and
a means for allowing each said master device which sent one of said colliding request packets to access the bus sequentially according to that priority.
Embodiment 149. The semiconductor device with the features of embodiment 143 wherein said priority is based on the physical location of each of said master devices.
Embodiment 150. The semiconductor device with the features of embodiment 143 wherein said priority is based on said master ID number of said master devices.

Claims

1. A semiconductor device, the semiconductor device comprising:

connection means adapted to connect the semiconductor device to an external bus which is a part of a semiconductor bus architecture including a plurality of semiconductor devices connected in parallel to the external bus, wherein the external bus includes a plurality of bus lines for carrying substantially all address, data, control and device select information needed by the semiconductor device for communication with substantially every other semiconductor device connected to the external bus, and wherein the external bus has substantially fewer bus lines than the number of bits in a single address, wherein each bus line of the plurality of bus lines is a terminated transmission line, and wherein the external bus carries device-select information for the semiconductor device without the need for a separate device-select line connected directly to the semiconductor device,
a bus line driver capable of producing a low-voltage-swing signal on one of the terminated transmission lines,

wherein the semiconductor device is a semiconductor transceiver device to interface multiple units to the bus, wherein the multiple units are DRAMs.

2. The semiconductor device of claim 1,
wherein the bus line driver is capable of producing a low-voltage-swing signal having a voltage swing of not more than 1 Volt on one of said terminated transmission lines.

3. The semiconductor device of claim 1 or 2,
wherein the bus line driver is capable of producing a low-voltage-swing signal having a voltage swing of at least 500 mV on one of said terminated transmission lines.

Petitioners
Ex. 1002, p. 391

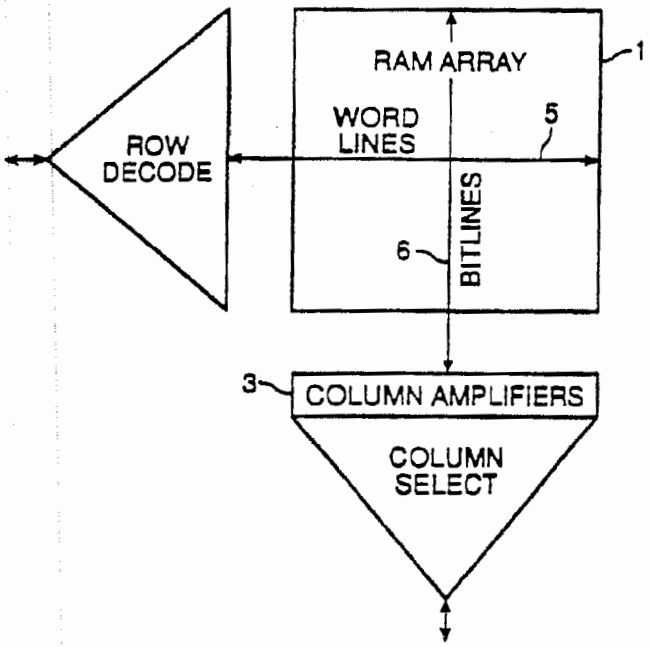


FIG. 1

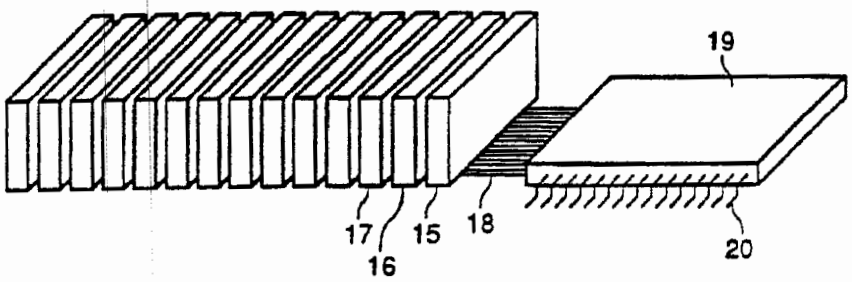


FIG. 3

Petitioners
Ex. 1002, p. 392

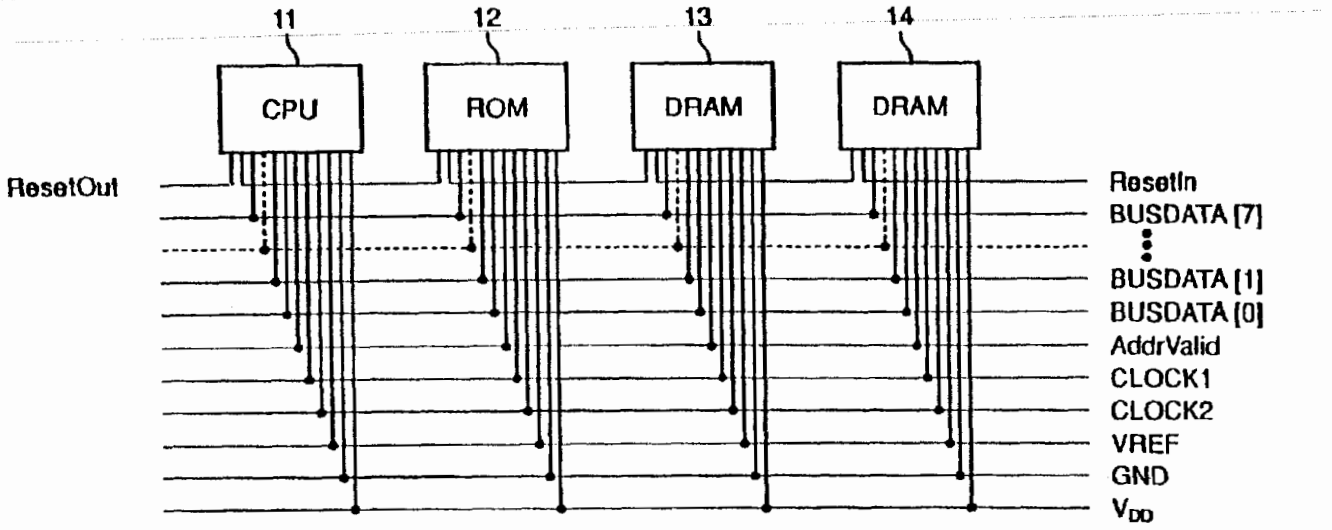


FIG. 2

Petitioners
Ex. 1002, p. 393

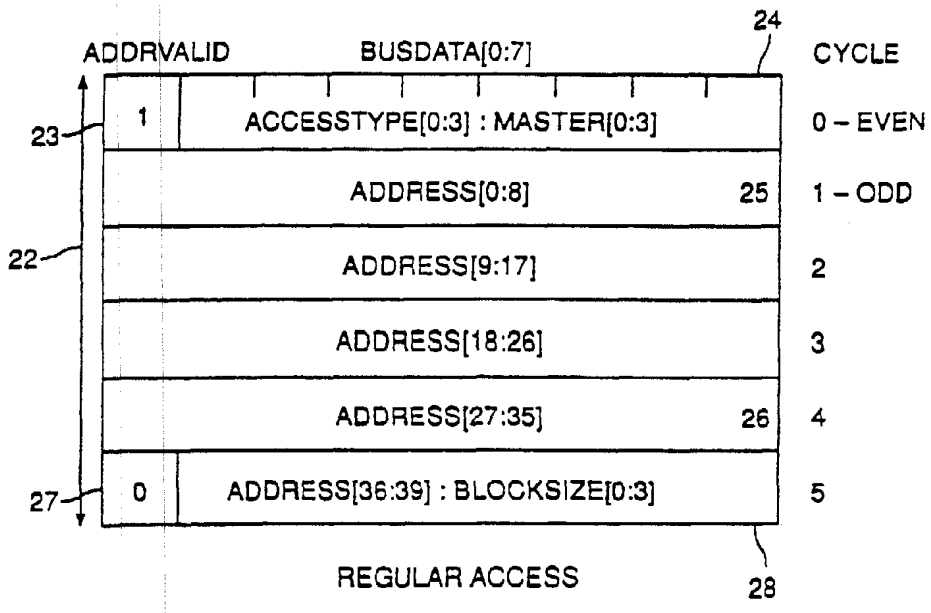


FIG. 4

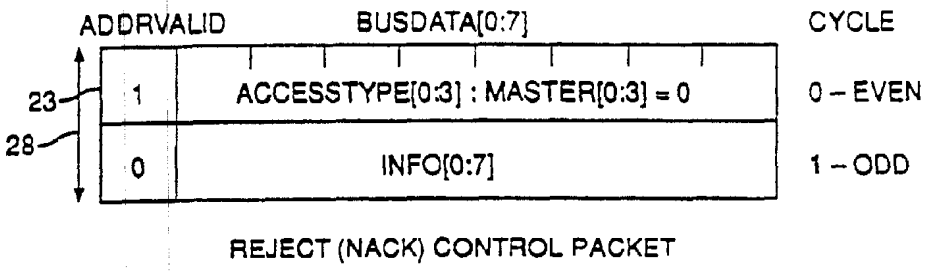


FIG. 5

Petitioners
Ex. 1002, p. 394

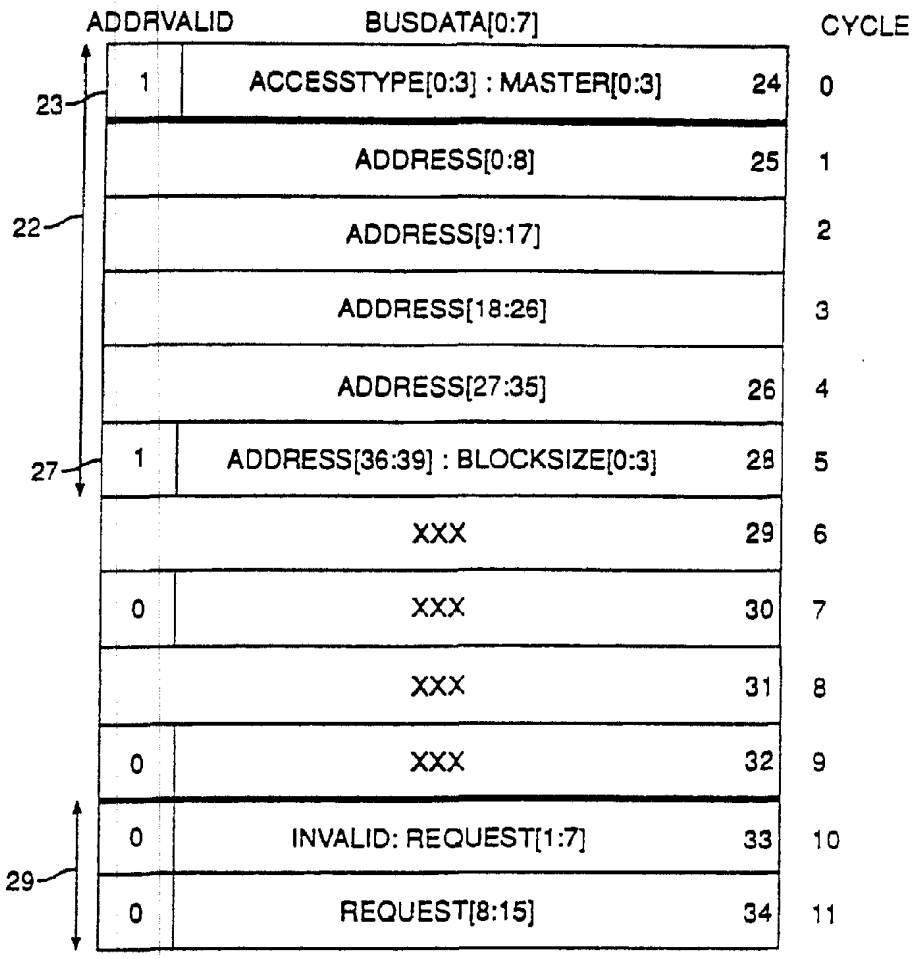


FIG. 6

Petitioners
Ex. 1002, p. 395



FIG. 7a

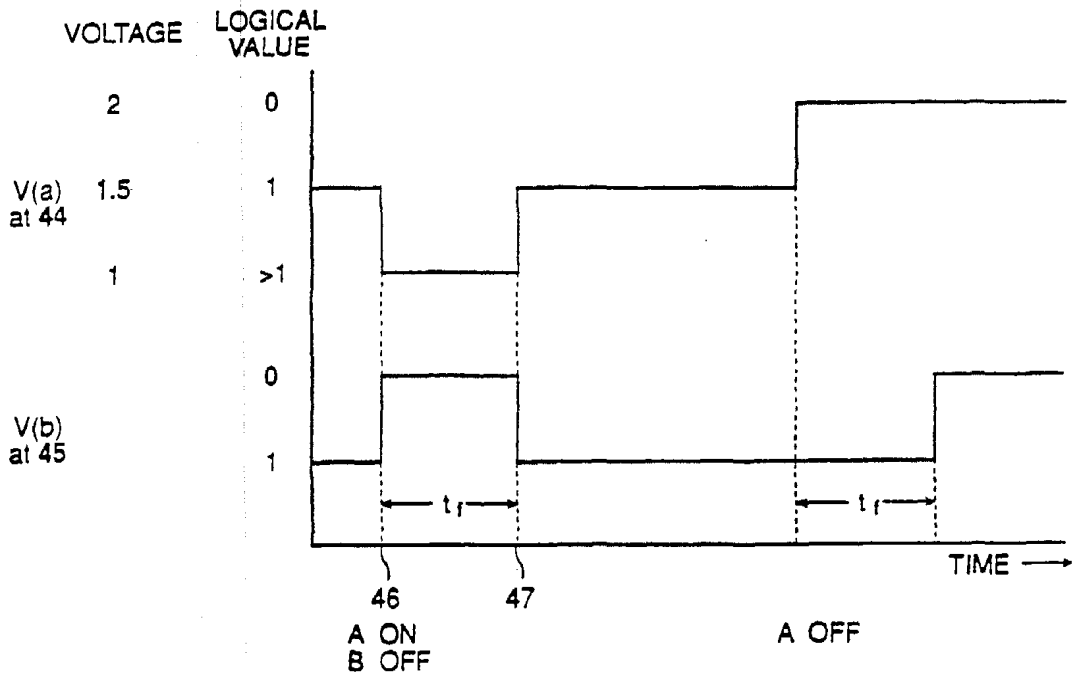


FIG. 7b

Petitioners
Ex. 1002, p. 396

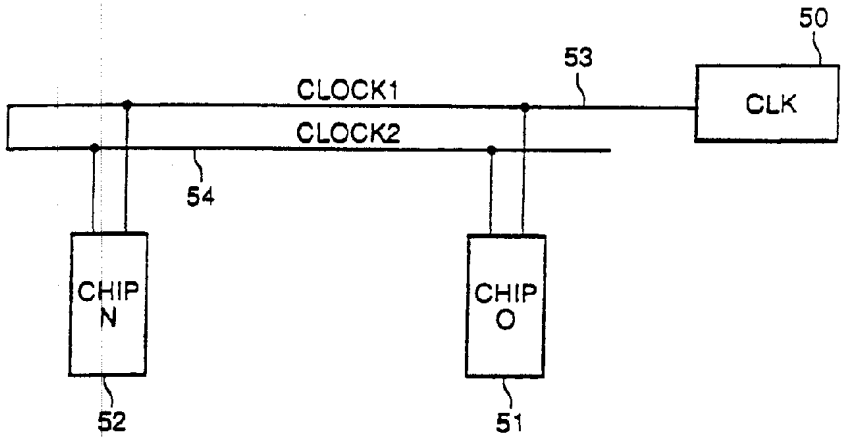


FIG. 8a

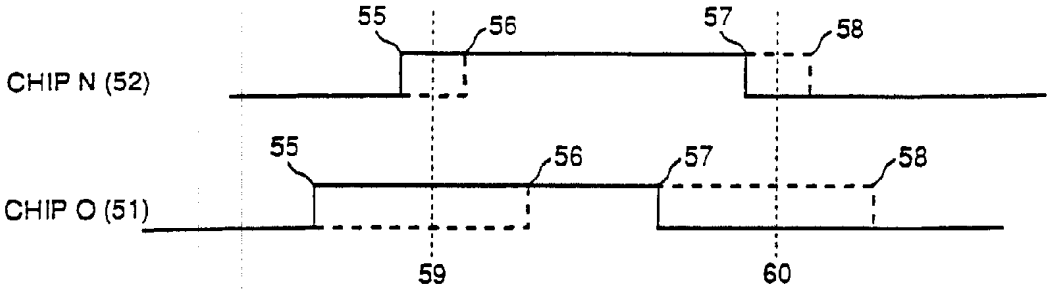


FIG. 8b

Petitioners
Ex. 1002, p. 397

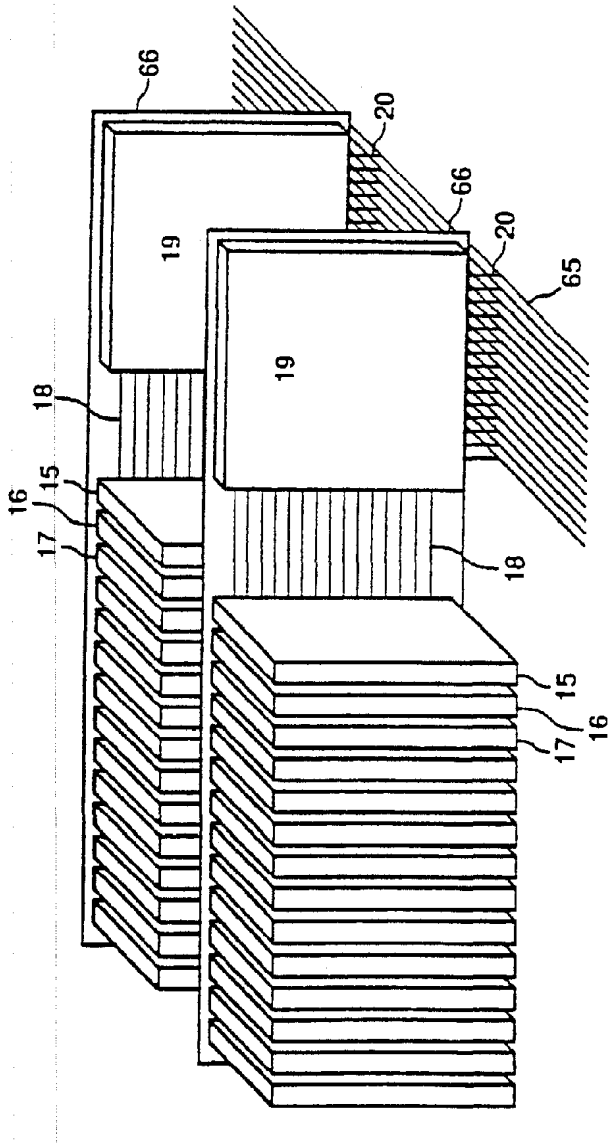


FIG. 9

Petitioners
Ex. 1002, p. 398

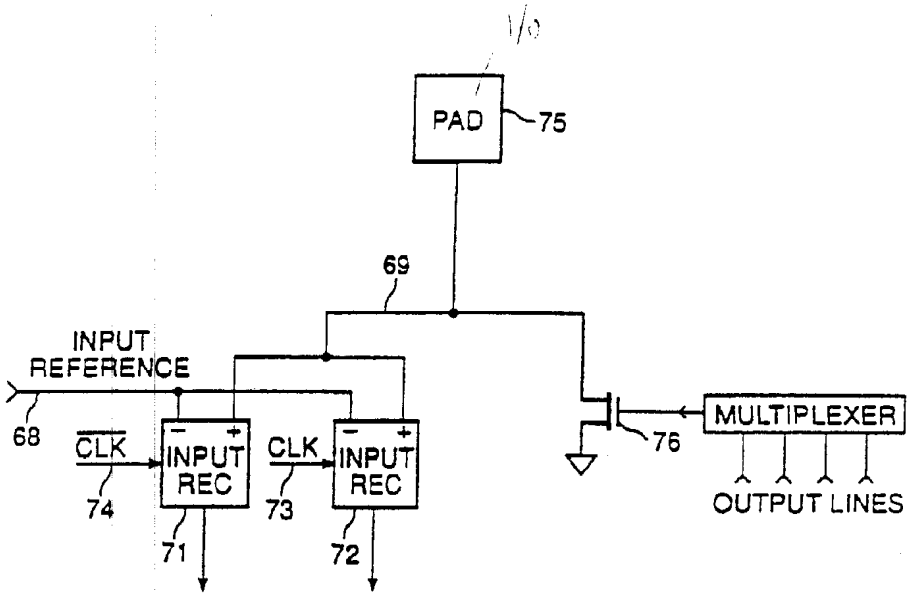


FIG. 10

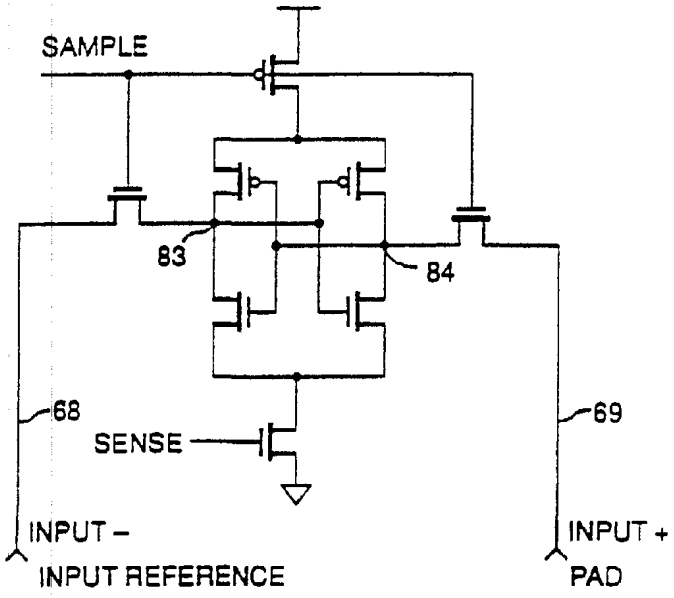


FIG. 11

Petitioners
Ex. 1002, p. 399

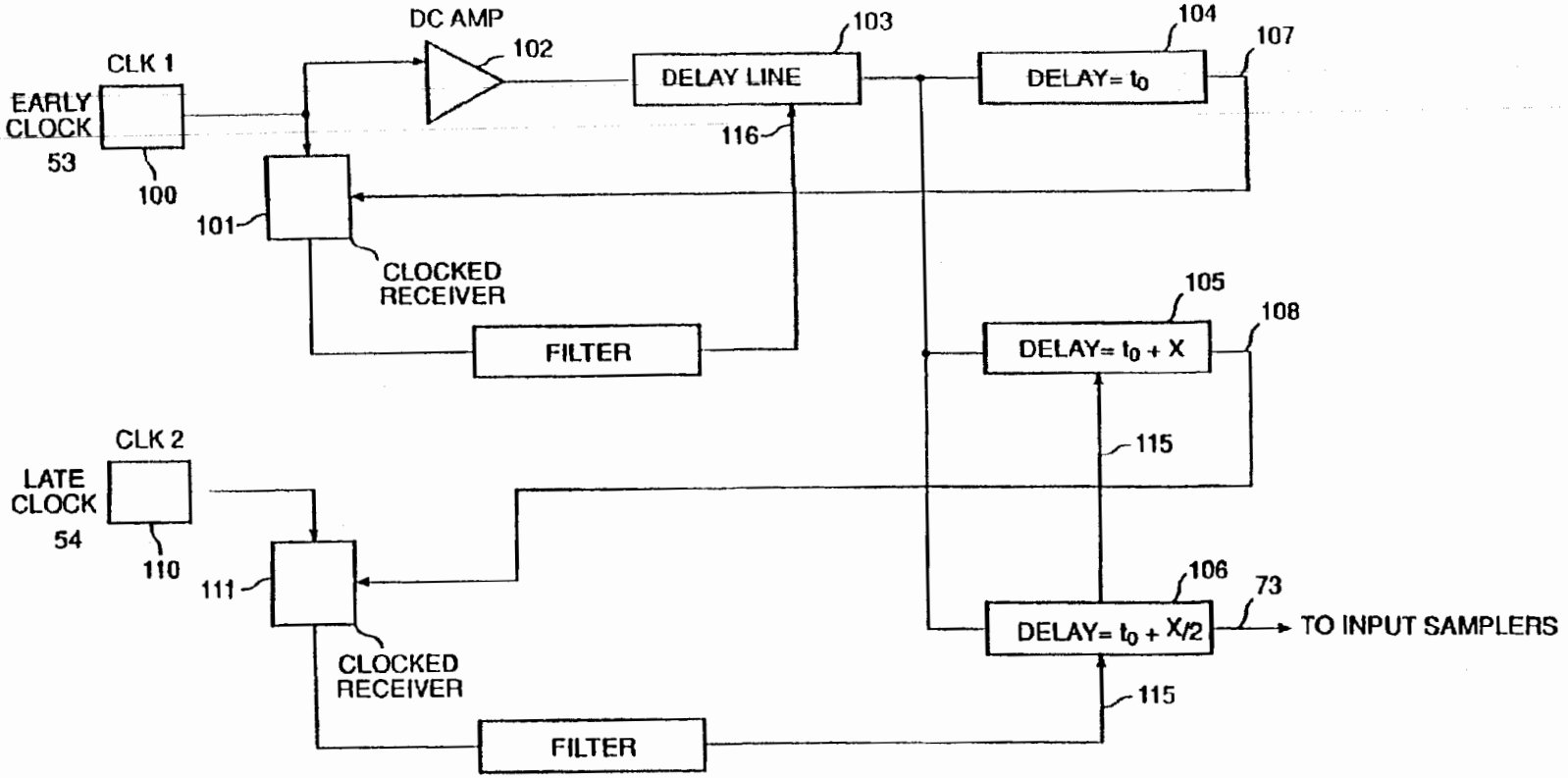

DC AMP
102
103
104
107
CLK 1
EARLY
CLOCK
53
100
DELAY LINE
DELAY= t_0
116
101
CLOCKED
RECEIVER
FILTER
105
108
DELAY= t_0 + X
115
CLK 2
LATE
CLOCK
54
110
111
106
73
DELAY= t_0 + X/2
TO INPUT SAMPLERS
CLOCKED
RECEIVER
115
FILTER

FIG. 12

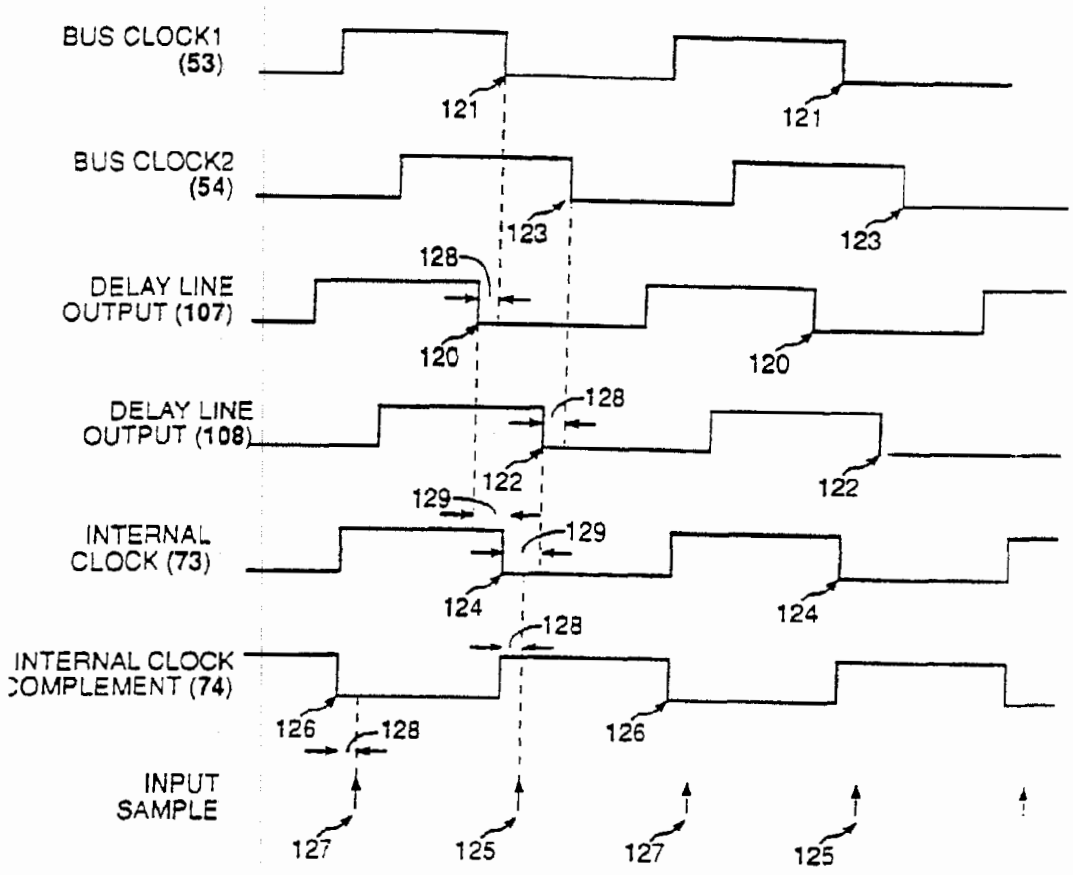


FIG. 13

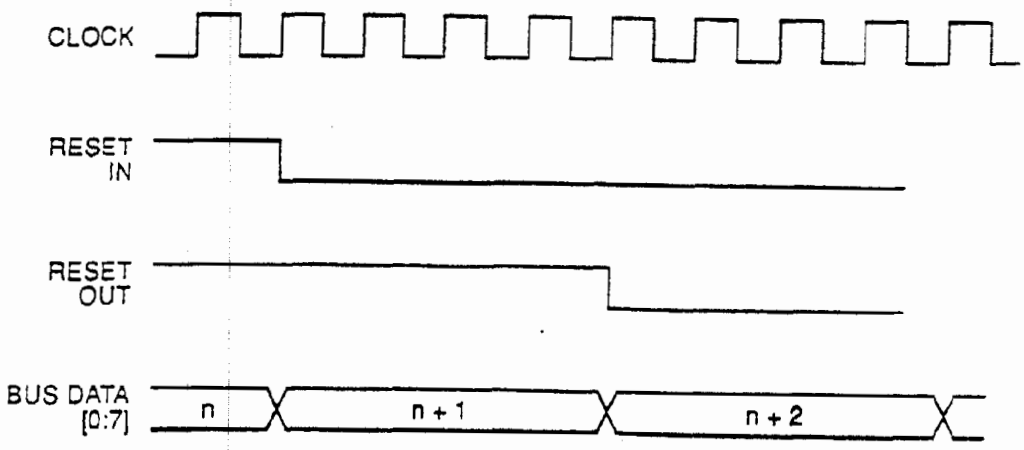


FIG. 14

Petitioners
Ex. 1002, p. 401

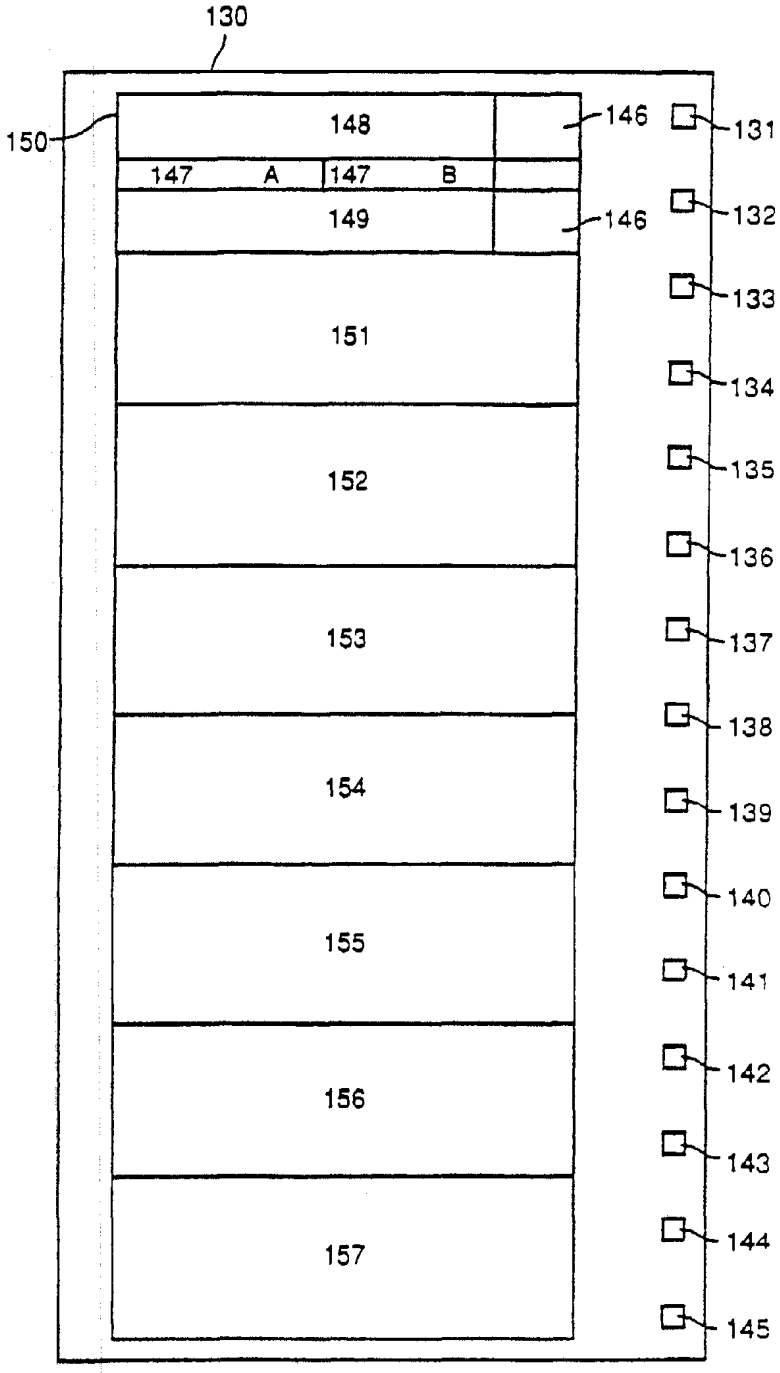


FIG. 15

Petitioners
Ex. 1002, p. 402

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 3821715 A, Hoff **[0005]**
- US 4315308 A, Jackson **[0006]**
- US 4449207 A, Kung **[0007]**
- US 4764846 A **[0008]**
- US 4706166 A, Go **[0008]**
- US 3969706 A, Proebsting **[0009]**
- US 4595923 A **[0010]**
- US 4481625 A, Roberts **[0010]**
- US 4837682 A, Culler **[0011]**
- US 4818985 A, Ikeda **[0011]**
- US 4779089 A, Theus **[0011]**
- US 4745548 A, Blahut **[0011]**
- US 4247817 A, Heller **[0012]**
- US 4646279 A, Voss **[0013]**

Non-patent literature cited in the description

- **ANANT AGARWAL.** An Analytic Cache Model. *ACM Transactions on Computer Systems,* May 1989, vol. 7 (2), 184-215 **[0055]**
- **ANANT AGARWAL et al.** Scaleable Directory Schemes for Cache Consistency. *15th International Symposium on Computer Architecture,* June 1988, 280-289 **[0083]**
- **HANS SCHUMACHER et al.** CMOS Subnanosecond True-ECL Output Buffer. *J. Solid State Circuits,* February 1990, vol. 25 (1), 150-154 **[0100]**

Petitioners
Ex. 1002, p. 403

Electronic Acknowledgement Receipt

EFS ID:	15607184
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	25-APR-2013
Filing Date:	03-NOV-2011
Time Stamp:	00:38:03
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	2013-04-25_IDS_COVER_NETL071A.pdf	55207 5a6529a4112b1e6aa82c4c94d0983927e9631b40	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Form (SB08)	2013-04-25_IDS_NETL071A.pdf	88030 6a2550e82a0107f134bbf69818ed09f37e52ddc3	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	Non Patent Literature	2013-01-30_PCT_Search_Report_Written_Opinion-NETL071WO.pdf	7352527 42ff1202e8251dc1bac40520f4e426a89bdc79f6	no	127
Warnings:					
Information:					
4	Foreign Reference	EP1816570A2.pdf	2964754 1dbe58fa7b9c63cb9028ab1e5723c8e4920f9424	no	43
Warnings:					
Information:					
		Total Files Size (in bytes):	10460518		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE STATEMENT

Applicant	:	Hyun Lee
App. No.	:	13/288,850
Filed	:	November 3, 2011
For	:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
Examiner	:	Amir Zarabian
Art Unit	:	2827
Conf No.	:	2466

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Enclosed for filing in the above-identified application is a Form PTO/SB/08a/b Equivalent listing 3 references. Copies of any foreign documents and/or their translation and non-patent literature are submitted herewith.

The applicants would appreciate the Examiner initialing and returning the Form PTO/SB/08a/b Equivalent indicating that the information has been considered and made of record herein.

The information contained in this Information Disclosure Statement is not to be construed as a representation that: (i) a complete search has been made; (ii) additional information material to this application does not exist; (iii) the information, protocols, results and the like reported by third parties are accurate or enabling; or (iv) the above information constitutes prior art to the subject invention.

This Information Disclosure Statement is being filed before the receipt of a first Office Action on the merits, and presumably no fee is required. If a first Office Action on the

merits was mailed before the mailing date of this Statement, the Commissioner is authorized to charge the fee set forth in 37 C.F.R. § 1.17(p) to Deposit Account No. 50-5963.

Respectfully submitted,

Dated: April 24, 2013

/Jamie J. Zheng/ 51167

Jamie J. Zheng (Reg. No.)
Customer No. 79141
(650) 273-6008

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

79141 7590 06/06/2013

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER	
HOANG, HUAN	

ART UNIT	PAPER NUMBER
2827	

NOTIFICATION DATE	DELIVERY MODE
06/06/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jz@jzpatent.com

Office Action Summary	Application No. 13/288,850	Applicant(s) LEE, HYUN	
	Examiner HUAN HOANG	Art Unit 2827	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 1 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *1-34* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☐ Claim(s) _____ is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☒ Claim(s) *1-34* are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some * c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Interim copies:

a)☐ All b)☐ Some c)☐ None of the: Interim copies of the priority documents have been received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

DETAILED ACTION

Election/Restrictions

1. Restriction to one of the following inventions is required under 35 U.S.C. 121:

   I. Claims 1-28, drawn to a memory package and a method for optimizing load in a memory package, classified in class 365, subclass 63.

   II. Claims 29-34, drawn to a memory module comprising a register device and a plurality of DRAM packages each comprising a plurality of DDR DRAM dies, classified in class 365, subclass 365/233.13.

2. The inventions are distinct, each from the other because of the following reasons:

Inventions II and I are related as combination and subcombination. Inventions in this relationship are distinct if it can be shown that (1) the combination as claimed does not require the particulars of the subcombination as claimed for patentability, and (2) that the subcombination has utility by itself or in other combinations (MPEP § 806.05(c)). In the instant case, the combination as claimed does not require the particulars of the subcombination as claimed because the memory module does not require die interconnects having specific functions as recited in claims 1-28. The subcombination has separate utility such as using die interconnects to transmit data to the plurality of array dies.

The examiner has required restriction between combination and subcombination inventions. Where applicant elects a subcombination, and claims thereto are subsequently found allowable, any claim(s) depending from or otherwise requiring all

the limitations of the allowable subcombination will be examined for patentability in accordance with 37 CFR 1.104. See MPEP § 821.04(a). Applicant is advised that if any claim presented in a continuation or divisional application is anticipated by, or includes all the limitations of, a claim that is allowable in the present application, such claim may be subject to provisional statutory and/or nonstatutory double patenting rejections over the claims of the instant application.

Restriction for examination purposes as indicated is proper because all these inventions listed in this action are independent or distinct for the reasons given above and there would be a serious search and/or examination burden if restriction were not required because at least the following reason(s) apply:

Different search;

Different classification; and

Recognized divergent subject matter.

Applicant is advised that the reply to this requirement to be complete must include (i) an election of a invention to be examined even though the requirement may be traversed (37 CFR 1.143) **and (ii) identification of the claims encompassing the elected invention**.

The election of an invention may be made with or without traverse. To reserve a right to petition, the election must be made with traverse. If the reply does not distinctly and specifically point out supposed errors in the restriction requirement, the election shall be treated as an election without traverse. Traversal must be presented at the time of election in order to be considered timely. Failure to timely traverse the requirement

will result in the loss of right to petition under 37 CFR 1.144. If claims are added after the election, applicant must indicate which of these claims are readable upon the elected invention.

Should applicant traverse on the ground that the inventions are not patentably distinct, applicant should submit evidence or identify such evidence now of record showing the inventions to be obvious variants or clearly admit on the record that this is the case. In either instance, if the examiner finds one of the inventions unpatentable over the prior art, the evidence or admission may be used in a rejection under 35 U.S.C. 103(a) of the other invention.

3. Any inquiry concerning this communication or earlier communications from the examiner should be directed to HUAN HOANG whose telephone number is (571)272-1779. The examiner can normally be reached on Tues-Fri 8:30AM-7:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Amir Zarabian can be reached on (571) 272-1852. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Huan Hoang/
Primary Examiner, Art Unit 2827

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	Art Unit
	HUAN HOANG	2827

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	06/02/2013								
	1	÷								
	2	÷								
	3	÷								
	4	÷								
	5	÷								
	6	÷								
	7	÷								
	8	÷								
	9	÷								
	10	÷								
	11	÷								
	12	÷								
	13	÷								
	14	÷								
	15	÷								
	16	÷								
	17	÷								
	18	÷								
	19	÷								
	20	÷								
	21	÷								
	22	÷								
	23	÷								
	24	÷								
	25	÷								
	26	÷								
	27	÷								
	28	÷								
	29	÷								
	30	÷								
	31	÷								
	32	÷								
	33	÷								
	34	÷								

Petitioners
Ex. 1002, p. 414

UNITED STATES PATENT AND TRADEMARK OFFICE
COMMISSIONER FOR PATENTS
P.O.BOX 1450
ALEXANDRIA VA 22313-1451

PRESORTED
FIRST-CLASS MAIL
U.S. POSTAGE PAID
POSTEDIGITAL
NNNNN

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

Courtesy Reminder for
Application Serial No: 13/288,850

Attorney Docket No: NETL.071A
Customer Number: 79141
Date of Electronic Notification: 06/06/2013

This is a courtesy reminder that new correspondence is available for this application. If you have not done so already, please review the correspondence. The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

An email notification regarding the correspondence was sent to the following email address(es) associated with your customer number:
jz@jzpatent.com

To view your correspondence online or update your email addresses, please visit us anytime at **https://sportal.uspto.gov/secure/myportal/privatepair.** If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov or call 1-866-217-9197.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	:	Hyun Lee
App. No.	:	13/288,850
Filed	:	November 3, 2011
For	:	ARCHITECTURE FOR MEMORY MODULE WITH PACKAGES OF THREE-DIMENSIONAL STACKED (3DS) MEMORY CHIPS
Examiner	:	ZARABIAN, AMIR
Art Unit	:	2827
Conf. No.	:	2466
Docket No.	:	NETL.071A

CERTIFICATE OF EFS WEB TRANSMISSION

I hereby certify that this correspondence and any other attachment noted on the Automated Acknowledgement Receipt, is being transmitted from within the Pacific Time zone to the Commissioner for Patents via the EFS Web Server on:

August 6, 2013

(Date)

/Jamie J. Zheng/

Jamie J. Zheng, Reg. No. 51167

AMENDMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

The enclosed Amendment is in response to the Restriction Requirement dated June 6, 2013 for the above-identified patent application.

The Commissioner is hereby authorized to charge any required fee(s) to Deposit Account No. 50-5963.

AMENDMENTS to the CLAIMS begin on Page 2 of this paper.

REMARKS begin on Page 11 of this paper.

IN THE CLAIMS:

Rewrite the pending claims as follows:

1. (**Currently Amended**) A memory package, comprising:

a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices;

a plurality of stacked array dies including a first group of array dies and a second group of at least one array die, each array die having data ports;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with ~~at least one data port of a first array die of the plurality~~the first group of array dies ~~and at least one data port of a second array die of the plurality of array dies~~ and not in electrical communication with the ~~data ports of at least a third~~second group of at least one array die ~~of the plurality of array dies~~, the second die interconnect in electrical communication with ~~at least one data port of the third~~the second group of at least one array die and not in electrical communication with ~~the data ports of~~ the first group of array dies ~~and the data ports of the second array die~~; and

a control die comprising at least a first data conduit ~~configured to transmit a data signal to~~between the first die interconnect and a first terminal of the plurality of input/output terminals~~and to not transmit the data signal to the second die interconnect~~, and at least a second data conduit ~~configured to transmit the data signal to~~between the second die interconnect and the first terminal ~~and to not transmit the data signal to the first die interconnect~~, the first terminal being a data terminal, the control die further comprising a control circuit to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals.

2. (**Currently Amended**) The memory package of claim 1, wherein the control signals include data path control signals for controlling the first and second data conduits ~~the first data conduit and the second data conduit are configured to receive the data signal from a common source as one another~~.

3. (**Currently Amended**) The memory package of claim 1, wherein the control circuit is configured to generate data path control signals for controlling the first and second data conduits in response to the received control signals ~~a first load on the first data conduit comprises a load of the first die interconnect, a load of the first array die, and a load of the~~

~~second array die, and wherein a second load on the second data conduit comprises a load of the second die interconnect and a load of the third array die~~.

4. (**Currently Amended**) The memory package of claim 3, wherein the control signals include command/address signals and wherein the control die is configured to provide the command/address signals to the plurality of stacked array dies ~~a difference between the first load and the second load is less than a load that would be on the first data conduit upon the first die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies~~.

5. (**Currently Amended**) The memory package of claim 3, wherein:

the first data conduit has a first load that is less than a load that would be on the first data conduit upon the first die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies; and

the second data conduit has a second load is less than a load that would be on the second data conduit upon the second die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies.

6. (**Currently Amended**) The memory package of claim 1, wherein the first die interconnect comprises a first ~~conducting rod~~through-silicon via and wherein the second die interconnect comprises a second ~~conducting rod~~through-silicon via.

7. (**Currently Amended**) The memory package of claim 1, wherein the control die further comprises chip-select conduits, ~~one of the first die interconnect and the second die interconnect is in electrical communication with at least one data port of a fourth array die of the plurality of array dies and the other one of the first die interconnect and the second die interconnect is not in electrical communication with the data ports of the fourth array die~~ the memory package further comprising:

third die interconnects coupled between respective chip-select conduits and respective ones of the plurality of stacked array dies.

8. (**Currently Amended**) The memory package of claim 1, wherein a first number of array dies in the first group of array dies and a second number of at least one array die in the second group of at least one array die are selected in consideration of a load of the first die

Petitioners
Ex. 1002, p. 418

interconnect and a load of the second die interconnect so as to reduce a different between a first load on the first data conduit and a second load on the second data conduit, the first load including a load of the first die interconnect, and a load of the first group of array dies, and the second load including a load of the second die interconnect and a load of the second group of at least one array die ~~is configured to transmit signals to the first array die and the second array die, and to receive signals from the first array die and the second array die, and the second die interconnect is configured to transmit signals to the third array die and to receive signals from the third array die~~.

9. (**Currently Amended**) The memory package of claim 1, further comprising at least a third die interconnect and a fourth die interconnect, the third die interconnect in electrical communication with ~~at least one data port of the first array die and at least one data port of the second array die~~ the first group of array dies and not in electrical communication with the ~~data ports~~ second group of at least ~~the third~~ one array die, the fourth die interconnect in electrical communication with the second group of at least one ~~data port of the third~~ array die and not in electrical communication with the ~~data ports of the first array die and the data ports of the second array die~~first group of array dies.

10. (**Currently Amended**) The memory package of claim 9, wherein the first die interconnect is configured to transmit signals to the first ~~array die and the second array die~~ group of array dies, the second die interconnect is configured to transmit signals to the ~~third~~ second group of at least one array die, the third die interconnect is configured to receive signals from the first ~~array die and the second array die~~ group of array dies, and the fourth die interconnect is configured to receive signals from the ~~third~~ second group of at least one array die.

11. (**Currently Amended**) A memory package, comprising:

a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices;

a plurality of array dies arranged in a stack, including a first group of array dies and a second group of at least one array die ~~having ports~~;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with ~~at least one port of a first array die of the plurality of~~ the first group of array dies ~~and at least one port of a second array die of the plurality of array dies~~ and not in electrical communication with the ~~ports of at least a third~~ second group of at

Petitioners
Ex. 1002, p. 419

least one array die ~~of the plurality of array dies~~, the second die interconnect in electrical communication with ~~at least one port of the third array die~~ the second group of at least one array die and not in electrical communication with the ~~ports of the first array die and the ports of the second array die~~ first group of array dies; and

a control die comprising at least a first data conduit between ~~configured to transmit a signal to~~ the first die interconnect and a first terminal of the plurality of input/output terminals~~to not transmit the signal to the second die interconnect, and~~ , at least a second data conduit ~~configured to transmit the signal to~~ between the second die interconnect and the first terminal ~~to not transmit the signal to the first die interconnect~~, and chip select conduits for providing chip select signals to respective array dies;

~~wherein a first load on the first conduit comprises a load of the first die interconnect, a load of the first array die, and a load of the second array die, and~~

wherein the control die further comprises a control circuit to control respective states of the first data conduit and the second data conduit to drive a data signal to an array die selected by at least one of the chip-select signals ~~a second load on the second conduit comprises a load of the second die interconnect and a load of the third array die~~.

12. (**Currently Amended**) The memory package of claim 11, wherein the chip select conduits pass through the control die ~~a difference between the first load and the second load is less than a load that would be on the first conduit upon the first die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies~~.

13. (**Currently Amended**) The memory package of claim 11, wherein the chip select conduits include drivers to drive the chip select signals to the respective array dies.~~:~~

~~the first load is less than a load that would be on the first conduit upon the first die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies; and~~

~~the second load is less than a load that would be on the second conduit upon the second die interconnect being placed in electrical communication with at least one port of each of the array dies of the plurality of array dies~~.

14. (**Currently Amended**) The memory package of claim 11, wherein the first die interconnect comprises ~~a first conducting rod~~ one or more through silicon vias and wherein

Petitioners
Ex. 1002, p. 420

the second die interconnect comprises one or more through silicon vias ~~a second conducting rod~~.

15. (**Currently Amended**) The memory package of claim 11, wherein: the first data conduit comprises at least a first driver, and the second data conduit comprises at least a second driver.

16. (**Currently Amended**) The memory package of claim 11~~15~~, wherein the ~~first driver and the second driver are configured to receive the signal from a common source as one another~~control die is configured to generate the chip select signals from control signals received via second terminals of the plurality of terminals.

17. (**Currently Amended**) The memory package of claim 11~~16~~, wherein the control circuit controls the respective states of the first data conduit and the second data conduit in response to control signals received via second terminals of the plurality of terminals ~~signal comprises a data signal~~.

18. (**Currently Amended**) The memory package of claim 16, wherein the control signals comprise ~~signal comprises~~ at least one of the following: a command signal, an address signal, and a data path control signal.

19. (**Currently Amended**) The memory package of claim 11, wherein the control die is configured to generate data path control signals from control signals received via second terminals of the plurality of terminals, and wherein the control circuit controls the respective states of the first data conduit in response to the data path control signals ~~16, further comprising a third driver and a fourth driver, wherein the third driver is configured to drive a first signal received from the first die interconnect, and wherein the fourth driver is configured to drive a second signal received from the second die interconnect~~.

20. (**Currently Amended**) A method for optimizing load in a memory package comprising a plurality of array dies arranged in a stack and including a first group of array dies and a second group of at least one array die, at least a first die interconnect and a second die interconnect, ~~and~~ a control die, and a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices ~~comprising at least a first driver and a second driver, the first driver configured to~~

Petitioners
Ex. 1002, p. 421

~~drive a signal along the first die interconnect, and the second driver configured to drive the signal along the second die interconnect~~, the method comprising:

receiving a data signal at a first terminal of the plurality of input/output terminals;

receiving control signals at second terminals of the plurality of input/output terminals;

providing chip select signals to respective array dies through the control die, the chip select signals being related to at least some of the control signals; and

selecting one of a first driver and a second driver in the control die to drive the data signal via a corresponding one of the first die interconnect and the second die interconnect to an array die selected by at least one of the chip select signals, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies.

~~selecting a first subset of array dies of the plurality of array dies and a second subset of array dies of the plurality of array dies, wherein the first subset of array dies and the second subset of array dies are exclusive of one another and are selected to balance a load on the first driver and on the second driver based at least in part on array die loads of array dies of the plurality of array dies and at least in part on die interconnect segment loads of segments of at least the first die interconnect and the second die interconnect;~~

~~forming electrical connections between the first die interconnect and the first subset of array dies; and forming electrical connections between the second die interconnect and the second subset of array dies.~~

21. (**Currently Amended**) The method of claim 20, further comprising:

selecting a first driver size for the first driver based, at least in part, on ~~the~~ a load on the first driver; and

selecting a second driver size for the second driver based, at least in part, on ~~the~~ a load on the second driver.

22. (**Currently Amended**) The method of claim 21, wherein the first driver size and the second driver size are both less than a driver size sufficient to drive ~~the~~ a signal along a die interconnect in electrical communication with each of the plurality of array dies ~~with a less than predetermined~~without significant signal degradation.

Petitioners
Ex. 1002, p. 422

23. (**Currently Amended**) The method of claim 20, further comprising generating the chip select signals from at least some of the control signals ~~wherein each array die of the plurality of array dies comprises a substantially equal array die load~~.

24. (**Currently Amended**) The method of claim 20, wherein the control signals include the chip-select signals ~~each segment of the first die interconnect and the second die interconnect comprises a substantially equal die interconnect segment load~~.

25. (**Currently Amended**) The method of claim 20, wherein the ~~segments of the die interconnect comprise one or more portions of a die interconnect between two adjacent array dies~~first and second die interconnects each comprises at least one through silicon vias.

26. (**Currently Amended**) The method of claim 20, wherein the chip select signals pass through through-silicon-vias in the control die ~~the segments of the die interconnect comprise one or more portions of a die interconnect between the control die and an adjacent array die~~.

27. (**Currently Amended**) The method of claim 20, further comprising generating data path control signals from at least some of the control signals, the data path control signals being used to select the one of the first driver and the second driver in the control die to drive the data signal~~wherein the first subset of array dies comprises at least two array dies~~.

28. (**Currently Amended**) The method of claim 20, wherein the one of the first driver and the second driver is selected using at least some of the control signals ~~balancing the load on the first driver and the load on the second driver comprises minimizing a load difference between the first driver and the second driver~~.

29. (**Currently Amended**) A memory module operable via a memory control hub, comprising:

a register device configured to receive command/address signals from the ~~a~~ memory control hub and to generate ~~data path~~ control signals; and

a plurality of DRAM packages, each DRAM package comprising:

a plurality of data terminals via which the DRAM package communicate data with the memory control hub, and a plurality of control terminals to receive the control signals;

a plurality of array dies arranged in a stack, including a first group of array dies and a second group of at least one array die;

Petitioners
Ex. 1002, p. 423

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die comprising at least a first data conduit between the first die interconnect and a first data terminal of the plurality of data terminals, at least a second data conduit between the second die interconnect and the first data terminal, and chip select conduits for providing chip select signals to respective array dies, the chip select signals being related to at least some of the control signals;

wherein the control die further comprises a control circuit to control respective states of the first data conduit and the second data conduit to drive a data signal to an array die selected by at least one of the chip-select signals.

~~a control die comprising a plurality of command/address buffers and a data path control circuit configured to control command/address time slots and data bus time slots, the control die configured to receive data signals from the memory control hub, the data path control signals from the register device, and command/address signals from the register device; and~~

~~a plurality of DDR DRAM dies operatively coupled to the control die to receive the data signals from the control die, wherein the memory module is selectively configurable into at least~~

~~two operational modes comprising:~~

~~a first operational mode in which the register device generates the data path control signals, and the control die uses the data path control signals to operate the data path control circuit; and~~

~~a second operational mode in which the control die operates the data path control circuit to provide the command/address signals to the plurality of DDR DRAM dies without decoding the command/address signals~~.

30. **(Currently Amended)** The memory module of claim 29, wherein the register device is further configured to perform rank multiplication by generating the chip select signals, and wherein the control signals include the chip select signals.

Petitioners
Ex. 1002, p. 424

31. (**Currently Amended**) The memory module of claim 29, wherein the control signals include data path control signals generated by the register device, the data path control signals being used to control the respective states of the first data conduit and the second data conduit ~~memory module does not perform rank multiplication while operating in the second operational mode~~.

32. (**Currently Amended**) The memory module of claim 29, wherein the control die is further configured to perform rank multiplication by generating the chip select signals from at least some of the control signals.

33. (**Currently Amended**) The memory module of claim 29, wherein the control signals include command/address signals, and the control die is configured to hold the command/address signals to control timing of the command/address signals ~~at least one of the plurality of DRAM packages comprises the memory package of claim 1~~.

34. (**Currently Amended**) The memory module of claim 29, wherein the control die is configured to generate data path control signals from at least some of the control signals, the data path control signals being used to control the respective states of the first data conduit and the second data conduit ~~at least one of the plurality of DRAM packages comprises the memory package of claim 11~~.

Petitioners
Ex. 1002, p. 425

REMARKS

This amendment responds to the Restriction Requirement mailed June 6, 2013. Applicant would like to elect without traverse the invention in Claims 1-28 as amended for examination.

I. McKesson Statement

The pending application is related to commonly owned U.S. Patent Nos. 7,289,386, 7,286,436, 7,442,050, 7,375,970, 7,254,036, 7,532,537, 7,636,274, 7,630,202, 7,619,893, 7,619,912, 7,811,097, 7,839,645, 7,864,627, 7,881,150, 8,001,434, 8.033.836, 8,154,901, 8,417,870, 8,489,837, and 8516185, and U.S. Patent Application 12/422,912. The Examiner is encouraged to review the art made of record, Office Action(s) and the Notice of Allowance in the above-mentioned related application, all of which are available on PAIR.

II. Amendments to the Claims

Claims 1-28 have been amended to better define the inventions which the Applicant intends to claim in the present application. Claims 29-34 have been amended to include the particulars of claims 11-19 so that they relate to the same subject matter as claims 11-19. Thus, claims 29-34 should no longer be subject to restriction requirement. No new matter is added.

The Examiner is invited to call the undersigned attorney at (650) 273-6008, if a telephone call could help resolve any remaining items.

Respectfully submitted,

Date: August 6, 2013

/Jamie J. Zheng/ 51167

Jamie J. Zheng (Reg. No.)

Customer No. 79141

Petitioners
Ex. 1002, p. 426

ASSERTION OF ENTITLEMENT TO SMALL ENTITY STATUS UNDER 37 CFR 1.27

Applicant	:	Hyun Lee
App. No.	:	13/288,850
Filed	:	November 3, 2011
For	:	ARCHITECTURE FOR MEMORY MODULE WITH PACKAGES OF THREE-DIMENSIONAL STACKED (3DS) MEMORY CHIPS
Examiner	:	ZARABIAN, AMIR
Art Unit	:	2827
Conf. No.	:	2466
Docket No.	:	NETL.071A

CERTIFICATE OF EFS WEB TRANSMISSION

I hereby certify that this correspondence and any other attachment noted on the Automated Acknowledgement Receipt, is being transmitted from within the Pacific Time zone to the Commissioner for Patents via the EFS Web Server on:

August 6, 2013

(Date)

/Jamie J. Zheng/

Jamie J. Zheng, Reg. No. 51167

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

The Assignee of all rights and title in the above-identified application, Netlist, Inc., is a small entity concern as defined in 37 C.F.R. 1.27(a). Assertion of entitlement to small entity status under 37 C.F.R. 1.27(c) is hereby made in this application.

Applicants acknowledge the duty to file notification of a loss of entitlement to small entity status in the application or patent prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate.

Respectfully submitted,

Date: August 6, 2013

/Jamie J. Zheng/ 51167
Jamie J. Zheng (Reg. No.)
Customer No. 79141

Electronic Patent Application Fee Transmittal

Application Number:	13288850
Filing Date:	03-Nov-2011
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Filer:	Jamie Jie Zheng
Attorney Docket Number:	NETL.071A

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	100	100

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**100**

Electronic Acknowledgement Receipt

EFS ID:	16517425
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	06-AUG-2013
Filing Date:	03-NOV-2011
Time Stamp:	17:55:51
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$100
RAM confirmation Number	5236
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1	Response to Election / Restriction Filed	2013-08-06_Amendment_Election_NETL071A.pdf	129620 ace2e7b9ea4fbb9918f44e827fb184244daf a14d	no	11
Warnings:					
Information:					
2	Assertion of entitlement to small entity status	2013-08-06_Assertion_Small_Entity_Status_NETL071A.pdf	77141 ae72308306482e40d976a257c8f1d08083f9 8329	no	1
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	30218 92b13efc56dc32eddbb14d1378309b8604e f0b6a	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	236979		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 13/288,850	Filing Date 11/03/2011	☐ To be Mailed

ENTITY: ☒ LARGE ☐ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☒ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	**380**
☒ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	**620**
☒ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	**250**
TOTAL CLAIMS (37 CFR 1.16(i))	34 minus 20 =	* 14	x $60 =	**840**
INDEPENDENT CLAIMS (37 CFR 1.16(h))	4 minus 3 =	* 1	x $250 =	**250**
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	**2340**

APPLICATION AS AMENDED – PART II

AMENDMENT	08/06/2013	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 34	Minus	** 34	= 0	x $80 =	0
	Independent (37 CFR 1.16(h))	* 4	Minus	***4	= 0	x $420 =	0
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	**0**

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	x $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	x $ =	
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

LIE
/AJAY R. DAVID/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

79141 7590 10/11/2013

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER	
HOANG, HUAN	

ART UNIT	PAPER NUMBER
2827	

NOTIFICATION DATE	DELIVERY MODE
10/11/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jz@jzpatent.com

Office Action Summary	Application No. 13/288,850	Applicant(s) LEE, HYUN	
	Examiner HUAN HOANG	Art Unit 2827	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *1-34* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1-7,9-21 and 23-34* is/are rejected.

8)☒ Claim(s) *8 and 22* is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *11/03/11* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some * c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *05/23/12, 04/25/13*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

The present application is being examined under the pre-AIA first to invent provisions.

DETAILED ACTION

Election/Restrictions

1. Applicant's election without traverse of claims 1-28 in the reply filed on 08/06/13 is acknowledged.

Claims 29-34 have been amended to include the particulars of claims 11-19 so that they relate to the same subject matter as claims 11-19. Accordingly, the restriction requirement is withdrawn.

Information Disclosure Statement

2. The document "A Quantitative Analysis of Performance Benefits of 3D Die Stacking on Mobile and Embedded SoC" in the information disclosure statement filed 05/23/2012 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because a date has not been provided. It has been placed in the application file, but the information referred to therein has not been considered as to the merits. Applicant is advised that the date of any re-submission of any item of information contained in this information disclosure statement or the submission of any missing element(s) will be the date of submission for purposes of determining compliance with the requirements based

on the time of filing the statement, including all certification requirements for statements under 37 CFR 1.97(e). See MPEP § 609.05(a).

3. The cite No.36 in the IDS filed on 05/23/12 has not been considered. It should be 2008/0025137 and has been cited in Form PPTO-892.

Claim Objections

4. Claim 5 is objected to because of the following informalities:

The word "dis" in claim 5, line 2 should be "is" and the term "different" in claim 8, line 4 should be "difference". Appropriate correction is required.

Claim Rejections - 35 USC § 112

5. The following is a quotation of 35 U.S.C. 112(b):

> (b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

6. Claim 5 is rejected under 35 U.S.C. 112(b) or 35 U.S.C. 112 (pre-AIA), second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor, or for pre-AIA the applicant regards as the invention.

The recitation of "the first data conduit has a first load that hat is less than a load that would be on the first data conduit upon the first die interconnect being placed in electrical communication with at least one data port of each of the array dies of the

plurality of array dies; and the second data conduit has a second load is less than a load that would be on the second data conduit upon the second die interconnect being placed in electrical communication with at least one data port of each of the array dies of the plurality of array dies." is confusing since the first die interconnect/the second die interconnect is only in electrical communication with the first group of array dies/the second group of at least one die; therefore, the first die interconnect/the second die interconnect cannot be in electrical communication with each of the array dies (at least one array die of the second group is not in electrical communication with the first die interconnect and the array dies of the first group are not in electrical communication with the with the second die interconnect).

Claim Rejections - 35 USC § 102

7. The following is a quotation of the appropriate paragraphs of pre-AIA 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

8. Claims 1-4, 7, 9-13, 15-21, 23, 24, 27-29, 31, 33 and 34 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Rajan et al. (US 2008/0025137).

Regarding claim 1: Rajan (Fig. 2B) shows a memory package, comprising:

a plurality of input/output terminals (address, Control, Clock and Data) via which the memory package communicate data and control/address signals with one or more external devices;

a plurality of stacked array dies (paragraph [0043], lines 1-3) including a first group of array dies (206A and 206C) and a second group (206B and 206D) of at least one array die, each array die having data ports;

at least a first die interconnect (Data connected to 206A and 206C) and a second die interconnect (Data connected to 206B and 206D), the first die interconnect in electrical communication with first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die (BUFFER CHIP) comprising at least a first data conduit between the first die interconnect and a first terminal (Data) of the plurality of input/output terminals, and at least a second data conduit between the second die interconnect and the first terminal, the first terminal being a data terminal, the control die further comprising a control circuit (paragraph [0044], lines 5-7) to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals (Address, Clock and Control).

Regarding claims 2-4: Rajan discloses the memory package of claim 1, wherein the control signals include data path control signals (Control, Address and Clock must be used to address the DRAM circuits and control the data between Data terminal to DRAM circuits) for controlling the first and second data conduits.

Regarding claims 7, 11, 12 and 16-19: Rajan discloses the memory package of claim 1, wherein the control die further comprises chip-select conduits, the memory package further comprising:

Regarding claims 9 and 10, the third die interconnect and the fourth die interconnect are considered the first die interconnect and the second die interconnect, respectively.

third die interconnects coupled between respective chip-select conduits and respective ones of the plurality of stacked array dies (paragraph [0057], lines 7-13).

Regarding claims 13 and 15: It is inherent that there are drivers to drive data from BUFFER CHIP to DRAM circuits.

Regarding claim 20: Rajan (Fig. 2B) discloses a method for optimizing load in a memory package comprising a plurality of array dies arranged in a stack and including a first group of array dies and a second group of at least one array die, at least a first die interconnect and a second die interconnect, a control die, and a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices, the method comprising:

receiving a data signal at a first terminal (Data) of the plurality of input/output terminals;

receiving control signals (Address, Control and Clock) at second terminals of the plurality of input/output terminals;

providing chip select signals (paragraph [0057], lines 7-13) to respective array dies through the control die, the chip select signals being related to at least some of the control signals; and

selecting one of a first driver and a second driver (it is inherent that there are drivers to drive data from BUFFER CHIP to DRAM circuits) in the control die to drive the data signal via a corresponding one of the first die interconnect and the second die interconnect to an array die selected by at least one of the chip select signals, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies.

are exclusive of one another and are selected to balance a load on the first driver and on the second driver based at least in part on array die loads of array dies of the plurality of array dies and at least in part on die interconnect segment loads of segments of at least the first die interconnect and the second die interconnect; forming electrical connections between the first die interconnect and the first subset of array dies; and forming electrical connections between the second die interconnect and the second subset of array dies

Regarding claim 21: It is inherent to select a drive size for a drive based on a load on the driver.

Regarding claims 23, 24, 27 and 28: Rajan discloses the method of claim 20, further comprising generating the chip select signals from at least some of the control

signals (the chip select signals are used to select the DRAM circuits; therefore, the chip select signals are generated from address bits, paragraph [0057], lines 7-13).

Regarding claims 29, 31, 33 and 34: Rjan discloses a memory module operable via a memory control hub, comprising:

a register device (paragraph [0076], lines 6-8) configured to receive command/address signals from the memory control hub and to generate control signals; and

a plurality of DRAM packages, each DRAM package comprising: a plurality of data terminals via which the DRAM package communicate data with the memory control hub, and a plurality of control terminals to receive the control signals; a plurality of array dies arranged in a stack, including a first group of array dies and a second group of at least one array die;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die comprising at least a first data conduit between the first die interconnect and a first data terminal of the plurality of data terminals, at least a second data conduit between the second die interconnect and the first data terminal, and chip select conduits for providing chip select signals to respective array dies, the chip select signals being related to at least some of the control signals;

wherein the control die further comprises a control circuit to control respective states of the first data conduit and the second data conduit to drive a data signal to an array die selected by at least one of the chip-select signals (paragraph [0057], lines 7-13).

Claim Rejections - 35 USC § 103

9. The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

10. Claims 6, 14, 25 and 26 rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Rajan in view of Kirby et al. (US 2011/0193226, hereinafter "Kirby").

The only difference between claims 6. 25 and 26 and Rajan is that the first and second die interconnects each comprising at least one through silicon vias. However, Kirby discloses the use of a through-silicon via for a conductive component for interconnecting with an external component (paragraph [0036], lines 7-12). It would have been obvious to one having ordinary skill in the art at the time the invention was made to use a through-silicon via as a die interconnect for interconnecting the memory package with an external component.

11. Claims 30 and 32 are rejected under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Rajan in view of Gervasi (US 2006/0259678).

Claims 30 and 32 differ from Rajan in reciting that the register device or the control die is configured to perform rank multiplication by generating the ship select from the control signals. However, Gervasi discloses a register distributing the address and controls to the RAM chips to perform rank multiplication (paragraph [0044], lines 1-5.) It would have been obvious to one having ordinary skill in the art at the time the invention was made to use a register device to perform rank multiplication from the control signals

Allowable Subject Matter

12. Claims 8 and 22 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

.The following is a statement of reasons for the indication of allowable subject matter:

.Regarding claim 8: The prior art made of record and considered pertinent to the applicant's disclosure does not teach the claimed limitation of "wherein a first number of array dies in the first group of array dies and a second number of at least one array die in the second group of at least one array die are selected in consideration of a load of the first die interconnect and a load of the second die interconnect so as to reduce a different between a first load on the first data conduit and a second load on the second data conduit, the first load including a load of the first die interconnect, and a load of the first group of array dies, and the second load including a load of the second die

interconnect and a load of the second group of at least one array die" in combination with the other limitations thereof as is recited in the claim.

Regarding claim 22: The prior art made of record and considered pertinent to the applicant's disclosure does not teach the claimed limitation of "wherein the first driver size and the second driver size are both less than a driver size sufficient to drive a signal along a die interconnect in electrical communication with each of the plurality of array dies without significant signal degradation." in combination with the other limitations thereof as is recited in the claim.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to HUAN HOANG whose telephone number is (571)272-1779. The examiner can normally be reached on Tues-Fri 8:30AM-7:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Amir Zarabian can be reached on (571) 272-1852. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Huan Hoang/
Primary Examiner, Art Unit 2827

Notice of References Cited	Application/Control No. 13/288,850	Applicant(s)/Patent Under Reexamination LEE, HYUN	
	Examiner HUAN HOANG	Art Unit 2827	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-2008/0025137	01-2008	Rajan et al.	365/239
*	B	US-2006/0259678	11-2006	Gervasi, William M.	711/002
*	C	US-2011/0193226	08-2011	Kirby et al.	257/738
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

PTO/SB/08a/b(07-09)
Approved for use through 7/31/2012. OMB 0651-0031
US Patent & Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	13/288,850
Filing Date	03-Nov-2011
First Named Inventor	Hyun Lee
Art Unit	2827
Examiner Name	~~Amir Zarabian~~ Huan Hoang
Attorney Docket Number	NETL.071A

US PATENT DOCUMENTS

Examiner Initials *	Cite No.[1]	Document Number / Number-Kind Code [2] (if known)	Publication Date (MM-DD-YYYY)	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	1	20080025137	01-31-2008	Natarajan et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials *	Cite No.[1]	Foreign Patent Document	Publication Date (MM-DD-YYYY)	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]
	2	EP 1 816 570 A2	08-08-2007	Rambus, Inc.		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	3	International Search Report and Written Opinion, PCT/US2011/059209, January 31, 2013	

EXAMINER	/Huan Hoang/	DATE CONSIDERED	10/06/2013

* EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patents Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST. 3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document and WIPO Standard ST. 16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	~~Amir Zarabian~~ Huan Hoang
SHEET 1 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT DOCUMENTS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	6,243,283 B1	06-05-2001	Bertin et al.	
	2	6,551,857 B2	04-22-2003	Leedy	
	3	7,098,541 B2	08-29-2006	Adelmann	
	4	7,200,021 B2	04-03-2007	Raghuram	
	5	7,254,036 B2	08-07-2007	Pauley et al.	
	6	7,269,042 B2	09-11-2007	Kinsley et al.	
	7	7,286,436 B2	10-23-2007	Bhakta et al.	
	8	7,289,386 B2	10-30-2007	Bhakta et al.	
	9	7,375,970 B2	05-20-2008	Pauley et al.	
	10	7,442,050 B1	10-28-2008	Bhakta et al.	
	11	7,532,537 B2	05-12-2009	Solomon et al.	
	12	7,619,893 B1	11-17-2009	Yu	
	13	7,619,912 B2	11-17-2009	Bhakta et al.	
	14	7,630,202 B2	12-08-2009	Pauley et al.	
	15	7,633,165 B2	12-15-2009	Hsu et al.	
	16	7,636,274 B2	12-22-2009	Solomon et al.	
	17	7,683,459 B2	03-23-2010	Ma et al.	
	18	7,811,097 B1	10-12-2010	Bhakta et al.	
	19	7,827,348 B2	11-02-2010	Lee et al.	
	20	7,830,692 B2	11-09-2010	Chung et al.	
	21	7,839,645 B2	11-23-2010	Pauley et al.	
	22	7,864,627 B2	01-04-2011	Bhakta et al.	
	23	7,881,150 B2	02-01-2011	Solomon et al.	
	24	7,894,229 B2	02-22-2011	Lahtinen et al.	
	25	7,894,230 B2	02-22-2011	Kim	
	26	7,978,721 B2	07-12-2011	Jeddeloh et al.	
	27	7,990,171 B2	08-02-2011	Chung et al.	
	28	7,999,367 B2	08-16-2011	Kang et al.	

Examiner Signature	Date Considered

*__Examiner:__ Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 2 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	29	8,001,434 B1	08-16-2011	Lee et al	
	30	8,019,589 B2	09-13-2011	Rajan et al.	
	31	8,033,836 B1	10-11-2011	Bhakta et al.	
	32	2006/0233012 A1	10-19-2006	Sekiguchi et al.	
	33	2007/0096332 A1	05-03-2007	Satoh et al.	
	34	2008/0025123 A1	01-31-2008	Rajan et al.	
	35	2008/0094808 A1	01-24-2008	Kanapathippillai et al.	
	36	2008/0025134 A1	01-31-2008	Rajan et al.	
	37	2008/0253085 A1	10-16-2008	Soffer	
	38	2008/0296779 A1	12-04-2008	Matsui et al.	
	39	2009/0070727 A1	03-12-2009	Solomon	
	40	2009/0103345 A1	04-23-2009	McLaren et al.	
	41	2009/0290442 A1	11-26-2009	Rajan	
	42	2010/0020583 A1	01-28-2010	Kang et al.	
	43	2010/0090338 A1	04-15-2010	Lee et al.	
	44	2010/0174858 A1	07-08-2010	Chen et al.	
	45	2011/0006360 A1	01-13-2011	Ikebuchi	
	46	2011/0016250 A1	01-20-2011	Lee et al.	
	47	2011/0016269 A1	01-20-2011	Lee et al.	
	48	2011/0050320 A1	03-03-2011	Gillingham	
	49	2011/0108888 A1	05-12-2011	Or-Bach et al.	
	50	2011/0125982 A1	05-26-2011	Choi et al.	
	51	2011/0156232 A1	06-30-2011	Youn et al.	
	52	2011/0169171 A1	07-14-2011	Marcoux	

Examiner Signature	Date Considered

*__Examiner:__ Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 3 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT APPLICATIONS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Filing Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	53	12/422,912 (Attorney Docket No.NETL.044A)	04-13-2009	Hyun Lee, et al.	
	54	12/422,853 (Attorney Docket No.NETL.045A)	04-13-2009	Hyun Lee, et al.	
	55	12/815,339 (Attorney Docket No.NETL.062A)	06-14-2010	Hyun Lee	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	56	WO 2010-138480	12-02-2010	Rambus Inc.		
	57	WO 2011-049710	04-28-2011	Rambus Inc.		
	58	WO 2011-094437	08-04-2011	Hewlett Packard Development Co.		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	59	AHMAD et al., "Modeling of peak-to-peak switching noise along a vertical chain of power distribution TSV pairs in a 3D stack of ICs interconnected through TSVs," Norchip Conference, Article Number 5669473, Nov. 15-16, 2010, IEEE Computer Society	
	60	BLACK et al., "Die Stacking (3D) Microarchitecture," MICRO-39, 39th Annual IEEE/ACM International Symposium on, Dec. 2006, 469-479, Orlando, FL	
	61	DANESHTALAB et al., "CMIT – A novel cluster-based topology for 3D stacked architectures," 3D Systems Integration Conference (3DIC), 2010 IEEE International, Nov. 16-18, 2010, pp. 1-5	
	62	FUNAYA et al., "Cache partitioning strategies for 3-D stacked vector processors," IEEE 3D System Integration Conference, article no. 5751453, Nov. 16-18, 2010, IEEE Computer Society	
	63	GHOSH et al., "Smart Refresh: An Enhancement Memory Controller Design for Reducing Energy in Conventional and 3D Die-Stacked DRAMs," Microarchitecture, 2007, 40th Annual IEEE/ACM International Symposium, Dec. 1-5, 2007, pp. 134-145.	
	64	KANG et al., "8 Gb 3-D DDR3 DRAM using through-silicon-via technology," IEEE Journal of Solid-State Circuits, v 45, n 1, 111-19, Jan. 2010, IEEE, USA	

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 4 OF 4	Attorney Docket No.	NETL.071A

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	65	KANG et al., "Signal integrity and reliability of a new Multi-Stack Package using a Pressure Conductive Rubber," Electrical Design of Advanced Packaging and Systems Symposium, 214-17, Dec. 2008, Seoul, South Korea	
	66	KAWANO, "A 3D Packaging Technology for High-Density Stacked DRAM," VLSI Technology, Systems and Applications, 2007, 23-25 April 2007, pp. 1-2	
	67	KIM et al., "A Quantitative Analysis of Performance Benefits of 3D Die Stacking on Mobile and Embedded SoC	
	68	KURITA et al., "A 3-D packaging technology with highly-parallel memory/logic interconnect," IEICE Transactions on Electronics, v E92-C, no. 12, pp. 1512-1522, 2009, Maruzen Col, Ltd.	
	69	KURITA et al., "Vertical Integration of Stacked DRAM and High-Speed Logic Device Using SMAFTI Technology," Advanced Packaging, IEEE Transactions, Aug. 2009, Volume 32 Issue 3, pp. 657-665	
	70	"Posts Tagged '3D Stacking'," from Chip Design Mag., http://chipdesignmag.com/lpd/blog/tag/3d-stacking/ (Printed October 13, 2011).	
	71	RUSSELL, Gill, "Intel Micron Hybrid Memory Cube: The Future of Exascale Computing," Bright Side of News. Sept. 19, 2011, <http://www.brightsideofnews.com/news/2011/9/19/intel-micron-hybrid-memory-cube-the-future-of-exascale-computing.aspx> Printed October 13, 2011 in 8 pages.	
	72	LOH, "3D-Stacked Memory Architectures for Multi-core Processors," Proceedings of the 35th Annual International Symposium on Computer Architecture (ISCA '08), 453-464, IEEE Computer Society, Washington DC, USA	
	73	VAL, "The 3D interconnection applications for mass memories and microprocessors," Proceedings of the Technical Conference, 1991 International Electronic Packaging Conference, 851-60, vol. 2, 1991, Int. Electron. Packaging Soc, Wheaton, IL, USA, 2008	
	74	WEIS et al., "Design space exploration for 3D-stacked DRAMs," Proceedings – Design, Automation and Test in Europe Conference and Exhibition, 2011	
	75	ZHANG et al. "A Customized Design of DRAM Controller for On-Chip 3D DRAM Stacking" Custom Integrated Circuits Conference (CICC), 2010 IEEE, Sept. 19-22 2010, pp. 1-4	

13081580

Examiner Signature	/Huan Hoang/	Date Considered	10/06/2013

*Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	71957	(die or dice or integrated adj circuit or ic or chip) with stack$3	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:26
L2	12869	1 and (die or dice or integrated adj circuit or ic or chip) with interconnect	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:27
L3	3356	1 and (die or dice or integrated adj circuit or ic or chip) with interconnect with (group or portion or number or part)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:27
L4	61	3 and control adj (die or chip or integrated adj circuit or ic)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:29
L5	1	4 and interconnect with through-silicon adj via	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:30
L6	59	3 and interconnect with through-silicon adj via	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:30
L7	3	1 and (die or dice or integrated adj circuit or ic or chip) near5 number with reduc$3 with load	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:51
L8	202	(die or dice or integrated adj circuit or ic or chip) near5 number with reduc$3 with load	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:59
L9	0	(die or dice or integrated adj circuit or ic or chip) near5 number with interconnect with reduc$3 with load	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 20:59
L10	0	(die or dice or integrated adj circuit or ic or chip) near5 number with interconnect$3 with reduc$3 with load	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:00
L11	7	(die or dice or integrated adj circuit or ic or chip) with rank adj multiplication	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:09
L12	66	"7289386"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:14

L13	1	12 and rank adj multiplication	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:15
L14	4	(register or control ajd die) with rank adj multiplication	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:21
L15	1	register with rank adj multiplication	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 21:23
L16	3	"20080025134"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 22:01
S1	5	"2008025137"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 16:31
S2	3	"20080025137"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2013/10/06 16:31

EAST Search History (Interference)

<This search history is empty>

10/6/2013 10:21:33 PM
C:\Users\hhoang1\Documents\EAST\Workspaces\default.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	Art Unit
	HUAN HOANG	2827

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	06/02/2013	10/06/2013							
	1	÷	✓							
	2	÷	✓							
	3	÷	✓							
	4	÷	✓							
	5	÷	✓							
	6	÷	✓							
	7	÷	✓							
	8	÷	O							
	9	÷	✓							
	10	÷	✓							
	11	÷	✓							
	12	÷	✓							
	13	÷	✓							
	14	÷	✓							
	15	÷	✓							
	16	÷	✓							
	17	÷	✓							
	18	÷	✓							
	19	÷	✓							
	20	÷	✓							
	21	÷	✓							
	22	÷	O							
	23	÷	✓							
	24	÷	✓							
	25	÷	✓							
	26	÷	✓							
	27	÷	✓							
	28	÷	✓							
	29	÷	✓							
	30	÷	✓							
	31	÷	✓							
	32	÷	✓							
	33	÷	✓							
	34	÷	✓							

Petitioners
Ex. 1002, p. 454

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	Art Unit
	HUAN HOANG	2827

CPC- SEARCHED		
Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner
365	63, 51, 52	10/6/13	HH

SEARCH NOTES		
Search Notes	Date	Examiner
EAST (text search)	10/6/13	HH

INTERFERENCE SEARCH			
US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

	/HUAN HOANG/ Primary Examiner.Art Unit 2827

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	:	Hyun Lee
App. No.	:	13/288,850
Filed	:	November 3, 2011
For	:	ARCHITECTURE FOR MEMORY MODULE WITH PACKAGES OF THREE-DIMENSIONAL STACKED (3DS) MEMORY CHIPS
Examiner	:	ZARABIAN, AMIR
Art Unit	:	2827
Conf. No.	:	2466
Docket No.	:	NETL.071A

CERTIFICATE OF EFS WEB TRANSMISSION

I hereby certify that this correspondence and any other attachment noted on the Automated Acknowledgement Receipt, is being transmitted from within the Pacific Time zone to the Commissioner for Patents via the EFS Web Server on:

January 13, 2014

(Date)

/Jamie J. Zheng/

Jamie J. Zheng, Reg. No. 51167

AMENDMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

The enclosed Amendment is in response to the Non-Final Office Action dated October 11, 2013 for the above-identified patent application.

The Commissioner is hereby authorized to charge any required fee(s) to Deposit Account No. 50-5963.

AMENDMENTS to the CLAIMS begin on Page 2 of this paper.

REMARKS begin on Page 9 of this paper.

IN THE CLAIMS:

Rewrite the pending claims as follows:

1. (**Previously Presented**) A memory package, comprising:

a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices;

a plurality of stacked array dies including a first group of array dies and a second group of at least one array die, each array die having data ports;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die comprising at least a first data conduit between the first die interconnect and a first terminal of the plurality of input/output terminals, and at least a second data conduit between the second die interconnect and the first terminal, the first terminal being a data terminal, the control die further comprising a control circuit to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals.

2. (**Previously Presented**) The memory package of claim 1, wherein the control signals include data path control signals for controlling the first and second data conduits.

3. (**Previously Presented**) The memory package of claim 1, wherein the control circuit is configured to generate data path control signals for controlling the first and second data conduits in response to the received control signals.

4. (**Previously Presented**) The memory package of claim 3, wherein the control signals include command/address signals and wherein the control die is configured to provide the command/address signals to the plurality of stacked array dies.

5. (**Cancelled**)

6. **(Previously Presented)** The memory package of claim 1, wherein the first die interconnect comprises a first through-silicon via and wherein the second die interconnect comprises a second through-silicon via.

7. **(Previously Presented)** The memory package of claim 1, wherein the control die further comprises chip-select conduits, the memory package further comprising:

third die interconnects coupled between respective chip-select conduits and respective ones of the plurality of stacked array dies.

8. **(Currently Amended)** The memory package of claim 1, wherein a first number of array dies in the first group of array dies and a second number of at least one array die in the second group of at least one array die are selected in consideration of a load of the first die interconnect and a load of the second die interconnect so as to reduce a difference ~~different~~ between a first load on the first data conduit and a second load on the second data conduit, the first load including a load of the first die interconnect, and a load of the first group of array dies, and the second load including a load of the second die interconnect and a load of the second group of at least one array die.

9. **(Cancelled)**

10. **(Cancelled)**

11. **(Previously Presented)** A memory package, comprising:

a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices;

a plurality of array dies arranged in a stack, including a first group of array dies and a second group of at least one array die;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die comprising at least a first data conduit between the first die interconnect and a first terminal of the plurality of input/output terminals, at least a second data conduit between the second die interconnect and the first terminal, and chip select conduits for providing chip select signals to respective array dies;

Petitioners
Ex. 1002, p. 458

wherein the control die further comprises a control circuit to control respective states of the first data conduit and the second data conduit to drive a data signal to an array die selected by at least one of the chip-select signals.

12. (**Previously Presented**) The memory package of claim 11, wherein the chip select conduits pass through the control die.

13. (**Previously Presented**) The memory package of claim 11, wherein the chip select conduits include drivers to drive the chip select signals to the respective array dies.

14. (**Previously Presented**) The memory package of claim 11, wherein the first die interconnect comprises one or more through silicon vias and wherein the second die interconnect comprises one or more through silicon vias.

15. (**Currently Amended**) The memory package of claim 11, wherein: the first data conduit comprises at least a first driver having a first driver size, and the second data conduit comprises at least a second driver having a second driver size, and wherein the first driver size and the second driver size are both less than a driver size sufficient to drive a signal along a die interconnect in electrical communication with each of the plurality of array dies without significant signal degradation.

16. (**Currently Amended**) The memory package of claim 11, wherein the control die is configured to generate the chip select signals from control signals received via second terminals of the plurality of terminals, and wherein the chip select signal is generated using an address signal in the control/address signals.

17. (**Currently Amended**) The memory package of claim 11, wherein the control circuit controls the respective states of the first data conduit and the second data conduit in response to at least some of the control/address signals received via second terminals of the plurality of terminals.

18. (**Currently Amended**) The memory package of claim 16, wherein the control/address signals comprise ~~at least one of the following: a~~ at least one command signal, ~~an~~ at least one address signal, and ~~a~~ at least one data path control signal.

19. (**Currently Amended**) The memory package of claim 11, wherein the control die is configured to generate data path control signals from the control/address signals received via

Petitioners
Ex. 1002, p. 459

second terminals of the plurality of terminals, and wherein the control circuit controls the respective states of the first data conduit in response to the data path control signals.

20. (**Previously Presented**) A method for optimizing load in a memory package comprising a plurality of array dies arranged in a stack and including a first group of array dies and a second group of at least one array die, at least a first die interconnect and a second die interconnect, a control die, and a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices, the method comprising:

receiving a data signal at a first terminal of the plurality of input/output terminals;

receiving control signals at second terminals of the plurality of input/output terminals;

providing chip select signals to respective array dies through the control die, the chip select signals being related to at least some of the control signals; and

selecting one of a first driver and a second driver in the control die to drive the data signal via a corresponding one of the first die interconnect and the second die interconnect to an array die selected by at least one of the chip select signals, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies.

21. (**Previously Presented**) The method of claim 20, further comprising:

selecting a first driver size for the first driver based, at least in part, on a load on the first driver; and

selecting a second driver size for the second driver based, at least in part, on a load on the second driver.

22. (**Previously Presented**) The method of claim 21, wherein the first driver size and the second driver size are both less than a driver size sufficient to drive a signal along a die interconnect in electrical communication with each of the plurality of array dies without significant signal degradation.

23. (**Previously Presented**) The method of claim 20, further comprising generating the chip select signals from at least some of the control signals.

Petitioners
Ex. 1002, p. 460

24. (**Previously Presented**) The method of claim 20, wherein the control signals include the chip-select signals.

25. (**Previously Presented**) The method of claim 20, wherein the first and second die interconnects each comprises at least one through silicon vias.

26. (**Previously Presented**) The method of claim 20, wherein the chip select signals pass through through-silicon-vias in the control die.

27. (**Previously Presented**) The method of claim 20, further comprising generating data path control signals from at least some of the control signals, the data path control signals being used to select the one of the first driver and the second driver in the control die to drive the data signal.

28. (**Previously Presented**) The method of claim 20, wherein the one of the first driver and the second driver is selected using at least some of the control signals.

29. (**Previously Presented**) A memory module operable via a memory control hub, comprising:

a register device configured to receive command/address signals from the memory control hub and to generate control signals; and

a plurality of DRAM packages, each DRAM package comprising:

a plurality of data terminals via which the DRAM package communicate data with the memory control hub, and a plurality of control terminals to receive the control signals;

a plurality of array dies arranged in a stack, including a first group of array dies and a second group of at least one array die;

at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and

a control die comprising at least a first data conduit between the first die interconnect and a first data terminal of the plurality of data terminals, at least a second data conduit between the second die interconnect and the first data terminal, and chip select conduits for providing chip select signals to respective array dies, the chip select signals being related to at least some of the control signals;

Petitioners
Ex. 1002, p. 461

wherein the control die further comprises a control circuit to control respective states of the first data conduit and the second data conduit to drive a data signal to an array die selected by at least one of the chip-select signals.

30. (**Previously Presented**) The memory module of claim 29, wherein the register device is further configured to perform rank multiplication by generating the chip select signals, and wherein the control signals include the chip select signals.

31. (**Previously Presented**) The memory module of claim 29, wherein the control signals include data path control signals generated by the register device, the data path control signals being used to control the respective states of the first data conduit and the second data conduit.

32. (**Currently Amended**) The memory module of claim 29, wherein the control die is further configured to perform rank multiplication by generating the chip select signals from at least some of the control signals that include at least one address signal.

33. (**Previously Presented**) The memory module of claim 29, wherein the control signals include command/address signals, and the control die is configured to hold the command/address signals to control timing of the command/address signals.

34. (**Previously Presented**) The memory module of claim 29, wherein the control die is configured to generate data path control signals from at least some of the control signals, the data path control signals being used to control the respective states of the first data conduit and the second data conduit.

35. (**New**) The memory package of claim 1, wherein the respective states of the first data conduit and the second data conduit are controlled by one or more data path control signals, wherein the control die is configurable to operate in any one of a first mode and a second mode, and wherein:

in the first mode, the control die receives the data path control signals from the one or more external devices; and

in the second mode, the control die generates the data path control signals from at least some of the control/address signals received from the one or more external devices.

Petitioners
Ex. 1002, p. 462

36. (**New**) The memory package of claim 1, wherein the control die further comprises command/address conduits configured to provide corresponding command/address signals to the array dies, the command/address signals including at least one memory cell address.

37. (**New**) The memory package of claim 1, wherein the control die further comprises one or more additional conduits configured to provide one or more of a supply voltage signal and a ground signal to the array dies.

Petitioners
Ex. 1002, p. 463

REMARKS

This amendment responds to the Non-Final Office Action dated October 11, 2013. In the Office Action, the Examiner:

- withdrew the restriction requirement of claims 29-34;
- noted two issues with the Information Disclosure Statement submitted on 5/23/2012;
- objected to informalities in claim 5 and claim 8;
- rejected claim 5 under 35 U.S.C. 112(b) or 35 U.S.C.112 (pre-AIA), second paragraph as being indefinite;
- rejected claims 1-4, 7, 9-13, 15-21, 23, 24, 27-29, 31, 33, and 34 under pre-AIA 35 U.S.C. 102(b) as being anticipated by Rajan et al. (US 2008/0025137);
- rejected claims 6, 14, 25, and 26 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Rajan in view of Kirby et al. (US 2011/0193226);
- rejected claims 30 and 32 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Rajan in view of Gervasi (US 2006/0259678); and
- objected to claims 8 and 22 as being dependent upon a rejected base claim.

Applicant appreciates the allowance of claims 8 and 22 if rewritten in independent form, but would like to request reconsideration of the rejected claims in light of the amendments and remarks.

I. CLAIM AMENDMENT

Claims 5, 9 and 10 are cancelled. Claims 8, 15-19 and 32 are amended. New claims 35-37 are presented. No new matter is added.

II. INFORMATION DISCLOSURE STATEMENT SUBMITTED ON 5/23/2012

The Examiner noted that the document "A Quantitative Analysis of Performance Benefits of 3D Die Stacking on Mobile and Embedded SoC" in the information statement filed 05/23/2012 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because a date has not been provided. Base on a recent Internet search, this article appears to have been published in 978-3-9810801-7-9/DATE11/©2011 EDAA (http://www.date-

Petitioners
Ex. 1002, p. 464

conference.com/proceedings/PAPERS/2011/DATE11/PDFFILES/10.7_4.PDF). So, the year of publication is 2011. The exact date of this publication, however, is unknown to the Applicant.

III. OBJECTION TO INFORMALITIES IN CLAIM 5 AND CLAIM 8

Claim 5 is cancelled. Claim 8 is amended to cure the informalities. Therefore, the objection to claim 5 is moot and the objection to claim 8 should be withdrawn.

IV. REJECTION OF CLAIM 5 UNDER 35 U.S.C. 112(B) OR 35 U.S.C.112 (PRE-AIA), SECOND PARAGRAPH

Claim 5 is cancelled. Therefore, the rejection thereof is moot.

V. REJECTION OF CLAIMS 1-4, 7, 9-13, 15-21, 23, 24, 27-29, 31, 33, AND 34 UNDER PRE-AIA 35 U.S.C. 102(B) AS BEING ANTICIPATED BY RAJAN ET AL. (US 2008/0025137)

Claim 1 recites:

> A memory package, comprising:
> a plurality of input/output terminals via which the memory package communicate data and control/address signals with one or more external devices;
> a plurality of stacked array dies including a first group of array dies and a second group of at least one array die, each array die having data ports;
> at least a first die interconnect and a second die interconnect, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies; and
> a control die comprising at least a first data conduit between the first die interconnect and a first terminal of the plurality of input/output terminals, and at least a second data conduit between the second die interconnect and the first terminal, the first terminal being a data terminal, the control die further comprising a control circuit to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals.

First of all, Rajan does not disclose "a plurality of stacked array dies." Rajan merely stacks DRAM circuits 206A-D, which are different from array dies. As a result, Rajan's buffer chip 202 also operates very differently from the control die in claim 1. In rejecting claim 1, the Examiner states that Rajan (FIG. 2B) shows a memory package as recited in

Petitioners
Ex. 1002, p. 465

claim 1. The Examiner further cites Paragraph [0044] as disclosing the control die as recited in claim 1. Paragraph [0044] in Rajan, however, merely states:

> As shown in each of such figures, the buffer chip 202 is placed electrically between an electronic host system 204 and a stack of DRAM circuits 206A-D. In the context of the present description, a stack may refer to any collection of memory circuits. Further, the buffer chip 202 may include any device capable of buffering a stack of circuits (e.g. DRAM circuits 206A-D, etc.). Specifically, the buffer chip 202 may be capable of buffering the stack of DRAM circuits 206A-D to electrically and/or logically resemble at least one larger capacity DRAM circuit to the host system 204. In this way, the stack of DRAM circuits 206A-D may appear as a smaller quantity of larger capacity DRAM circuits to the host system 204.

Thus, the statements in this paragraph [0044] of Rajan says nothing about the existence of a first data conduit and a second data conduit in the buffer chip 202 and "a control circuit to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals." According to this paragraph [0044], all that is required of the buffer chip 202 is the capability of buffering the stack of DRAM circuits 206A-D to electrically and/or logically resemble at least one larger capacity DRAM circuit to the host system, and this requirement does not necessitate the use of a first data conduit and a second data conduit in the buffer chip 202 and "a control circuit to control respective states of the first data conduit and the second data conduit in response to control signals received via one or more second terminals of the plurality of terminals."

Furthermore, according to paragraphs [0043] and [0044], this same capability of buffering the stack of DRAM circuits 206A-D to electrically and/or logically resemble at least one larger capacity DRAM circuit to the host system is required for each of the FIGS. 2A-2E, which show various configurations of a buffered stack of DRAM circuits 206A-D with the buffer chip 202. Since FIG. 2D of Rajan shows a single data bus between the buffer chip and all of the stacked DRAM circuits 206A-D (*Rajan, paragraph [0051]*), this capability of buffering the stack of DRAM circuits 206A-D, which also applies to the configuration in FIG. 2D, can not be said to imply the existence of first and second data conduits in the buffer chip 202 and a control circuit controlling the respective states of the first data conduit and the second data conduit in response to any received control signals, because the first and second data conduits would be coupled to a same data line in the single data bus in FIG. 2D and it would make no sense to try to control respective states or the first data conduit and the second data conduit.

Petitioners
Ex. 1002, p. 466

Therefore, Rajan fails to disclose each and every limitation in claim 1, and claim 1 is patentable over Rajan.

Claims 2-4 and 7 depend from claim 1 and include further limitations in addition to the limitations in claim 1. Therefore, claims 2-4 and 7 are patentable for at least the same reasons claim 1 is patentable.

Claims 2-4 and 7 are also patentable for the additional features recited therein. For example, claim 2 includes the further limitations that the control signals received via the one or more second terminals of the plurality of terminals include data path control signals for controlling the first and second data conduits. Rajan does not disclose any data path control signals. In Rajan, besides the data signals, the signals the buffer chip receives are address/control/clock signals (*Rajan, Paragraph [0025]*). These are signals used for memory operations, not for controlling any data conduits in the buffer chip 202.

As a further example, claim 3 recites "the control circuit is configured to generate data path control signals for controlling the first and second data conduits in response to the received control signals." As stated above, Rajan does not disclose any data path control signals. Therefore, Rajan cannot be said to disclose anything about generating the data path control signals in its buffer chip 202.

Claims 9-10 are cancelled. Therefore, the rejections thereof are moot.

The arguments regarding claim 1 applies to claim 11. Therefore, claim 11 is also patentable over Rajan.

Claims 12-13, and claims 15-19 as amended depend from claim 11 and include further limitations in addition to the limitations in claim 11. Therefore, claims 12-13, and claims 15-19 as amended are patentable for at least the same reasons claim 11 is patentable.

Claims 12-13, and claims 15-19 as amended are also patentable for the additional features recited therein. For example, claim 15 is amended to include certain features in claim 22, which are considered allowable.

Claim 16 as amended recites that the control die is configured to generate the chip select signals from control signals received via second terminals of the plurality of terminals, and wherein the chip select signal is generated using at least one address signal in the

Petitioners
Ex. 1002, p. 467

control/address signals. Rajan does not teach generating chip select signals in the buffer chip 202, let alone generating chip select signals using a received address signal.

Claim 18 as amended recites that the control/address signals comprise: at least one command signal, at least one address signal, and at least one data path control signal. Again Rajan does not teach data path control signal, nor does Rajan teaches receiving such signals at the buffer chip 202.

Claim 19 as amended recites that the control die is configured to generate data path control signals from the control/address signals received via second terminals of the plurality of terminals, and wherein the control circuit controls the respective states of the first data conduit in response to the data path control signals. Again, Rajan does not teach data path control signals, let alone generating such signals as the buffer chip from received control/address signals.

Claim 20 recites, among other things:

> selecting one of a first driver and a second driver in the control die to drive the data signal via a corresponding one of the first die interconnect and the second die interconnect to an array die selected by at least one of the chip select signals, the first die interconnect in electrical communication with the first group of array dies and not in electrical communication with the second group of at least one array die, the second die interconnect in electrical communication with the second group of at least one array die and not in electrical communication with the first group of array dies

Rajan does not teach these limitations. As stated above, Rajan only requires its buffer chip 202 to buffer the stack of DRAM circuits 206A-D to electrically and/or logically resemble at least one larger capacity DRAM circuit to the host system, and this same requirement is applicable for each of the FIGS. 2A-2E (*Rajan, paragraphs [0043] and[0044]*). Since FIG. 2D shows that all of the DRAM circuits 206A-D are coupled to the buffer chip 202 via a same data bus, this requirement cannot be used to infer that the buffer chip 202 has to select "one of a first driver and a second driver to drive the data signal via a corresponding one of the first die interconnect and the second die interconnect to an array die selected by at least one of the chip select signals." Therefore, Rajan does not disclose each and every limitations in claim 20, and claim 20 is patentable over Rajan.

Claims 21, 23-24, and 27-28 depend from claim 20 and include further limitations in addition to the limitations in claim 20. Therefore, claims 21, 23-24, and 27-28 are patentable for at least the same reasons claim 20 is patentable.

Petitioners
Ex. 1002, p. 468

The arguments regarding claim 1 is applicable to claim 29. Therefore, claim 29 is also patentable over Rajan.

Claims 31 and 33-34 depend from claim 29 and include further limitations in addition to the limitations in claim 29. Therefore, claims 31 and 33-34 are patentable for at least the same reasons claim 20 is patentable.

VI. REJECTION OF CLAIMS 6, 14, 25-26, 30 AND 32 UNDER PRE-AIA 35 U.S.C. 103(A)

Claims 6 and 14 depend from claims 1 and 11, respectively, and include further limitations in addition to the limitations in claims 1 and 11, respectively. Therefore, claims 6 and 14 are patentable for at least the same reasons claims 1 and 11, respectively, are patentable.

Claims 25-26 depend from claim 20 and include further limitations in addition to the limitations in claim 20. Therefore, claims 25-26 are patentable for at least the same reasons claim 20 is patentable.

Claim 30 and claim 32 as amended depend from claim 29 and include further limitations in addition to the limitations in claim 29. Therefore, claim 30 and claim 32 are patentable for at least the same reasons claim 20 is patentable.

VII. NEW CLAIMS 35-37

New claims 35-37 depend from claim 1, and include further limitations in addition to the limitations in claim 1. Therefore, new claims 35-37 are patentable for at least the same reasons claim 1 is patentable.

VIII. NO DISCLAIMERS OR DISAVOWALS

Although the present communication may include alterations to the application or claims, or characterizations of claim scope or referenced art, Applicant is not conceding in this application that previously pending claims are not patentable over the cited references. Rather, any alterations or characterizations are being made to facilitate expeditious prosecution of this application. Applicant reserves the right to pursue at a later date any previously pending or other broader or narrower claims that capture any subject matter

Petitioners
Ex. 1002, p. 469

supported by the present disclosure, including subject matter found to be specifically disclaimed herein or by any prior prosecution. Accordingly, reviewers of this or any parent, child or related prosecution history shall not reasonably infer that Applicant has made any disclaimers or disavowals of any subject matter supported by the present application.

IX. SUMMARY

At least for the foregoing reasons, Applicant submits that claims 1-4, 6-8, and 11-37 are in condition for allowance, and Applicant respectfully requests such action. The Examiner is invited to call the undersigned at (650) 273-6008 if such phone call would help resolve any remaining issues.

Respectfully submitted,

Date: January 13, 2014

/Jamie J. Zheng/ 51167

Jamie J. Zheng (Reg. No.)

Customer No. 79141

Petitioners
Ex. 1002, p. 470

Electronic Acknowledgement Receipt

EFS ID:	17902931
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	14-JAN-2014
Filing Date:	03-NOV-2011
Time Stamp:	02:46:11
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		2014-01-13_Amendment_NFOA_NETL071A.pdf	168355 b33ca64fdc3320438b5b3ac2786db1fd6b4608e0	yes	15

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment/Req. Reconsideration-After Non-Final Reject	1	1
	Claims	2	8
	Amendment/Req. Reconsideration-After Non-Final Reject	9	15

Warnings:	
Information:	
Total Files Size (in bytes):	168355

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 13/288,850	Filing Date 11/03/2011	☐ To be Mailed

ENTITY: ☒ LARGE ☐ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☒ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	380
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	380

APPLICATION AS AMENDED – PART II

AMENDMENT 01/14/2014

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 34	Minus	** 34	= 0	X $80 =	0
Independent (37 CFR 1.16(h))	* 4	Minus	***4	= 0	X $420 =	0
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					TOTAL ADD'L FEE	0

AMENDMENT

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

LIE
/TER JACKSON WOODRUFF/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

79141 7590 03/26/2014
The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER
HOANG, HUAN

ART UNIT	PAPER NUMBER
2827	

DATE MAILED: 03/26/2014

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

TITLE OF INVENTION: METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$960	$0	$0	$960	06/26/2014

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

Petitioners
Ex. 1002, p. 474

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

79141 7590 03/26/2014
The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

TITLE OF INVENTION: METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$960	$0	$0	$960	06/26/2014

EXAMINER	ART UNIT	CLASS-SUBCLASS
HOANG, HUAN	2827	365-063000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:
- ❑ Issue Fee
- ❑ Publication Fee (No small entity discount permitted)
- ❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ❑ A check is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ❑ Applicant certifying micro entity status. See 37 CFR 1.29 — NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ❑ Applicant asserting small entity status. See 37 CFR 1.27 — NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ❑ Applicant changing to regular undiscounted fee status. — NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

Petitioners
Ex. 1002, p. 475

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

79141 7590 03/26/2014

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER	
HOANG, HUAN	

ART UNIT	PAPER NUMBER
2827	

DATE MAILED: 03/26/2014

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)

(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 151 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 151 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No. 13/288,850	Applicant(s) LEE, HYUN	
	Examiner HUAN HOANG	Art Unit 2827	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *the Amendment filed on 01/14/14*.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on_____.
2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.
3. ☒ The allowed claim(s) is/are *1-4,6-8 and 11-37*. As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov .
4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:
   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
5. ☒ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/HUAN HOANG/
Primary Examiner, Art Unit 2827

The present application is being examined under the pre-AIA first to invent provisions.

EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

The application has been amended as follows:

In the claims:

Claim 1, line 2, replace "communicate" with –communicates--;

Claim 11, line 2, replace "communicate" with –communicates--;

Claim 20, line 5, replace "communicate" with –communicates--; and

Claim 29, line 6, replace "communicate" with –communicates--.

2. Any inquiry concerning this communication or earlier communications from the examiner should be directed to HUAN HOANG whose telephone number is (571)272-1779. The examiner can normally be reached on Tues-Fri 8:30AM-7:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Amir Zarabian can be reached on (571) 272-1852. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/HUAN HOANG/
Primary Examiner, Art Unit 2827

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	**Art Unit**
	HUAN HOANG	2827

CPC				
Symbol			**Type**	**Version**
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		
		/		

CPC Combination Sets						
Symbol			**Type**	**Set**	**Ranking**	**Version**
		/				
		/				

		Total Claims Allowed:	
NONE (Assistant Examiner)	(Date)	34	
/HUAN HOANG/ Primary Examiner.Art Unit 2827 (Primary Examiner)	03/18/2014 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 2

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	Art Unit
	HUAN HOANG	2827

US ORIGINAL CLASSIFICATION		INTERNATIONAL CLASSIFICATION										
CLASS	SUBCLASS	CLAIMED					NON-CLAIMED					
365	63	G	1	1	C	5 / 06 (2006.01.01)						
CROSS REFERENCE(S)												
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)											
365	51											

NONE		Total Claims Allowed:	
(Assistant Examiner)	(Date)	34	
/HUAN HOANG/ Primary Examiner.Art Unit 2827	03/18/2014	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	2

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	**Art Unit**
	HUAN HOANG	2827

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	17	17	33	33										
2	2	18	18	34	34										
3	3	19	19	8	35										
4	4	20	20	9	36										
	5	21	21	10	37										
5	6	22	22												
6	7	23	23												
7	8	24	24												
	9	25	25												
	10	26	26												
11	11	27	27												
12	12	28	28												
13	13	29	29												
14	14	30	30												
15	15	31	31												
16	16	32	32												

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 34	
/HUAN HOANG/ Primary Examiner.Art Unit 2827 (Primary Examiner)	03/18/2014 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 2

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	13288850	LEE, HYUN
	Examiner	Art Unit
	HUAN HOANG	2827

CPC- SEARCHED		
Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner
365	63, 51, 52	10/6/13	HH
Above	Updated	3/18/14	HH

SEARCH NOTES		
Search Notes	Date	Examiner
EAST (text search)	10/6/13	HH
EAST (text search)	3/18/14	HH
Inventor search	3/18/14	HH

INTERFERENCE SEARCH			
US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner
	EAST (US-PGPUB, UPAD)	3/18/14	HH

	/HUAN HOANG/ Primary Examiner.Art Unit 2827

Petitioners
Ex. 1002, p. 484

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 2466

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
13/288,850	11/03/2011	365	2827	NETL.071A

APPLICANTS

INVENTORS

Hyun Lee, Residence Not Provided;

**** CONTINUING DATA ***************************

This appln claims benefit of 61/409,893 11/03/2010

**** FOREIGN APPLICATIONS ***************************

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****

11/14/2011

Foreign Priority claimed ☐ Yes ☑ No 35 USC 119(a-d) conditions met ☐ Yes ☑ No Verified and Acknowledged /HUAN HOANG/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
			8	34	4

ADDRESS

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306
UNITED STATES

TITLE

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

FILING FEE RECEIVED		
2470	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	83	"7289386"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:19
L2	2	("7286436").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:21
L3	2	("7442050").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:21
L4	2	("7375970").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:21
L5	2	("7254036").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:22
L6	2	("7532537").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:22
L7	2	("7636274").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:23
L8	2	("7630202").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:24
L9	2	("7619893").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:24
L10	2	("7619912").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:25

L11	3	("7811097").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:25
L12	2	("7839645").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:25
L13	2	("7864627").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:26
L14	2	("7881150").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:26
L15	7	("8001434").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:26
L16	2	("8033836").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:27
L17	2	("8154901").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:27
L18	2	("8417870").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:27
L19	2	("8489837").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:28
L20	2	("8516185").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT	OR	OFF	2014/03/18 18:28
L22	0	"12422912"	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:29
L23	700863	(memopry or stor$3) and ("input/output" or "i/o")	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:32
L24	6676	23 and stack$3 with (die or dice or ic or integrated adj circuit)	US-PGPUB; USPAT; EPO;	OR	ON	2014/03/18 18:36

			JPO; DERWENT			
L25	1318	23 and stack$3 with (die or dice or ic or integrated adj circuit) with (group or set or subset or number)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:38
L26	1936	23 and stack$3 with (die or dice or ic or integrated adj circuit) with (group or set or subset or number or plurality)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:38
L27	693	26 and (interconnect$3 or connect$3) with (die or dice or ic or integrated adj circuit) with (group or set or subset or number)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:39
L28	80	27 and control$4 near3 (die or dice or ic or integrated adj circuit) with (interconnect$3 or connect$3)	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:40
L29	1	28 and data with conduit	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:41
L30	4	28 and control$4 near3 (die or dice or ic or integrated adj circuit) with (interconnect$3 or connect$3) with ("input/output" or "i/o")	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:42
L31	11	28 and chip adj select$3	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:55
L32	9	28 and data with ("input/output" or "i/o") and chip adj select$3	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:58
L33	1	32 and driv$3 with chip adj select$3	US-PGPUB; USPAT; EPO; JPO; DERWENT	OR	ON	2014/03/18 18:58

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L35	0	(memory and package and terminal and communicate and external and die and group and electrical and control and interconnect and conduit).clm.	US-PGPUB; USPAT; UPAD	OR	ON	2014/03/18 19:59
L36	0	(memory and package and terminals and communicate and external and die and stack and electrical and control and interconnect and conduit).clm.	US-PGPUB; USPAT; UPAD	OR	ON	2014/03/18 20:00
L37	0	(memory and package and external and die and stack and electrical and control and interconnect and conduit).clm.	US-PGPUB; USPAT; UPAD	OR	ON	2014/03/18 20:00
L38	0	(memory and package and external and die and stack and interconnect and conduit).clm.	US-PGPUB; USPAT; UPAD	OR	ON	2014/03/18 20:01

3/ 18/ 2014 8:02:24 PM
C:\ Users\ hhoang1\ Documents\ EAST\ Workspaces\ default.wsp

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

79141 7590 05/30/2014

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER	
HOANG, HUAN	

ART UNIT	PAPER NUMBER
2827	

NOTIFICATION DATE	DELIVERY MODE
05/30/2014	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jz@jzpatent.com



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Application No. : 13288850
Applicant : Lee
Filing Date : 11/03/2011
Date Mailed : 05/30/2014

NOTICE TO FILE CORRECTED APPLICATION PAPERS

Notice of Allowance Mailed

This application has been accorded an Allowance Date and is being prepared for issuance. The application, however, is incomplete for the reasons below.

Applicant is given two (2) months from the mail date of this Notice within which to respond. This time period for reply is extendable under 37 CFR 1.136(a) for only TWO additional MONTHS.

The informalities requiring correction are indicated in the attachment(s). If the informality pertains to the abstract, specification (including claims) or drawings, the informality must be corrected with an amendment in compliance with 37 CFR 1.121 (or, if the application is a reissue application, 37 CFR 1.173). Such an amendment may be filed after payment of the issue fee if limited to correction of informalities noted herein. See Waiver of 37 CFR 1.312 for Documents Required by the Office of Patent Publication, 1280 Off. Gaz. Patent Office 918 (March 23, 2004). In addition, if the informality is not corrected until after payment of the issue fee, for purposes of 35 U.S.C. 154(b)(1)(iv), "all outstanding requirements" will be considered to have been satisfied when the informality has been corrected. A failure to respond within the above-identified time period will result in the application being ABANDONED.

See attachment(s).

A copy of this notice MUST be returned with the reply. Please address response to "Mail Stop Issue Fee, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450".

/Kam Sin/
Publication Branch
Office of Data Management
(571) 272-4200

Application No. 13288850

IDENTIFICATION OF APPLICATION DEFICIENCIES

☐ Applicant must provide legible text for the following item(s).

- ☐ Specification filed , page(s) .
- ☐ Claims filed , claim(s) .
- ☐ Oath/declaration filed .
- ☐ Other: .

☐ Applicant must provide missing information on the following page(s) of the specification by amending the specification to add the missing text. No new matter may be added.

☐ The specification refers to one or more applications by attorney docket number and does not show the U.S. application number(s). Applicant must supply the U.S. application number in place of each attorney docket number.

☐ Applicant must provide an Abstract of the Disclosure.

☐ Applicant has submitted a DECLARATION (37 CFR 1.63) FOR A UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76) (e.g., form PTO/SB/01A). The Application Data Sheet, however, is not present with the filed application. Applicant must submit an Application Data Sheet or file a new oath or declaration (e.g., PTO/SB/01) executed by the inventors and containing the information required in 37 CFR 1.63.

☐ Applicant must provide an executed declaration.

☐ Applicant must provide the missing page(s) of the oath/declaration or Application Data Sheet filed

☐ Applicant must provide a declaration signed by inventor(s) .

☐ The oath/declaration filed shows non-initialed and/or non-dated alterations. Applicant must file a new oath/declaration in compliance with 37 CFR 1.67(a).

☒ Applicant(s) Hyun Lee in the latest-filed oath/declaration or Application Data Sheet (ADS) did not show the inventor's residence at all, or did not show both a city and state in the U.S. inventor's residence, or did not show both a city and country in the non-U.S. inventor's residence. Applicant must supply an oath/declaration or Application Data Sheet (ADS) that shows each U.S. inventor's city and state of residence and each non-U.S. inventor's city and country of residence.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

79141 7590 05/30/2014

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

EXAMINER	
HOANG, HUAN	

ART UNIT	PAPER NUMBER
2827	

NOTIFICATION DATE	DELIVERY MODE
05/30/2014	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jz@jzpatent.com

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Application No. : 13288850
Applicant : Lee
Filing Date : 11/03/2011
Date Mailed : 05/30/2014

NOTICE TO FILE CORRECTED APPLICATION PAPERS

Notice of Allowance Mailed

This application has been accorded an Allowance Date and is being prepared for issuance. The application, however, is incomplete for the reasons below.

Applicant is given two (2) months from the mail date of this Notice within which to respond. This time period for reply is extendable under 37 CFR 1.136(a) for only TWO additional MONTHS.

The informalities requiring correction are indicated in the attachment(s). If the informality pertains to the abstract, specification (including claims) or drawings, the informality must be corrected with an amendment in compliance with 37 CFR 1.121 (or, if the application is a reissue application, 37 CFR 1.173). Such an amendment may be filed after payment of the issue fee if limited to correction of informalities noted herein. See Waiver of 37 CFR 1.312 for Documents Required by the Office of Patent Publication, 1280 Off. Gaz. Patent Office 918 (March 23, 2004). In addition, if the informality is not corrected until after payment of the issue fee, for purposes of 35 U.S.C. 154(b)(1)(iv), "all outstanding requirements" will be considered to have been satisfied when the informality has been corrected. A failure to respond within the above-identified time period will result in the application being ABANDONED.

See attachment(s).

A copy of this notice MUST be returned with the reply. Please address response to "Mail Stop Issue Fee, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450".

/Kam Sin/
Publication Branch
Office of Data Management
(571) 272-4200

Application No. 13288850

IDENTIFICATION OF APPLICATION DEFICIENCIES

☐ Applicant must provide legible text for the following item(s).

- ☐ Specification filed , page(s) .
- ☐ Claims filed , claim(s) .
- ☐ Oath/declaration filed .
- ☐ Other: .

☐ Applicant must provide missing information on the following page(s) of the specification by amending the specification to add the missing text. No new matter may be added.

☐ The specification refers to one or more applications by attorney docket number and does not show the U.S. application number(s). Applicant must supply the U.S. application number in place of each attorney docket number.

☐ Applicant must provide an Abstract of the Disclosure.

☐ Applicant has submitted a DECLARATION (37 CFR 1.63) FOR A UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76) (e.g., form PTO/SB/01A). The Application Data Sheet, however, is not present with the filed application. Applicant must submit an Application Data Sheet or file a new oath or declaration (e.g., PTO/SB/01) executed by the inventors and containing the information required in 37 CFR 1.63.

☐ Applicant must provide an executed declaration.

☐ Applicant must provide the missing page(s) of the oath/declaration or Application Data Sheet filed

☐ Applicant must provide a declaration signed by inventor(s) .

☐ The oath/declaration filed shows non-initialed and/or non-dated alterations. Applicant must file a new oath/declaration in compliance with 37 CFR 1.67(a).

☒ Applicant(s) Hyun Lee in the latest-filed oath/declaration or Application Data Sheet (ADS) did not show the inventor's residence at all, or did not show both a city and state in the U.S. inventor's residence, or did not show both a city and country in the non-U.S. inventor's residence. Applicant must supply an oath/declaration or Application Data Sheet (ADS) that shows each U.S. inventor's city and state of residence and each non-U.S. inventor's city and country of residence.

Electronic Acknowledgement Receipt

EFS ID:	19191286
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	03-JUN-2014
Filing Date:	03-NOV-2011
Time Stamp:	01:34:14
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	2014-06-02_Datasheet_NETL071A.pdf	1425270 38609c862bbdea43d7d6d94697a0951e5c2b07b7	no	4

Warnings:	
Information:	

2	Miscellaneous Incoming Letter	2014-05-30_Misc_Comm_NETL 071A.pdf	112660 982298cf26b79b7a8c0a1a43535f0dc34bf7f9f5	no	3

Warnings:	
Information:	
Total Files Size (in bytes):	1537930

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Applicant Information:

Applicant 1 Remove

Applicant Authority ◉ Inventor ○ Legal Representative under 35 U.S.C. 117 ○ Party of Interest under 35 U.S.C. 118

Prefix	Given Name	Middle Name	Family Name	Suffix
	Hyun		Lee	

Residence Information (Select One) ◉ US Residency ○ Non US Residency ○ Active US Military Service

City	Ladera Ranch	State/Province	CA	Country of Residence i	US

Citizenship under 37 CFR 1.41(b) i: US

Mailing Address of Applicant:

Address 1	21 Thalia St.		
Address 2			
City	Ladera Ranch	State/Province	CA
Postal Code	92694	Country i	US

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	79141		
Email Address	jz@jzpatent.com	Add Email	Remove Email

Application Information:

Title of the Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE		
Attorney Docket Number	NETL.071A	Small Entity Status Claimed	☒
Application Type	Nonprovisional		
Subject Matter	Utility		
Suggested Class (if any)		Sub Class (if any)	
Suggested Technology Center (if any)			
Total Number of Drawing Sheets (if any)	8	Suggested Figure for Publication (if any)	

PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Enter either Customer Number or complete the Representative Name section below. If both sections are completed the Customer Number will be used for the Representative Information during processing.

Please Select One:	◉ Customer Number	○ US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	79141		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78(a)(2) or CFR 1.78(a)(4), and need not otherwise be made part of the specification.

Prior Application Status	Expired		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
13288850	non provisional of	61409893	2010-11-03

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. Add

Foreign Priority Information:

This section allows for the applicant to claim benefit of foreign priority and to identify any prior foreign application for which priority is not claimed. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55(a).

Remove

Application Number	Country [i]	Parent Filing Date (YYYY-MM-DD)	Priority Claimed
			◉ Yes ○ No

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. Add

Assignee Information:

Providing this information in the application data sheet does not substitute for compliance with any requirement of part 3 of Title 37 of the CFR to have an assignment recorded in the Office.

Assignee 1 Remove

PTO/SB/14 (11-08)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	NETL.071A
	Application Number	
Title of Invention	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE	

If the Assignee is an Organization check here. ☒			
Organization Name	Netlist, Inc.		
Mailing Address Information:			
Address 1	175 Technology, Suite 150		
Address 2			
City	Irvine	**State/Province**	CA
Country [i]	US	Postal Code	92618
Phone Number		Fax Number	
Email Address			
Additional Assignee Data may be generated within this form by selecting the **Add** button.		Add	

Signature:

A signature of the applicant or representative is required in accordance with 37 CFR 1.33 and 10.18. Please see 37 CFR 1.4(d) for the form of the signature.

Signature	/Jamie J. Zheng/			Date (YYYY-MM-DD)	2014-06-02
First Name	Jamie	Last Name	Zheng	Registration Number	51167

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
13/288,850	11/03/2011	2827	2470	NETL.071A	34	4

CONFIRMATION NO. 2466

CORRECTED FILING RECEIPT

79141
The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

OC00000068828135

Date Mailed: 06/05/2014

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Hyun Lee, Ladera Ranch, CA;

Applicant(s)

Hyun Lee, Ladera Ranch, CA;

Assignment For Published Patent Application

NETLIST, INC., Irvine, CA

Power of Attorney: The patent practitioners associated with Customer Number 79141

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/409,893 11/03/2010

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

If Required, Foreign Filing License Granted: 11/14/2011

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 13/288,850**

Projected Publication Date: Not Applicable

Non-Publication Request: No

Early Publication Request: No

Petitioners
Ex. 1002, p. 502

Title

METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE

Preliminary Class

365

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications:

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

Petitioners
Ex. 1002, p. 503

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

Customer No. 79141

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	11/03/2011	Hyun Lee	NETL.071A	2466

TITLE OF INVENTION:

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	06/26/2014

EXAMINER	ART UNIT	CLASS-SUBCLASS

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list

(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,

(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 JAMIE J. ZHENG, Ph.D, Esq.
2
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Netlist, Inc.

(B) RESIDENCE: (CITY and STATE OR COUNTRY): Irvine, CA

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ■ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:

☑ Issue Fee

☐ Publication Fee (No small entity discount permitted)

☐ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**

☐ A check is enclosed.

☑ Payment by credit card. Form PTO-2038 is attached.

☐ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number 50-5863 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)

☑ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.

☐ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature /Jamie J. Zheng/ Date June 9, 2014

Typed or printed name Jamie J. Zheng Registration No. 51,167

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Electronic Patent Application Fee Transmittal

Application Number:	13288850
Filing Date:	03-Nov-2011
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Filer:	Jamie Jie Zheng
Attorney Docket Number:	NETL.071A

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Utility Appl Issue Fee	2501	1	480	480
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**480**

Electronic Acknowledgement Receipt

EFS ID:	19253466
Application Number:	13288850
International Application Number:	
Confirmation Number:	2466
Title of Invention:	METHOD AND APPARATUS FOR OPTIMIZING DRIVER LOAD IN A MEMORY PACKAGE
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NETL.071A
Receipt Date:	09-JUN-2014
Filing Date:	03-NOV-2011
Time Stamp:	18:51:55
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$480
RAM confirmation Number	5677
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1	Issue Fee Payment (PTO-85B)	2014-06-09_Issue_Fee_Transmittal_NETL071A.pdf	74147 2e517c33b694cddba51ef32d2c663465d1a436f8f	no	2
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	30189 0a2f95766f7d423b949c8301f272c7fb22ea4876	no	2
Warnings:					
Information:					
			Total Files Size (in bytes):	104336	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	13/288850
	Filing Date	November 3, 2011
	First Named Inventor	Hyun Lee
	Art Unit	2827
(Multiple sheets used when necessary)	Examiner	Amir Zarabian
SHEET 2 OF 4	Attorney Docket No.	NETL.071A

U.S. PATENT DOCUMENTS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	29	8,001,434 B1	08-16-2011	Lee et al	
	30	8,019,589 B2	09-13-2011	Rajan et al.	
	31	8,033,836 B1	10-11-2011	Bhakta et al.	
	32	2006/0233012 A1	10-19-2006	Sekiguchi et al.	
	33	2007/0096332 A1	05-03-2007	Satoh et al.	
	34	2008/0025123 A1	01-31-2008	Rajan et al.	
	35	2008/0094808 A1	~~01-24-2008~~	Kanapathippillai et al.	April 24, 2008
Change(s) applied to document, /R.T.H./ 6/12/2014	**36**	~~2008/0025134 A1~~	~~01-31-2008~~	~~Rajan et al.~~	
	37	2008/0253085 A1	10-16-2008	Soffer	
	38	2008/0296779 A1	12-04-2008	Matsui et al.	
	39	2009/0070727 A1	03-12-2009	Solomon	
	40	2009/0103345 A1	04-23-2009	McLaren et al.	
	41	2009/0290442 A1	11-26-2009	Rajan	
	42	2010/0020583 A1	01-28-2010	Kang et al.	
	43	2010/0090338 A1	04-15-2010	Lee et al.	
	44	2010/0174858 A1	07-08-2010	Chen et al.	
	45	2011/0006360 A1	01-13-2011	Ikebuchi	
	46	2011/0016250 A1	01-20-2011	Lee et al.	
	47	2011/0016269 A1	01-20-2011	Lee et al.	
	48	2011/0050320 A1	03-03-2011	Gillingham	
	49	2011/0108888 A1	05-12-2011	Or-Bach et al.	
	50	2011/0125982 A1	05-26-2011	Choi et al.	
	51	2011/0156232 A1	06-30-2011	Youn et al.	
	52	2011/0169171 A1	07-14-2011	Marcoux	

Examiner Signature	Date Considered

*__Examiner:__ Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

__T__[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.H./

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/288,850	07/22/2014	8787060	NETL.071A	2466

79141 7590 07/02/2014

The Law Office of Jamie Zheng, Ph.D Esq.
P.O. Box 60573
Palo Alto, CA 94306

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 103 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Hyun Lee, Ladera Ranch, CA;

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY.DOCKET NO./TITLE	REQUEST ID
13/288,850	11/03/2011	Hyun Lee	N3044.10015US02	55680

Acknowledgement of Loss of Entitlement to Entity Status Discount

The entity status change request below filed through Private PAIR on 01/22/2018 has been accepted.

CERTIFICATIONS:

Change of Entity Status:

X Applicant changing to regular undiscounted fee status.

NOTE: Checking this box will be taken to be notification of loss of entitlement to small or micro entity status, as applicable.

This portion must be completed by the signatory or signatories making the entity status change in accordance with 37 CFR 1.4(d)(4).

Signature:	/Jamie J. Zheng/
Name:	Jamie Zheng
Registration Number:	51167