PTO/SB/14 (07-07)
Approved for use through 06/30/2010. 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.096PR
	Application Number	
Title of Invention	HIGH DENSITY DIMMS	

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
	Hyun		Lee	

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

City	Ladera Ranch	State/Province	CA	Country of Residence	US

Citizenship under 37 CFR 1.41(b) US

Mailing Address of Applicant:

Address 1	21 Thalia Street
Address 2	

City	Ladera Ranch	State/Province	CA
Postal Code	92694	Country	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	20995		
Email Address	efiling@kmob.com	Add Email	Remove Email

Application Information:

Title of the Invention	HIGH DENSITY DIMMS		
Attorney Docket Number	NETL.096PR	Small Entity Status Claimed	☐
Application Type	Provisional		
Subject Matter	Utility		
Suggested Class (if any)		Sub Class (if any)	
Suggested Technology Center (if any)			
Total Number of Drawing Sheets (if any)	10	Suggested Figure for Publication (if any)	

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.096PR
	Application Number	
Title of Invention	HIGH DENSITY DIMMS	

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			Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)

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[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.

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

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.096PR
	Application Number	
Title of Invention	HIGH DENSITY DIMMS	

If the Assignee is an Organization check here. ☐				
Prefix	**Given Name**	Middle Name	**Family Name**	Suffix
Mailing Address Information:				
Address 1				
Address 2				
City		**State/Province**		
Country		Postal Code		
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)	2011-07-28
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.**

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.

HIGH DENSITY DIMMS

BACKGROUND

[0001] As technology advances and the usage of tablet notebook computer increases, more data needs to be transferred among the data centers and to/from the end users. In many cases, the data centers are built by clustering many servers that are networked to increase the performance.

[0002] Although there are many types of networking that are based on the applications, the basic concept is to increase the performance by dynamically allocating the computing resources. In recent years, the server technology has been evolved to serve many specific applications such as 'finance transaction' (e.g., POS, inter-bank transaction, stock market transaction), 'scientific computation' (e.g., fluid dynamic for automobile and ship design, weather prediction, oil and gas expeditions), 'medical diagnostics' (e.g., diagnostics based on the fuzzy logic, medical data processing), 'simple information sharing and searching' (e.g., web search, retail store website, company home page), 'email' (information distribution and archive), 'security service', 'entertainment' (e.g., video-on-demand), etc. However, all these applications share the same information transfer bottleneck, which resides in the CPU efficiently being able to transfer the data. This is because all data transfers are going though the CPU I/O channels.

[0003] The data transfer limitation by the CPU also applies to the data transfer from/to either HD or SSD to/from the memory subsystem that is connected to the FSB (Front Side Bus).

[0004] There have been various configurations to increase the data transfer from/to the main storage, such as HD or DDS, to the local DRAMs for the faster data throughput rate. Especially, the EMM that is supported with EcoRAM (Spansion) (see, Appendix B) presents an idea of constructing a fast SSD that assumes the physical form factor of a DIMM. Spansion populates the EcoRAM with a large quantity of Flash memories and a limited amount of DRAMs that serves as a data buffer. This arrangement delivers a much higher throughput rate than standard SSD since this EcoRAM is connected to the CPU via the HT interface while SSD is connected via SATA or USB. The read random access

Petitioners
Ex. 1008, p. 5

throughput rate of EcoRAM is near 3GB/s compared with the standard PCIe-based NAND SSD which is about 400MB/s. This gives a 7.5X improvement. The write random access throughput rate of EcoRAM over the standard PCIe-based NAND SSD is less than 2X (197MBs vs. 104MBs) due to the fact that the write speed is predominated by the Flash write access time. The idea of EcoRAM is to build a SSD with the form factor of a standard DIMM such that it can be connected to the Front Side bus. However, due to the interface protocol difference between DRAM and Flash, Spansion developed a new interface device (Spansion EcoRAM Accelerator) that occupies one of the CPU sockets. Spansion EcoRAM is well suited for the "Read Access Intensive" applications, such as 'medical diagnostics, simple information sharing and searching, and entertainment. However, it does not fit to serve the applications that require extensive data manipulations, such as 'scientific computation', 'security service' and 'finance transactions'.

[0005] As it has been stated in the previous paragraphs, the two deficiencies of EcoRAM are: 1) it requires an interface device that occupies a CPU socket, thus less CPU, and in turn it reduces the overall computation efficiency; and 2) the high disparity in the throughput rate between read and write operations limits the utilization of this solution.

[0006] In addition to these deficiencies, the throughput rate of EcoRAM is much lower than DRAM, such that the read access throughput rate is only 25% of the DRAM read access rate; the write access throughput rate is less than 2% of the DRAM access rate.

SUMMARY

[0007] This application presents certain embodiments that overcome the shortcomings of the prior art: the significantly slower access time compare to the DRAM, and the high parity in the throughput rate between read and write throughput rate (15 to 1 ratio).

[0008] This application discloses certain embodiments of a Flash-DRAM-hybrid DIMM with integrated address spaces of Flash and DRAM. The application also presents a method of controlling such devices. In certain embodiments, the device is able to configure the actual memory density of the DIMM and/or the ratio of DRAM memory to Flash memory for optimal use with a particular application (e.g., POS, inter-bank transaction, stock market transaction, 'scientific computation,' such as fluid dynamics for automobile and ship design, weather prediction, oil and gas expeditions, 'medical diagnostics,' such as diagnostics based

on the fuzzy logic, medical data processing, 'simple information sharing and searching,' such as web search, retail store website, company home page, 'email' or information distribution and archive, 'security service,' 'entertainment,' such as video-on-demand).

[0009] In certain embodiments, the device contains a high density Flash memory with a low density DRAM, where DRAM is used as a data buffer for read/write operation. The Flash serves as the main memory.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Figure 1 shows the data path in a conventional memory arrangement.

[0011] Figure 2 shows the general architecture of existing non-volatile memory module products. A part of SSD/HD is migrated into the Front Side Bus (FSB) to support the RAID (*Redundant Array of Independent Disks*) operation.

[0012] Figure 3 shows a currently available concept of connecting a Flash DIMM to the FSB.

[0013] Figure 4 shows the general architecture of a currently available product that uses the Flash DIMM with a DIMM controller.

[0014] Figure 5 shows an example memory module that contains DRAMs, Flash, and Central DIMM Controller (CDC), Data Manager (DMgr) in accordance with certain embodiments disclosed herein. The CDC provides two main functions: the standard RDIMM register function, and the data transfer management function between DRAM and Flash.

[0015] Figure 6 shows an example connection of the CDC to the register and DMgr, to MCH, DRAM, and Flash in accordance with certain embodiments described herein.

[0016] Figure 7 shows an example of the functional blocks in the CDC in accordance with certain embodiments described herein.

[0017] Figure 8A shows the general concept of an existing Flash-DRAM hybrid DIMM.

[0018] Figure 8B shows example functional blocks in a DRAM controller that is in the CDC in accordance with certain embodiments described herein.

[0019] Figure 9 shows an example state diagram of data transfer between the Flash memory and the DRAM memory in accordance with certain embodiments described herein.

[0020] Figure 10 shows an example of how the MCH addresses the Flash area in accordance with certain embodiments described herein.

[0021] Figure 11 is a table that shows example amounts of DRAMs that are used to support different applications in accordance with certain embodiments described herein.

[0022] Figures 12A-12C show a conventional NVDIMM, a conventional FDHDIMM, and an example FDHDIMM compatible with certain embodiments described herein, respectively, to illustrate some differences between the three modules.

BRIEF DESCRIPTION OF THE APPENDICES

[0023] Appendix A: "HyperVault: Multi-Terra Bytes of FSB Memory for HPC," description regarding example device in accordance with certain embodiments described herein.

[0024] Appendix B: "Spansion® EcoRAM™ Solution: IOZone Solution Performance White Paper," Version 0.4, description of currently available product.

DETAILED DESCRIPTION

[0025] A Flash and DRAM hybrid DIMM (FDHDIMM) that interfaces with MCH operates in two modes; as a standard DDR DIMM with SSD, or as a high density DIMM.

[0026] As a standard DDR DIMM with SSD, the FDHDIMM supports two views to the MCH. First, the MCH views the FDHDIMM as a combination of DRAM DIMM and SSD. In this case the MCH needs to manage two address spaces, one for the DRAMs and one for the Flash. The advantage of this mode is that the CPU does not need to be in the data path when data is moved from Flash to DRAM or vice versa. Second, the MCH views the FDHDIMM as an On-DIMM Flash with the SSD in an extended memory space that is behind the DRAM space.

[0027] Thus, for the normal standard DDR DIMM with SSD operation, the MCH physically fetches data from the SSD to the DDR. This action will be much quicker than the

data going through the CPU, which happens when the current data move from SSD to DIMMs occurs, since all data movement occurs on a DIMM.

[0028] In the high density DIMM mode, the MCH recognizes the physical density of DIMM, which is the density of Flash alone. The DRAM serves as a data buffer that buffers read data, and serves as a temporary storage for the write date. The MCH has two options to execute a read operation.

[0029] The first read option is to send an active command (actually RAS) for the FDHDIMM to pre-fetch the read data from the Flash to the DRAMs, where the pre-fetch size is a page (1KB or 2KB, optional). This option will allow the data to transfer from the Flash to DRAMs, and provides the data latency and the throughput rate that is the same as any DRAM operation as long as the operations are done onto the pages that are opened with the pre-active command. However, this option requires much a longer separation period between the RAS and the first CAS command.

[0030] The other read option is executing the RAS and CAS (active and read/write) commands just as the standard DRAM. However, this option requires the initial Flash read time after the first CAS command (i.e. extensive CAS latency). This option supports the same throughput rate as the standard DRAM DIMM as long as the accesses occur from the open pages.

[0031] Certain embodiments described herein overcome the shortcomings of having the long separation period between the RAS and the first CAS command, and of having the extensive CAS latency.

[0032] In Figure 1, the SSD/HD and DRAM DIMM of a conventional memory arrangement are connected to CPU via separated memory control ports. This figure shows the data flow path between the server main memory (SSD or HD) to the DRAM DIMMs. Since the SSD/HD data I/O and the DRAM DIMM data I/O are controlled by the CPU, the CPU needs to allocate its process cycles to control these I/Os, which includes the IRQ (Interrupt) service that the CPU performs periodically. Furthermore, since all the data traffic goes through the CPU, the data move from/to SSD/HD to/from the DRAM DIMM occupies the CPU internal data bus, thus further hindering the overall CPU performance.

[0033] Figure 2 shows the general architecture of a non-volatile memory module that is currently being exercised. Non-Volatile DIMM (NVDIMM) is a Flash-DRAM hybrid DIMM (FDHDIMM), which uses a backup power (either battery or charged capacitor) to copy the data in the DRAM memory into the Flash memory when power loss occurs while the system in operation. In this architecture, a part of SSD/HD is migrated into the FSB in the form of Flash. However, the density of the Flash is rather low since the Flash memory size does not have to be any larger than the DRAM memory size. The purpose of this type of architecture is to provide a non-volatile storage that is connected to the FSB to support the RAID (*Redundant Array of Independent Disks*) type of operation.

[0034] Figure 3 shows the system level architecture of a currently available product, which includes a Flash DIMM controller chip and Flash DIMM. In this figure, the Flash DIMM controller chip occupies a CPU socket and communicates with the CPU via HT (Hyper Transport). Although this product utilizes the HT interface, the actual interconnection between the controller chip and CPU is not limited to HT. Although this architecture does not address the shortcomings that associate with the traditional way (Figure 1) of connecting the SSD/HD and DRAM DIMMs, it allows increase of the density of non-volatile memory space that is connected to the virtual FSB since the CPU views the Flash DIMM controller chip as another processor with a large density of memory on its FSB.

[0035] There are two shortcomings with this architecture. First, it occupies one or two CPU sockets. Thus, it reduces the utilization of the server Mother Board (MB). Second, the Flash access speed is limited by 4 items: 1) the read/write speed of the Flash memory, 2) the FSB bus size, 3) the Flash DIMM controller chip latency, and 4) the HT interconnect efficiency, which is dependent on the HT interface controller in the CPU and Flash DIMM controller chip.

[0036] The published results indicate that these shortcomings are evident in that the maximum throughput rate is 1.56 GBs for the read operation and 104 MBs for the write operation. These access rates are 25% of the DRAM read access speed, and 1.7% of the DRAM access speed at 400MHz operation. The parity in the access speed (15 to 1) between the read operation and write operation clearly shows that this architecture is limited by the read/write speed of the Flash memory. The discrepancy of the access speed between this type

of architecture and JEDEC standard DIMM grows wider as the memory technology advances, since as the memory technology advances, the Flash access rate stays relatively constant while the DRAM access speed increases.

[0037] Figure 4 shows the general architecture of the currently available FDHDIMM product, which receives two sets of control signals from the MCH/CPU while a single data bus interfaces with the MCH. In this architecture, since the MCH/CPU does not have a view to the data path between the DRAM memory and the Flash memory, it requires extra complication in designing the MCH.

[0038] Figure 5 shows the basic structure of an example memory module in accordance with certain embodiments described herein.

[0039] One major difference from the prior art is that the example memory module contains two on-DIMM controllers: the Central DIMM Controller (CDC) and the Data Manager (DMgr) to manage the interface between Flash, DRAM, and the MCH. This architecture addresses all the shortcomings that are associated with the prior art, such as the MCH views of the example memory module as a very high density DRAM DIMM.

[0040] The CDC controls the read/write access to/from the Flash memory to the DRAM memory, and DRAM memory to MCH/CPU. Since there is no direct access from the MCH/CPU to the Flash memory, the Flash access speed has very minimal impact to the overall access speed. The CDC performs the memory address translation/mapping and address domain conversion, Flash access control, error correction, data width modulation between the Flash memory and DRAM, etc. The CDC ensures that the example memory module operates with the industry standard DDR, DDR2, DDR3, DDR4 protocols.

[0041] The DMgr controls the data flow rate, the data transfer size, data buffer size, error monitoring and correction.

[0042] Figure 6 shows an example interoperability among the components of certain embodiments described herein. After the initialization, the CDC receives a command from CPU to prepare for a certain memory space, which is configurable or programmable by the CPU. The CPU can partition or parse the memory space in various ways that are optimum for the particular application that it needs to execute. The CDC maps the actual physical DRAM and Flash memory space to support the CPU view of the memory space. The CDC

also manages to support any booting operation by copying the necessary information from the Flash memory to the DRAM memory during system power up. This activity may utilize the CDC to initiate a data copying operation from the Flash memory to the DRAM memory just as the CPU copies the boot information from the SSD/HD to the DRAM memory. The DMgr controls the data traffic between the Flash memory and the DRAM memory. The DMgr receives the data transfer size, formatting information (540), direction of data flow (511, 512, 521, 522), and the starting time of the actual data transfer from CDC. The DMgr also receives the data flow path from the CDC to correctly establish the data transfer fabric. The DMgr also functions as a bi-directional data transfer fabric such that it has more than 2 sets of data ports facing the Flash and the DRAM memory. The multiplexors 511 and 512 provide the data paths from the DRAMs (DRAM-A and DRAM-B) to the MCH and the Flash memory, the multiplexors 521 and 522 provide the data paths from the MCH and the Flash memory to the DRAMs (e.g., DRAM-A and DRAM-B).

[0043] Two sets of multiplexors (511, 512, 521, 522) allow independent data transfer to the Flash memory from the DRAM memory. For example, the DMgr transfers data from DRAM-A to the MCH (via 511) at the same time as DRAM-B to the Flash memory (via 512), or it transfers data from DRAM-B to the MCH (via 511) at the same time as the Flash memory to DRAM-A (via 521). This special arrangement of data transfer connect also allows direct data transfer from the MCH to the Flash memory (e.g., via 521 and 512), and allows direct data transfer from the Flash memory to the MCH (e.g., via 521 and 511).

[0044] The Data Format (540) also contains a data buffer to support a wide data bus facing the Flash memory, and to match the data flow between the DRAM memory and the Flash memory.

[0045] Figure 7 shows example functional blocks of the CDC in accordance with certain embodiments described herein. The CDC has four major functional blocks: DRAM control block (DRAMCtrl), Flash control block (FlashCtrl), MCH/CPU command interpreter (CmdInt), DRAM-Flash interface scheduler (Scheduler), and the DMgr control block (DMgrCtrl).

[0046] DRAMCtrl (710) generates DRAM commands that are independent from the DRAM commands from the MCH. However, if the MCH initiates a read/write operation from/to the same DRAM that is currently executing a command from the DRAMCtrl (710), the DRAMCtrl (710) aborts the operation. The FlashCtrl (720) generates Flash commands for the proper read/write operations. The CmdInt (730) constantly intercepts the MCH command and provides the correct interpretation. For example, CmdInt issues an interrupt signal to the DRAMCtrl (710) if the MCH issues a command that collides (conflicts) with the currently executing or pending commands that DRAMCtrl initiated. The Scheduler (740) schedules the Flash –DRAM interface operation such that there is no resource conflict in the DMgr. The Scheduler assigns the time slots for the DRAMCtrl, FlashCtrl operation based on the current status and the command from the MCH. The DMgrCtrl (750) controls the data paths between the DRAM memory and the Flash memory, between the DRAM memory and the MCH, and between the MCH and the Flash memory.

[0047] The memory space mapping in Figure 8A for an existing Flash-DRAM hybrid DIMM shows two separate and independent address spaces, one for the Flash memory and the other for the DRAM memory. The CPU treats these two address spaces no differently than the conventional address spaces of DRAM and SSD/HD. The memory in each address space is controlled by the conventional method of controlling the DRAM and SSD/HD. However, the on-DIMM data path between Flash and DRAM allows direct data transfer to occur between the Flash memory and the DRAM memory. This method of transferring data bypasses the conventional method of transferring data from SSD/HD to the FSB memory via the CPU, thus reducing the data transfer time. Since this method also removes the CPU from the data path, it allows the over efficiency of the CPU usage, and thus improves the overall system performance. However, the shortcoming of this method are that: 1) the MCH needs to manage two address spaces, 2) the MCH needs to manage two different memory protocols simultaneously, 3) the MCH needs to map the DRAM memory space into the Flash memory space, 4) the data interface time cannot be controlled due to the difference in the data access time between the Flash memory and the DRAM memory, and 5) the MCH is completely blind about the data transfer activity between the Flash memory and the DRAM memory, and

as the result, controlling the data transfer between the Flash memory and the DRAM memory is practically impossible.

[0048] The memory space mapping in Figure 8B for a memory module in accordance with certain embodiments described herein shows an overlapping memory address space between the DRAM memory and the Flash memory. In this case, the MCH does not see any difference between the DRAM memory space and the Flash memory space. The MCH sees the large Flash memory space as the DRAM memory space, and accesses this memory space with the standard DDR protocol. The DRAM memory space holds the copy of data from the selected Flash memory space such that the MCH can access these data via FSB.

[0049] Figure 9 shows an example transfer of data from the Flash memory to the DRAM and vice versa in accordance with certain embodiments described herein. The data transfer from the Flash memory to the DRAM memory occurs when the CDC receives a command to move data from the Flash memory to the DRAM memory. The data transfer from the DRAM memory to the Flash memory occurs when all the pages in a particular block are closed. The CDC partitions the DRAM memory space into number of blocks for the purpose of supporting the optimum operation. The block size is configurable by the CPU, such that the number pages in a block can be optimized to support a particular application or a task. For example, when CPU concurrently executes more than one task, the CPU can configure a different block size for each task based on the DRAM space it needs for the particular task.

[0050] The MCH can access the memory module of certain embodiments described herein using the standard FSB DIMM access protocol (e.g., JEDEC). Since any page (or block) that does not reside in the DRAM memory has to be copied from the Flash memory, and the copying process takes time in the order of tens of microseconds, the MCH initiates the copying of the page ahead by writing into two specific RC registers along with the valid block address. When CDC receives these particular RC write commands from the MCH, it recognizes it as the block address that it needs to prepare, and it will copy the block to the DRAM memory. Once the block is in the DRAM memory space, the MCH can access the data with the standard DRAM DIMM operation protocol.

[0051] In Figure 9, after power up (910) and initialization (920), CPU/MCH initiates the data move from the Flash memory to the DRAM memory by writing to the control register in CDC (930), which copies a block of data from the Flash to the DRAM and erases the block in the Flash (840). Once the block of data is moved to the DRAM, the MCH can operate on this data block with the standard RDIMM access protocol (950). The CDC checks if any of the copied blocks are closed (960). If there are any closed blocks, the CDC writes the updated block from the DRAM memory to the Flash memory (970). If the CPU/MCH reopens (980) the closed block that is currently being written back into the Flash memory, the CDC stops the Flash write and erases the block being written (955). Otherwise, the CDC finishes the Flash write (990).

[0052] The dotted activity in Figure 9 is an independent/parallel activity of the CDC. When the CDC, at any time, receives the DRAM load command (935), it starts the Load DRAM w/RC (930), and initiates another branch of activity that includes 940 – 990. As it is shown in this figure, the CDC controls the DRAM memory to the Flash memory operation, such that the Flash is completely hidden from the MCH. The MCH accesses the memory module using the standard DDR protocol.

[0053] Figure 10 shows the definition of the block address field which includes the address bits that are not used by the JEDEC standard (A0 – A4) in accordance with certain embodiments described herein. In this figure, two sets of address bits AD6 to AD17 are allocated for the Block address.

[0054] For example, with the minimum block size of 256K bytes, 24 bit block address (two sets of AD6 to AD17) allows access to 4TB of Flash memory space. If the memory module has 1GB of DRAMs, it can hold 4K blocks of data in the DRAM memory.

[0055] The 4K blocks can be assigned to different virtual ranks and banks, where the number of virtual ranks and banks can be configured. For example if the memory module is configured to be 2 ranks with 8 banks/rank, then each bank holds 250 blocks or 62K 1KB pages.

[0056] The time to copy a block of data from the Flash memory to the DRAM memory is around 400μs with 8 DMgr. Figure 10 shows the DMgr configuration to support this block of data transfer in accordance with certain embodiments described herein.

[0057] The time to copy a block of data from the DRAM memory to the Flash memory is about 22ms. The CDC manages the block write operation by queuing the blocks that are ready to be written back to the Flash memory. If any page in the block is reopened, then the CDC will stop the block write-back operation until all the pages in the block are closed, and then restart the write-back operation.

[0058] Since the write-back operation takes 22ms while any copying operation takes 400μs, to keep the read and write access time to be similar, the read to write ratio is 55 to 1. This means that, if the block usage time (i.e. the duration that block is being open) in the DRAM memory is average 22ms, then the block write-back operation is completely hidden from the MCH. If the block usage time is 11ms, then the CDC makes sure that, on average over 100ms, there are no more than 9 closed blocks in the queue to be written-back to the Flash memory. Otherwise, the memory module cannot support the standard DDR operation. Thus, the number of closed blocks in the queue to be written-back to the Flash memory varies with the average block usage time.

[0059] Max number of closed blocks to be written-back to Flash = ((#of blocks per bank)/(ratio of 'flash_block_write_time' to 'flash_read_time'))*((Block usage time)/('flash_block_write_time'))

[0060] This equation indicates that if the block usage time is the same as the 'flash_block_write_time', then each bank should not have more than 4 closed blocks, on average over 88ms (22ms * 4) that need to be written back to the Flash.

# of block per bank = 250 (for 1GB DRAM memory)
The ratio of 'flash_block_write_time' to 'flash_read_time' = 55
Floor [250/55] = 4.

[0061] The above equation also indicates that the bigger DRAM memory space can support shorter block usage times. For example, 2GB of DRAM memory allows the 8 closed blocks to be written-back to Flash.

[0062] Table 1 shows the maximum allowed closed blocks in the queue to be written back to the Flash memory.

[0063] Figure 12A schematically illustrates a conventional NVDIMM having a switch that connects to the MCH for normal operation. When power loss occurs, the switch

connects to the Flash for saving the DRAM contents to the Flash. When power restoration occurs, the switch connects to the Flash for retrieving the DRAM contents from the Flash.

[0064] Figure 12B schematically illustrates a conventional FDHDIMM. The MCH interfaces with DRAM. The interface is controlled by the FDHDIMM controller. Although the MCH interfaces with the DRAM, the DRAM access time is governed by the Flash access time. The DRAM is only used as a data buffer: data from the Flash memory is transferred to the DRAM at the Flash access speed, and buffered (collected) into the DRAM, which transfers the buffered data to the MCH at the rate of DRAM I/O speed. Since the Flash memory access speed is very slow (a few hundred microseconds), the average data throughput rate is limited by the Flash access speed.

[0065] Figure 12C schematically illustrates an example embodiment of a FDHDIMM in accordance with certain embodiments described herein. The CDC controller of this example embodiment receives the standard DDR commands from the MCH, and interprets, and produces commands to control the Flash memory and the DRAM memory. The CDC also produces a set of switch control signals for the data path router amongst DRAMs, Flash and MCH. The switch control signals are independently operated without any exclusivity.

[0066] Various embodiments of the present invention have been described above. Although this invention has been described with reference to these specific embodiments, the descriptions are intended to be illustrative of the invention and are not intended to be limiting. Various modifications and applications may occur to those skilled in the art without departing from the true spirit and scope of the invention.

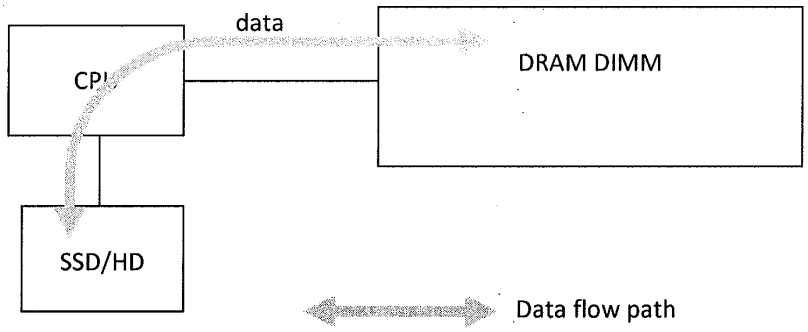


Figure 1: Conventional memory arrangement

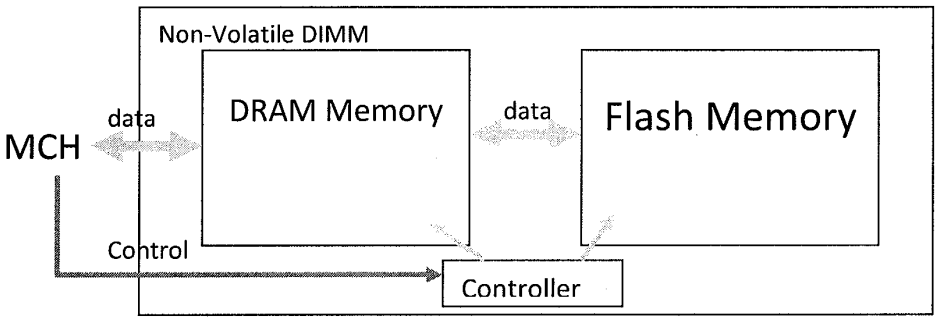
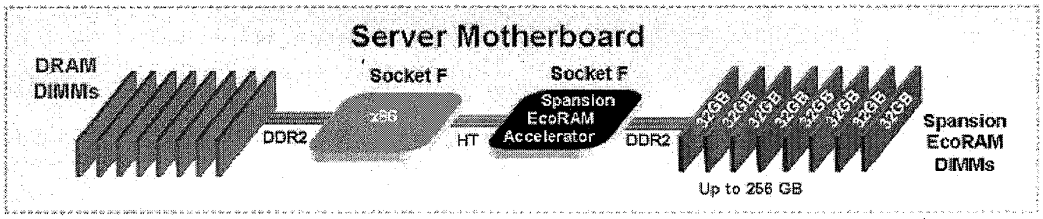


Figure 2: General Architecture of Non-Volatile DIMM (NVDIMM)
(density of DRAM memory = density of Flash memory)

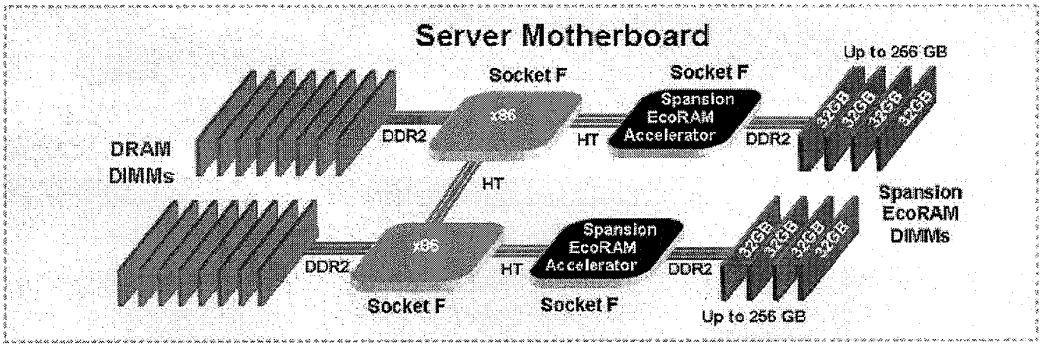
Spansion EcoRAM Configurations

256GB Spansion EcoRAM Solution – Single Accelerator



256GB Single Accelerator Spansion EcoRAM Solution

256 GB Spanison EcoRAM Solution – Dual Accelerator



256GB Dual Accelerator Spansion EcoRAM Solution

Figure 3: system level architecture of a currently available Flash DIMM product

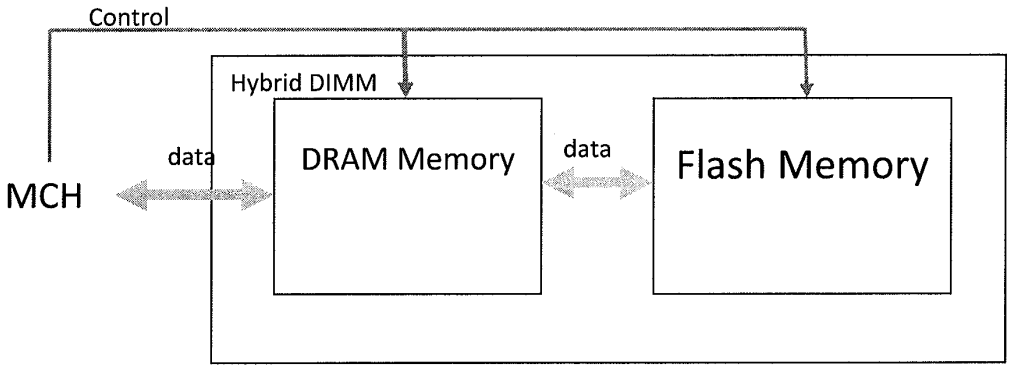


Figure 4: General Architecture of current Flash-DRAM Hybrid DIMM
(density of DRAM memory << density of Flash memory)

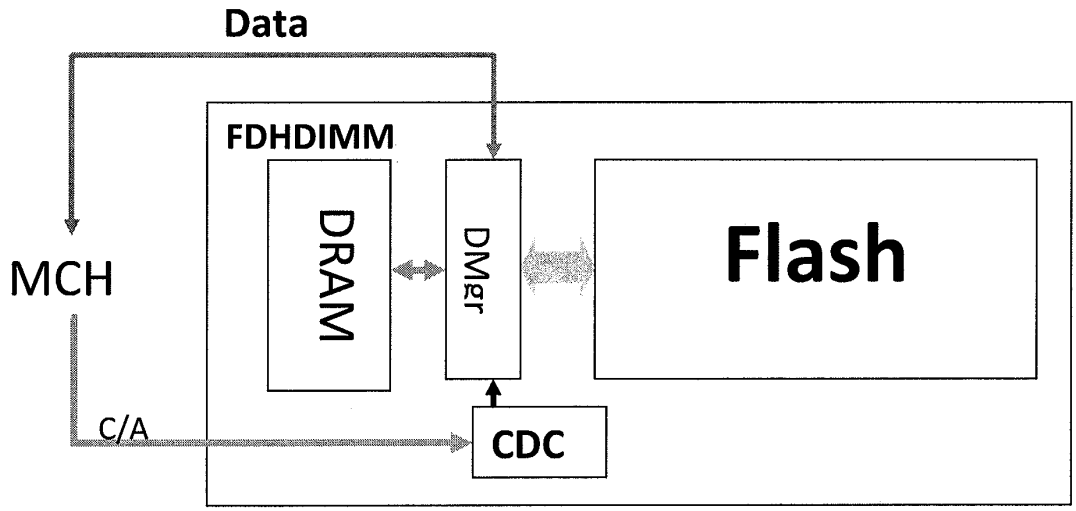


Figure 5: Embodiment that contains DRAMs, Flash, and Central DIMM Controller (CDC), Data Manager (DMgr)

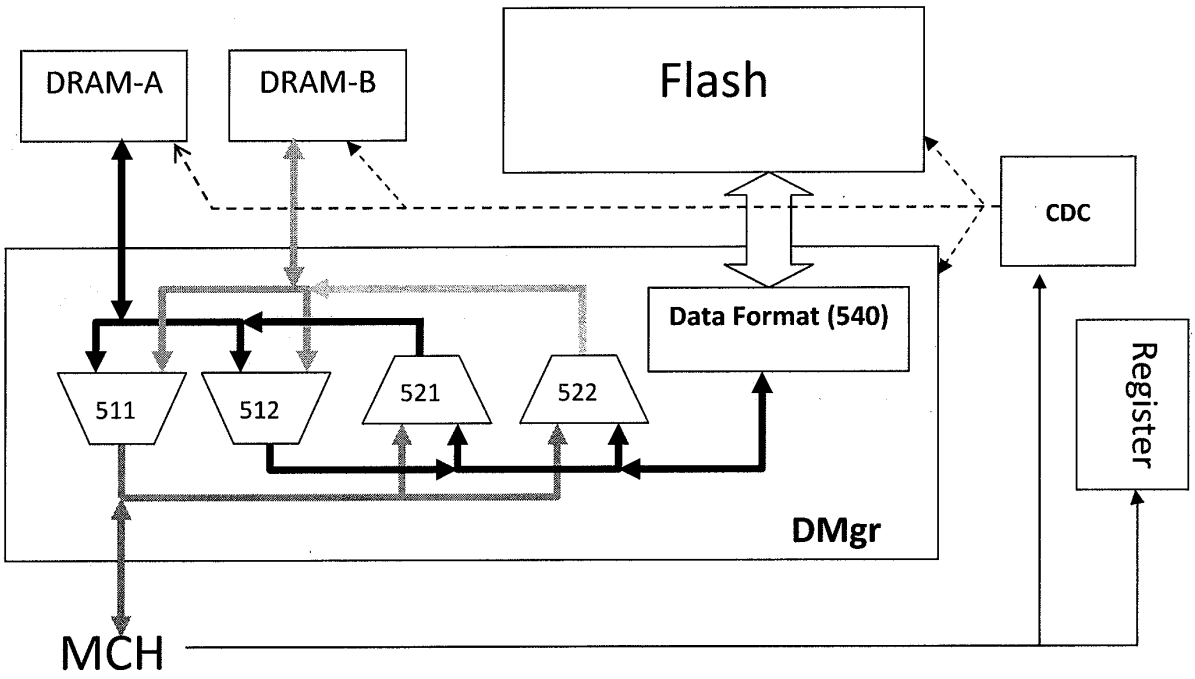


Figure 6: Connection of CDC to the register and DMgr, to MCH, DRAM, and Flash

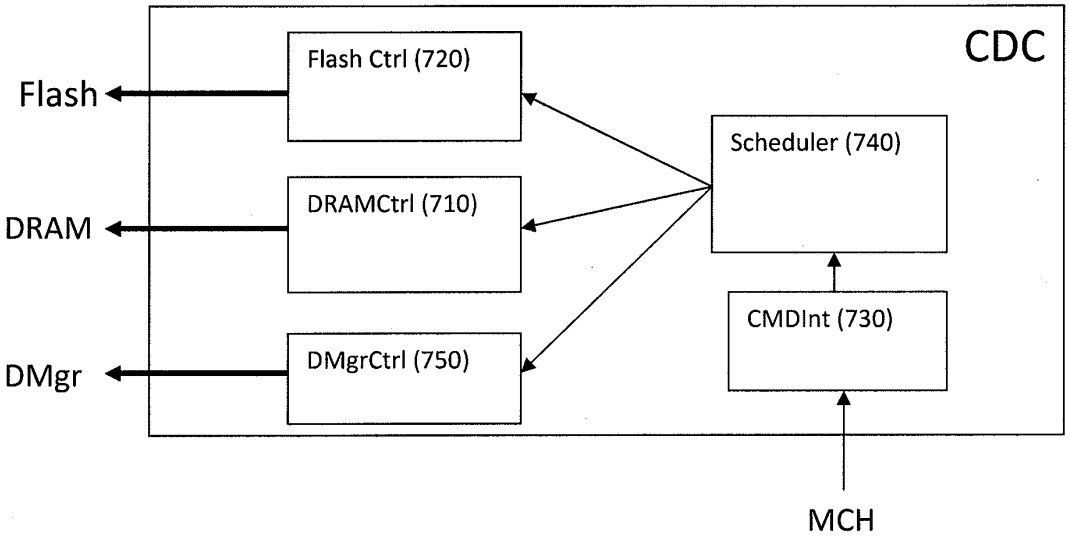


Figure 7: Functional blocks in CDC

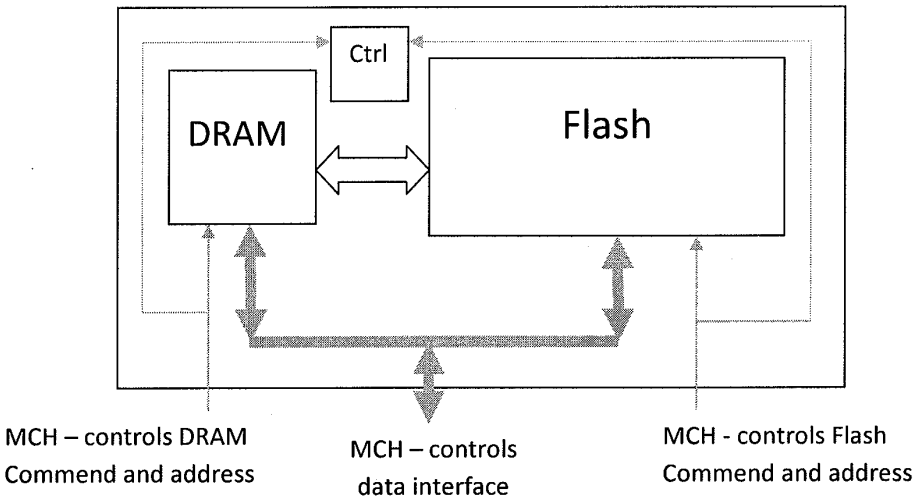


Figure 8A: General concept of the current Flash-DRAM Hybrid

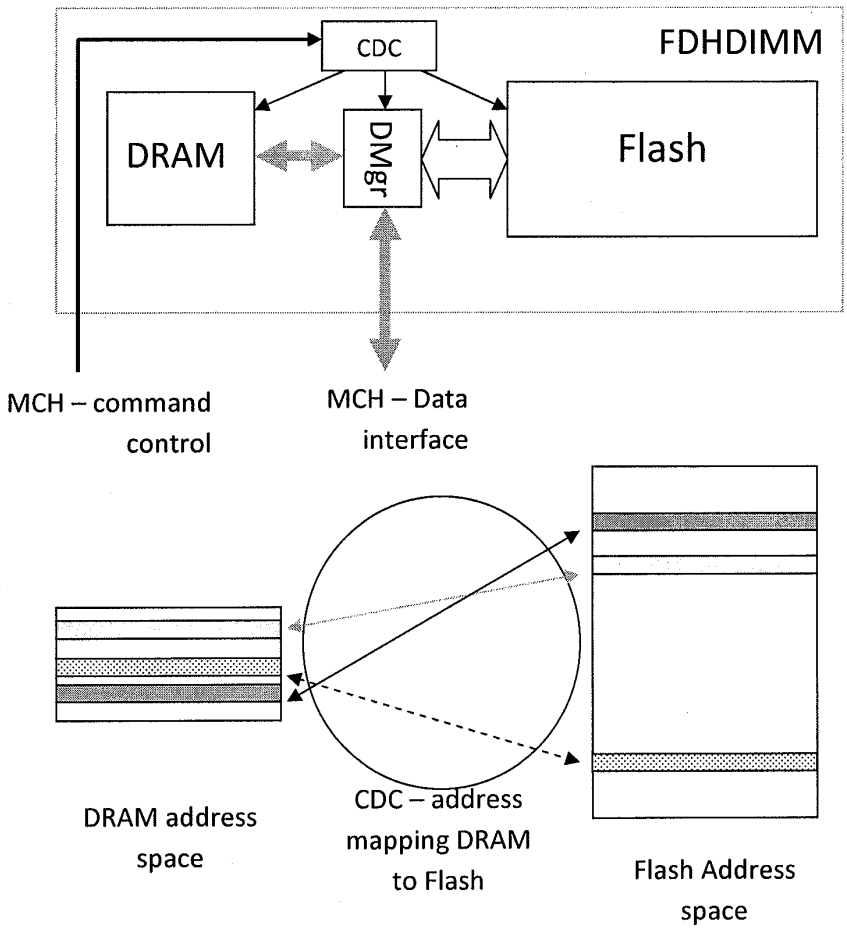


Figure 8B: Embodiment with single memory space with single control

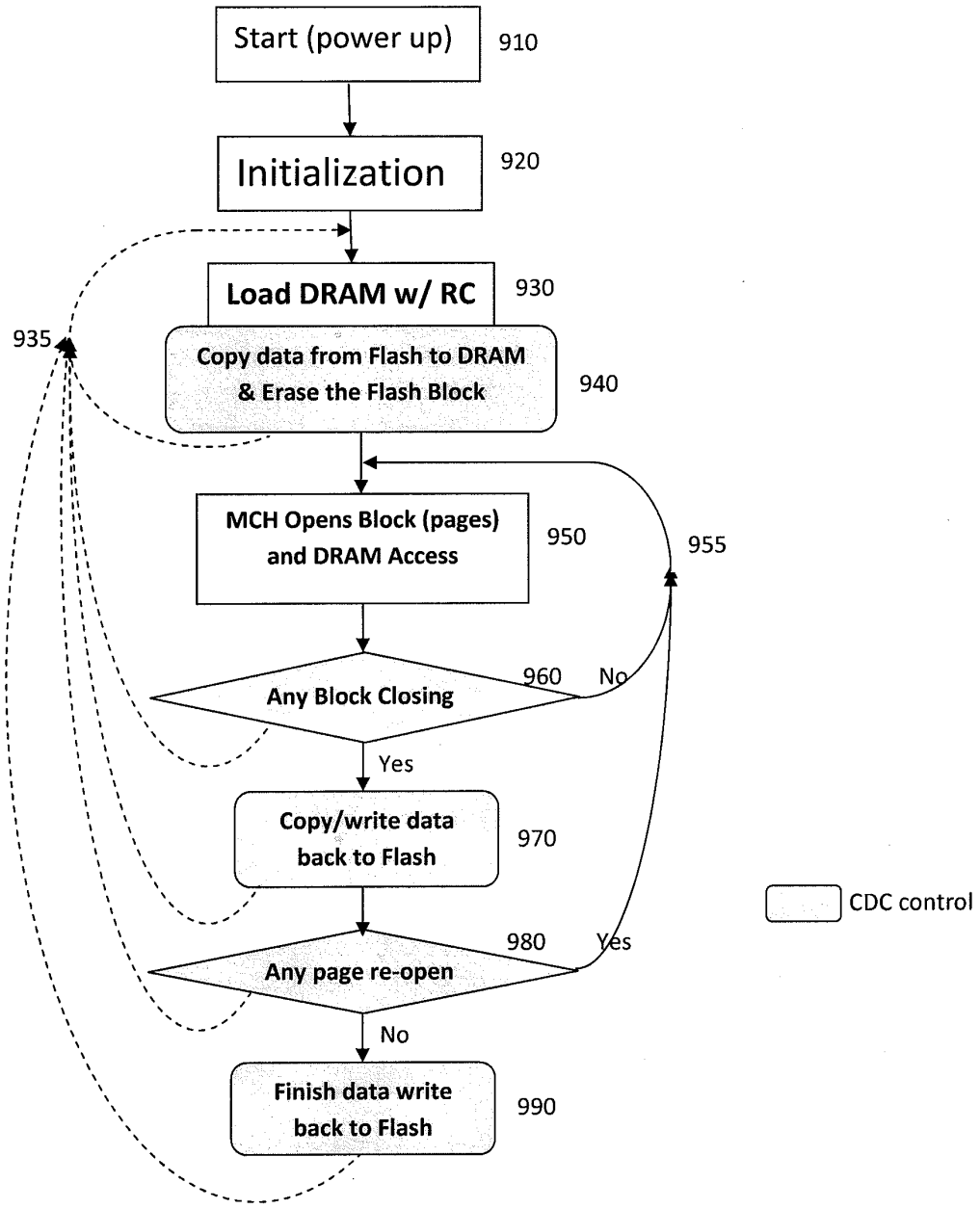


Figure 9: Operational state diagram of an embodiment

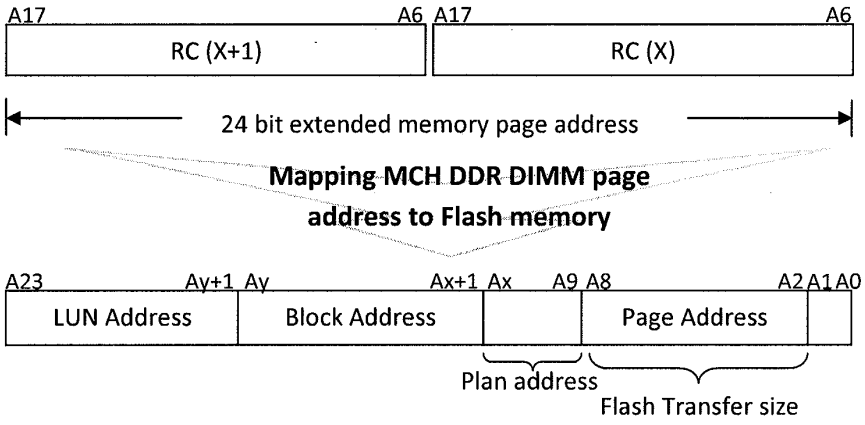


Figure 10: Mapping the DRAM page address to the Flash memory address field

DRAM density (GB)	# of blocks per bank	Flash wr-time to rd-time ratio	Avg block use time (sec)	Flash write tine (sec)	Max allowed Closed Blk in queue to be written back to Flash
1	250	55	1.00E-03	2.00E-02	0
1	250	55	1.00E-02	2.00E-02	2
1	250	55	2.00E-02	2.00E-02	5
1	250	55	5.00E-02	2.00E-02	11
2	500	55	1.00E-03	2.00E-02	0
2	500	55	1.00E-02	2.00E-02	5
2	500	55	2.00E-02	2.00E-02	9
2	500	55	5.00E-02	2.00E-02	23
4	1000	55	1.00E-03	2.00E-02	1
4	1000	55	1.00E-02	2.00E-02	9
4	1000	55	2.00E-02	2.00E-02	18
4	1000	55	5.00E-02	2.00E-02	45

Figure 11: example maximum allowed closed blocks in the queue to be written into Flash memory

Figure 12A:

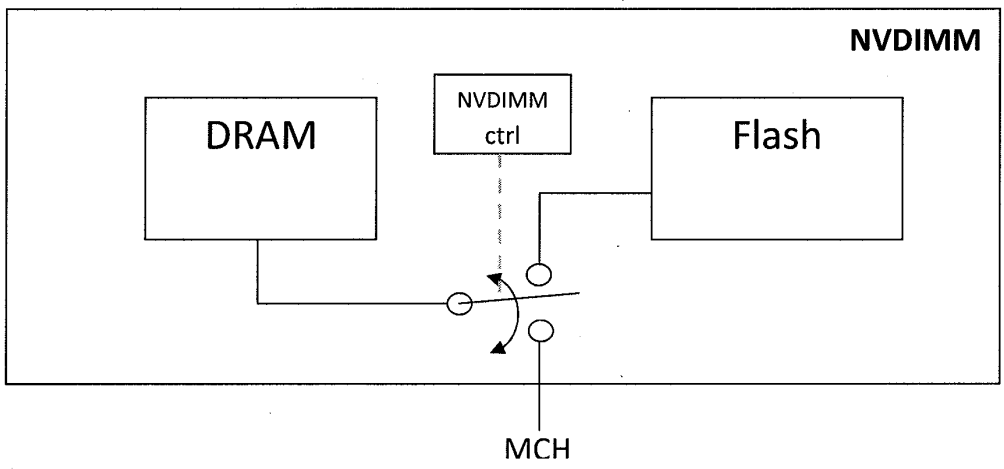


Figure 12B:

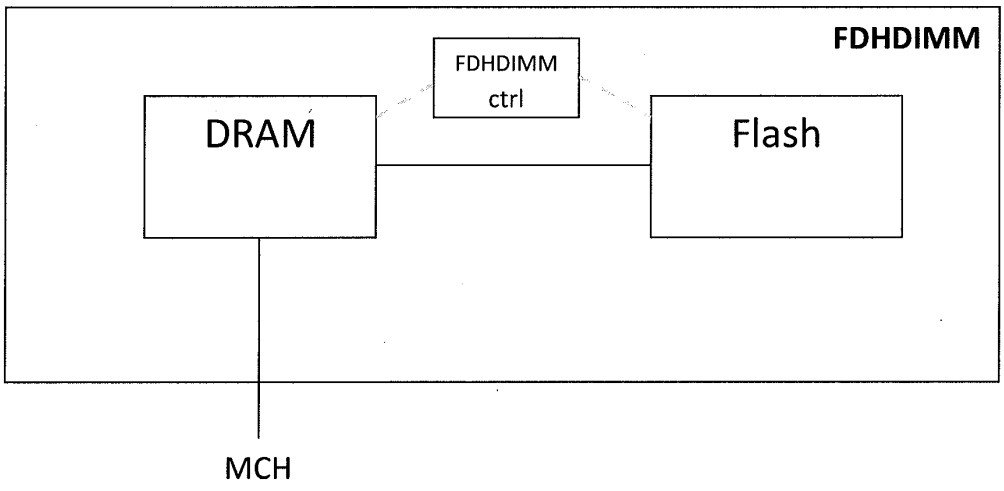
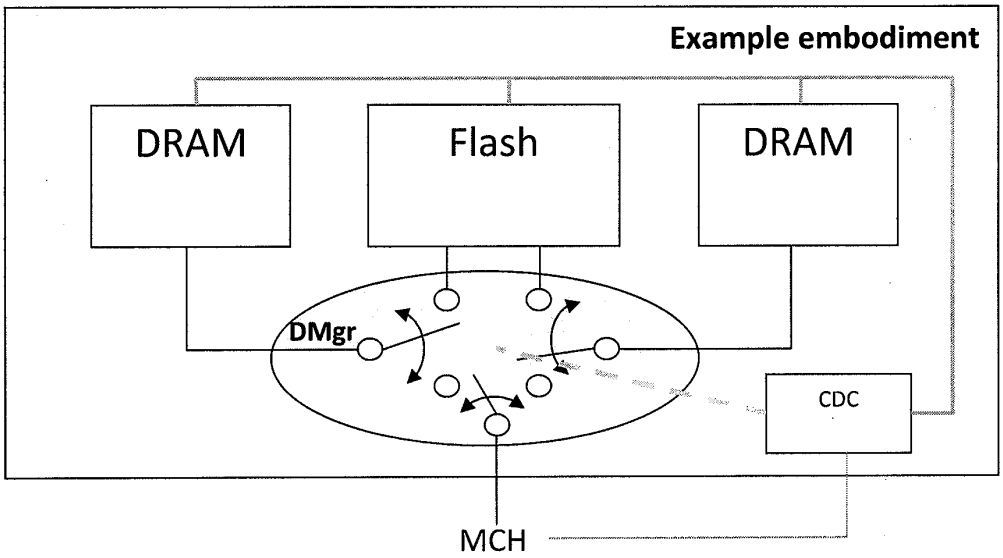


Figure 12C:



Appendix A

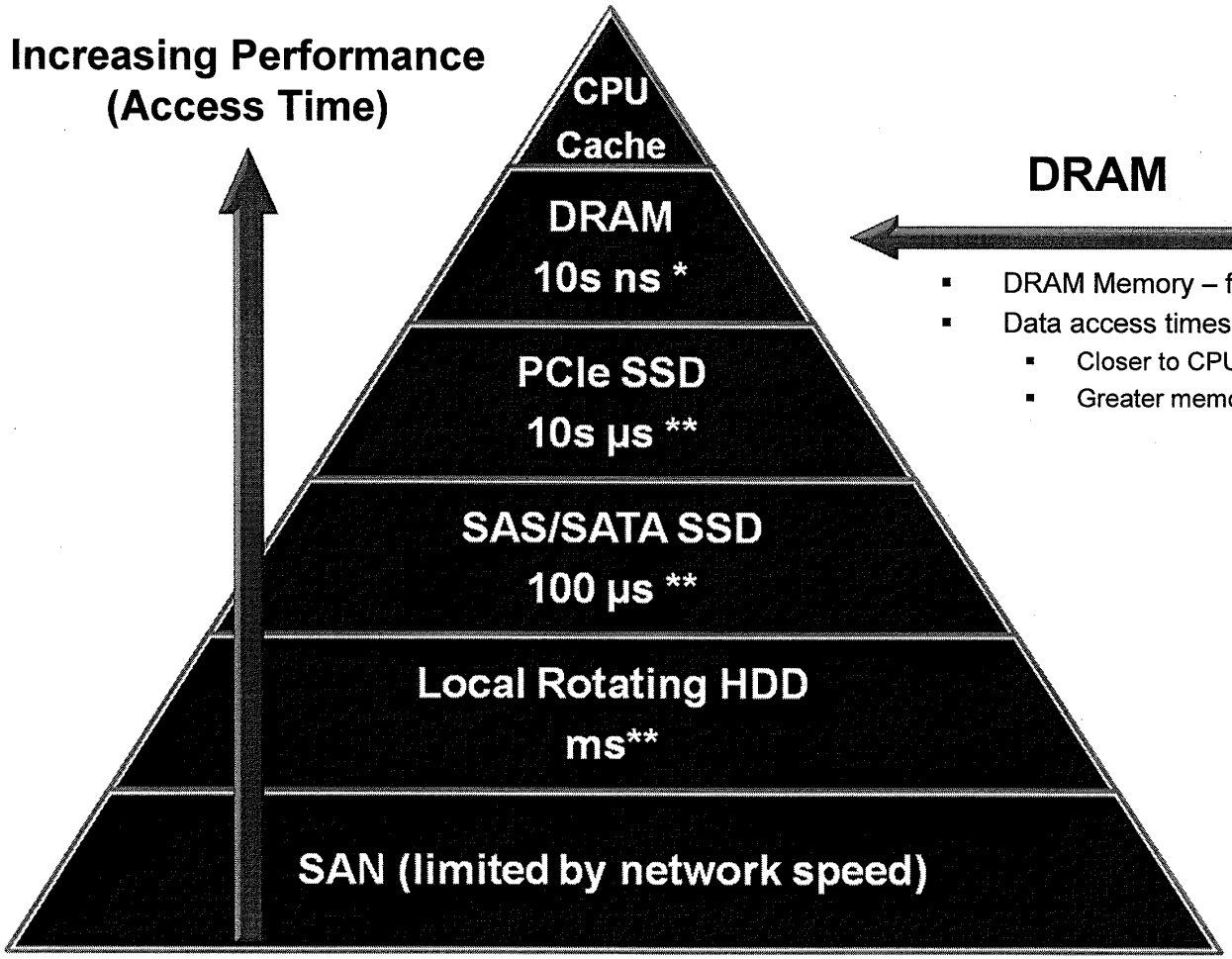
Hyper Vault

Multi-Terra Bytes of FSB Memory for HPC

NETLIST

DRAM in Tiered Storage Hierarchy

SCALING THE CLOUD



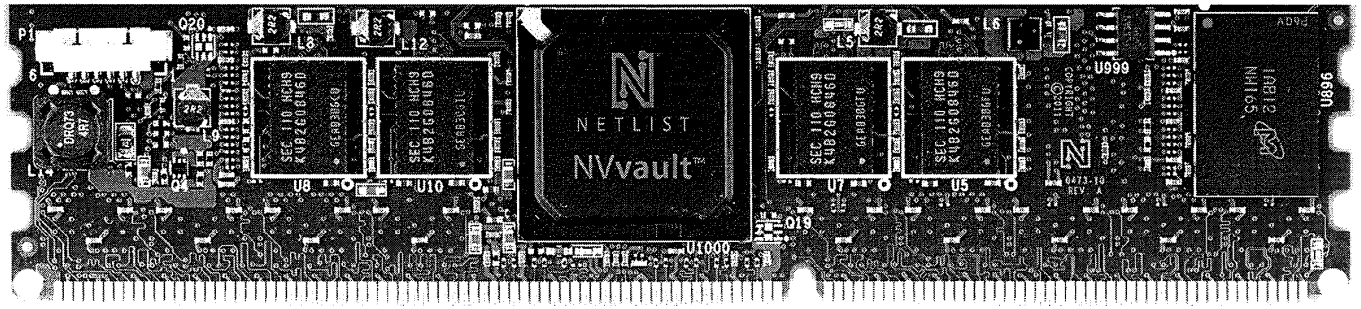
DRAM

- DRAM Memory – fastest data access times
- Data access times (speed of storage)
  - Closer to CPU the faster the access times
  - Greater memory density and speed is key

Notes:
[*] Based on Netlist calculations
[**] Based on product survey

NVvault DDR3 Product

SCALING THE CLOUD



Features

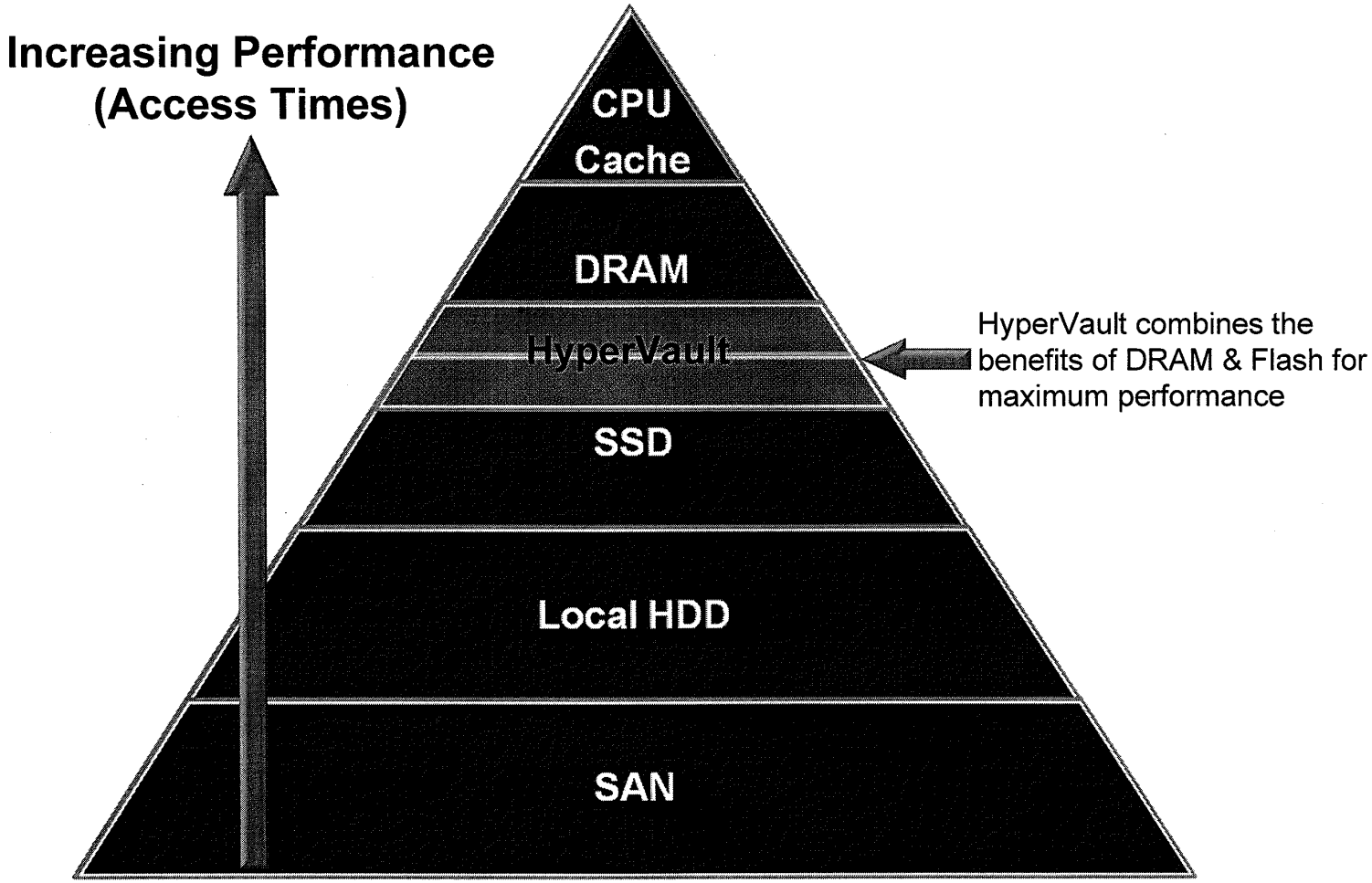
- 1333 MT/s throughput
- Fast R/W access time performance
- Fast backup/restore times (seconds)
- Standard DDR3 memory slot interface
- In-field programmability
- Programmable backup size
- Ultracapacitor for backup power
- 1GB/2GB capacity 7/11 (4GB/8GB Q4'11)

Benefits

- Accelerated write speeds for fast cache backup
- Low latency interface
- Plug 'n play hardware
- Eliminates batteries & reduces TCO
- No write life cycle limitations or performance degradation over time – writes to flash only upon power loss or system failure

HyperVault – Hybrid RDIMM Memory

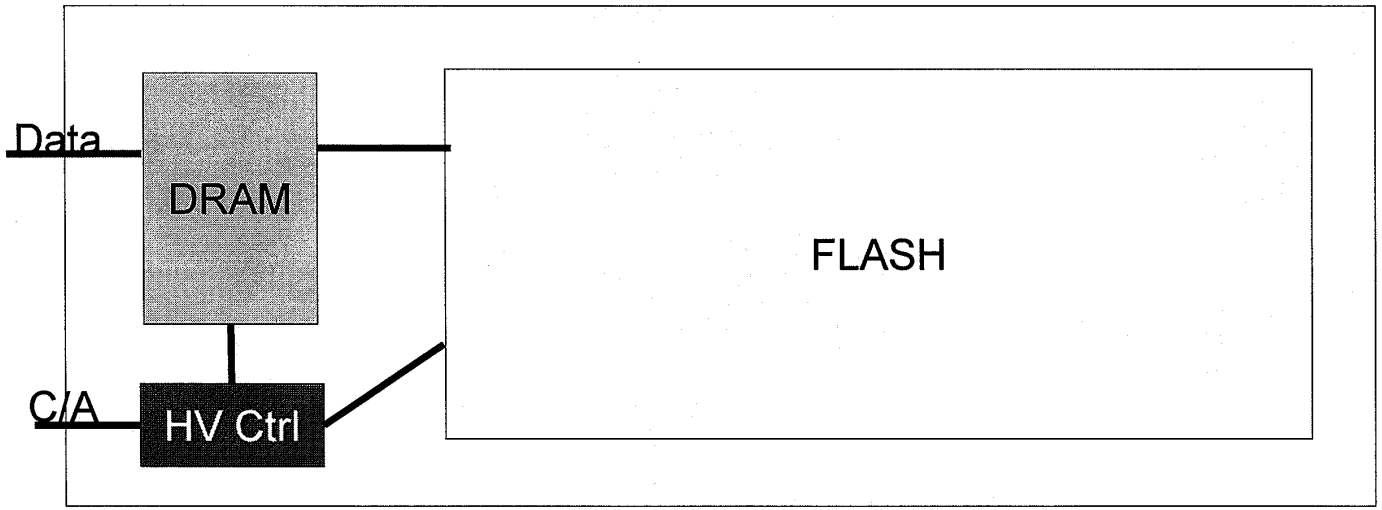
SCALING THE CLOUD



HyperVault – Hybrid RDIMM Memory

SCALING THE CLOUD

- What is HyperVault
  - Hybrid RDIMM with Flash and DRAM
  - Very High Density DIMM (1/4 to 1 TB)
  - Supports standard DDR DIMM interface



HyperVault – Hybrid RDIMM Memory

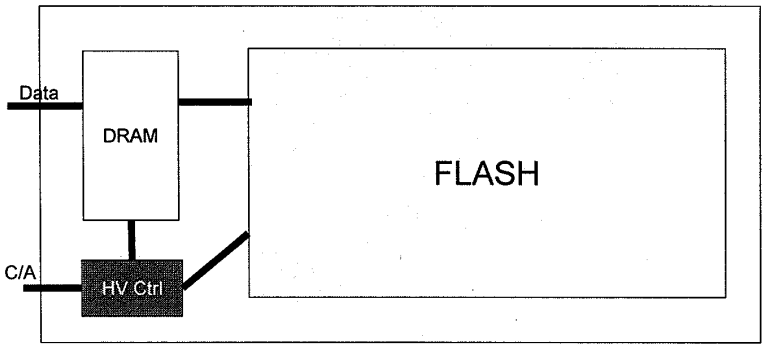
SCALING THE CLOUD

- What is HyperVault
  - No parity in Read Access vs. Write Access speed
  - Low power operation
  - Configurable page sizes
  - Configurable number of ranks per DIMM
  - Configurable number of banks per rank
  - Extended addressable memory space by
    - Concatenated RAS commands
    - Concatenated RAS and 1st CAS commands
  - Configurable NV memory space to support RAID operation
    - Automatic Data backup on the event of power loss
    - Automatic Data restoration as power comes back

HyperVault – Hybrid RDIMM Memory

SCALING THE CLOUD

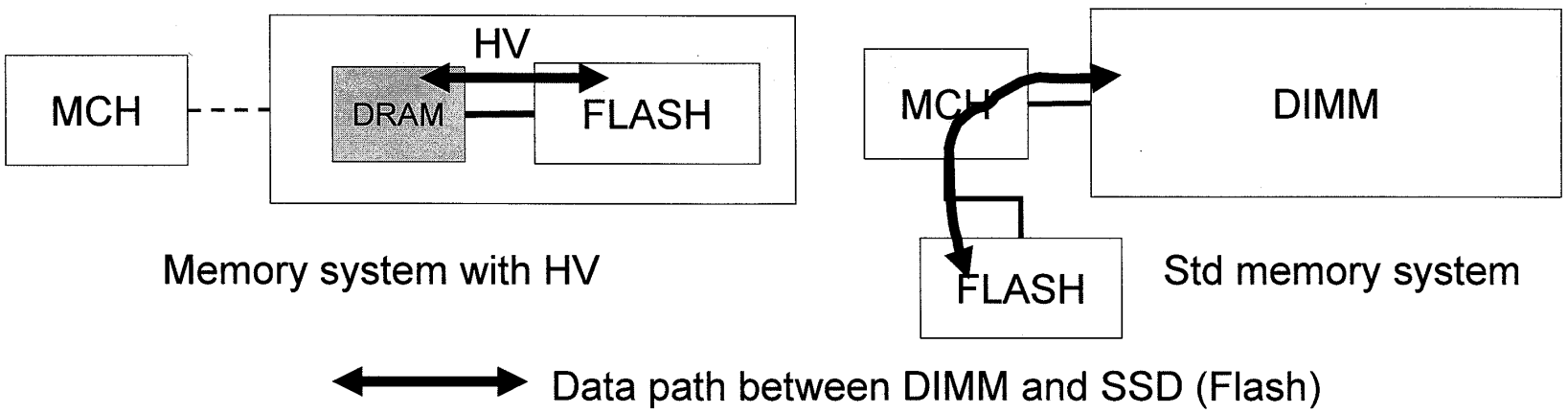
- What is HV Ctrl
  - HyperVault Controller
  - Data path manager
    - Between MCH and DRAM
    - Between DRAM and FLASH
  - Data format manipulator
    - Reformats the Data (by burst-size and by page-size)
    - Reformats the Data (by the number of banks per rank, by the number of ranks)
  - Memory address manager
    - Manages the virtual FLASH memory address space, and the physical DRAM address space
    - Address Mapping between DRAM and FLASH
  - Error manager (correction and recovery)
  - Flow rate controller



HyperVault – Hybrid RDIMM Memory

SCALING THE CLOUD

- Why HyperVault
  - Integration of SSD into the FSB memory
  - Improves CPU usage by minimizing the data transfer through CPU between SSD and FSB memory
  - Improves the Data transfer time between FLASH and DRAM
  - Expandable Memory concept
  - Field / User configurable SSD density based on the target application
  - Adjustable SSD (Flash) over DRAM ratio for optimum performance
  - Automatic power loss protection



Appendix B

Spansion® EcoRAM™ Solution
IOZone Solution Performance White Paper

Version 0.4

For

The Spansion EcoRAM Solution Configurations

Copyright © 2009 SPANSION® INC.

This document contains information proprietary to SPANSION® INC.
Use or disclosure without written permission
of an officer of SPANSION® INC. is expressly forbidden.

Overview

The Spansion® EcoRAM™ Solution consists of one or two Spansion EcoRAM Accelerators, 8 or 16 Spansion EcoRAM DIMMs (16GB or 32GB each) and associated driver software working together to provide a memory-expansion solution for AMD "Socket F" server motherboards.

Functionally, the Spansion EcoRAM solution initially appears as a POSIX-compliant device and operates like an ultra-fast SSD (Solid State Disk). In addition to the operating mode as an ultra-fast SSD, the Spansion EcoRAM solution also has a mode of operating as extremely high density main memory expansion.

In Expanded Memory Mode (EMM), Spansion EcoRAM memory supports read operations at cache-line granularity from any CPU in the system: data is returned in response to a CPU load instruction. The CPU accesses Spansion EcoRAM memory directly without first requiring a copy to a DRAM buffer and the initial access latency is a mere few hundred nanoseconds regardless of the size of data accessed. Like DRAM, Spansion EcoRAM memory supports ECC at cache-line granularity providing DRAM-class reliability levels.

Due to the extremely low read access latency of the Spansion EcoRAM's Expanded Memory Mode, there is no need to copy any portion of the data-set into DRAM to achieve higher performance. Low-latency, direct read accesses to Spansion EcoRAM occur as a cache line fill (64 Bytes). This capability is well suited to small, random read accesses on extremely large data sets where performance would traditionally require an equally large amount of DRAM to cache the IO accesses. Application workload throughput increases as much as two orders of magnitude with Spansion EcoRAM memory vs. Traditional IO, depending on the application workload's dataset size to DRAM ratio, memory access locality of reference and memory access granularity.

Testing Summary

Platform Configuration

- Server: Super Micro H8QME+2, 16GB DDR2 DRAM
- Operating System: CentOS 5.0 (2.6.18.3 kernel)

Spansion EcoRAM Configurations

256GB Spansion EcoRAM Solution – Single Accelerator

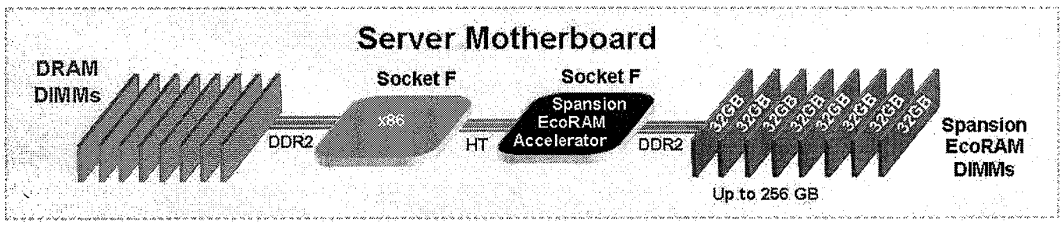


Figure 1: 256GB Single Accelerator Spansion EcoRAM Solution

256 GB Spanison EcoRAM Solution – Dual Accelerator

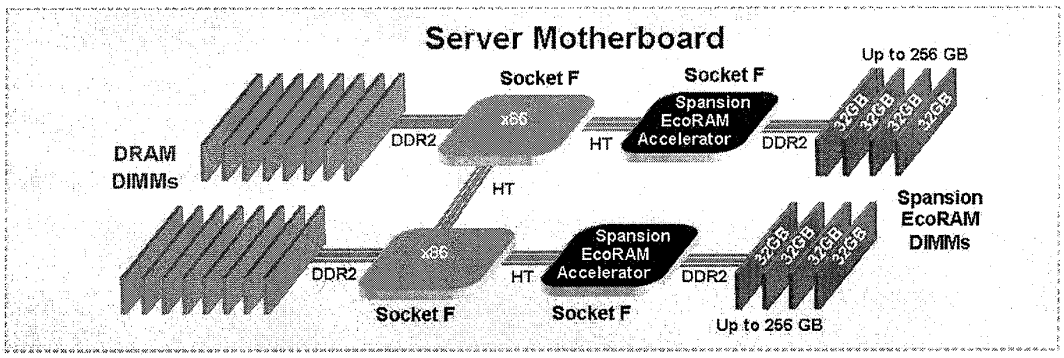


Figure 2: 256GB Dual Accelerator Spansion EcoRAM Solution

512 GB Spanison EcoRAM Solution – Dual Accelerator

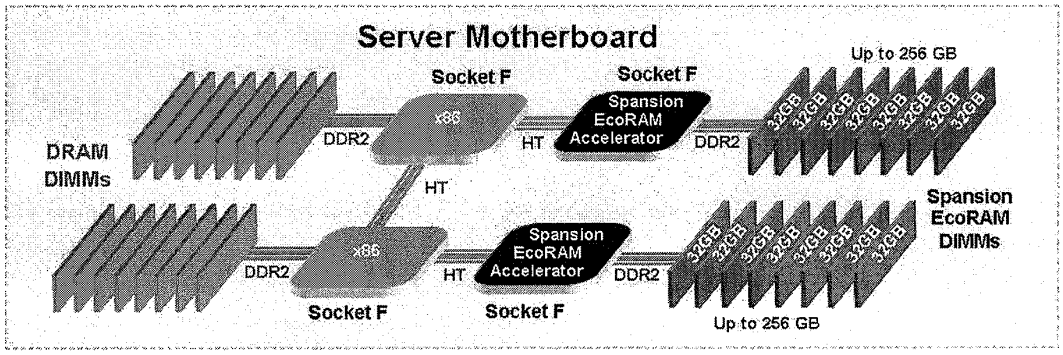


Figure 3: 512GB Dual Accelerator Spansion EcoRAM Solution

Test Tools

The performance analysis tool used in this evaluation is the open-source file system benchmark, IOZone v3.291 compiled for 64bit. In all tests, the number of threads used equals the number of cores available.

For write testing, two instances of IOZone were loaded and NUMA controlled.

Multiple passes of each test were run and averaged maintaining a Standard Deviation of <2%. PCIe-Based SSD results are from obtained from published product literature and have not been independently verified.

Test Details

File Sizes are extra large at 4GB because smaller files may not provide realistic performance. All Tests run each thread on a unique 4GB File.

The 256GB Spansion EcoRAM configurations use a single 2382 2.60GHz Opteron CPU, a single accelerator and 8 EcoRAM DIMMs. The 512GB configurations use dual 2382 2.60 GHz Opteron CPUs, dual accelerators and 16 EcoRAM DIMMs.

IOZone Command Line for Expanded Memory Mode Tests on Single Spansion EcoRAM Accelerator

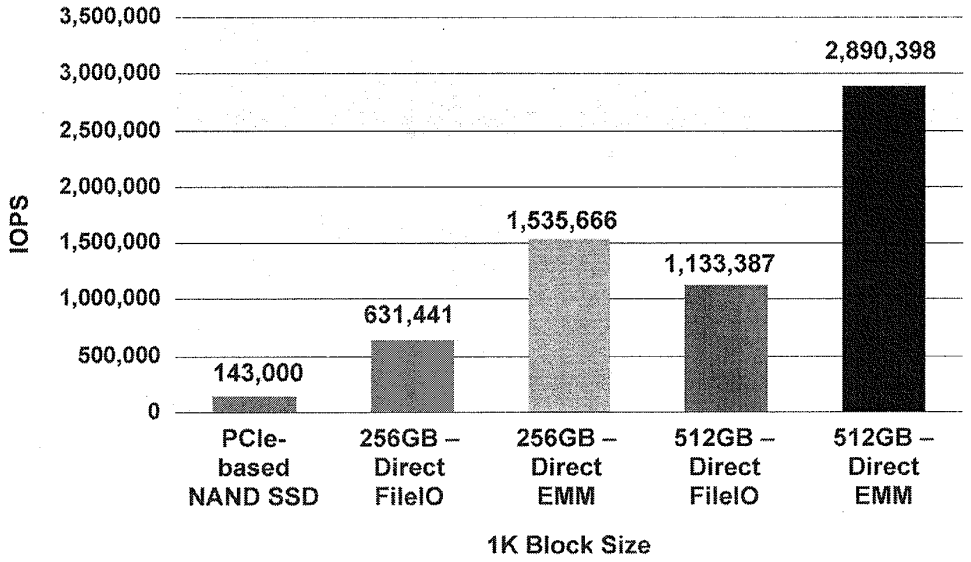
```
iozone -P 0 -L 64 -t 4 -F /uvfs1/file0 /uvfs1/file1 /uvfs1/file2 /uvfs1/file3 -R -i 1 -w -s 4g -r $BS –B
```

IOZone Command Line for FileIO Tests on Single Spansion EcoRAM Accelerator

```
iozone -P 0 -L 64 -t 4 -F /uvfs1/file0 /uvfs1/file1 /uvfs1/file2 /uvfs1/file3 -R -i 1 -w -s 4g -r $BS
```

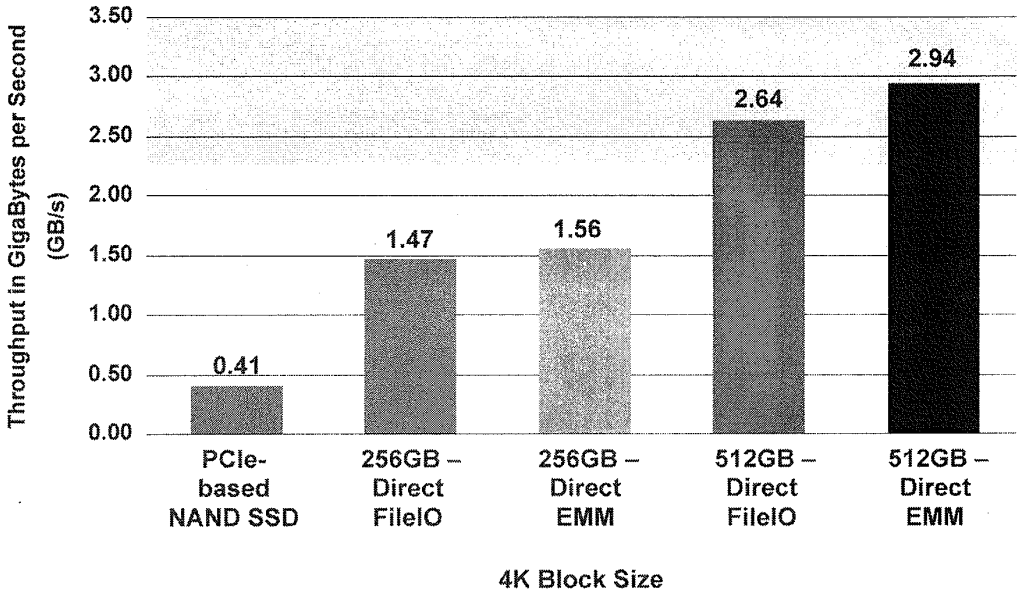
Results

IOZone Small Block Random Read IOPS – 1KB Block Size

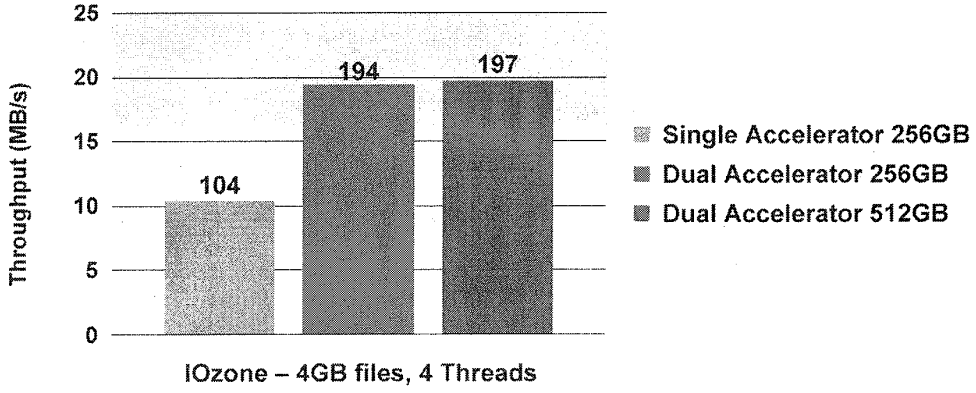


The extreme value of EcoRAM is in small block, random read accesses, particularly when using EcoRAM's EMM (Expanded Memory Mode).

IOZone Small Block Random Read Throughput Performance



IOZone Maximum Sustained Write Throughput – Direct Mode



Conclusions

This paper publishes the raw random read and sequential write throughput of a server equipped with Spansion EcoRAM using the IOZone Benchmarking Suite. Spansion EcoRAM's FileIO mode provides outstanding throughput for traditional IO access methods achieving significantly higher bandwidth and IOPS than the leading PCIe-Based NAND SSDs. For very small read accesses (as small as 64 bytes), the direct access capability of Spansion EcoRAM's Expanded Memory Mode maintains superior throughput at much smaller access sizes, due to the very low device latency (100-200ns.) and the elimination of buffer copies and system calls when accessing data. These performance capabilities, coupled with Spansion EcoRAM's ultra-high density and DRAM-class device reliability enables the expansion of a server's addressable memory space by incorporating Spansion EcoRAM for read-centric workloads.

Revision History

Revision	Date	Notes
.10	03/09/2009	First draft released
.20	07/15/2009	Results Update
.30	07/22/2009	Formatting
.40	10/08/2009	Results Update

Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	HIGH DENSITY DIMMS
First Named Inventor/Applicant Name:	Hyun Lee
Filer:	Bruce S. Itchkawitz/Jessica Noda
Attorney Docket Number:	NETL.096PR

Filed as Large Entity

Provisional Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Provisional application filing	1005	1	220	220
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**220**

Electronic Acknowledgement Receipt	
EFS ID:	10622596
Application Number:	61512871
International Application Number:	
Confirmation Number:	8113
Title of Invention:	HIGH DENSITY DIMMS
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	20995
Filer:	Bruce S. Itchkawitz/Valerie Jones
Filer Authorized By:	Bruce S. Itchkawitz
Attorney Docket Number:	NETL.096PR
Receipt Date:	28-JUL-2011
Filing Date:	
Time Stamp:	20:12:49
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$220
RAM confirmation Number	6325
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	Application Data Sheet	ads.pdf	240305 8619667969127842 1b532762c787e1a5f84c81db	no	4
Warnings:					
Information:					
This is not an USPTO supplied ADS fillable form					
2	Specification	specification.pdf	722158 840eae8ab61f3fb6ba1ccde8d448cc638b6525a8	no	13
Warnings:					
Information:					
3	Drawings-only black and white line drawings	drawings.pdf	320706 7bea41a4ca8883a2c23361de330a9b41674122d9	no	10
Warnings:					
Information:					
4	Appendix to the Specification	appendix_a.pdf	2523426 3474b0bf6e2acc29cee072df9471de4342ff9a0c	no	9
Warnings:					
Information:					
5	Appendix to the Specification	appendix_b.pdf	570775 807409f3e16244857ca06f9fbe0e18434ed6fdb5	no	8
Warnings:					
Information:					
6	Fee Worksheet (SB06)	fee-info.pdf	29458 f73572d8b63725e793e38bf764208dca3dd906ad	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	4406828		

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.