VOLUME II OF II (Appx5117-12151)
Nos. 24-2240, 24-2241

IN THE

United States Court of Appeals

FOR THE FEDERAL CIRCUIT

NETLIST, INC.,

Appellant,

v.

SAMSUNG ELECTRONICS CO., LTD., MICRON TECHNOLOGY INC., MICRON SEMICONDUCTOR PRODUCTS, INC., MICRON TECHNOLOGY TEXAS, LLC,

Appellees.

On Appeal from the United States Patent and Trademark Office, Patent Trial and Appeal Board, Nos. IPR2022-01427, IPR2022-01428, IPR2023-00882, IPR2023-00883

JOINT APPENDIX

Michael Hawes
Counsel of Record
Lori Ding
BAKER BOTTS L.L.P.
910 Louisiana Street
Houston, TX 77002
(713) 229-1234
michael.hawes@bakerbotts.com

Eliot Williams
BAKER BOTTS L.L.P.
700 K Street, N.W.
Washington, D.C. 20001

Counsel for Appellee Samsung Electronics Co., Ltd.

Jeffrey A. Lamken
Counsel of Record
Rayiner Hashem
Jennifer Elizabeth Fischell
Lidiya Mishchenko
Kayvon Ghayoumi
MOLOLAMKEN LLP
The Watergate, Suite 500
600 New Hampshire Avenue, N.W.
Washington, D.C. 20037
(202) 556-2000
jlamken@mololamken.com

Counsel for Appellant Netlist, Inc.

(Additional Counsel Listed on Inside Cover)

Netlist EX2028
Samsung v. Netlist
IPR2025-01402

Theodore W. Chandler
BAKER BOTTS L.L.P.
1801 Century Park East, Suite 2400
Los Angeles, CA 90067

Counsel for Appellee Samsung Electronics Co., Ltd.

Michael R. Rueckheim
WINSTON & STRAWN LLP
225 Shoreline Drive, Suite 520
Redwood City, CA 94065

Juan Yaquian
WINSTON & STRAWN LLP
800 Capitol Street, Suite 2400
Houston, TX 77002

Counsel for Appellees Micron Technology Inc., Micron Semiconductor Products, Inc., and Micron Technology Texas, LLC

Elizabeth Kathleen Clarke
Bonnie K. St. Charles
MOLOLAMKEN LLP
300 N. LaSalle Street, Suite 5350
Chicago, IL 60654

Sara Margolis
Catherine Martinez
MOLOLAMKEN LLP
430 Park Avenue
New York, NY 10022

Counsel for Appellant Netlist, Inc.

TABLE OF CONTENTS

VOLUME I

Date Filed	Dkt. No.	Document	First Appx Page of Document
Required Materials			
4/1/2024	Paper 48 (IPR2022-01427)	Final Written Decision in IPR2022-01427	Appx1
4/1/2024	Paper 48 (IPR2022-01428)	Final Written Decision in IPR2022-01428	Appx69
6/17/2024	Paper 53 (IPR2022-01427), Paper 52 (IPR2022-01428)	Decision Denying Patent Owner's Request on Rehearing	Appx144
8/26/2022	Ex. 1001 (IPR2022-01427)	U.S. Patent No. 9,318,160, Lee	Appx147
8/26/2022	Ex. 1001 (IPR2022-01428)	U.S. Patent No. 8,787,060, Lee	Appx171
10/3/2024	Appellate Dkt. 16	Certified List	Appx195
IPR2022-01427, U.S. Patent No. 9,318,160			
8/26/2022	Paper 1	Petition for Inter Partes Review of U.S. Patent No. 9,318,160	Appx203
7/23/2023	Paper 20	Patent Owner Response	Appx555
10/19/2023	Paper 24	Petitioner Reply to Patent Owner Response	Appx691

Date Filed	Dkt. No.	Document	First Appx Page of Document
10/26/2023	Paper 25	Decision Granting Institution of Inter Partes Review and Granting Motion for Joinder	Appx747
11/21/2023	Paper 28	Patent Owner Sur-Reply	Appx761
1/11/2024	Paper 39	Patent Owner Demonstratives for Oral Hearing	Appx907
1/11/2024	Paper 46	Record of Oral Hearing Held January 11, 2024	Appx1292
5/1/2024	Paper 50	Patent Owner Request for Director Review	Appx1392
10/26/2023	Paper 51	Decision Granting Institution of Inter Partes Review and Granting Motion for Joinder in IPR2023-00883	Appx1411
8/19/2024	Paper 53	Patent Owner Notice of Appeal	Appx1418
8/26/2022	Ex. 1003	Declaration of Dr. Andrew Wolfe Regarding U.S. Patent No. 9,318,160	Appx2009
8/26/2022	Ex. 1011	U.S. Patent Application No. 2008/0025137 A1, Rajan et al.	Appx3556
8/26/2022	Ex. 1014	U.S. Patent Application No. 2011/0103156 A1, Kim et al.	Appx3622
8/26/2022	Ex. 1015	U.S. Patent No. 8,041,881 B2, Rajan et al.	Appx3634
8/26/2022	Ex. 1016	U.S. Patent Application No. 2011/0026293 A1, Riho	Appx3670
8/26/2022	Ex. 1017	U.S. Patent No. 7,969,192, Wyman et al.	Appx3695
8/26/2022	Ex. 1019	JEDEC Standard JESD79-3C, DDR3 SDRAM Specification (Nov. 2008)	Appx3731

Date Filed	Dkt. No.	Document	First Appx Page of Document
8/26/2022	Ex. 1022	Bruce L. Jacob et al., Memory Systems: Cache, DRAM, Disk (2008)	Appx4299
8/26/2022	Ex. 1023	Bruce L. Jacob, Synchronous DRAM Architectures, Organizations, and Alternative Technologies (2002)	Appx4595
8/26/2022	Ex. 1024	U.S. Patent No. 7,796,446 B2, Ruckerbauer et al.	Appx4617
8/26/2022	Ex. 1025	U.S. Patent No. 8,258,619 B2, Foster et al.	Appx4639
8/26/2022	Ex. 1026	U.S. Patent Application No. 2006/0277355 A1, Ellsberry et al.	Appx4656
8/26/2022	Ex. 1031	U.S. Patent No. 7,532,537 B2, Solomon et al.	Appx4851

VOLUME II

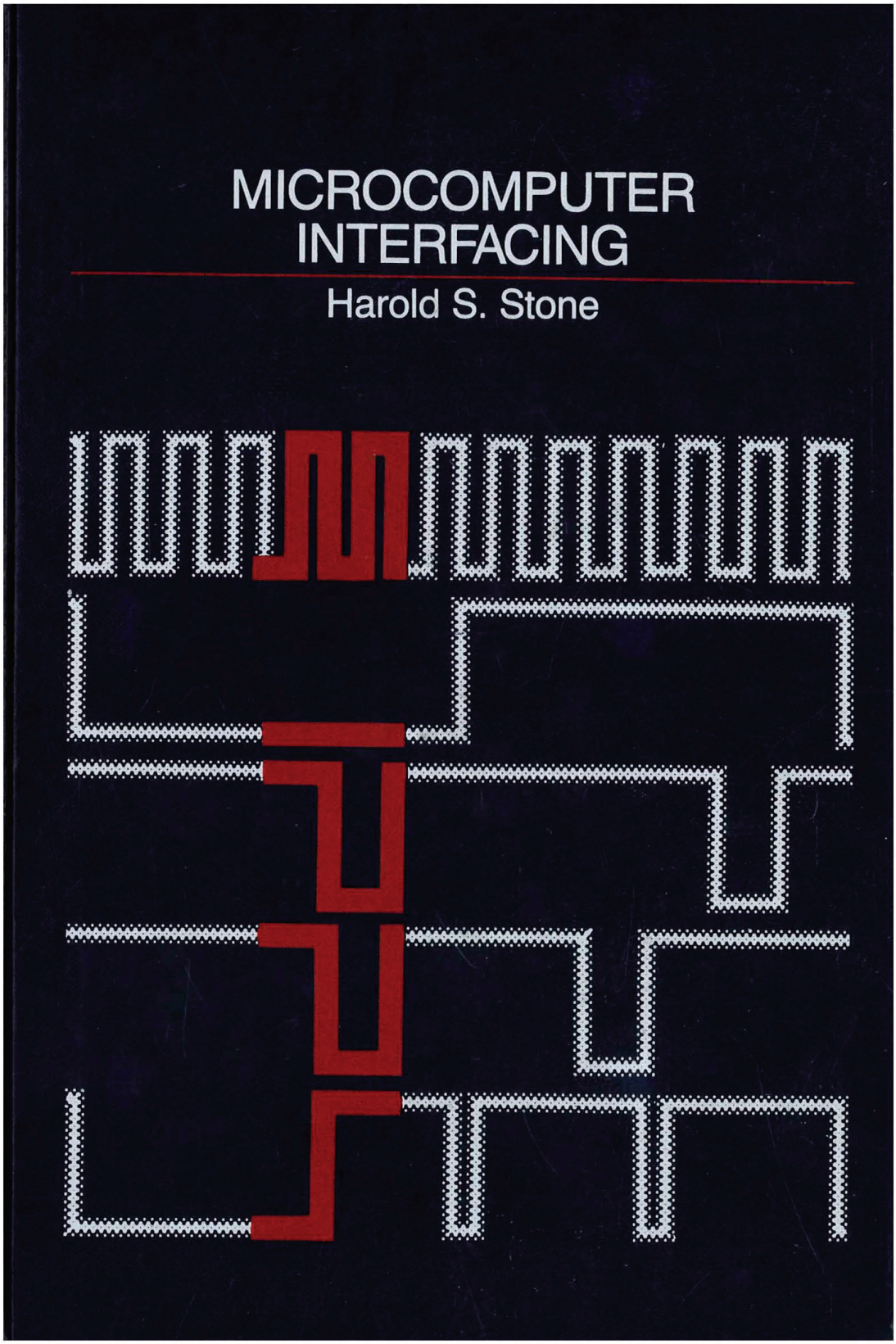
Date Filed	Dkt. No.	Document	First Appx Page of Document
IPR2022-01427, U.S. Patent No. 9,318,160 (cont'd)			
8/26/2022	Ex. 1038	Harold S. Stone, Microcomputer Interfacing (1982)	Appx5117
8/26/2022	Ex. 1044	First Amended Complaint, *Netlist, Inc. v. Samsung Elecs. Co., Ltd.*, *Samsung Elecs. Am., Inc. & Samsung Semiconductor, Inc.*, Dkt. 23, No. 2:21-cv-463 (E.D. Tex. May 3, 2022)	Appx5601

Date Filed	Dkt. No.	Document	First Appx Page of Document
1/19/2023	Ex. 2007	Jin Hee Cho et al., A 1.2V 64Gb 341GB/s HBM2 Stacked DRAM with Spiral Point-to-Point TSV Structure and Improved Bank Group Data Control	Appx6113
7/27/2023	Ex. 2023	Declaration of Dr. Michael C. Brogioli	Appx6374
7/27/2023	Ex. 2025	Deposition with Errata of Dr. Andrew Wolfe (June 27, 2023)	Appx6856
7/27/2023	Ex. 2027	Transcript of First Day of Jury Trial in *Netlist, Inc. v. Samsung Elecs. Co., Ltd.*, *Samsung Elecs. Am., Inc. & Samsung Semiconductor, Inc.*, No. 2:21-cv-463 (E.D. Tex. Apr. 14, 2023)	Appx7398
7/27/2023	Ex. 2035	Samsung, *Samsung Begins Mass Producing World's Fastest DRAM – Based on Newest High Bandwidth Memory* (HBM) Interface (Jan. 19, 2016)	Appx7503
7/27/2023	Ex. 2038	HyperCloud and N3DS Move to High-Density for Cloud Computing (Feb. 2011)	Appx7528
10/19/2023	Ex. 1052	Deposition and Errata of Dr. Michael C. Brogioli (Sept. 27, 2023)	Appx9601
10/19/2023	Ex. 1054	Netlist's Demonstratives for Technology Tutorial in Inv. No. 337-TA-1023 (May 8, 2017)	Appx10308
1/8/2024	Ex. 1067	Petitioner's Demonstratives for Oral Argument	Appx10674

Date Filed	Dkt. No.	Document	First Appx Page of Document
IPR2022-01428, U.S. Patent No. 8,787,060			
8/26/2022	Paper 1	Petition for Inter Partes Review of U.S. Patent No. 8,787,060	Appx11014
7/27/2023	Paper 20	Patent Owner Response	Appx11382
10/19/2023	Paper 24	Petitioner Reply to Patent Owner Response	Appx11519
11/21/2023	Paper 28	Patent Owner Sur-Reply	Appx11588
5/1/2024	Paper 50	Patent Owner Request for Director Review	Appx11841
8/19/2024	Paper 53	Patent Owner Notice of Appeal	Appx11867
8/26/2022	Ex. 1003	Declaration of Dr. Andrew Wolfe Regarding U.S. Patent No. 8,787,060	Appx11952

MICROCOMPUTER INTERFACING

Harold S. Stone



Samsung Electronics Co., Ltd.
Ex. 1038, p. i

Appx5117

Microcomputer Interfacing

HAROLD S. STONE
UNIVERSITY OF MASSACHUSETTS, AMHERST

ADDISON-WESLEY PUBLISHING COMPANY
READING, MASSACHUSETTS
MENLO PARK, CALIFORNIA
LONDON
AMSTERDAM
DON MILLS, ONTARIO
SYDNEY

Samsung Electronics Co., Ltd.
Ex. 1038, p. ii

Appx5118

This book is in the **ADDISON-WESLEY SERIES IN ELECTRIC ENGINEERING**

SPONSORING EDITOR: *Tom Robbins*
PRODUCTION EDITOR: *Marilee Sorotskin*

TEXT DESIGNER: *Herb Caswell*
ILLUSTRATOR: *Jay's Publishers Service Inc.*
COVER DESIGN AND ILLUSTRATOR: *T. A. Philbrook*
ART COORDINATOR: *Joseph Vetere*

PRODUCTION MANAGER: *Sue Zorn*
PRODUCTION COORDINATOR: *Helen Wythe*

The text of this book was composed in Times Roman on a Mergenthaler 202 by Information Sciences Corporation and was printed by R. R. Donnelley and Sons.

Library of Congress Cataloging in Publication Data

Stone, Harold S., 1938–
Microcomputing interfacing.

Bibliography: p.

1. Interface circuits. 2. Microcomputers—Circuits.
I. Title. TK7868.158S76 621.3819'5835 81-17619
ISBN 0-201-07403-6 AACRZ

Copyright © 1982 by Addison-Wesley Publishing Company, Inc.

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America. Printed simultaneously in Canada.

ISBN 0-201-07403-6
ABCDEFGHIJK-DO-898765432

Samsung Electronics Co., Ltd.
Ex. 1038, p. iii

Appx5119

2 / SHIELDING, GROUNDING, AND TRANSMISSION-LINE TECHNIQUES

In the vast majority of microcomputer applications, the microprocessor serves as an "intelligent" controller in addition to any other functions it performs. As a controller, the micro senses the activity taking place around it and issues signals periodically to alter the external activity. Micros in automobiles, video terminals, communications links, for example, are almost completely dedicated to a control function. In the automobile, the micro positions mechanical actuators that govern air/fuel mixture and other critical parameters of the internal-combustion engine. In the video terminal, the micro alters no physical or mechanical devices but rather manages a reasonably complex, real-time I/O system that includes at least a keyboard, video display, and serial data link. In communications systems, the control function includes message storing and forwarding, error detection and recovery, and routing-table maintenance. Word processors and small-business systems use the microcomputer for substantially more than the control function, in that the text manipulation and financial record-management computations that characterize these systems are strongly computational rather than control in nature. Nevertheless, the control function is present, and the ability to connect the microcomputer to external peripherals such as printer, disk, and communications link is a critical ingredient in the system's effectiveness.

With the microprocessor almost universally operating as a controller of external devices, there is an inherent need to understand how to interconnect a microcomputer to other system components. This chapter treats the interconnection problem at the electrical level where we see what is necessary for reliable and safe interconnection. Chapter 1 gives a view of the interfaces between the micro and memory and I/O ports in terms of the information that flows across the interfaces. That information has to be converted into a physical form such as voltage, current, or light when it is transported from one module to another. This chapter treats these electrical aspects of the interconnections. The next chapter deals with timing and logical control of the signals on the interconnections.

The electrical issues raised in this chapter are quite real and extremely important. Interfacing is so challenging in its nonelectrical aspects that we would much prefer to ignore the electrical problems if only we could. Ideally we wish to connect points A and B on the microcomputer with points C and D on a peripheral device, knowing that the voltages are compatible, and then to ignore all other electrical aspects of this connection. But the underlying physical laws that dictate what happens over this connection can work against us. When connections that look deceptively simple are made improperly, the connections can introduce excessive noise, lower reliability, and serious hazards. Certain techniques for interconnections have evolved over the years, giving the system designer a wealth of

Samsung Electronics Co., Ltd.
Ex. 1038, p. 42

Appx5173

sound ways for implementing interconnections. Standards exist for some important types of interconnections so that at least in these instances, the designer can freely connect components, following the rules of the standard, with relative assurance that the system will function safely and reliably. But even with standard interconnections, unskilled designers can and do make mistakes that negate the value of the standard.

This chapter provides the essential background in the physical principles from which all methods of interconnection evolve. In the latter part of the chapter we look at some basic techniques in current use to illustrate the principles.

2.1 GROUNDING AND SHIELDING

Interconnections are a major of source of noise in a computer system. Typically, interconnections are made through cables or backplanes whose lengths run from roughly 10 cm to 10 m. When cabling is done improperly, electrical noise from motors, electrical storms, or from nearby electronic equipment can easily couple into the cabling where it is then indistinguishable from data and control signals. Electronic transmission on interconnecting cables, especially on long cables, simply does not behave the same as transmission on short interconnections. The basic physical principles are the same for both short and long wires, but many effects are not noticeable on short interconnections, so these connections can be made in the most straightforward possible manner with little risk. The differences in behavior are second-order effects that are essentially indiscernible on short conductors. It is very tempting to ignore these effects for long interconnections as well, but if one does one finds that the effects are quite visible and that they can severely upset the integrity of the interconnections if the interface fails to deal with the effects properly.

As an example of one such situation, consider the interconnection of two chassis, each of which has a separate connection to ground, as shown in Fig. 2.1. The ground point shown in the figure is the electrical ground on the AC distribution system. (This ground is accessible through the large round prong of 3-prong AC connectors required for grounded equipment by the electrical code in the United States.) The signal ground within each of the two chassis is connected to the chassis, and each chassis is connected to the electrical ground of the AC system through the ground prong of the AC connector. If the two chassis are located in the same cabinet or are in close physical proximity, then it is possible to make a safe electrical connection between the chassis ground points on the two chassis. This is shown in Fig. 2.1(a), and is common practice for chassis that are mounted in the same equipment rack. When all of the assumptions on physical proximity are met, the two chassis are at the same potential, so that no current flows through the conductor that interconnects the two chassis ground points. (Actually, the potentials of the chassis could vary a little, and a small current might flow through the conductor equalizing the potentials.)

Fig. 2.1(b) shows a situation in which the chassis are separated by a large physical distance. The ground points in the figure are grounds on different AC connectors. Even when the connectors lie in the same room, if the AC connectors are controlled by different circuit breakers, the ground conductors may be totally separate and distinct conductors

Samsung Electronics Co., Ltd.
Ex. 1038, p. 43
Appx5174

the power of the direct wave. When stubs are terminated with an infinite impedance, any wave that enters the stub is reflected back to the bus. At the bus the wave splits, and propagates in both directions on the bus, and reflects back into the stub. We have seen earlier that subsequent reflections in the stub repeat this process until all of the stub energy is returned to the transmission line, half in the source direction and half in the load direction.

As long as reflections are small compared to the direct wave, the peaks and troughs visible in the direct wave arising from stub reflections will be small and will die out as the line charges to the impressed voltage. The depth of the troughs depend on the impedances that terminate the stubs, and the decay time depends on the length of the stubs and the length of the line. In critical applications, the transmission line should be routed to each receiver input, so that stub length is essentially zero, and the wave impedance of the tap should be as high as possible.

Another alternative used in practice is to eliminate all taps from buses. The signal is transmitted using point-to-point driver and receiver pairs. Each receiver has a corresponding driver that forwards the signal to the next point on the bus. This method increases the delay in propagation time because signals pass through many receiver/driver pairs in moving from one end of the bus to the other, but each segment of the line is point to point and can be terminated properly for noise-free operation.

The most complex situation encountered in practice is the one depicted in Fig. 2.22. This figure shows a line with multiple taps in which each tap has both a driver and receiver. In this case, neither sources nor loads can be terminated because all termination impedances are in parallel and will draw excessive power. This type of bus is used commonly in multiboard applications, where each board is a tap on the bus line with a receiver and driver for each bus wire. Systems function satisfactorily without termination in the 1 to 10 MHz range when the bus is contained in a single chassis, and has a length not exceeding 0.2 to 0.3 m. As frequencies or distance or both increase, this busing method breaks down because of the noise on the transmission line caused by all of the reflections. The IEEE-488 bus standard covers a bus for peripheral devices that could be up to 20 m at 500 K-bytes per second, and up to 10 m long to attain 1 M-bytes per second. To achieve this type of performance, the standard recommends that each tap on the bus be that shown in Fig. 2.22(b). Each device has a pull-up and pull-down resistor whose parallel equivalent is about 2 kΩ. The resistors act as a local voltage divider that sets the quiescent voltage at about 3.3 V. Hence, when no active device drives the bus line, the line goes to a logic 1. The termination at each device is too large to match the line impedance, but it does absorb some of the energy that reaches it, and reduces the energy returned in reflected waves. As more devices are added to the bus, the effect of the termination resistors is to create a greater burden on a gate that drives the bus. The IEEE standard states that the bus can achieve a data rate of 500 K-bytes per second for buses up to 20 m in length, with up to one device per 2 m of cable. An even higher data rate is achievable if the maximum length of the cable is limited to 10 m with up to one device per meter of cable. The higher speed requires devices to be placed closer together, but the maximum number of devices on the bus is not increased in spite of the closer proximity of devices. Loading effects caused by additional devices on the bus tend to reduce the maximum speed.

Samsung Electronics Co., Ltd.
Ex. 1038, p. 66

Appx5197

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

NETLIST, INC.	)	
	)	
Plaintiff,	)	
	)	
v.	)	
	)	
SAMSUNG ELECTRONICS CO., LTD., SAMSUNG ELECTRONICS AMERICA, INC., SAMSUNG SEMICONDUCTOR, INC.	)	Civil Action No. 2:21-CV-463-JRG
	)	
	)	JURY TRIAL DEMANDED
Defendants.	)	

FIRST AMENDED COMPLAINT

1. Plaintiff Netlist, Inc. ("Netlist"), by its undersigned counsel, for its First Amended Complaint against defendants Samsung Electronics Co., Ltd. ("SEC"), Samsung Electronics America, Inc. ("SEA"), and Samsung Semiconductor, Inc. ("SSI") (collectively, "Samsung" or "Defendants"), states as follows, with knowledge as to its own acts, and on information and belief as to the acts of others:

2. Netlist filed its initial complaint against Defendants on December 20, 2021. Dkt. 1 ("Initial Complaint"). In its Initial Complaint, Netlist presented allegations regarding Defendants' infringement of three of Netlist's patents: U.S. Patent Nos. 10,860,506 (the "'506 Patent," Ex. 1), 10,949,339 (the "'339 Patent," Ex. 2), and 11,016,918 (the "'918 Patent," Ex. 3). Netlist also indicated its intent to assert then pending U.S. Pat. No. 11,232,054 (the "'054 Patent," Ex. 4) upon issuance. On December 28, 2021, Defendants filed an unopposed motion confirming

the parties' agreement to waive service for SEC, and set the deadline for Defendants' answer to the Initial Complaint as April 12, 2022. Dkt. 10. Defendants collectively filed an answer on April 12, 2022 (Dkt. 16) without moving to dismiss the Initial Complaint. Netlist now enters its First Amended Complaint as a matter of course pursuant to Fed. R. Civ. Proc., Rule 15(a)(1)(B).

3. Upon the entry of this First Amended Complaint, this action now involves six of Netlist's patents: the '506 Patent, the '339 Patent, the '918 Patent, the '054 Patent, and U.S. Patent Nos. 8,787,060 (the "'060 Patent," Ex. 5), and 9,318,160 (the "'160 Patent," Ex. 6) (collectively, the "Patents-in-Suit").

I. THE PARTIES

4. Plaintiff Netlist is a corporation organized and existing under the laws of the State of Delaware, having a principal place of business at 111 Academy Drive, Suite 100, Irvine, CA 92617.

5. On information and belief, SEC is a corporation organized and existing under the laws of the Republic of Korea, with its principal place of business at 129 Samsung-ro, Yeongtong-gu, Suwon, Gyeonggi, 16677, Republic of Korea. On information and belief, SEC is the worldwide parent corporation for SEA and SSI, and is responsible for the infringing activities identified in this First Amended Complaint. On information and belief, SEC's Device Solutions division is involved in the design, manufacture, use, offering for sale and/or sales of certain semiconductor products, including the Accused Instrumentalities as defined below. On information and belief, SEC is also involved in the design, manufacture, and provision of products sold by SEA.

6. On information and belief, SEA is a corporation organized and existing under the laws of the State of New York. On information and belief, SEA, collectively with SEC, operates the Device Solutions division, which is involved in the design, manufacture, use, offering for sale

Samsung Electronics Co., Ltd.
Ex. 1044, p. 2

and/or sales of certain semiconductor products, including the Accused Instrumentalities as defined below. Defendant SEA maintains facilities at 6625 Excellence Way, Plano, Texas 75023. SEA may be served with process through its registered agent for service in Texas: CT Corporation System, 1999 Bryan Street, Suite 900, Dallas, Texas 75201. SEA is a wholly owned subsidiary of SEC.

7. On information and belief, SSI is a corporation organized and existing under the laws of the State of California. On information and belief, SSI, collectively with SEC, operates the Device Solutions division, which is involved in the design, manufacture, use, offering for sale and/or sales of certain semiconductor products, including the Accused Instrumentalities as defined below. On information and belief, Defendant SSI maintains facilities at 6625 Excellence Way, Plano, Texas 75023. Defendant SSI may be served with process through its registered agent National Registered Agents, Inc., 1999 Bryan St., Ste. 900, Dallas, TX 75201-3136. On information and belief, SSI is a wholly owned subsidiary of SEA.

8. On information and belief, Defendants have used, sold, or offered to sell products and services, including the Accused Instrumentalities, in this judicial district.

II. JURISDICTION AND VENUE

9. Subject matter jurisdiction is based on 28 U.S.C. § 1338, in that this action arises under federal statute, the patent laws of the United States (35 U.S.C. §§ 1, *et seq.*).

10. Each Defendant is subject to this Court's personal jurisdiction consistent with the principles of due process and/or the Texas Long Arm Statute.

11. Personal jurisdiction exists generally over the Defendants because each Defendant has sufficient minimum contacts and/or has engaged in continuous and systematic activities in the forum as a result of business conducted within the State of Texas and the Eastern District of Texas. Personal jurisdiction also exists over each Defendant because each, directly or through

Samsung Electronics Co., Ltd.
Ex. 1044, p. 3

12.3 A 1.2V 64Gb 341GB/s HBM2 Stacked DRAM with Spiral Point-to-Point TSV Structure and Improved Bank Group Data Control

Jin Hee Cho, Jihwan Kim, Woo Young Lee, Dong Uk Lee, Tae Kyun Kim, Heat Bit Park, Chunseok Jeong, Myeong-Jae Park, Seung Geun Baek, Seokwoo Choi, Byung Kuk Yoon, Young Jae Choi, Kyo Yun Lee, Daeyong Shim, Jonghoon Oh, Jinkook Kim, Seok-Hee Lee

SK hynix, Gyeonggi, Korea

With the recent increasing interest in big data and artificial intelligence, there is an emerging demand for high-performance memory system with large density and high data-bandwidth. However, conventional DIMM-type memory has difficulty achieving more than 50GB/s due to its limited pin count and signal integrity issues. High-bandwidth memory (HBM) DRAM, with TSV technology and wide IOs, is a prominent solution to this problem, but it still has many limitations: including power consumption and reliability. This paper presents a power-efficient structure of TSVs with reliability and a cost-effective HBM DRAM core architecture.

HBM's major obstacle for achieving high bandwidth and low power is the heavy capacitive load due to the thousands of TSVs in 8Hi stacks. The previous HBM (with multi-drop TSV) had limited capabilities for managing the 8Hi heavy TSV loading, which is one of the major obstacles for achieving both bandwidth and density. To reduce the heavy loading of a conventional multi-drop TSV structure, a spiral point-to-point (P2P) TSV structure is proposed. Figure 12.3.1 shows a conventional multi-drop TSV structure [1] and the proposed spiral P2P TSV structure. Compared to a multi-drop TSV structure, where each core die has its own TX, RX, and a 4-to-1 MUX, the proposed spiral P2P TSV structure has only three sets of TX and RX for an 8Hi stack. Furthermore, it eliminates the need for a channel selection MUX and its complicated routing. The current consumption to drive TSVs was reduced by 30%, due to the reduced capacitance, and its slew rate was increased from 3.4 to 4.9V/ns as shown in Fig. 12.3.1.

Another challenge for HBM development is achieving high yield for TSVs and micro-bumps (μ-bump): the total yield of a stacked chip is obtained by squaring the yield of one TSV by the number of TSVs. Furthermore, in case of an 8Hi-stack HBM, one failed TSV connection causes 9 dies to be discarded. Therefore, a TSV repair technique is essential to compensate for TSV yield. Unlike conventional TSV repair techniques [2] that need test equipment to test for the open/short-state of a TSV connection, the proposed automatic TSV self-repair technique, shown in Fig. 12.3.2, performs open/short tests during the boot-up sequence, without the need for test equipment or fuses. To detect a weak TSV connection, a core-side strong PMOS and a base-side NMOS, for which the leakage current is controlled by a bias voltage, are turned on at the same time. The quantized voltage level of the TSV is stored in a latch. A base-side strong PMOS and core-side NMOS are also turned on to confirm the TSV connectivity. By sequentially reading out the latched results, the locations of TSV failures can be determined. And the TSV self-repair operation can be performed by changing core die to find exact positions of failed TSVs. The slice ID signal sent from the base die to the core die is changed using the test mode to make each core die behave like the top slice. The proposed architecture also supports a conventional current scan, without additional circuitry, by jointly using the PMOS' as a current source and the DFF as a switch-enable signal shifter. The TSVs used to control the repair operations cannot make use of self-repair, but instead they exist in pairs to ensure robust operation. The repair procedure is performed during the boot-up sequence, so that users do not need to execute post-package repair.

HBM2 has a pseudo-channel function [1], which decreases the page size in half while doubling the effective number of banks to improve DRAM core timing such as t_{FAW} (four-active-window) and t_{CCDL} (column-to-column access timing). Prior work [1] used four channels in a slice, whereas the proposed architecture includes only two channels per core die via a spiral TSV structure. Each channel is divided into two pseudo channels: a pseudo channel has 16 independent banks and 64 IOs. Because of the large number of IO lines, the area overhead of HBM is significant, thus leveraging increased pre-fetch, which is a common method in conventional DRAM, is restricted. The improved bank group control, shown in Fig. 12.3.3, is proposed to mitigate speed and area penalties. For a 4b pre-fetch operation, each IO has 4 internal data lines, and there should be 64×4b×4 bank group IOs (BG_IO) and 256 global IOs (GIO) per each pseudo channel. However, the proposed architecture has BG_IOs with a 2b pre-fetch (feasible due to a relaxed t_{CCDL}) and GIO has a 4b pre-fetch, that is divided into left and right, to keep the effective line similar to the 128 GIOs. Figure 12.3.3 shows the timing diagram of the proposed architecture where column commands and Y-pulses are divided into even and odd groups by the order of the command input (BL0, 1 and BL2, 3). Therefore, the core die effectively has twice the number of data lines and column addresses, and it has twice the core timing margin without significant area penalty.

The power distribution in the HBM core die is also a challenge because significant power is consumed in a small area. The IR drop in a 3D stack structure causes t_{CDDL} degradation in the DRAM core operation. To mitigate this issue, additional bank-power TSVs are placed in-between row decoders: these directly supply power from base die to power-hungry core areas. Another merit of bank power TSVs is the ability to share the power distribution network among each core die, which greatly reduces the peak IR drop for an 8HI-stack as shown in Fig. 12.3.4. Based on PDN simulation results, more than 50% of IR drops are diminished in IDD4W (gapless write; worst pattern for core IR drop) compared to a previous version without bank power TSVs.

The base die and core dies in HBM2 all have temperature sensors. Memory controllers can read out the 8b of highest temperature code among the core dies, 8b for the base die temperature code, and 1b of catastrophic trip threshold (CATTRIP). Therefore, 4Hi and 8Hi stacks require 36 and 72 TSVs. The proposed serial temperature read-out scheme uses only 2 TSVs, one for the temperature code and the other for the CATTRIP. Figure 12.3.5 shows the core die temperature read-out scheme. The core die temperature codes are shifted, in descending order of core dies, by the shift clock that is generated in base die. The base die stores 8b of code in the CTEMP register, and compares it with the code stored in the CTEMP_MAX register, which stores the maximum temperature code observed. The CATTRIP scheme is depicted in the right side of Fig. 12.3.5. Since CATTRIP must reflect all information from every core die and base die, all dies share one TSV using wired-OR logic. The base die always turns on a pull-down transistor to drive the CATTRIP μ-bump LOW. When any die reaches the limit temperature (e.g. 125°C), it generates a CATTRIP flag to make the TSV HIGH, which is driven onto the CATTRIP μ-bump.

Since the PHY μ-bump cannot be probed, all tests of the HBM were performed using a direct access ball (DA). However, because of the operational characteristics of the PHY IO and the necessity for system implementation verification, an active interposer package (AIP), depicted in Fig. 12.3.6, is proposed. HBM DA and PHY operation can be verified, and 2-channel interleaving technology can be applied to determine the influence of the independent operations between channels. The AC characteristics of the HBM PHY can be measured: such as the input setup/hold, the 1-pin input setup/hold, and t_{DV} (data valid window). Since the signal integrity between the controller and the HBM in the SiP is reflected on the AIP, the LFSR/MISR can be operated between the active interposer and the HBM under similar conditions. Test flexibility is improved by applying a serial-test-mode input technique similar to IEEE1500 [3].

Several key technologies have been introduced to address impediments to increasing bandwidth for HBM memories. The spiral-P2P scheme and the TSV self-repair scheme both provide good solutions for managing the heavy 8Hi TSV loading. The improved bank-group data control optimizes area overhead and DRAM core speed. Additional bank power TSVs reduce the IR drop at the bank side by 50%. The HBM shmoo results, shown in Fig. 12.3.6, shows a 341GB/s 8Hi known-good-stack dies (KGSD) gapless-read operation at 1.2V and 105°C, and 320GB/s at 1.15V and 105°C. The chip micrographs for the 8Hi-stacked 8Gb DRAM dies and the base die are shown in Fig. 12.3.7.

References:

[1] J. C. Lee, et al., "A 1.2V 64Gb 8-channel 256GB/s HBM DRAM with peripheral-base-die architecture and small-swing technique on heavy load interface," *ISSCC*, pp. 318-319, 2016.

[2] D. U. Lee, et al., "An Exact Measurement and Repair Circuit of TSV Connections for 128GB/s High-Bandwidth Memory(HBM) Stacked DRAM," *IEEE Symp. VLSI Circuits*, 2014.

[3] JEDEC Standard High Bandwidth Memory (HBM) DRAM Specification, 2015.

Netlist EX2007
Samsung v Netlist
IPR2022-01427

 978-1-5090-4940-0/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: Irell & Manella LLP. Downloaded on October 03,2022 at 19:09:34 UTC from IEEE Xplore. Restrictions apply.

Appx6113

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.
Petitioner,

v.

NETLIST, INC.
Patent Owner.

Case No. IPR2022-01427
Patent No. 9,318,160

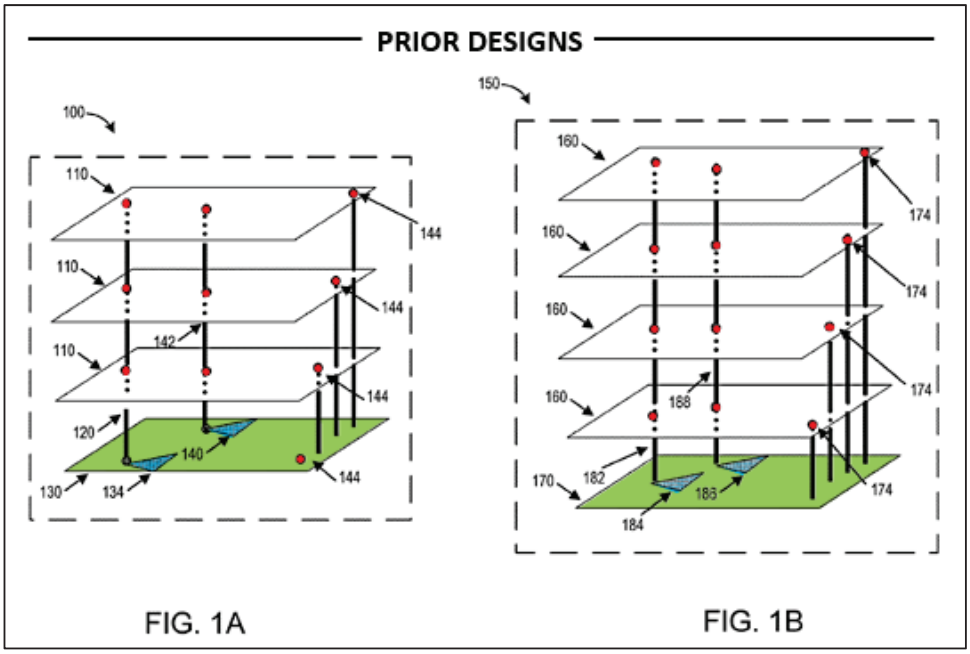
DECLARATION OF MICHAEL C. BROGIOLI, PH.D. IN SUPPORT OF PATENT OWNER'S RESPONSE

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

V. Overview of the '060 Patent

33. I discuss below pertinent background technology, including terminology, as related to the '060 and '160 inventions.

34. As shown in the specification, then proposed 3D-stacked memory packages each contained array dies and a control die interconnected using a multi-drop configuration, whereby each data die interconnect is in electrical communication with each array die.



('060 patent, Figs. 1A-1B annotated)

35. Of note, in these configurations, to avoid data collisions, only one of the array dies 110 or 160 would receive or output data via a commonly shared TSV (120 or 182) at a given time, with the data ports on the other array dies disabled.

Nevertheless, all dies, even ones with disabled data ports, are considered to be in electrical communication with the TSV.

36. The inventor realized and taught that the prior art multi-drop configurations typically required a large driver to drive the signals along a die interconnect, but "generally a larger driver not only consumes more space on the control die, but also consumes more power." '060 patent, 2:8-15.

> 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 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.

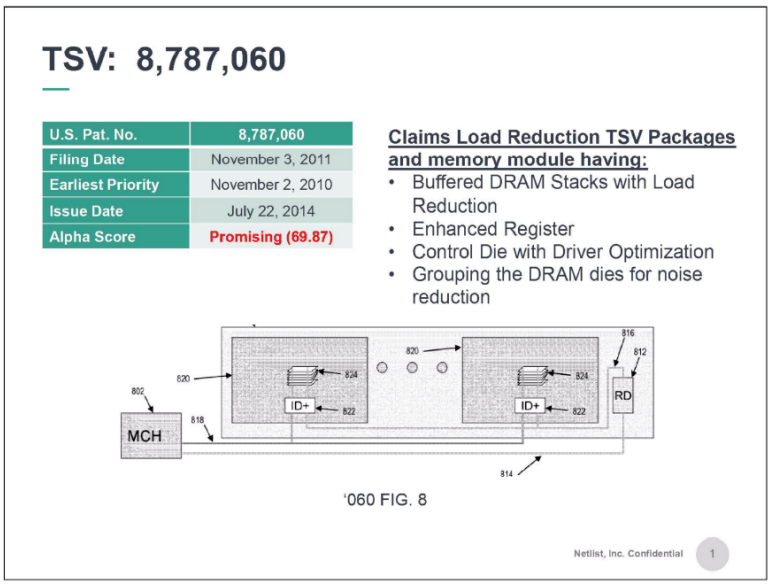
37. To overcome the above problem and to reduce the power consumption as well as the driver size, the inventor teaches that

> 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.

'060 patent, 4:22-29.

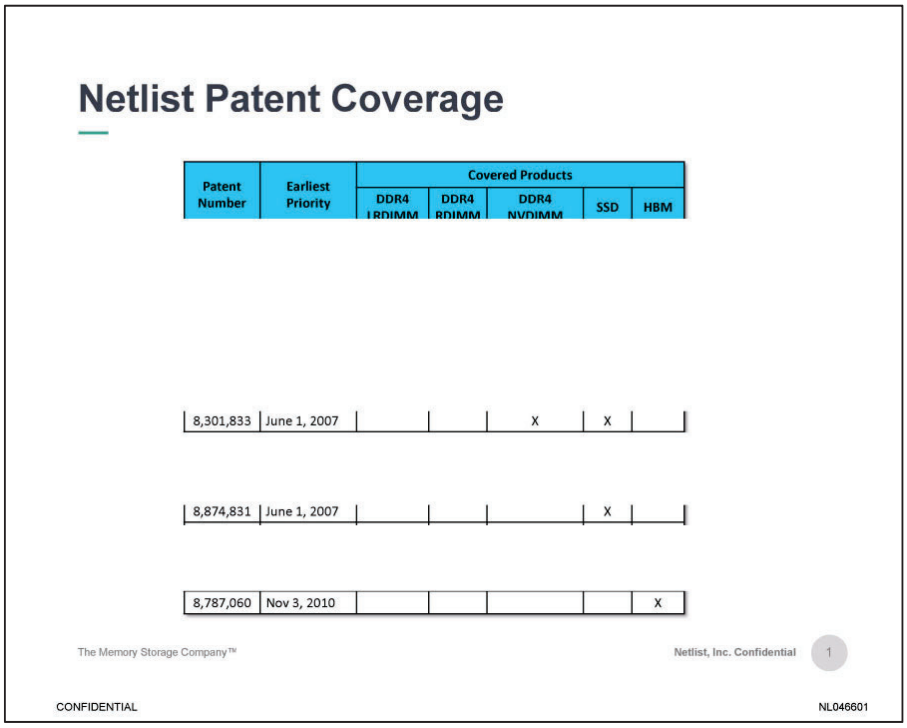
38. Thus, as shown in Figures 2 and 3, instead of electrically communicating with every array die, the die interconnects are in electrical

54. I also understand that in April, 2015, Netlist shared a presentation regarding Netlist's technologies and patents with Samsung. *See* EX2032 [PX1778]-0004; EX2027 [2023-04-17 Trial Tr. (Milton)], 220:17-222:1. The presentation disclosed Netlist's TSV design covered by the '060 patent which claims "Load Reduction TSV Packages and memory module having: Buffered DRAM Stacks with Load Reduction[;] Enhanced Register[;] Control Die with Driver Optimization[;] Grouping the DRAM dies for noise reduction." *See*, *e.g.*, EX2032 [PX1778]-0004; EX2031 [PDX2]-58.[3]



[3] Netlist's corporate witness, Mr. Scott Milton, testified at trial that Samsung was given a presentation of this document. EX2027 at 220:17-221:1, 221:13-19.

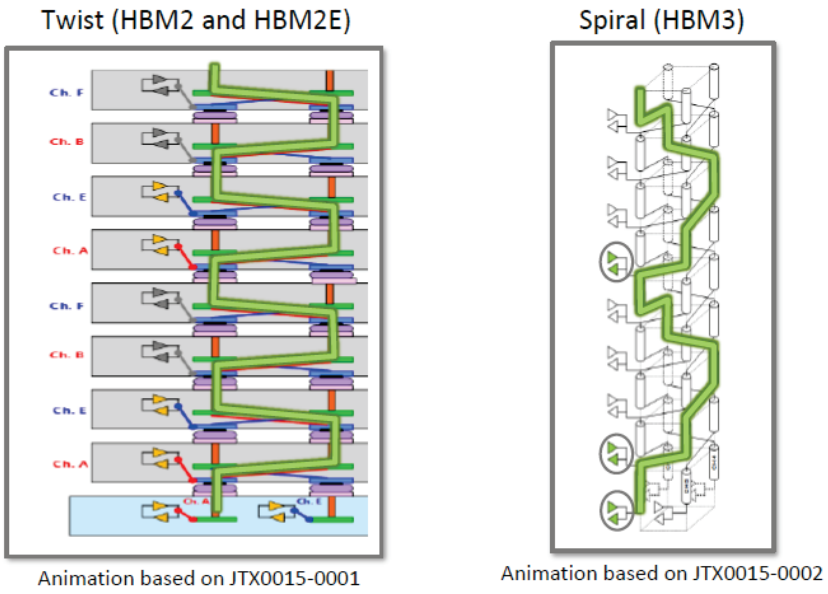
55. I understand that, on November 8, 2016, Mr. Ji Bum Kim of Netlist sent Samsung a summary of Netlist patent portfolio, which explicitly stated that the '060 patent covers HBM Products. PX0446-0002; EX2027 [2023-04-17 Trial Tr. (Milton)] at 222:8-24.



56. Samsung launched its first generation HBM2 product, Flarebolt in 2016. EX2034 at 8 ("Samsung launched HBM2 in Q1 2016.") (https://developer.download.nvidia.com/video/gputechconf/gtc/2019/presentation/s91028-2-pivotal-memory-technologies-enabling-new-generation-of-ai-workloads-presented-by-samsung.pdf); EX2035 (https://hothardware.com/news/samsung-begins-mass-production-of-the-worlds-fastest-dram-hbm2-launched); EX2055

(https://semiconductor.samsung.com/news-events/tech-blog/high-bandwidth-memory-is-catching-on-for-data-intensive-computing/).

57. Samsung's HBM2 products as well as later generation HBM products utilize the patents' teachings of reduced capacitive load for optimized power consumption and driver size. *See*, *e.g.*, EX2030, DDX4-53.



VI. Claim Construction

58. I have been asked to analyze whether the asserted combinations render the claims of the '060 patent and '160 patent obvious according the plain and ordinary meaning of all claim terms except where the District Court has construed the terms at Samsung's or Micron's urge. I understand that Samsung contends that no express constructions are needed for this position. Pet. 21.

> input/output circuits cannot be separately provided for the respective first and second ranks Rank0 and Rankl.

Kim, [0011].

65. Because of this inability to provide a shared input/output circuit that could select one of the GIO lines, Kim notes that in the typical semiconductor memory apparatus, "in the case where the write operation of the first rank Rank0 is performed and then the read operation of the second rank Rank1 is performed, a data collision is unavoidable." *Id.*, [0042]. Kim attributes the cause of the problem to the fact that "the first and second ranks Rank0 and Rank1 cannot share a data input/output circuit." *Id.* Of note, the shared data input/output circuit in the prior art did not have the ability to select between GIO1 and GIO2 lines.

66. According to Kim, the data collusion occurs because in Kim's memory system, the memory device would output read data immediately on a global input/output line (*e.g.*, GIO1 or GIO2) while it takes time for write data inputted through the data pad (*e.g.*, DQ0 or DQ1) to appear on the global input/output line because write data has to propagate through and be processed by the input/output section (30 or 40), which includes a data buffer (which holds data), a data aligning unit (which may align the data to data strobe or clock) and a write sense amplifier (which amplifies aligned write data):

> Assuming that the first and second ranks Rank0 and Rank1 share a data input/output circuit, first and second data input/output lines operate in the same manner. Since it takes a time for data inputted through the data pad DQ to pass through the data input/output section, when the write operation of the first rank Rank0 is performed, a time A at which the data is actually loaded on the data input/output line is after a certain time interval is lapsed. That is to say, the write data is loaded on the data input/output line GIO1 or GIO2 at the time A.

Kim, [0042].

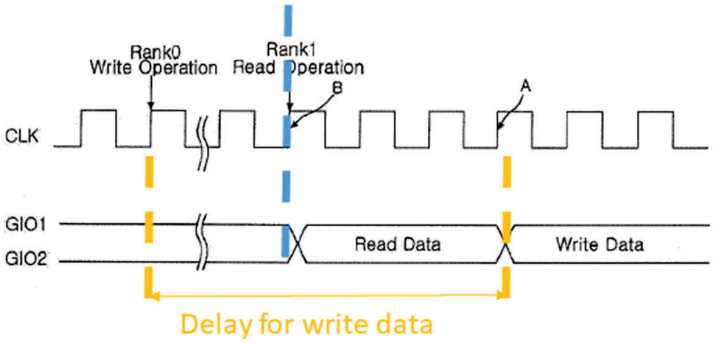
67. As a result,

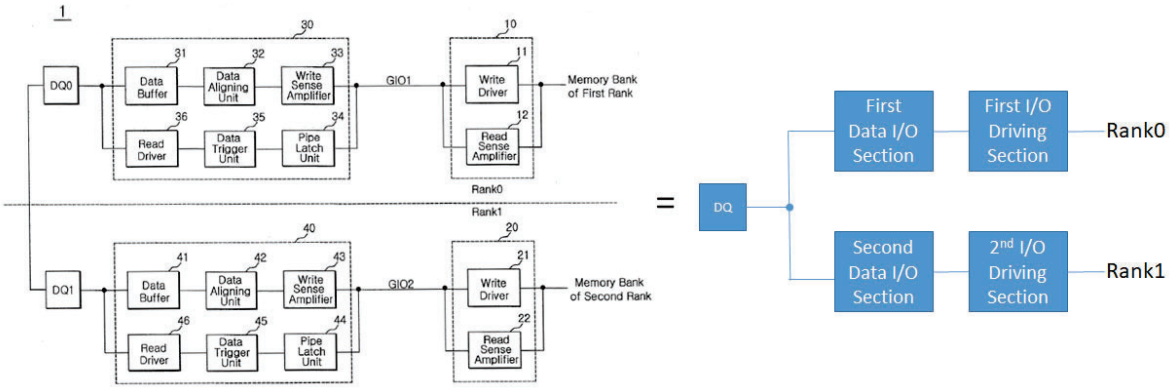
> [w]hen the read operation of the second rank Rank1 is performed following the write operation of the first rank Rank0, data of the second rank Rank1 is immediately loaded on the data input/output line GIO1 or GIO2. As a result, because the time A at which the write data is loaded on the data input/output line GIO1 and GIO2 is later than a time B at which the read data is loaded on the data input/output line GIO1 and GIO2, the data to be outputted to the data pad DQ is changed from the read data to the write data, resulting in a data collision.

Kim, [0043].

68. This is depicted below in annotated Figure 4A, where it can be seen that read data from Rank1 appears on GIO2 line at time B, almost immediately

after the read operation starts. In contrast, there is a delay for write data for a write operation to Rank0 to appear on the GIO lines at time A.





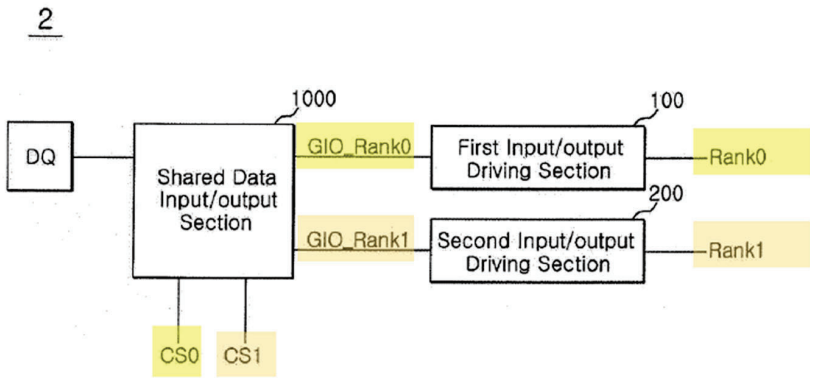
69. It should be noted that because in the typical semiconductor memory apparatus, DQ0 and DQ1 are connected, any data inputted on DQ0 (for Rank0 write operation) will not only propagate to GIO1, but will also appear on DQ1, and from DQ1, propagate via input/output section 40 to GIO2 at a later time A. Because the write data intended for Rank0 appears on GIO2 later in time than read data from Rank1, the later appeared Rank0 write data would displace the earlier appeared Rank1 read data, resulting in error to the data outputted to DQ1 in such

cross-rank read-after-write operations. The same collision would not occur in the same-rank read-after-write operations because "the read operation and the write operation for one rank are performed with a predetermined time interval set therebetween." Kim, [0042]. But apparently, in Kim's system, applying such time interval was insufficient to solve the data collision problem in cross-rank read-after-write operations, including apparently in situations where "two chips operate as one semiconductor memory apparatus" (*e.g.*, Kim, [0027]).

70. To address the cross-rank read-after-write data collision problem, Kim proposes "a data input/output circuit compris[ing] a rank selecting section configured to be selectively connected to one of a first and second ranks in response to a chip selection signal, and output data to a connected rank or receive data from the connected rank; and a data input/output section configured to output the data transmitted from the rank selecting section through a data pad to an [sic] device during a read operation, and output the data inputted to the data pad to the rank selecting section during a write operation." Kim, [0013].

71. An overview of the schematic is shown in Figure 2, reproduced below with annotation.

FIG.2



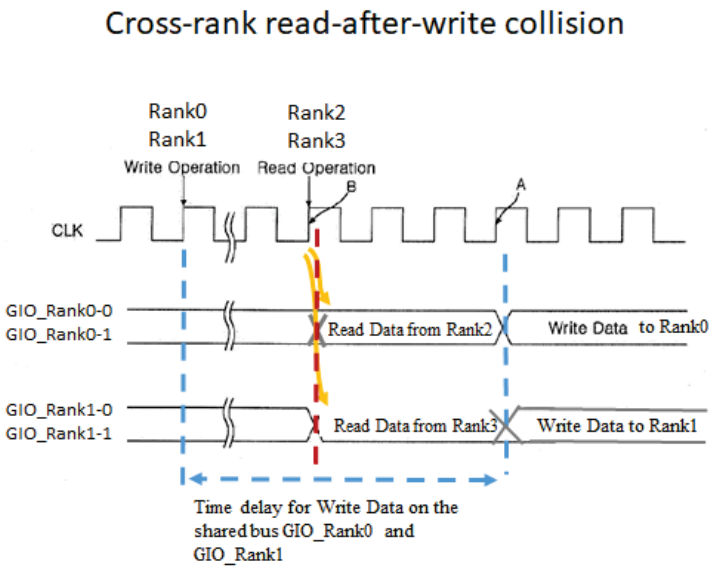
72. As shown in Figure 2, a shared data input/output section 1000 is placed between data pad DQ and data input/output lines GIO_Rank0 and GIO_Rank1. GIO_Rank0 and GIO_Rank 1 are connected, respectively, to first and second input/output driving sections 100 and 200, which are in turn connected respectively to Rank0 and Rank1 memories. *See* Kim, [0025].

73. According to Kim:

> The first input/output driving section 100 stores data transmitted through the first data input/output line GIO_Rank0 in a memory bank of a first rank Rank0 during a write operation, amplifies data stored in the memory bank of the first rank Rank0 and outputs the amplified data to the first data input/output line GIO_Rank0 during a read operation. The second input/output driving section 200 stores data transmitted through the second data input/output line GIO_Rank1 in a memory bank of a second rank Rank1 during a write operation and amplifies data stored in the memory bank of the second rank

129. As a result, between Rank0 and Rank2 (or between Rank1 and Rank3), the unavoidable cross-rank read-after-write data collisions would occur. This is because, in the absence of a rank selecting unit that selects between GIO0-0 and GIO0-1, write data placed on the shared GIO_Rank0 path would appear on both GIO0-0 and GIO0-1. Hence, write data intended for Rank0 would also appear on GIO0-1 line associated with Rank2 and write data intended for Rank2 would also appear on GIO0-0 line associated with Rank0. Similarly, in the absence of a rank selecting unit that selects between GIO1-0 and GIO1-1, write data placed on the shared GIO_Rank1 path would appear on both GIO1-0 and GIO1-1. Hence, write data intended for Rank1 would also appear on GIO1-1 line associated with Rank3 and write data intended for Rank3 would also appear on GIO1-0 line associated with Rank1. This would reintroduce the data collision problem experienced by a "typical semiconductor apparatus."

130. For instance, as in a "typical semiconductor memory apparatus," when a write operation for Rank0 occurs, it takes time for the write data to propagate through the data I/O section and shared GIO_rank0 path, before being loaded on both GIO0-0 and GIO0-1 lines. *See* Kim, [0042].

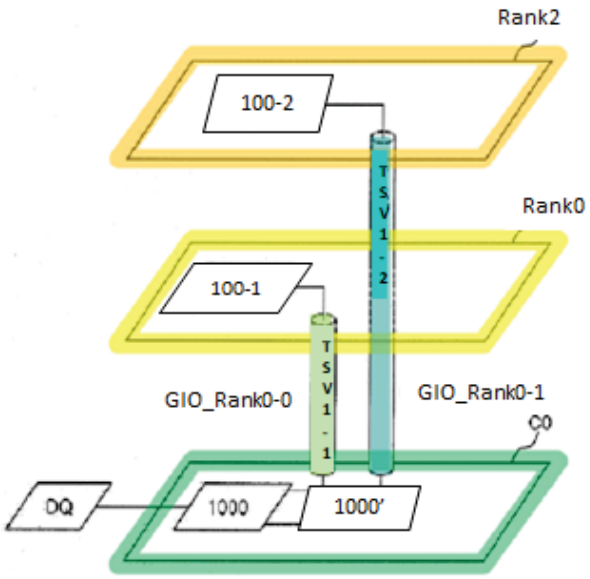


133. Similar unavoidable data collisions would occur when there is a Rank0 read following a Rank2 write, a Rank3 read following a Rank1 write, or a Rank1 read following a Rank3 write.

134. To prevent such unavoidable data collisions to occur in Kim's memory systems, a person of ordinary skill in the art would understand that the different ranks need to be placed on non-shared TSVs with rank-selecting units or other circuits between the DQ pad and the TSVs to make sure that write data wouldn't be loaded onto more than one TSV at a time. Put another way, in light of Kim's express teachings on the constraints on read/write timing for Kim's memory systems and the reasons why a different GIO is needed for each rank, a person of ordinary skill in the art would understand that there cannot be two ranks sharing

the same GIO line. They thus would not have modified Kim in the manner suggested by Samsung and Dr. Wolfe because this would re-introduce the unavoidable data collision that Kim tried to solve.

135. Just to be clear, data collisions in the cross-rank read-after-write operations can probably be avoided by using a rank-selecting unit such as the one shown for the four-rank-four-GIO configuration in Section 3, or by placing additional rank selecting units between GIO_Rank0 (GIO_Rank1) on the one hand and GIO0-0 and GIO1-0 (GIO1-0 and GIO1-1) on the other hand. But in either implementation, the four ranks would be connected to separate (and non-shared) GIO lines to prevent write data intended for one rank to appear on the GIO line(s) associated with another rank or other ranks. In such implementations, because each TSV would be in electrical communication with only one chip, while the claims require that at least one die interconnect be in electrical communication with a group of "array dies" (plural), the modification would not result in the claimed inventions. *See* below which does not include the branch/part associated with Rank1 and Rank3.



1000' = rank selection between GIO0-0 and GIO0-1 paths and paths associated with Ranks1 and Rank3 not shown

136. In summary, Dr. Wolfe has overlooked the inoperability issues that would arise if multiple dies connected to the same TSV belong to different physical ranks (*e.g.*, in a rank multiplication implementation), in particular, the "unavoidable" data collision issue in cases where a read in one rank is performed after a write in a different rank as noted by Kim in [0042]-[0043]. Kim solved the problem by including a rank selection unit 1100 that responds to chip selection signals CS0 and CS1. *Id*., [0044]. However, in Kim, the rank selection unit 1100 is before TSV1 (and TSV2) such that it wouldn't be able to control the timing for Rank0 versus Rank2 (or Rank1 versus Rank3) that are connected to the other side

of TSV1 (or TSV2). As explained, the reintroduction of the unavoidable data collision issue would dissuade a person of ordinary skill in the art from making the proposed changes to Kim.

137. I am also asked to address the possibility of solving the data collision problem in the cross-rank read-after-write operations by changing the latency for a read operation such that the read data for *e.g.*, Rank0 read operation would appear later in time than the write data for Rank2 write operation. The read timing in Kim —that is, read data would be loaded almost immediately after a read command— would be an inherent property of Kim's memory devices. Hence, the change in latency for the read operations would require a redesign of Kim's memory chips in order to meet the new timing constraint, something that Dr. Wolfe and Samsung do not suggest. This is not surprising because changing the latency could mean complete redesign of the electrical circuitry and characteristics of the memory devices, an undertaking beyond the skill level of a person of ordinary skill in the art with 2-3 years of practical experience (and those experience would also likely be for JEDEC-based DRAM devices that Dr. Wolfe said were not the basis of the invalidity theories). Moreover, the lengthening of latency would affect not only the latency in cross-rank read-after-write cases, but also all read operations (which can number in the tens of thousands or even millions per second; and the read-to-

write ratio can range from single digits to the order of hundreds). This significant increase in read latency would lead to significant reduction in data throughput, which would deter a person of ordinary skill in the art to do at a time when the normal market trend is improved, not reduced, data throughput, and faster—not slower—data speed where the memory speed was already too slow for CPUs. Another difficulty is it is impossible for a memory controller and a memory device to accurately define the boundary of the read-after-write operations to add in the appropriate delay just for a read-after-write operation, especially with varying length burst modes that cannot be interrupted. Hence, Samsung and Dr. Wolfe have neither demonstrated the technical feasibility nor benefits for increasing the read latency for Kim's memory devices. Furthermore, if lengthening read latency could have solved Kim's unavoidable data collision problem, there would be no need for the improvement that Kim describes through its Figures 2, 3 or 5 implementations. That is, there would not even be the base Kim design that Samsung is now relying on.

C. No Motivation to Combine

1. No Motivation to Combine Kim and Rajan

138. At his deposition, Dr. Wolfe admits that expanding Kim from two ranks to four ranks would be "complicated," but alleges that the changes would be

Netlist EX2023-p.81
Samsung v Netlist
IPR2022-01427

140. While Samsung alleged that a person of ordinary skill in the art would "have been familiar with the structure and circuitry used in stacked memory devices, including structure and circuitry in JEDEC-proposed three-dimensional stacking and circuitry used to access and control computer memories, including sophisticated circuits such as ASICs, FPGAs, and CPLDs, and more low-level circuits such as tri-state buffers and their corresponding electrical loads," (Pet. 5), it is not clear what particular structure and circuitry Samsung has in mind. But based on my experience and familiarity of the level of the skills for people with the education and skill level that Samsung defines, they would not have had sufficient knowledge or experience to know how to modify them while maintaining the correct timing and electrical characteristics, including to make modified Kim operable in light of Kim's timing constraint.

141. Kim, for example, does not suggest solutions other than an intervening rank selecting unit that selectively places data onto one of the two GIO lines. It does not suggest modifying latency or otherwise adjusting the electrical or timing characteristics of its memory systems.

142. A person of ordinary skill in the art would recognize that this is a multi-dimensional problem that requires changing the inherent timing and electrical properties of Kim's memory apparatuses, which could not be easily

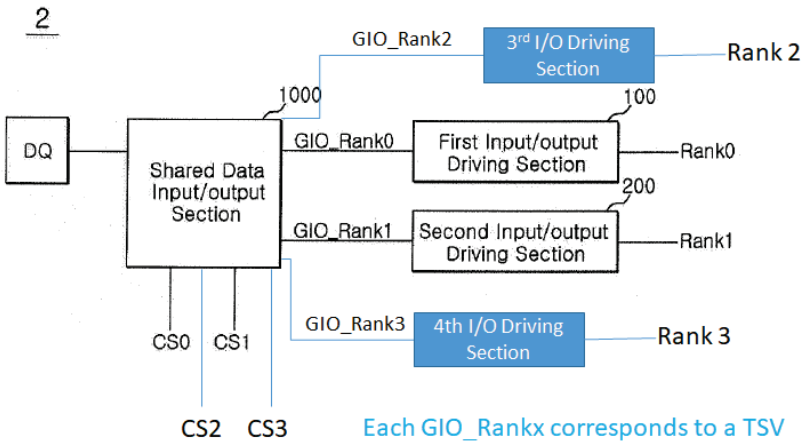
Netlist EX2023-p.83
Samsung v Netlist
IPR2022-01427

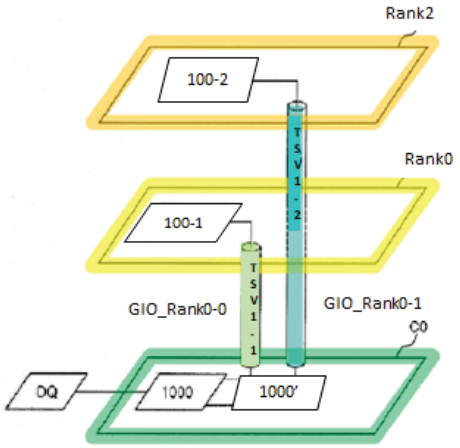
solved. This is particularly true given the relatively low level of skill proposed by Petitioner. I note that Dr. Wolfe, who has been active in the field for over 30 years, has far more experience than what is required to be a person of ordinary skill in the art. It is my opinion that a person of ordinary skill in the art, at the time of invention of the '060 and '160 patents would not be able to address the significant hurdles I have discussed above and below, and would not be motivated to try. In my opinion, Dr. Wolfe's opinions are driven by implicit hindsight bias, and not the approach that a person of ordinary skill would actually take.

143. I note that the Petition points cursorily to Rajan, the other reference involved in Ground 1. Rajan discloses DRAM circuits that are different from the claimed array dies, and the Petition contains no explanation why a POSITA would have looked to Rajan to solve the issues identified above. This is because a POSITA would understand that Rajan's grouping is not what is recited in the claims. In Rajan, DRAM circuits are grouped not by what data conduit and interconnect they are attached to, but based on whether they share a data bus. A data bus can be shared without sharing interconnects, especially in Rajan. This is because Rajan uses wire bonds for connecting its DRAM circuits to the data busses, and wire bonds ordinarily cannot be shared in the way that TSVs are shared. *See* Rajan137, [0050] ("data signals may be wired as one common bus,

Netlist EX2023-p.84
Samsung v Netlist
IPR2022-01427

be met. Hence, the requirement that at least one die interconnect be in connection with a group of array die***s*** (plural) would be incompatible with Kim's system.





1000' = rank selection between GIO0-0 and GIO0-1 paths and paths associated with Ranks1 and Rank3 not shown

150. For those reasons, a POSITA would understand that having Kim's TSV in electrical communication with more than one memory chip would result in data collisions and the resulting structure would be inoperable. A person of

ordinary skill in the art would not have had a reason or a reasonable expectation of success to modify Kim so that multiple dies share the same TSV signaling path. That is, Samsung has not shown that modifying Kim in the manner suggested in the Petition or newly at the deposition by Dr. Wolfe would have yielded the recited configuration in an operable manner.

151. As such, Petitioner has not shown that a person of ordinary skill in the art would have modified Kim in light of Rajan for a shared bus approach or would have been able to do so with a reasonable expectation of success.

2. Other "die grouping" patents also do not provide a motivation to modify Kim to arrive at the claimed inventions

a) Ruckerbauer (EX1024)

152. Samsung also cites to U.S. Patent No. 7,796,446 ("Ruckerbauer"), Figure 5 for the general concept of die grouping. Pet. 37-38.

Netlist EX2023-p.91
Samsung v Netlist
IPR2022-01427

IX. The Combination Proposed in Ground 3 Does Not Disclose or Suggest All Limitations or Render Obvious the Inventions (Also Applies to the Last Element of Claim 1 of the '160 Patent)

210. Ground 3 incorporates Wyman into Ground 1, asserting that the combination also renders claims 15 and 20-28 obvious. Samsung cites Wyman for the alleged disclosure that shorter paths have less load and thus require less drive, while longer paths have more load and require a larger drive. Pet., 84, 88-89, 90 n.8. In particular, Samsung alleges that a person of ordinary skill in the art would incorporate Wyman in Kim "to improve power efficiency by using a smaller driver … for the shorter TSV1 … and a larger driver … for the longer TSV2 …, without using the full strengths of a driver which Wyman teaches would be 'wasteful and inefficient.'" Pet. 85.

211. Wyman provides a single driver circuit with "a set of selectable drivers each having an individual drive capability." Wyman, 1:36-37. In Wyman, "depending on the necessary signal strength, one need only select the corresponding tap which supplies no less than the amount necessary for the particular element of device." *Id.*, 2:66-3:1. For example, a driver circuit shown in Figure 5 with T0-T3 taps are used to drive loads associated with through-chip via connections shown in Figures 8A-8D. *Id.*, 6:15-50. Each tap provides a drive strength sufficient to drive a range of loads. For example, T1 tap could drive loads

Netlist EX2023-p.127
Samsung v Netlist
IPR2022-01427

in a range between a load associated with the connection shown in Figure 8B and that associated with the connection shown in Figure 8C. *Id.*, 6:38-43.

212. Put another way, Wyman does not teach providing individual drivers to different TSVs. Instead, Wyman discloses a single driver with sub-driver components that can be selectively combined to drive loads falling within given ranges. Wyman, Fig. 4, 4:46–51 ("It should now be understood that the approach provides, as an advantage, a level of design flexibility because it enables a designer to adjust and minimize power consumption through reduced requirement runs by specifically tailoring driver outputs according to the particular element or device's requirements, specification or application."). That is a dissimilar approach from assigning different drivers to different TSVs, and my opinion is that a person of ordinary skill in the art would not arrive at the inventions. Specifically, given Wyman's disclosures, a person of ordinary skill in the art would not have even used different drivers, let alone different driver sizes, to drive TSV1 and TSV2 as asserted. *Contra* Pet. 85.

213. First, Kim's TSV1 and TSV2 are used at different times, and Wyman discloses the use of a single driver circuitry to drive different connections. That would suggest to a person of ordinary skill in the art the use of Wyman's allegedly more "efficien[t]" driver circuitry to drive TSV1 and TSV2, instead of using

Netlist EX2023-p.128
Samsung v Netlist
IPR2022-01427

separate drivers for each of TSV1 and TSV2. For instance, Wyman expressly teaches that its driver circuitry can be used to drive a connection "between [a] mother chip with a fan out to both daughter 2 … and daughter 3" chips. Wyman, 6:44-50, Figs 5, 8D. In other word, a person of ordinary skill in the art would have used the Wyman-style driver circuitry, including connection to appropriate taps as needed, to drive TSV1 and TSV2. (As noted below, TSV1 and TSV2 can be driven from the same tap if they have sufficiently close loads, in which case, a driver circuitry like that shown in Wyman, Figure 2 may suffice).

214. Implementing a Wyman-style driver circuitry in Kim would therefore not result in a first driver for TSV1 and a second driver for TSV2, because a single driver circuitry is used. And because no first or second drivers are used, there would also be no drivers of different sizes.

215. Second, Wyman's driver circuitry, once fabricated, would be regarded by a person of ordinary skill in the art as a single driver of a given size, regardless of the number of driver components used. For example, Dr. Stone was asked a question regarding the physical size of a driver that has ten transistors but uses a subset of the transistors for some of the applications. Stone, 149:24-150:8. He confirmed that it would be the same driver of the same physical size, even if only a subset was used:

Netlist EX2023-p.129
Samsung v Netlist
IPR2022-01427

Q So the -- it's -- it's the same example as before where you have a ten-transistor driver, right, where a subset of them can be used to drive the load, or all ten can be used to drive the load; right?

A I got it.

Q Okay. And in -- in that example, you're saying that it's the same driver of the same physical size, whether you use the subset or whether you use all ten of them; right?

A That's correct.

Id.; *see also* Stone, 141:18-145:14, 145:24-146:7. In other words, even if different taps of a Wyman-style driver are used to drive TSV1 and TSV2, that would still involve the use of a driver of the same size.

216. Third, Wyman and Kim do not suggest that the two TSVs with different lengths would result in sufficiently different load to even warrant using different taps in a Wyman-style driver. For instance, load contribution from TSV "may be negligible compared to load contribution from the array dies." *See* EX1001, 4:36-38. In such situations, same driver strengths can be used and a driver circuitry like that shown in Wyman, Figure 2 will suffice.

217. I also note that a person of ordinary skill in the art would understand that it would not be desirable to have individual taps for every different load, no matter how small the difference, in a Wyman-style driver as each additional tap would add control and signaling complexity to the design. Using a same tap for a

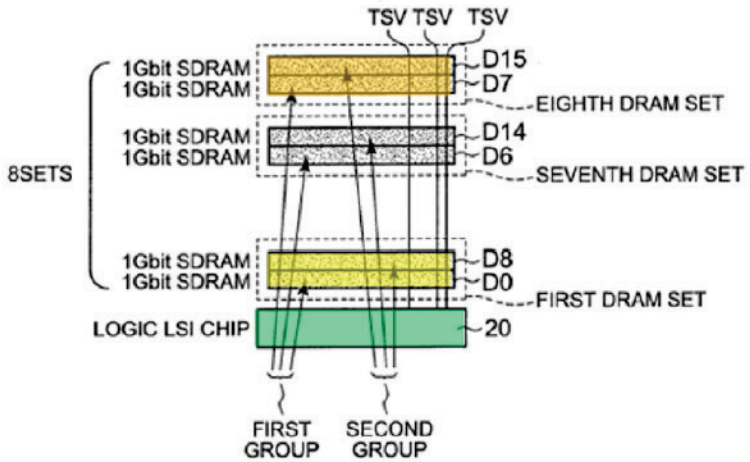
range of load would ease manufacturing variability and improve yields, especially given the number of TSVs and associated drivers.

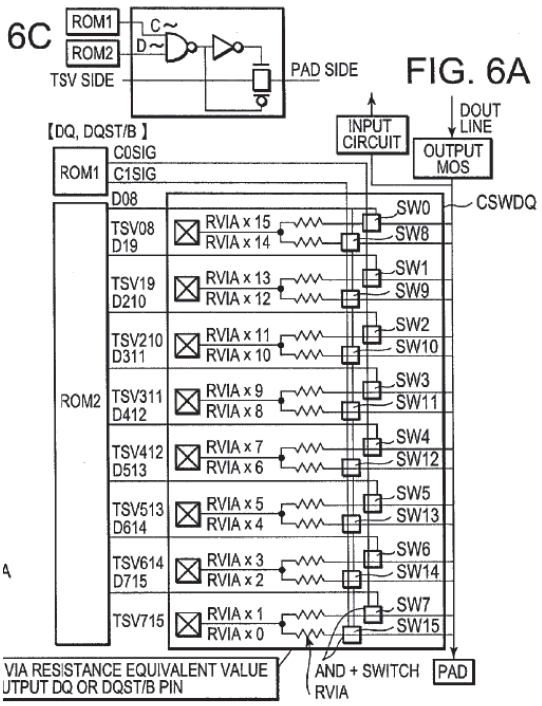
X. The Combination Proposed In Ground 4 Does Not Disclose or Suggest All Limitations or Render Obvious the Inventions

A. Overview of Riho

218. Riho discloses a memory package having sixteen SDRAM chips D0-D15 placed on top of a logic chip. Riho divides the 16 SDRAM chips into two "chip selection groups": D0-D7 and D8-D15. EX1016, [0028] and [0033]. In each access cycle, only the eight dies from the same chip selection group are accessed. *Id.*, [0033]. "Since the first and second groups are mutually exclusively controlled by the control chip 20 at the first and second access cycles, each of the TSVs corresponding to one I/O bit is shared by the first and second groups." *Id.*

FIG. 1





225. The above disclosures make clear that Riho's TSV structures adopt a multi-drop structure in which each of the TSVs is electrically coupled to and electrically communicate with each of the switch circuit of each die. *See* Riho, Figs. 4-6. For instance, as shown in Figure 4, each of the TSVs penetrates through each of the SDRAMs, even though segments of the TSVs are not used. *See id*., Fig. 4, [0045].

226. But, as shown in Figures 5 and 6A, each TSV is connected to circuits in SDRAMs that do not use that TSV. For instance, as shown in Figure 5, TSV715 is connected in similar fashions to SDRAMs D0-D15. Riho, [0019], Fig. 5.

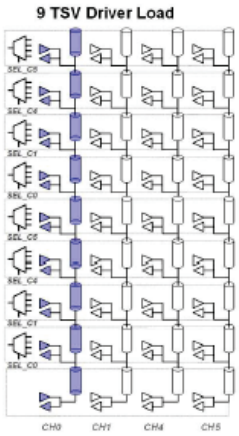
TSV715 does not send or receive data to SDRAMs D0-D6 or D8-D14, not because the electrical connections between it and these SDRAMs differ from the electrical connections between it and D7 or D15. Rather, no data is exchanged between TSV715 and D0-D6 or D8-D14 because the switch circuits sa's and sb's in D0-D6 and D8-D14 are all turned off in response to the values outputted by ROM1 (deciding which chip selection group) and ROM2 (deciding which TSV to use). See id., [0072]-[0075], [0091]. In this arrangement, each of Riho's TSVs is connected to a data port of every chip in the stack even though some of the ports are turned off.

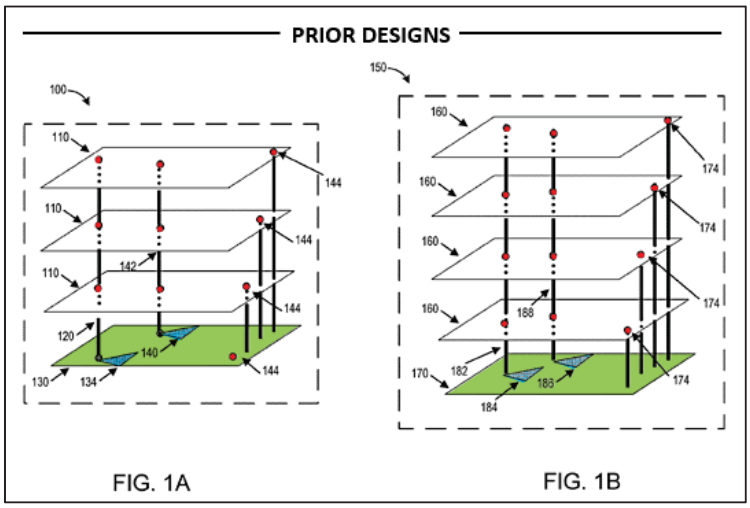
227. A POSITA would also recognize that there is electrical coupling between each data signal TSV and each SDRAM from Figure 6A, which illustrates the circuit for a SDRAM. *See id*., [0020]. Each SDRAM is supposed to have the same layout. *Id*., [0080] ("Referring now to FIG. 6A, the SDRAM chips forming the semiconductor device shown in Fig. 5 will be described in more detail"). As shown in Figure 6A, "the switch elements SW0 to SW15 respectively connected to the data signal DQ TSV08 to TSV715 on each SDRAM chip are collectively given as a chip switch circuit portion CSWDQ." *Id*., [0083].

228. A POSITA would recognize that Riho's limited, multi-drop topology structure does not differ from a conventional multi-drop architecture (see structure

Netlist EX2023-p.138
Samsung v Netlist
IPR2022-01427

on the left below and Figure 1A of the '060 patent), which does not fall within the scope of the challenged claims.





See, e.g., EX2007 at 2; EX1001, Figs. 1A-1B.

229. This is because, in the multi-drop structure, even though some of the driver circuits can be turned off, each TSV is still connected to a port on each die, thus experiencing a corresponding increase in load. *Id*., 1 ("The previous HBM (with multi-drop TSV) had limited capabilities for managing the 8Hi heavy TSV loading, which is one of the major obstacles for achieving high bandwidth and density.").

be excluded from the "plain and ordinary meaning" of "chip select signals." *See* EX2004 at 33-34.

285. In addition, Petitioner's expert also testified that he was not relying on a combination that used SDRAM/DRAM chips, only non-DRAM ones. Wolfe, 94:10-18. In that case, as with the Kim grounds (the discussion for which is incorporated), chip-select signals need not be and should not be sent to the respective memory dies because the non-DRAM devices do not need those signals for operation and it would be wasteful to add manufacturing complexity to construct TSVs for unneeded signals.

F. Claim 7

286. For the same reasons stated for Ground 2, Samsung has not shown that Riho discloses or renders obvious claim 7 limitation.

287. The reasons include: (1) under Samsung's theory that "only one of the two SDRAMs in the pair [is] connected to [an associated] TSV at a time," if true (which I have shown above is not correct because Riho's TSVs are each in electrical communication with every die in the stack), the load of the data conduit would include only the load of a single chip and not the load of the group of array dies (allegedly two) as required for element 7(b); (2) under Samsung's theory that "only one of the two SDRAMs in the pair [is] connected to [an associated] TSV at

Netlist EX2023-p.177
Samsung v Netlist
IPR2022-01427

a time," each TSV would be in electrical communication with only a single chip. This does not satisfy the requirement in 1.d.1-1.d.2 that at least one die interconnect be in electrical communication with a group of array dies; and (3) under Samsung's theory that "only one of the two SDRAMs in the pair [is] connected to [an associated] TSV at a time," there would be no need to consider the number of array dies in a group in consideration of the respective load of the TSVs. That is because regardless of the number of dies in the group or the respective load of the TSVs, under Samsung's theory, the total load on the data conduit would only include the load of a single chip (and that of the load attributable to TSV segments which would be the same for all TSVs). There would therefore be no need to select the number of dies in a memory chip group/set in consideration of the load on the associated TSVs.

G. The Combination Proposed In Ground 5 Does Not Disclose or Suggest All Limitations or Render Obvious the Inventions

288. Ground 5 incorporates Riho2 into Ground 4, asserting that the combination also renders claims 15 and 20-28 obvious. However, the addition of Riho2 does not cure any of the deficiencies identified with respect to Ground 4, nor does the Petition assert as much.

Netlist EX2023-p.178
Samsung v Netlist
IPR2022-01427

introduced by having to divide addresses by a number that is not power of 2, and to have word sizes that are not a power of two, which would generally be difficult and complicated and beyond the skill level of a person of ordinary skill in the art. *Cf.*, Wolfe, 44:21-45:4.

296. As to the Riho combination, as noted above, all the data TSVs are of the same length and as disclosed, each TSV is shared by all the chips in the stack (or two chips under Samsung's theory). So there is no disclosure or suggestion of different TSV lengths and balancing the potential load difference with a different number of chips in electrical communication with the TSV.

297. Furthermore, given the requirement that "array dies" be different from DRAM circuits and Dr. Wolfe's position that he was not relying on SDRAM/DRAM chips, Samsung has not explained how the requirement for "DRAM packages" in claims 10-20 is to be met.

XIII. Other Considerations

298. In my opinion, the Petition fails to demonstrate the obviousness of any of the challenged claims. However, I have been instructed that certain factors may further support a finding of non-obviousness.

299. For example, the development path of the industry highlights the inventive aspect of the '060 patent. The first generation of high bandwidth memory

(HBM) solutions all had multi-drop configurations. *See, e.g.*, ISSCC 2018 Hynix Paper at 1 (noting previous generation HBM had "multi-drop TSV"). SK Hynix developed HBM1 in 2014. https://news.skhynix.com/the-story-of-sk-hynixs-hbm-development/. Likewise, the 3D-stacked DRAMs were and still are constructed with multi-drop configurations.

300. If the inventions were so obvious, there is no explanation for why the industry did not hit on the idea when they were developing those products at or years after the invention. In my opinion, this reflects what a POSITA would have done, and not the hindsight bias applied in the Petition.

Netlist EX2023-p.187
Samsung v Netlist
IPR2022-01427

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Samsung Electronics Co. Ltd.	Case No. IPR2022-01427
Petitioner,	Patent No. 9,318,160
vs.	
Netlist, Inc.,	Case No. IPR2022-01428
Patent Owner.	Patent No. 8,787,060

VIDEOCONFERENCE DEPOSITION OF ANDREW WOLFE, Ph.D.

Tuesday, June 27, 2023

Volume I (Pages 1 - 191)

Reported Remotely and Stenographically by:

JANIS JENNINGS, CSR No. 3942, CLR, CCRR

Job No. 5965315

Netlist EX2025-p.1
Samsung v Netlist
IPR2022-01427

REMOTE APPEARANCES:

For the Patent Owner Netlist, Inc.:

IRELL & MANELLA LLP

BY: H. ANNITA ZHONG, ESQ.

1800 Avenue of the Stars

Suite 900

Los Angeles, California 90067

310.277.1010

hzhong@irell.com

For the Petitioner Samsung Electronics Co., Ltd.:

BAKER BOTTS LLP

BY: THEODORE W. CHANDLER, ESQ.

FERENC PAZMANDI, Patent Agent

101 California Street

Suite 3200

San Francisco, California 94111

415.291.6200

ted.chandler@bakerbotts.com

Netlist EX2025-p.3
Samsung v Netlist
IPR2022-01427

embodiments. 11:13
Q. I think we're talking about two things. I'm 11:13
trying to figure out what kind of memory system is 11:13
at issue in Kim. 11:13
A. Oh, in Kim? 11:14
Q. I think -- yeah. And I think we started 11:14
discussing in which memory systems the problems in 11:14
paragraph 42 -- described in Kim 42 would appear. 11:14
So I was asking you, are you aware of any 11:14
NAND memory systems in which the NAND memories 11:14
are all 9 and ranked and one NAND -- and the 11:14
memory -- the NAND memory in Rank1 would perform a 11:14
read operation after a write operation on Rank1 is 11:14
performed? 11:15
MR. CHANDLER: Objection. Form. 11:15
THE WITNESS: Both Riho and Rajan disclosed 11:15
NAND memories or nonvolatile memories, which one 11:15
would equate at that time with NAND flash memory. 11:15
BY MS. ZHONG: 11:15
Q. Okay. 11:15
A. And they are organized in ranks and they 11:15
would be able to perform a read following a write or 11:15
a write following a read. 11:15
Q. Okay. And in those systems, is data 11:15
collision avoidable as described in 42? 11:15

Netlist EX2025-p.69
Samsung v Netlist
IPR2022-01427

A. Data collision is always avoidable. The question is what constraints the system designer wants to place on time. If one carefully reads 42, paragraph 42 and 43 of Kim, they're not saying that that problem is unavoidable, but merely that they think that they have an improvement. 11:15 11:15 11:15 11:15 11:16 11:16

Q. Okay. So I don't understand. Why would there be a data collision problem when you do a read operation after a write operation across the rank? 11:16 11:16 11:16

MR. CHANDLER: Objection. Incomplete. 11:16

THE WITNESS: Because in this one very narrow circumstance that is being discussed, these are burst operations and burst operations require multiple data transfers. And in this one particular example, the command to data delays have not been properly selected for optimal performance on a two-rank system. But there are many, many known solutions to that problem. 11:16 11:16 11:17 11:17 11:17 11:17 11:17

BY MS. ZHONG: 11:17

Q. Okay. Do SRAMs output data in bursts? 11:17

A. Some do. 11:17

Q. What about NAND? 11:17

A. Most do. 11:17

Q. Okay. And when you say they output a data in burst, what do you mean exactly? 11:17 11:17

Netlist EX2025-p.70
Samsung v Netlist
IPR2022-01427

drive circuits but with different taps; correct? 14:07
MR. CHANDLER: Objection. 14:07
BY MS. ZHONG: 14:07
Q. It is just that, for example, for the TSV 14:07
shown in Figure 8A, Wyman is suggesting that it's 14:08
connected to either tap T0 or tap T1; isn't that 14:08
correct? 14:08
MR. CHANDLER: Objection. 14:08
THE WITNESS: No. I disagree with both 14:08
parts of it. 14:08
BY MS. ZHONG: 14:08
Q. Okay. 14:08
A. One, the suggestions also include taps T2 14:08
and taps T3 for the particular example in Figure 8, 14:08
but also the taps that I identified previously 14:08
explains that Figure 5 can be embodied or 14:08
implemented in many different ways, including one 14:08
driver supporting multiple taps, one driver 14:08
supporting one tap, or multiple drivers supporting 14:08
each tap. So all of those are contemplated by 14:08
Wyman. 14:08
Q. Okay. But when the circuit -- what is a 14:08
driver that's being provided to, for example, 14:08
Figure 8A tap? Is it an entire 500, like the driver 14:09
portion shown in 500, or it's only a portion of that 14:09

Netlist EX2025-p.128
Samsung v Netlist
IPR2022-01427

circuit that's being provided? 14:09
A. Again, understanding that there are many 14:09
variations disclosed of circuit 500, the particular 14:09
examples indicate that you can use the taps 14:09
independently, or you can provide a separate copy of 14:09
circuit 500 for each thing that needs to be driven 14:09
after you analyze the current requirements. 14:09
And there are a number of examples given 14:09
where you can use just a portion of the total 14:09
current and you can disable other portions of 500, 14:09
or you can aggregate taps to get a larger current, 14:10
or you can supply portions of the current to 14:10
different wiring networks if they both need to see 14:10
the same signal. 14:10
Q. Okay. But the circuit size for the 500, 14:10
that is set; right? The silicon -- the space it 14:10
occupies on the silicon, that is the same regardless 14:10
of which path you happen to be using or which of the 14:10
multiple combination taps you will be using; right? 14:10
MR. CHANDLER: Objection to form. 14:10
BY MS. ZHONG: 14:10
Q. It is still going to take the entire space 14:10
for the 500, isn't it? 14:10
MR. CHANDLER: Objection. 14:10
THE WITNESS: Yes. But there's also the 14:10

Netlist EX2025-p.129
Samsung v Netlist
IPR2022-01427

teaching in Wyman that you can make 500 at various different sizes based on how you want to balance out the currents and how your analysis indicates that the currents are to be used. 14:10 14:11 14:11 14:11

So the teaching of Wyman teaches you how to make 500 in many different sizes based on what any particular net uses. 14:11 14:11 14:11

BY MS. ZHONG: 14:11

Q. So how is that different than, say, I make many different driver circuits for different loads? 14:11 14:11

A. It's a minor improvement. What this says essentially is that you can change the capability over time. So, for example, one could connect taps T2 and T3 together, and then enable T2 or T3 or both if somehow you had different current requirements at different times. 14:11 14:11 14:11 14:11 14:12

Also, this is not limited to a hardwired system. It is contemplated that you could change from one tap to another in a reasonably easy way. 14:12 14:12 14:12

And also, there is kind of -- there's a teaching here that perhaps you may want to just build this circuit at an estimated size and then make a decision later as to which tap to use after you have done some more analysis. So all those teachings are present. 14:12 14:12 14:12 14:12 14:12

Q. Okay. But in order to do that, you have to have all the -- all the possible combinations there; right?

So, for example, if you want to accommodate a range of the load -- a range of possible loads, then you need to have all four of the drivers in 500 there. You can't just build three of them and hope that somehow you will be able to accommodate a load that requires four taps; correct?

MR. CHANDLER: Objection. Form.

THE WITNESS: That is correct for a particular wire. But there is a clear teaching that you can make a determination for one wire that you might require two taps, like Figure 3, and for another wire, you might require a three-buffer system like Figure 4, and for another wire, you might require a system like Figure 5, and even within each of those, you can make them in different sizes.

So for different wires, there's a teaching that you can have different sized drivers within. There is also a teaching as to how to determine what to choose, and certainly to the extent that the claim only requires that you have a different current capability, that's disclosed even within an

Netlist EX2025-p.131
Samsung v Netlist
IPR2022-01427

already implemented version of 500. 14:14
So all those things are taught. Again, as 14:14
you mentioned, everywhere in the prior art, it was 14:14
known to be able to select the proper size driver 14:14
for a particular current load. That part in itself 14:14
was not new. 14:14
BY MS. ZHONG: 14:14
Q. Okay. So if you have two driver circuits as 14:14
shown in Figure 5, one uses only taps 0 through 2, 14:14
and the other uses all four of the drivers, do they 14:14
take up the same or different silicon space? 14:14
MR. CHANDLER: Objection. Form. 14:15
THE WITNESS: If you were to make up that -- 14:15
make that choice, they would use the same silicon 14:15
space. But, of course, if you used any commonsense 14:15
and you were only -- and you knew you were only 14:15
going to use T0, T1 and T2, you could instead 14:15
implement Figure 4, which would use less silicon 14:15
space. 14:15
BY MS. ZHONG: 14:15
Q. Okay. But if you implement, say, Figure 5, 14:15
even if you use only a subset of the drivers shown 14:15
here, would the size of the driver portion 500 14:15
differ or they would remain the same? 14:15
MR. CHANDLER: Objection. Form. 14:15

THE WITNESS: The physical size doesn't depend on how many taps you use. But as I mentioned, there are many teachings about how many taps you implement and the size of the individual taps. 14:15 14:15 14:15 14:16 14:16

BY MS. ZHONG: 14:16

Q. Okay. That's what I'm trying to understand. 14:16

So once you have implemented the circuit shown on 500, the size of the driver doesn't really change based on the number of taps you use; is that correct? 14:16 14:16 14:16 14:16 14:16

MR. CHANDLER: Objection. Form. 14:16

THE WITNESS: I think that's correct. But, again, you -- because you've already established the size. It's because you've already chosen a size. I think that that would be true of any circuit is, after you decide what size it should be and build it, then the size just doesn't change. But the size has still been calculated based on the load. 14:16 14:16 14:16 14:16 14:16 14:16 14:16 14:16 14:16

BY MS. ZHONG: 14:16

Q. Okay. I understand where you're coming from. I just want to confirm that once you have built a four-driver circuit as shown in Figure 5, the size of the driver circuit doesn't really change based on whether you choose one of the drivers, two 14:16 14:16 14:17 14:17 14:17

Netlist EX2025-p.133
Samsung v Netlist
IPR2022-01427

of the drivers, or all the drivers were used; is that correct?

MR. CHANDLER: Objection. Form.

THE WITNESS: Once you implement it in silicon, the size of the driver has already been selected and will not change.

BY MS. ZHONG:

Q. Okay.

A. But the current output could change.

Q. Understood. So if you --

DEPOSITION REPORTER: Clarification. Did you say "could change" or "would change"?

THE WITNESS: Could change.

BY MS. ZHONG:

Q. So if you want to change the size of the driver, you cannot do so by programming the driver circuitry by selecting a subset of the active transistors that were used; is that right?

A. There is a little bit of a judgment call there in terms of semantics. Right. So there definitely are programmable drivers. They were known well before this.

Q. Okay.

A. The size of the driver being used changes, as it does in Figure 5.

Q. Okay. Uh-huh.

A. But not the size that the total circuit takes up on the silicon. Right. So it's very common to use not only this design, but also to use designs where you had multiple transistors in parallel. And then if you went from using four transistors, eight transistors, it would be very common there to say that the size of the driver had changed because --

Q. Okay.

A. -- the amount of silicon that's driving has changed, but the physical space at the eight transistors, or in this case, these would be four transistors each, so the sixteen transistors take up doesn't change.

So the physical space of the entire circuit doesn't change after you've built it, but the portion of it that is being used as a driver can change.

Q. Okay. So the silicon space doesn't change, but the portion that you use may change in the programmable driver; is that right?

MR. CHANDLER: Objection. Form.

THE WITNESS: The silicon used by all the transistors doesn't change after it's been built,

Veritext Legal Solutions
Calendar-CA@veritext.com 866-299-5127

Netlist EX2025-p.135
Samsung v Netlist
IPR2022-01427

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

NETLIST, INC.,	(	CAUSE NO. 2:21-CV-463-JRG
	)	
Plaintiff,	(	
	)	
vs.	(	
	)	
SAMSUNG ELECTRONICS CO., LTD., et al.,	(	MARSHALL, TEXAS
	(	APRIL 14, 2023
Defendants.	)	9:00 A.M.

VOLUME 1

TRIAL ON THE MERITS

BEFORE THE HONORABLE RODNEY GILSTRAP
UNITED STATES CHIEF DISTRICT JUDGE
and a jury

SHAWN McROBERTS, RMR, CRR
100 E. HOUSTON STREET
MARSHALL, TEXAS 75670
(903) 923-8546
shawn_mcroberts@txed.uscourts.gov

Netlist EX2027
Samsung v Netlist
IPR2022-01427

A P P E A R A N C E S

FOR THE PLAINTIFF:

IRELL & MANELLA, LLP - LOS ANGELES
1800 AVENUE OF THE STARS
SUITE 900
LOS ANGELES, CA 90067-4276
(310) 203-7096
BY: MR. JASON SHEASBY
MS. LISA GLASSER

McKOOL SMITH, P.C. - MARSHALL
104 E. HOUSTON ST., SUITE 300
MARSHALL, TEXAS 75670
(903) 923-9000
BY: MR. SAM BAXTER
MS. JENNIFER TRUELOVE
MR. KEVIN BURGESS

FOR THE DEFENDANTS:

FISH & RICHARDSON, PC - WASHINGTON DC
1000 MAINE AVE. SW, SUITE 1000
WASHINGTON, D.C. 20024
(202) 783-5070
BY: MR. RUFFIN CORDELL
MR. MICHAEL McKEON

GILLAM & SMITH, LLP
303 SOUTH WASHINGTON AVENUE
MARSHALL, TEXAS 75670
(903) 934-8450
BY: MS. MELISSA SMITH

OFFICIAL REPORTER:

SHAWN M. McROBERTS, RMR, CRR
100 E. HOUSTON STREET
MARSHALL, TEXAS 75670
(903) 923-8546

Netlist EX2027
Samsung v Netlist
IPR2022-01427

INDEX

EXAMINATION

Witness Name	**Page**
SCOTT MILTON	
Direct By MR.SHEASBY	189
Cross By MR. McKEON	230

Netlist EX2027
Samsung v Netlist
IPR2022-01427

DIRECT EXAMINATION

BY MR. SHEASBY:

Q. Good afternoon, Mr. Milton.

A. Good afternoon, Mr. Sheasby.

Q. Can you introduce yourself to Judge Gilstrap and the jury, please?

A. Definitely. Your Honor, ladies and gentlemen of the jury, my name is Scott Milton, and I am the vice president of engineering at Netlist. And that is our highest technical position at the company.

Q. Now, do you have a personal connection to this case separate from your role as the -- an officer of Netlist?

A. Yes, I do. I am an inventor on one of the patent families, and I was also the technical interface with Samsung on the JDLA.

Q. And what is the history of Netlist?

A. So Netlist was founded in the year 2000. And at that time there were a couple of very large companies, Samsung and SK hynix, that were really dominant in the memory industry, selling modules and the raw components.

We at Netlist found that there was some room to innovate in the area of the modules, which is what the raw memory components go on, and we founded a company to -- to do that.

Q. Can you show us an example of a Netlist module?

A. Yes. So it's a little hard to see, but this is an

Netlist EX2027
Samsung v Netlist
IPR2022-01427

Q. And this is PX 1663. Is this another Samsung document?
A. Yes, sir.
Q. And what does it say?
A. So here we're talking about we moved ahead to enter into an agreement for technical collaboration to resolve the patent risk related to LRDIMM.
Q. The third family you talked about is the '060 and '160 Patents. Is that correct?
A. That's correct.
MR. SHEASBY: Mr. Huynh, can we have slide 57, please?
Q. (BY MR. SHEASBY) And these are the two patent families. Is that correct?
A. The two patents in that family, yes.
Q. And this is JTX 5 and 6. Is that correct?
A. That is correct.
Q. So I want to show you a document. This is PX 1778. What is PX 1778?
A. So this document is a description of the '060 Patent that was provided to Samsung.
Q. And was the actual document provided to Samsung or was a presentation given?
A. A presentation was given.
Q. And I want to show you -- so this describes the '060 Patent as relating to TSVs. Is that correct?

Netlist EX2027
Samsung v Netlist
IPR2022-01427

A. That is correct.
Q. And now you have to explain in detail what a TSV is.
A. Okay. So if we look at the screen here, what this is showing is we've got DRAM die, and those DRAM die are made out of silicon. And the vertical columns there are what we call vias. They are basically ways to get from one layer to another, and they're going through the silicon. That's the 'through silicon' part, and the 'via' is the path. And what that is is a way to interconnect those dies together.
Q. And in your research at Netlist, are you now working with HBMs?
A. Yes.
Q. And do you have personal knowledge of the -- I want to show you another document. So this is the email that PX 1778 is attached to, and it says "additional slides for Samsung." Is that correct?
A. Yes, sir.
Q. And was this presentation given to Samsung?
A. Yes, it was.
Q. You have personal knowledge regarding the high bandwidth memory through silicon via market space?
A. Yes.
Q. What was Samsung's position in the market when they asked for a presentation on your technology in April 2015?
A. In 2015, Samsung was behind their biggest competitor

Netlist EX2027
Samsung v Netlist
IPR2022-01427

SK hynix in this product space.
Q. And when did Samsung launch its HBM product with TSV?
A. When was the -- I'm drawing a blank on the date.
Q. It's okay. If you don't know, you should just say you don't remember.
A. I don't remember.
Q. No problem at all.
And I'm showing you PX 446. What is this document?
A. This is a document in response to a request to get information on our patents and what products they cover.
Q. And who requested this information?
A. This was requested by Samsung.
Q. And what does this page of the document show?
A. So this page shows that the '060 Patent is related to the HBM product family.
Q. About how many times -- this document's from 2006. From 2006 through the duration of your commercial relationship with Samsung, about how many meetings have you -- did you have with Samsung?
A. We had several. I mean, I don't know--10, 20. I -- yeah.
Q. At any point in time did Samsung ever claim that its HBM designs were not using your patents?
A. No.
Q. How does Netlist's design work? And specifically I want

Netlist EX2027
Samsung v Netlist
IPR2022-01427

in the HBM stack we see the 'through silicon vias'.
Q. Is the internal organization of the DRAM circuits on the left and the die arrays on the right different as well?
A. They are different.
Q. Can a DRAM circuit be used in place of the TSV array dies?
MR. McKEON: Your Honor, objection.
THE COURT: I'll sustain that. That calls for an opinion.
Q. (BY MR. SHEASBY) I want to show you another Passage from your specification. This is column 5, lines 34 through 40. What does this describe?
A. So this is talking about the fact that with our teachings we can go from the four array dies to eight array dies, maybe even 16 array dies.
Q. Does the patent present any techniques for dealing with testing for defects in these complex products?
A. Yes, it does.
Q. I'm showing you column 17, lines 36 through 49. What role does this passage relate -- play in testing, if any?
A. Yeah. So what this talks about is driver sizes, and we teach that you can size the drivers--and that's the electrical component that actually sends the signals--and you can size that driver based on what it needs to do.
For the test mode you don't have to test it at the full

Netlist EX2027
Samsung v Netlist
IPR2022-01427

speed, for example, because you want to see if it has connectivity. So, as a result, you can use a smaller driver, and what that does for you is reduce the amount of area it takes up on the die as well as reduces power.

Q. And I'm turning to page 4 -- column 4, lines 8 through 16 of your specification, and it says you incorporated by reference application 12/422,915. Do you see that?

A. I do.

Q. First off what does 'incorporation by reference' mean?

A. So what that means, we're taking all of the words in that document and essentially saying that they're in that patent application.

Q. And what does the 12/422,925 application describe?

A. It talks about tests.

Q. Mr. Milton, after termination and before this lawsuit was filed, did Samsung ever approach Netlist and seek permission to use Netlist's patents?

A. They did not.

MR. SHEASBY: I pass the witness, Your Honor.

THE COURT: Cross examination, please.

MR. McKEON: May I begin, Your Honor?

THE COURT: Looks like your team has binders to pass out.

MR. McKEON: Okay. Thank you.

THE COURT: No, Mr. Sheasby. It's Defendants' job

Netlist EX2027
Samsung v Netlist
IPR2022-01427

SAMSUNG

Samsung Begins Mass Producing World's Fastest DRAM – Based on Newest High Bandwidth Memory (HBM) Interface

First 4-Gigabyte HBM2 Dynamic Random Access Memory (DRAM) enables data transmission speed over seven times that of the previous fastest DRAM

SEOUL, Korea – January 19, 2016 – Samsung Electronics Co., Ltd., announced today that it has begun mass producing the industry's first 4-gigabyte (GB) DRAM package based on the second-generation High Bandwidth Memory (HBM2) interface, for use in high performance computing (HPC), advanced graphics and network systems, as well as enterprise servers. Samsung's new HBM solution will offer unprecedented DRAM performance – more than seven times faster than the current DRAM performance limit, allowing faster responsiveness for high-end computing tasks including parallel computing, graphics rendering and machine learning.

"By mass producing next-generation HBM2 DRAM, we can contribute much more to the rapid adoption of next-generation HPC systems by global IT companies," said Sewon Chun, senior vice president, Memory Marketing, Samsung Electronics. "Also, in using our 3D memory technology here, we can more proactively cope with the multifaceted needs of global IT, while at the same time strengthening the foundation for future growth of the DRAM market."

The newly introduced 4GB HBM2 DRAM, which uses Samsung's most efficient 20-nanometer process technology and advanced HBM chip design, satisfies the need for high performance, energy efficiency, reliability and small dimensions making it well suited for next-generation HPC systems and graphics cards.

Following Samsung's introduction of a 128GB 3D TSV DDR4 registered dual inline memory module (RDIMM) last October, the new HBM2 DRAM marks the latest milestone in TSV (Through Silicon Via) DRAM technology.

The 4GB HBM2 package is created by stacking a buffer die at the bottom and four 8-gigabit (Gb) core dies on top. These are then vertically interconnected by TSV holes and microbumps. A single 8Gb HBM2 die contains over 5,000 TSV holes, which is more than 36 times that of a 8Gb TSV DDR4 die, offering a dramatic improvement in data transmission performance compared to typical wire-bonding based packages.

Samsung's new DRAM package features 256GBps of bandwidth, which is double that of a HBM1 DRAM package. This is equivalent to a more than seven-fold increase over the 36GBps bandwidth of a 4Gb GDDR5 DRAM chip, which has the fastest data speed per pin (9Gbps) among currently manufactured DRAM chips. Samsung's 4GB HBM2 also enables enhanced power efficiency by doubling the bandwidth per watt over a 4Gb-GDDR5-based solution, and embeds ECC (error-correcting code) functionality to offer high reliability.

In addition, Samsung plans to produce an 8GB HBM2 DRAM package within this year. By specifying 8GB HBM2 DRAM in graphics cards, designers will be able to enjoy a space savings of more than 95 percent, compared to using GDDR5 DRAM, offering more optimal solutions for compact devices that

require high-level graphics computing capabilities.

The company will steadily increase production volume of its HBM2 DRAM over the remainder of the year to meet anticipated growth in market demand for network systems and servers. Samsung will also expand its line-up of HBM2 DRAM solutions to stay ahead in the high-performance computing market and extend its lead in premium memory production.

About Samsung Electronics Co., Ltd.

Samsung Electronics Co., Ltd. inspires the world and shapes the future with transformative ideas and technologies that redefine the worlds of TVs, smartphones, wearable devices, tablets, cameras, digital appliances, printers, medical equipment, network systems, and semiconductor and LED solutions. We are also leading in the Internet of Things space with the open platform SmartThings, our broad range of smart devices, and through proactive cross-industry collaboration. We employ 319,000 people across 84 countries with annual sales of US $196 billion. To discover more, and for the latest news, feature articles and press material, please visit the Samsung Newsroom at news.samsung.com.

HyperCloud and N3DS

Move to High-Density for Cloud Computing



NETLIST

February 2011

Appx7528

Netlist EX2038
Samsung v Netlist
IPR2022-01427

Agenda

SCALING THE CLOUD

- Update on HyperCloud
  - Market adoption
  - Competitive products analysis
- HyperCloud Alliance Project
  - Achieve mainstream industry adoption
  - Altering the economic model of DRAMs
  - Nanya roadmap with and without HyperCloud
- Overview of N3DS
  - HyperCloud framework for N3DS
  - Comparison of 3DS alternatives
  - Benefits of N3DS

Appx7529

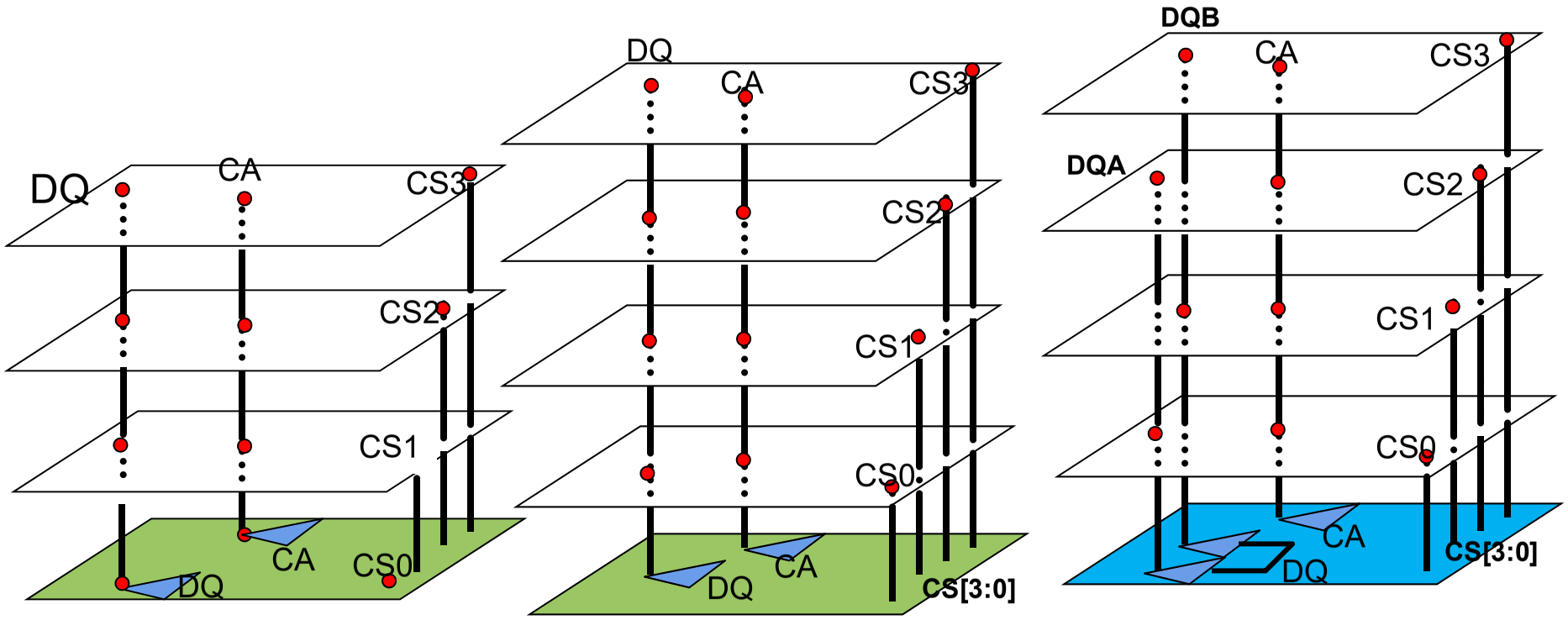
Optimized Architecture for 3DS SDRAM

Appx7530

N3DS - Optimized Architecture to Support 3DS SDRAM

SCALING THE CLOUD

DRAM – die inter-connect options



4-H
1 Master with Array +3 Slaves
1 DQ Fanout

3DS SDRAM Option 1

5-H
1 Master + 4 Slaves
1 DQ Fanout

3DS SDRAM Option 2

5-H
1 Master + 4 Slaves
Two DQ Fanout - reduce slave load

N3DS SDRAM

Appx7531

Scaling the Cloud with Advanced Memory

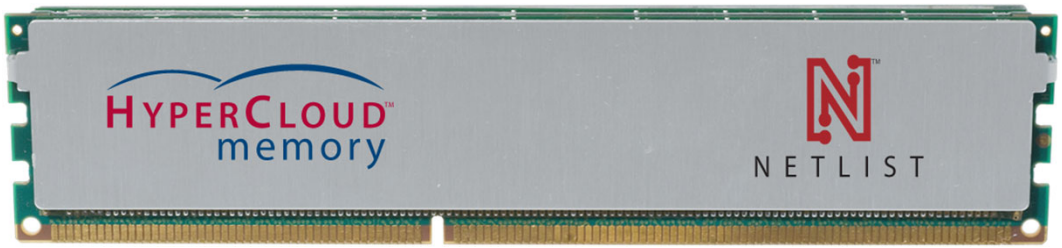
NETLIST

Appx7532

Q & A

Thank You

NETLIST



Appx7533

UNITED STATES PATENT AND TRADEMARK OFFICE

_____

BEFORE THE PATENT TRIAL AND APPEAL BOARD

_____

SAMSUNG ELECTRONICS CO., LTD.,

Petitioner,

v.

NETLIST, INC.,

Patent Owner

_____

IPR2022-01427

Patent 9,318,160

_____

IPR2022-01428

Patent 8,787,060

_____

September 27, 2023

Remote deposition of MICHAEL BROGIOLI, Ph.D. was taken pursuant to notice, beginning at 9:01 a.m. on the above date before Gail Verbano, Registered Diplomate Reporter, Certified Realtime Reporter, Certified Shorthand Reporter, and Notary Public, there being present via videoconferencing:

A P P E A R A N C E S:

On behalf of Samsung:
THEODORE W. CHANDLER, ESQ.
ted.chandler@bakerbotts.com
FERENC PAZMANDI, ESQ.
ferenc.pazmandi@bakerbotts.com
BAKER BOTTS, LLP
101 California Street, Suite 3200
San Francisco, California 94111-5802
415.291.6259

On behalf of Netlist, Inc.:
ANNITA ZHONG, ESQ.
HZHONG@irell.com
IRELL & MANELLA, LLP
1800 Avenue of the Stars, Suite 900
Los Angeles, California 90067-4276
310.277.1010

ALSO PRESENT:

CHRIS JORDAN, Legal Technician

- - -

P R O C E E D I N G S

- - -

CONTINUED EXAMINATION

THE VIDEOGRAPHER: Back on the record. The time is 12:51.

BY ATTORNEY CHANDLER:

Q. Dr. Brogioli, do you have any clarifications or corrections you would like to make to any of your testimony so far?

A. Not that I recall, no.

Q. Could you turn back to Exhibit 1019, the JEDEC standard for DDR3, and let me know when you have it in front of you, and go to page 23.

A. Okay.

Q. Toward the bottom right of page 23 of Exhibit 1019, the JEDEC standard for DDR3, there is a figure with a column for CAS latency; correct?

A. Yes, I see that.

Q. What is CAS latency?

A. I believe it refers to the column address strobe latency. And it's an aspect of

timing related to -- in certain types of memory devices. And it looks like this section talks about some of the mode registers for DDR3 that relate to that and some other fields.

Q. Under the JEDEC standard for DDR3 memory devices, CAS latency could be set to 5, 6, 7, 8, or 9 clock cycles, as shown on page 23 of Exhibit 1019; is that correct?

A. That looks like at least some of the values that the -- excuse me -- that the CAS latency could be set to. I think you mentioned 5 through 9.

Q. Or 10 clock cycles; right? That's also specified in the standard?

A. I see that, yes. Those are the -- that's the CAS latency. You set the bits accordingly to get a different value.

Q. So the JEDEC standard permits a range of different CAS latency. It doesn't mandate which of those CAS latencies you have to use; instead, the JEDEC standard permits you to choose one of the possible CAS latencies between, for example, five clock cycles up through 10 clock cycles? Is that fair?

A. It looks like it's reporting the writing

of this mode register to handle the different modes of operation for the DDR3 SDRAMs themselves. And so you can set the field accordingly as to which value you're using in your system.

Q. At the bottom of the next page, which is page 24 of Exhibit 1019, the JEDEC standard for DDR3, it says that, in the third or fourth line from the bottom, that read latency is defined as additive latency plus CAS latency; correct?

And it uses the abbreviations RL for read latency, AL for additive latency, and CL for CAS latency, resulting in a little equation, RL equals AL plus CL; correct?

A. That looks to me what it's saying in this CAS latency subsection.

Q. So the JEDEC standard for DDR3 memory devices permits a range of different read latencies that you can choose from rather than mandating only one possible read latency value; is that correct?

ATTORNEY ZHONG: Objection; form.

THE WITNESS: Looks like it's saying the read latency here, the overall read latency is comprised of two things, one being the CAS

latency, CL; and then you'd say, if you want additional information here, you can refer to the standard speed bin section on page 157.

BY ATTORNEY CHANDLER:

Q. But the read latency will be equal to or greater than the CAS latency. And as shown on page 23, the CAS latency can be set to the value that you get to choose between 5 clock cycles and at least 10 clock cycles; correct?

ATTORNEY ZHONG: Objection; form.

THE WITNESS: It looks like CAS latency is a parameter that you can control by this mode register based on, it sounds like, operating frequency and speed of your devices. And then as we go back to the bottom of 24, it refers us back to more of the standard speed bins, section for additional information.

And if I go there, it looks like it's showing a bunch of speed bin and table modes, so relating to how fast your device is operating or successfully operating, et cetera.

BY ATTORNEY CHANDLER:

Q. Could you turn to page 62 of Exhibit 1019, the DDR3 standard. Let me know when you're there.

three NoOps and a write at clock T4.

Q. In Exhibit 1019 on page 74 the JEDEC DDR3 standard provides examples of timing diagrams for a write command followed by a read command; using the burst length and latencies specified in note 1 of Figure 53 and note 1 of Figure 54; correct?

A. It looks like a write-to-read scenario where the write happens at time T-zero; and then there's NoOps for, I guess, 12 cycles or 12 time clots until a read in time 13. And the parameters used here are denoted in the four notes below the table.

Q. So on page -- I'm sorry.

On page 74 of Exhibit 1019, for the timing diagrams in Figure 53 and 54 where there's a write followed by a read command -- in both of those examples, the read latency is 5 clock cycles -- that's abbreviated RL; the CAS latency, CL, is also 5 clock cycles; the additive latency, AL, is zero clock cycles; and the write latency, WL, is 5 clock cycles; correct?

A. That seems to be what note 1 is saying, yes.

Q. And a person of ordinary skill in the

art would understand these timing diagrams in the DDR3 standard marked as Exhibit 1019 that we've been discussing on page 74, 73, 63, 62; is that fair?

ATTORNEY ZHONG: Objection to form.

THE WITNESS: I think that's right. In the context of this document, they would understand these timing -- this scenario that these timing diagrams are describing.

BY ATTORNEY CHANDLER:

Q. And we discussed earlier, a person of ordinary skill in the art would understand that it's important to follow the timing specifications in the JEDEC standard because the data lines for the DDR3 memory devices are bidirectional, meaning the same data lines used for a write command, that can be used for a read command. And if the timing isn't correct, there's a risk of a data collision; correct?

ATTORNEY ZHONG: Objection; form, vague.

THE WITNESS: I would say if a person of skill in the art is attempting to build a JEDEC-compliant device, they would look to the relevant JEDEC specification for timing that they -- they need to follow for that device. An

BY ATTORNEY CHANDLER:

Q. And would it be fair to say that, in general, collisions are bad, something that you would try to avoid?

A. In this context, it looks like there's collisions happening from multiple types of commands. It looks a write and an activate or a write and precharge, and you don't want to do that. That's something to avoid, I would say.

Q. Figure 14 of Rajan teaches how to avoid such a collision, as explained, for example, starting at Column 11, line 12, of Exhibit 1015; is that fair?

ATTORNEY ZHONG: Objection; form, scope.

THE WITNESS: It looks like Figure 14 says it's a timing diagram illustrating methods of avoiding such collisions. These look like command collisions of things like activate commands, read and write commands.

BY ATTORNEY CHANDLER:

Q. For example, Rajan teaches at Column 11, lines 23 through 25, of the Exhibit 1015 that:

"The buffer chip can delay the activate and precharge operations to

avoid collisions in a manner that's transparent to the host system."

Correct?

A. That in this certain example, starting in the description on line 20, if the DRAM chips are what it says, 4 clock cycles each, then the buffer chip tells the hosts that there are six clock cycles, and then delays -- the buffer chip delays various activate and precharge commands or operations to, it looks like, avoid collisions on the command side.

Q. Could you go to your expert declaration and turn to page 97, and let me know when you're there.

A. Okay.

Q. At the bottom of page 97 of your expert declaration marked as Exhibit 2023, in the last sentence, you express the opinion that, quote:

"The presence of significant load to a TSV attributable to a disabled buffer that is described as not electrically coupled to the TSV confirms that the buffer is still in electrical communication with the TSV," end quote.

Did I read that accurately?

of skill in the art would understand that there's -- in this context, there's a problem there. That's problematic. 11:42:02

Q. Do you agree a person of ordinary skill in the art knew about various techniques to avoid collisions when memory devices in two different ranks share a data line? 11:42:12

A. I would say, in the general sense, there were probably examples yesterday. It might have been from JEDEC, or something of that nature, showing some solutions to that problem that may have been acceptable in that context. But I -- I don't -- I wouldn't say that they were aware of sort of a generally accepted solution for -- for all problems in this -- in this space. 11:42:40 11:42:57

Q. What technique does the '912 patent marked as Exhibit 1058 state, here at column 23, line 60 through 67, for avoiding collisions when memory devices in two different ranks share a data line? 11:43:22

MR. LINDSAY: Objection. Outside the scope of direct.

THE DEPONENT: It looks like, in this paragraph, the '912 is talking about memory boundaries between ranks in a memory module and 11:43:40

that in the case of what it refers to as a back-to-back adjacent read command, it, at least from the language of this paragraph, says the memory controllers can insert wait-time intervals or clock cycles to avoid collisions or interference in that scenario. 11:43:50 11:44:07

Q. (By Mr. Chandler) Would it be a fair to say a person of ordinary skill in the art knew that memory controllers could insert wait-time intervals or clock cycles to avoid collisions? 11:44:26

MR. LINDSAY: Objection. Outside the scope of direct.

THE DEPONENT: I would answer that that there may have been, I think, references, even yesterday, that mention this really naive solution that may be entirely unacceptable to just insert -- insert cycles of idle time or idle activity to just sort of dumbly wait and sit there. But that isn't -- isn't generally -- in a general sense, isn't desirable in high-performance memory systems for a number of reasons, either performance or just standards compliance perhaps. But I think there were references that mentioned that somewhat naive solution. 11:44:40 11:45:08

Q. (By Mr. Chandler) In the '060 and '160 11:45:34

patents, are there any limitations of the claims that require high performance? 11:45:36

A. I would say the inventions, overall in the '060 and '160, are striving for very high storage capacity in terms of bits -- bytes of storage and high-performance and high-bandwidth memory performance. 11:45:57

If you are asking does the phrase, quote, "high performance" explicitly present itself in the claims, I don't recall that it does. 11:46:17

Q. Do you regard Netlist '912 patent marked as Exhibit 1058 as claiming a high-performance memory module?

MR. LINDSAY: Objection. Outside the scope of direct. 11:46:33

THE DEPONENT: I don't know that I have analyzed that aspect of it, of the '912, and high performance has to probably be characterized in terms of the -- whatever the performance is that you are trying to make high versus low, if that -- if that makes sense, if it's speed or bandwidth or storage capacity or things like that. 11:47:00

Q. (By Mr. Chandler) Would it be fair to say that if there's data collisions, that could ruin the performance of any memory module? 11:47:24

standard marked as Exhibit 1019 and let me know when you are there. 12:30:46

A. Okay.

Q. And I want focus, for now, on the Figure 2 in the middle of that page. 12:31:02

You got it?

A. Got it.

Q. Figure 2 on page 12 of the JEDEC DDR3 standard marked as Exhibit 1019 illustrates a memory package, including four stacked DDR3 memory chips, correct? 12:31:16

A. Looks like this is talking about a DDR3 SDRAM package and various aspects of it, and it shows -- it looks like it's intending to show the qual-stacked or quad-die DDR3 SDRAM with various -- it looks like various -- some of the signals that go to these various devices. 12:31:45

Q. So there would be four chips or four dies, correct?

A. It looks like the language of Figure 2 is "Qual-stacked/Quad-die DDR3 SDRAM," and it shows the -- the various signaling you would expect for a -- a device. 12:32:13

Q. And "quad" means four, correct?

A. That's correct, yes. 12:32:34

Q. In Figure 2 of page 12 of the DDR3 standard marked as Exhibit 1019, each chip in the stack has its own chip select signal labeled at the bottom as "CS0," "CS1," "CS2," and "CS3" respectively, correct? 12:32:39 12:33:02

A. I see those four -- those four chip select signals, yes.

Q. And each chip in the stack has its own circuitry to output data or to receive data on eight lines labeled at the top as "DQ0-7," correct? 12:33:27

MR. LINDSAY: Objection. Form.

THE DEPONENT: It looks like each of these DDR3 SDRAM devices has the DQ0-7 input/output on it for -- for transmitting data.

Q. (By Mr. Chandler) And those input/output data lines, DQ0 through 7 for each chip, are connected to a common set of data lines that are also labeled "DQ0-7" at the top left of Figure 2, correct? 12:33:58

A. That's how this figure is illustrated; that's correct. 12:34:24

Q. And a person of ordinary skill in the art would understand that this arrangement of stacked memory devices shown on page 2 of the DDR3 standard marked as Exhibit 1019 would be expected to work 12:34:50

without causing collisions on the common set of data lines labeled "DQ0-7" -- and if I said "page 2," I meant page 12 of Exhibit 1019 -- is that correct? 12:34:58

MR. LINDSAY: Objection. Form. 12:35:14

THE DEPONENT: I would say the intent of this Figure 2 on page 12, I would say, would be -- my understanding of the section would be it's informing the reader of how the devices operate when taken in context of this section and signaling and so forth. 12:35:33

Q. (By Mr. Chandler) And a person of ordinary skill in the art, reviewing the stacked memory devices shown on page 12 of the DDR3 standard marked as Exhibit 1019, would expect those stacked memory devices to operate correctly, without collisions, on the set of data lines labeled "DQ0-7," correct? 12:35:54

MR. LINDSAY: Objection. Form.

THE DEPONENT: I think they would understand this figure to be describing one way of implementing this D- -- DDR3 implementation within the confines of JEDEC and, taken as a whole, to provide an operating -- or a correctly operating device -- a functioning operational device per 12:36:15 12:36:41

JEDEC's definition. 12:36:46

Q. (By Mr. Chandler) As a general matter, would you agree that the DDR3 standard, in the time frame of the '060 and '160 patents, was the leading standard for DRAM memory devices commonly used in 12:37:07 the industry for lots and lots of memory devices sold throughout the world?

MR. LINDSAY: Objection. Form.

THE DEPONENT: I would say, at that time frame, that DDR3 devices were used in many, many, 12:37:23 many -- or JEDEC-compliant DDR devices were used in many, many systems and products.

Q. (By Mr. Chandler) Around the entire world, correct?

A. That -- that sounds fair to me, yeah. 12:37:39

Q. All right. Let's look at Exhibit 1022, which I will put in the chat. I think it's a little bit bigger than some of the other exhibits, so it may take a little while to come through, but let me know when you've got it. 12:38:01

A. Okay.

Q. Exhibit 1022 is the textbook by Professor Jacob, correct?

A. That looks correct, yes.

Q. It's titled "Memory Systems: Cache, 12:38:31

skill in the art would not have shared a TSV for data signals connected to multiple chips in the same rank, correct? 01:00:17

A. I discuss as to why a person of skill would not do that in the context of Kim, given the solution that Kim -- or in Dr. Wolfe's proposed modifications to Kim, given the -- what Kim discloses to begin with, that it wouldn't work. 01:00:42

Q. And the reason that you state it wouldn't work is because of collisions, correct? 01:01:06

A. As I recall, that's part of the reason I discuss, in light of at least one of Dr. Wolfe's modifications to Kim.

Q. In your opinion, it was known at the time that sharing a TSV for data signals among multiple chips in the same rank could lead to data collisions, correct? 01:01:33

A. That's correct, in the context I discuss in my declaration.

Q. It was also known to a person of ordinary skill in the art that multiple chips in different ranks could share the same data line without causing collisions, correct? 01:02:00

A. In certain contents, correct.

Q. Please turn to page 64. 01:02:32

Do you have page 64 in front of you? 01:02:51

A. Yes.

Q. At the top of page 64 of Exhibit 2023, you created a modified version of Figure 2 of the Kim reference marked as Exhibit 1014, correct? 01:03:16

A. Yes, that's correct.

Q. And, to be clear, the figure at the top of page 64 of Exhibit 2023 was created by you, not by Dr. Wolfe or by Samsung, correct?

A. That's correct, yes. 01:03:47

Q. In your modified version of Figure 2 of Kim shown on page 64 of your declaration marked as Exhibit 2023, you have added four chips labeled "chip A," "chip B," "chip C," and "chip D," correct? 01:04:06

A. Correct.

Q. And the way that you have modified Figure 2 of Kim at the top of page 64 of your declaration marked as Exhibit 2023 is inconsistent with the disclosure of Kim because you have chip C in Rank0 connecting to GIO_Rank1, correct? 01:04:27

A. I would say that this figure is intended to visualize one of the modifications to Kim that Dr. Wolfe presented in his deposition.

Q. Well, to be clear, just so there's no 01:05:02

misunderstanding, Dr. Wolfe did not present the figure shown at the top of page 64 of your declaration marked as Exhibit 2023, but, rather, that figure is your characterization of what Dr. Wolfe was talking about, correct? 01:05:05 01:05:18

A. I would say it's my visualization of what I'm understanding him to say in his deposition.

Q. Kim, marked as Exhibit 1014, does not teach having a chip in Rank0 connected to GIO_Rank1 and does not teach having a chip in Rank1 connected to GIO_Rank0 as you have shown at the top of page 64 of your declaration marked as Exhibit 2023, correct? 01:05:44

A. I don't believe Kim explicitly teaches that, no. 01:06:05

Q. In your opinion, if Kim were modified such that a chip in Rank0 were connected to GIO_Rank1 and a chip in Rank1 were connected to GIO_Rank0, the result of that could be collisions, correct? 01:06:47

A. As I discuss in my declaration, there are a few different adverse-effect scenarios that occur. I think collisions was one. Perhaps loss of data was another one, depending on how things are operating. 01:07:04

Q. And, in general, a person of ordinary skill in the art would be motivated to avoid collisions and loss of data; is that fair? 01:07:08

A. As we have talked about earlier today and yesterday, in a general sense, I think that's -- that's fair. 01:07:20

Q. Could you please turn to page 68 of your declaration and let me know when you are there.

A. Okay.

Q. In the middle of page 68 of your declaration marked as Exhibit 2023, you have created another modified version of Figure 2 of Kim marked as Exhibit 1014, correct? 01:07:50

A. That is correct. The intent was for me to visualize another modification to Kim that Dr. Wolfe expressed in his deposition, as I recall. 01:08:13

Q. And, to be clear, the figure in the middle of page 68 in your declaration marked as Exhibit 2023 was created by you, not by Dr. Wolfe or by Samsung, correct? 01:08:35

A. That's correct.

Q. And you did not show this figure to Dr. Wolfe and ask him if it's an accurate representation of his opinions, correct?

A. I -- I, personally, haven't communicated 01:09:01

was created by me; you're right. And to the second part of your question, similar to before, I don't know one way or the other if this was presented to Dr. Wolfe at some point in time. 01:11:05

Q. In your four-rank version of Figure 2 of Kim shown at the top of page 70 of your expert declaration marked as Exhibit 2023, you assume that Kim's circuits would operate the same way as in Kim's original Figure 2, correct? 01:11:18

A. That is correct, as I recall figure -- the -- the earlier figures of Kim. 01:11:46

Q. For example, in your modification shown at the top of page 70 of Exhibit 2023, even though you modify Figure 2 of Kim to have four ranks of memory devices, you only show two chip select signals, CS0 and CS1, received by the shared data I/O circuit 1000, and you don't show anywhere in this figure that there would be four chips of select signals, correct? 01:12:10

MR. LINDSAY: Objection. Form. 01:12:36

THE DEPONENT: I don't -- as I recall, Dr. Wolfe didn't articulate anything about additional chip select signals that would relate to what this figure is visualizing.

Q. (By Mr. Lindsay) And you have not shown 01:12:56

or addressed whether or not there would be four chip select signals in your four-rank modified version of Figure 2 of Kim shown at the top of page 70 of your declaration marked as Exhibit 2023, correct? 01:12:59 01:13:15

A. Dr. Wolfe didn't offer that in his modification to Kim. I am simply addressing what he proposed.

Q. Well, to be clear, you are addressing what you are characterizing he proposed. 01:13:27

You never asked him -- and I don't believe Netlist ever asked him -- whether there would be four chip select signals if there are four ranks, correct?

MR. LINDSAY: Objection. Form. 01:13:44

THE DEPONENT: I'm going -- my -- my visualization here is what Dr. Wolfe, my understanding, has testified to under oath. I don't recall the exact line of question and answers. If I've mischaracterized it, I would be happy to bring up his deposition transcript, perhaps -- if you think he mentioned additional chip select. I don't recall that. 01:13:57

Q. (By Mr. Chandler) Do you recall anyone asking him how many chip select signals would be 01:14:09

have any corrections or clarifications you would like to make to your earlier testimony? 02:15:22

A. Not at this time, no.

Q. Could you pull up Exhibit 1017, which is the Wyman reference, and I put it in the chat at the beginning of the day. 02:15:47

A. Okay.

Q. And please turn to Figure 5. Let me know when you are there.

A. Okay. 02:16:21

Q. In Figure 5 of the Wyman reference marked as Exhibit 1017, the signal will come from the left side of the figure and can be output at any of the tap-off locations, T0, T1, T2 or T3, as well as 508 on the right, correct? 02:16:49

A. That's right, the -- the five tap-off locations that Wyman refers to.

Q. In Figure 5 of Wyman marked as Exhibit 1017, if only the driver 502-1 is enabled, the input signal is driven only by driver 502-1 to tap T1, correct? 02:17:26

A. I am just rereading the section of Wyman.

So it looks like if only the 502-1 driver is off, you would be using the T1, or 506-2 tap, to get that output value. 02:18:20

Q. And the transcript, I think, didn't accurately get your testimony. I think you said if only 502-1 is "on," not "off" but "on," correct? 02:18:26

A. That -- that's correct, yes.

Q. So for clarity, could you give your answer again. 02:18:47

A. I was saying that if -- if only the 502-1 driver is on, one would use the -- you could use 506-2 tap to get a -- a value output.

Q. So in the situation where only driver 502-1 is enabled, the signal is driven only by driver 502-1 to tap T1, which is also labeled "506-2"; is that correct? 02:19:15

A. That appears to be at least what Figure 5 is illustrating. 02:19:41

Q. And if Figure 5 of Wyman marked as Exhibit 1017, if only drivers 502-1 and 502-2 are enabled, the input signal is driven by both driver 502-1 to tap T1 marked as "506-2," and driver 502-2 will drive to tap T2 marked as "506-3," correct? 02:20:10

A. I think that sounds right, in -- Wyman talked about different configurations or -- different configurations of the different drivers to divide up the driver level.

Q. So in Figure 5 of Wyman marked as 02:20:47

Exhibit 1017, tap T2 labeled "506-3" can be selected as the output driven both drivers 502-1 and 502-2, correct? 02:20:49

A. That's my -- that is my recollection; that's right. You could pull the T2 or 506-3 tap which would be including a signal coming from ENB1 and ENB0. 02:21:21

Q. For clarity, I think you just misspoke. The -- the ENB is referring to the --

A. You're right. 02:22:07

Q. -- a name --

A. You're right. I -- I -- let me reanswer.

If -- if you're -- the output at tap T2, or 506-3, would be a signal compromising the -- the driver triangles we see in 502-2 and 502-1. 02:22:21

Q. And just to do one more example from Figure 5 of Wyman marked as Exhibit 1017, if only drivers 502-1, 502-2 and 502-3 are enabled, then the same input signal is driven by driver 502-1 to tap T1, driver 502-2 to tap T2, and driver 502-3 to tap T3, correct? 02:22:47

A. There are -- I would say that those drivers are used to generate a signal available on those various taps, as I recall.

Q. Thus, in that example of Figure 5 of 02:23:16

Wyman marked as Exhibit 1017, tap T3 can be selected as the output driven by drivers 502-1, 506-2, and 502-3, correct? 02:23:19

A. The driver -- those driver components within Figure 5, yes. 02:23:36

Q. In Figure 5 of Wyman marked as Exhibit 1017, in the example where the output is tapped off at T2, the driver size is larger than when the output is tapped off at T1, correct?

A. No. The driver size of Wyman is the fixed driver that Wyman talks about. It's -- Figure 5 just gives you various output level selections that you can select from its Figure 5 circuit. 02:24:16

Q. Take a look at Figures 3 and 4 of Wyman and let me know when you are there. 02:24:35

A. Okay.

Q. The driver size in Figure 4 of Wyman, in your opinion, is larger than the driver size in Figure 3 of Wyman, correct? 02:24:54

A. It looks like, at least the way these two figures are visualized, Figure 4 has more stages to it than Figure 3 --

Q. And would you consider --

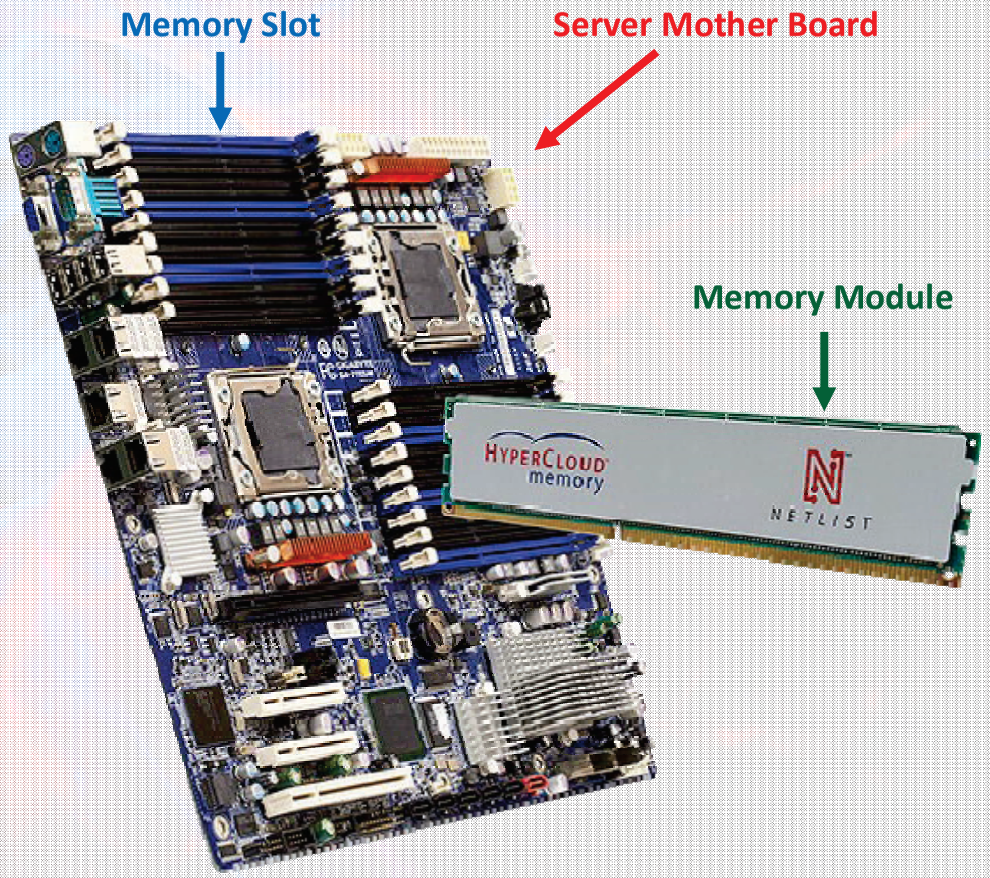
A. -- so -- 02:25:14


NETLIST
Technology Tutorial
In the Matter of Certain Memory Modules and Components Thereof and Products Containing Same
Inv. No. 337-TA-1023
May 8, 2017
CDX-200.1

Appx10308

Technology Overview: Memory Module

- The Netlist Patents relate to memory modules in a server
  - A memory module is inserted in a memory slot in a server's mother board
  - Memory modules are called Dual In-Line Memory Modules (DIMMs) because there are in-line connections on each side of the module

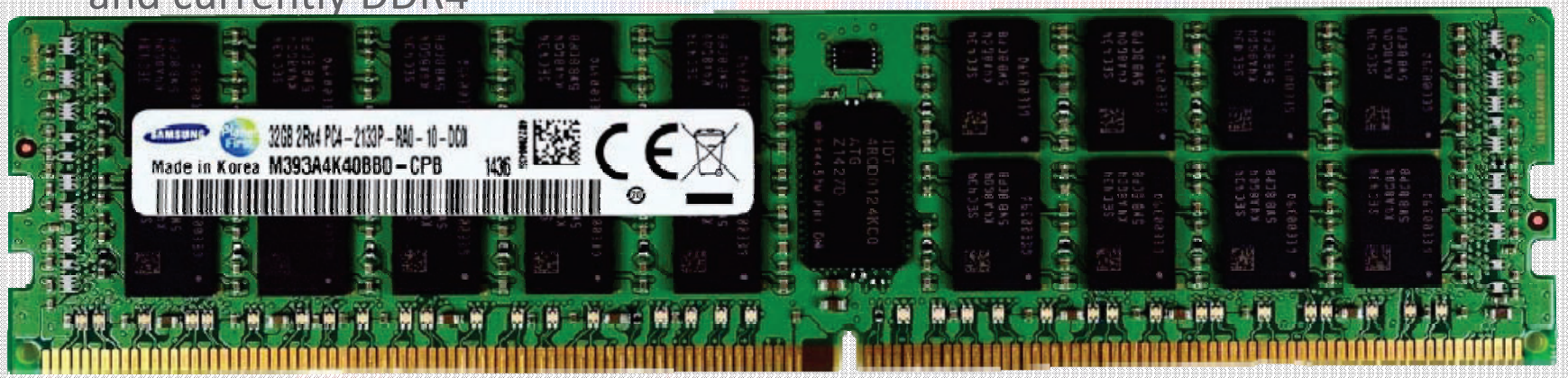


Appx10309

CDX-200.3

Technology Overview: What is Memory?

- Memory is used to store information for fast access by the processor
- Memory commonly used today is SDRAM or synchronous dynamic random access memory
  - SDRAM emerged in the 1990s and has been standardized by the JEDEC (Joint Electron Devices Engineering Council)
  - SDRAM has evolved in several generations, including DDR, DDR2, DDR3 and currently DDR4

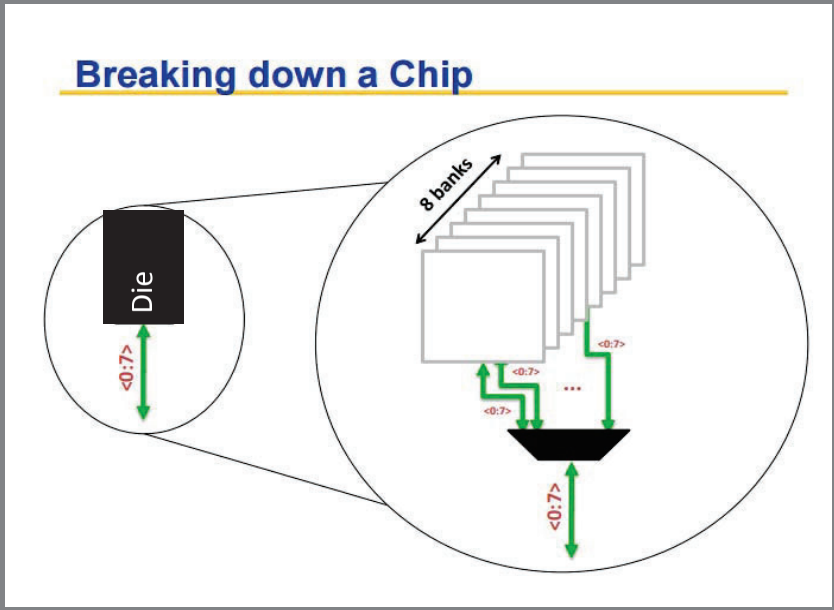


CDX-200.4

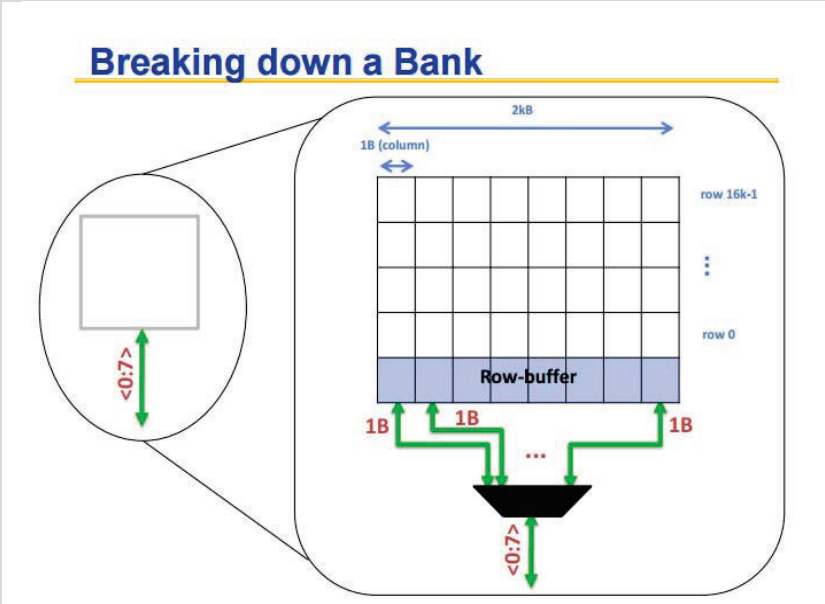
Appx10310

Technology Overview: SDRAM Fundamentals

SDRAM contains several sections called banks



Each bank consists of rows and columns of memory

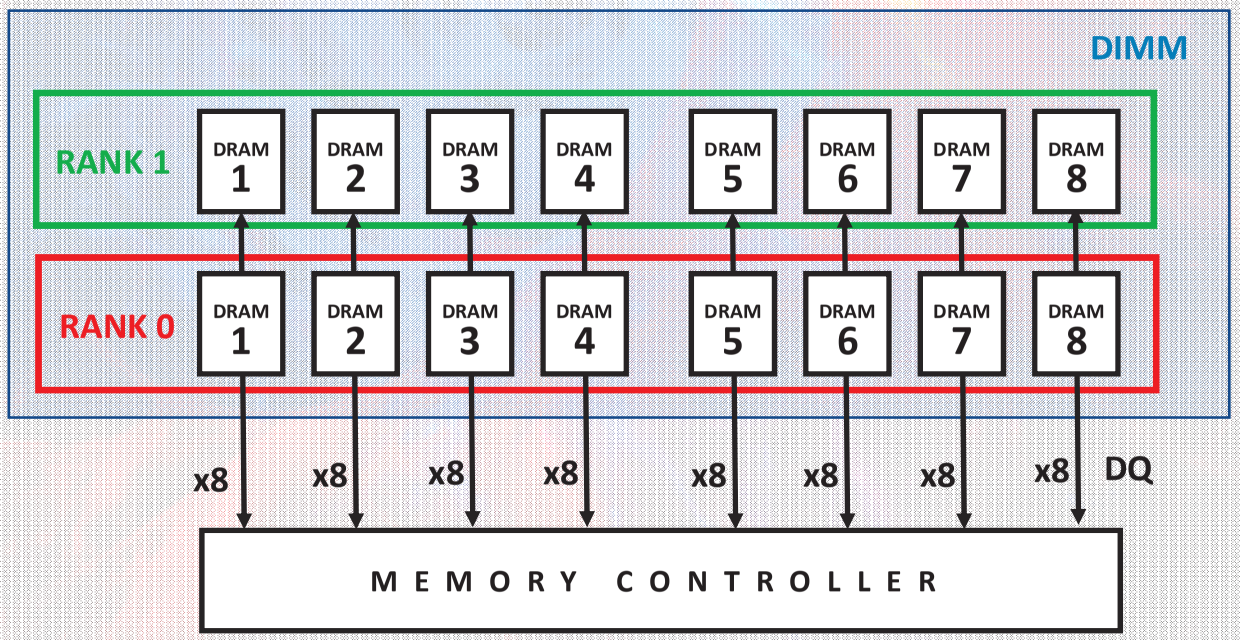


Appx10311

CDX-200.5

Technology Overview: Memory Ranks

- A rank is a set of DRAMs that are written to or read from together
- For example, a memory controller reads or writes 64 DQ (data) bits at a time
  - Each DRAM reads/writes 8 DQ bits (x8, or by eight)
  - Need 8 DRAMs in a rank to provide the 64 bits at once



Appx10312

Samsung Ex. 1054, p. 5
Samsung Electronics Co., Ltd. v. Netlist, Inc., IPR2022-01427

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD., MICRON TECHNOLOGY, INC., MICRON SEMICONDUCTOR PRODUCTS, INC., and MICRON TECHNOLOGY TEXAS LLC,[†]

Petitioners,

v.

NETLIST, INC.,

Patent Owner

IPR2022-01427
Patent 9,318,160 B2

EXHIBIT 1067:

Petitioner's Demonstratives for Oral Argument on January 11, 2024

[†] Micron Technology, Inc., Micron Semiconductor Products, Inc., and Micron Technology Texas LLC filed a motion for joinder and a petition in IPR2023-00883 and have been joined as petitioners in this proceeding.

Petitioner Samsung Electronics Co., Ltd.'s Oral Argument for PTAB Hearing

Samsung Electronics Co., Ltd. v. Netlist, Inc.,

IPR2022-01427 and IPR2022-01428 (U.S. Patent Nos. 9,318,160 and 8,787,060)

January 11, 2024

© Copyright Baker Botts 2024. All Rights Reserved.

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

BAKER BOTTS

'060/'160 Patent: Optimizing Load on "Die Interconnects"

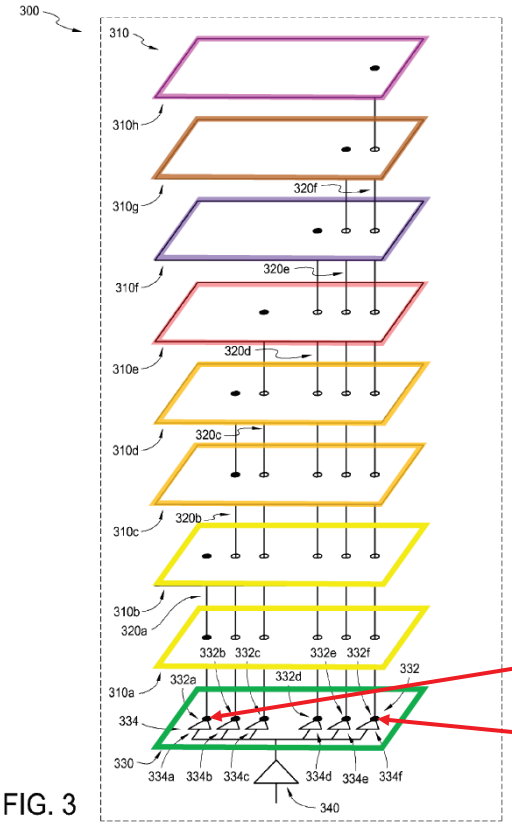


FIG. 3

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 **320***a* may be equivalent to one quarter of the load of an array die **310**. Thus, for example, the load of the conduit **332***a* 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 **320***a* and a load contribution of 0.25L per segment (two in this case) of the die interconnect **320***a*. As a second example, the load of the conduit **332***f* may be 3L assuming the same load values as the previous example and a load contribution from one array die **310***h* and eight die interconnect **320***f* segments. Table 1 specifies the capacitive load values for each conduit **332**

1L load / die | **0.25L load / segment of die interconnect**

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

'060: EX1001 at Fig. 3, 12:51-:66 & Table 1; Paper 1 (Pet.) at 11-14; EX1003 at ¶¶ 61-65
'160: EX1001 at Fig. 3, 12:46-:61 & Table 1; Paper 1 (Pet.) at 11-16; EX1003 at ¶¶ 61-65

Appx10681

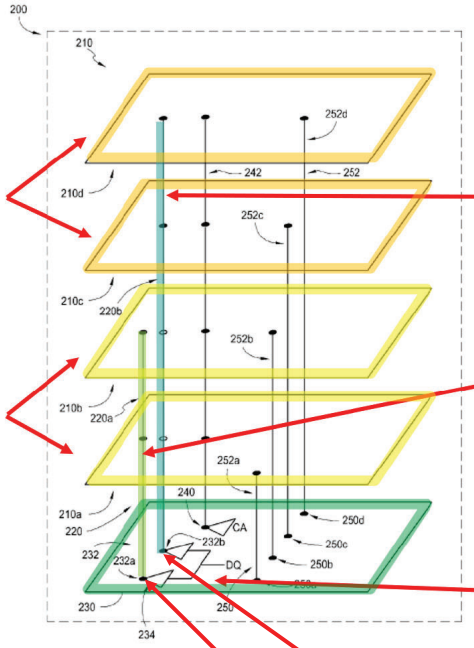
'060/'160 Claims: "Control Die" with "Data Conduits" to "Die Interconnects"

'060 Patent, Claim 1:

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.

'160 Patent, Claim 1:

a control die comprising first data conduits between the first die interconnects and the data terminals, and second data conduits between the second die interconnects and the data terminals, the first data conduit including first



Array Dies 210c and 210d (orange)

Array Dies 210a and 210b (yellow)

Die interconnect 220b (teal)

Die Interconnect 220a (light green)

Drivers 234 (for data)

FIG. 2

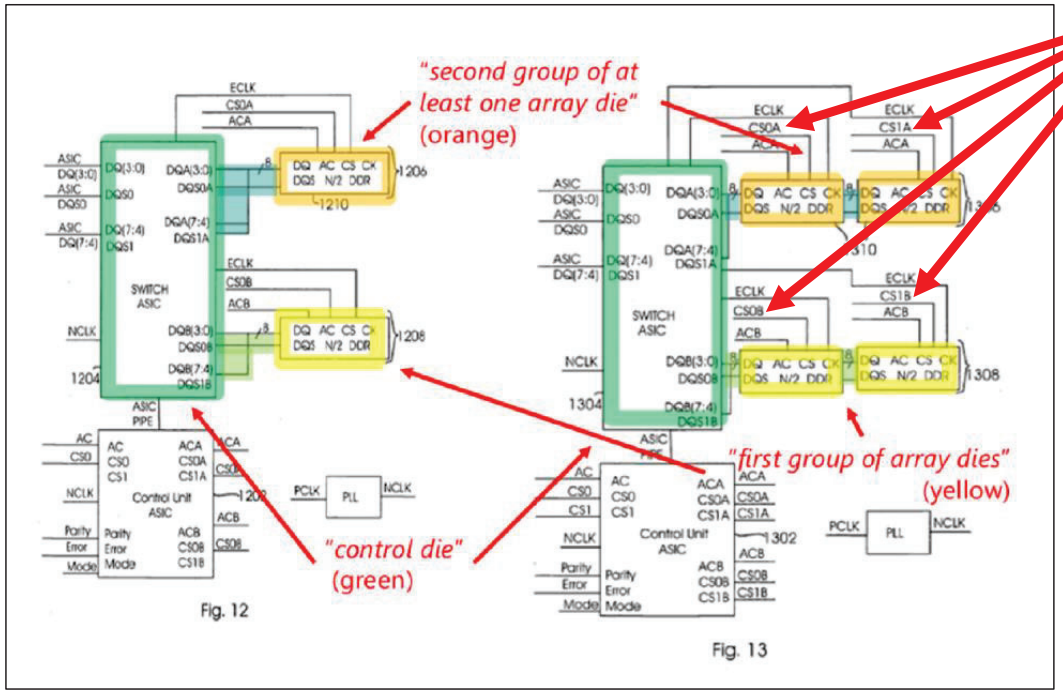
The control die **230** may include a number of data conduits **232**, which includes data conduits **232***a* and **232***b*. Each of these data conduits **232** may be configured to transmit a data signal to a single die interconnect **220**. For example, the data

'060: EX1001 at 6:48-:51
'160: EX1001 at 6:51-:54

'060: Paper 1 (Pet.) at 11-14; EX1003 at ¶¶ 61-65
'160: Paper 1 (Pet.) at 11-16; EX1003 at ¶¶61-65

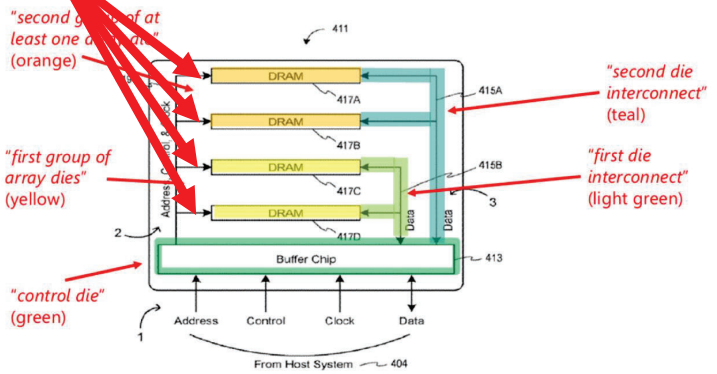
Appx10684

Rajan's separate chip selects and shared data buses (on the right) were well known for rank multiplication (without data collisions) as shown by Ellsberry (on the left)



EX1026 (Ellsberry) at Figs. 12-13

Four separate chip selects (CS0A, CS1A, CS0B, CS1B), one for each rank (four ranks total)



the extra address bits may be decoded by the buffer chip to individually select the DRAM chips, utilizing separate chip select signals (not shown) to each of the DRAM chips in the stack.

EX1015 (Rajan) at 6:34-:38 & Fig. 4

'060: Paper 1 (Pet.) at 27, 29-30; Paper 24 (Reply) at 8-9
'160: Paper 1 (Pet.) at 29, 21-32; Paper 24 (Reply) at 9-10

Appx10710

Undisputed: JEDEC standards taught how to avoid collisions on shared data lines (green) using properly timed chip-select signals (red) for commands



Figure 2 — Qual-stacked / Quad-die DDR3 SDRAM x8 rank association

EX1019 at 12

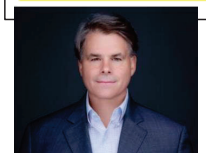
Table 6 — Command Truth Table

Function	Abbreviation	CKE		CS#	RAS#	CAS#	WE#	BA0-BA2	A13-A15	A12-BC#	A10-AP	A0-A9, A11	Notes
		Previous Cycle	Current Cycle										
Mode Register Set	MRS	H	H	L	L	L	L	BA	OP Code				
Write (Fixed BL8 or BC4)	WR	H	H	L	H	L	L	BA	RFU	V	L	CA	
Read (Fixed BL8 or BC4)	RD	H	H	L	H	L	H	BA	RFU	V	L	CA	
Device Deselected	DES	H	H	H	X	X	X	X	X	X	X	X	11

EX1019 at 33

Q. (By Mr. Chandler) And a person of ordinary skill in the art, reviewing the stacked memory devices shown on page 12 of the DDR3 standard marked as Exhibit 1019, would expect those stacked memory devices to operate correctly, without collisions, on the set of data lines labeled "DQ0-7," correct?

THE DEPONENT: I think they would understand this figure to be describing one way of implementing this D- -- DDR3 implementation within the confines of JEDEC and, taken as a whole, to provide an operating -- or a correctly operating device -- a functioning operational device per JEDEC's definition.



EX1052 (Brogioli) at 303:12-304:1

'060: Paper 1 (Pet.) at 46; Paper 24 (Reply) at 4-5
'160: Paper 1 (Pet.) at 37-38; Paper 24 (Reply) at 4-5

Appx10711

JEDEC standard taught using programmable timing parameters (e.g., CAS latency, CL, of 5 to 10 clock cycles) to avoid collisions

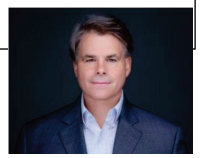
Q. Under the JEDEC standard for DDR3 memory devices, CAS latency could be set to 5, 6, 7, 8, or 9 clock cycles, as shown on page 23 of Exhibit 1019; is that correct?

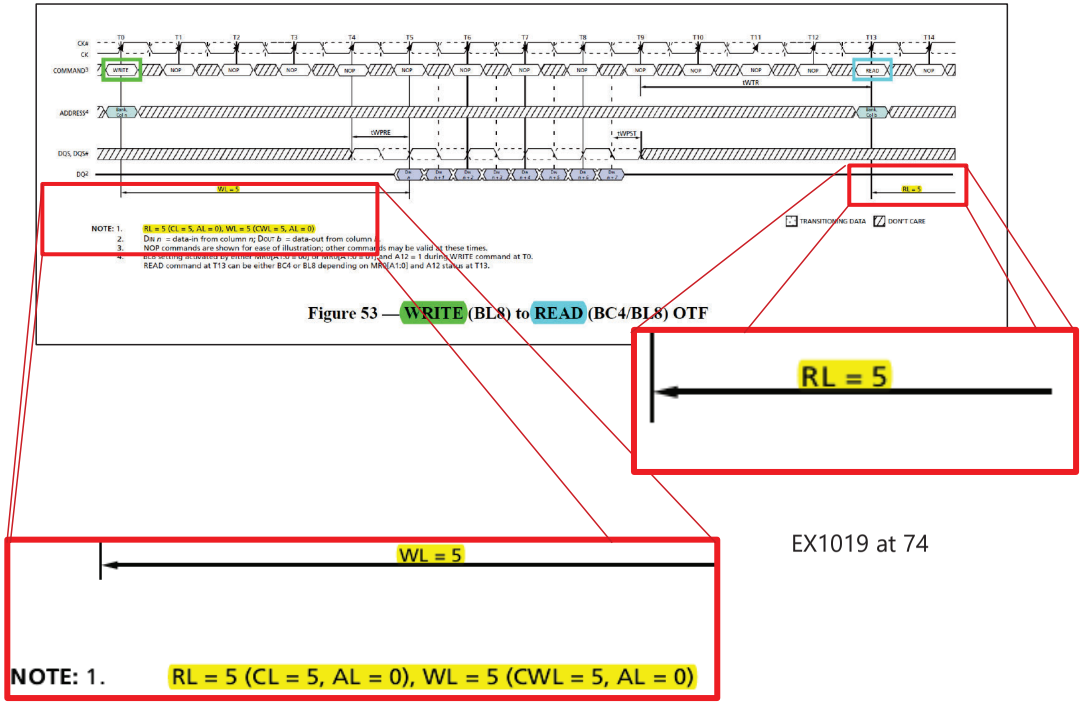
A. That looks like at least some of the values that the -- excuse me -- that the CAS latency could be set to. I think you mentioned 5 through 9.

Q. Or 10 clock cycles; right? That's also specified in the standard?

A. I see that, yes. Those are the -- that's the CAS latency. You set the bits accordingly to get a different value.

EX1052 (Brogioli) at 110:5-:17;
see also id. at 116:2-117:9



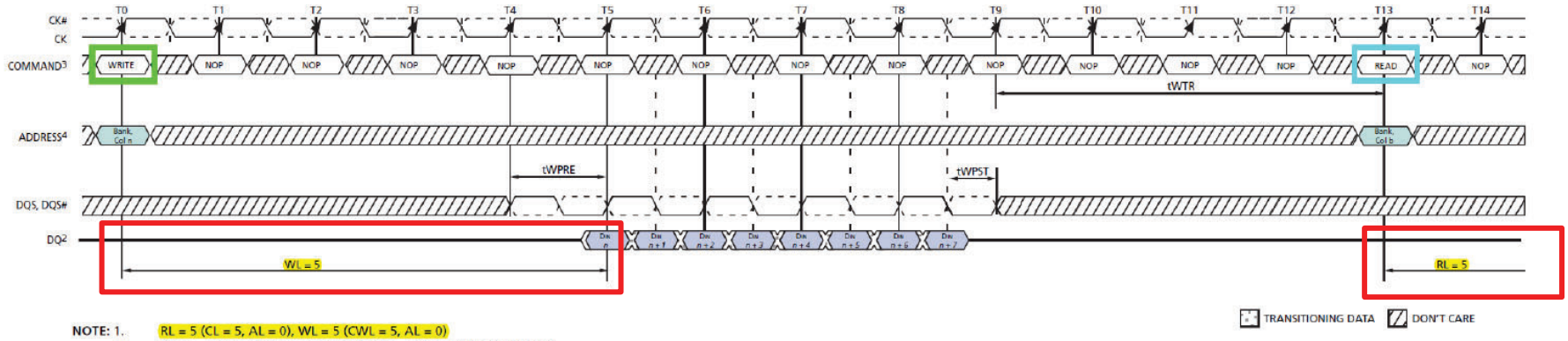


EX1019 at 74

Appx10712

'060: Paper 24 (Reply) at 4-5
'160: Paper 24 (Reply) at 4-5

JEDEC standard teaches that READ latency (RL) can be increased independently of WRITE latency (WL), so no collision



NOTE: 1. RL = 5 (CL = 5, AL = 0), WL = 5 (CWL = 5, AL = 0)
2. DIN *n* = data-in from column *n*; DOUT *b* = data-out from column *b*.
3. NOP commands are shown for ease of illustration; other commands may be valid at these times.
4. BL8 setting activated by either MR0[A1:0 = 00] or MR0[A1:0 = 01] and A12 = 1 during WRITE command at T0. READ command at T13 can be either BC4 or BL8 depending on MR0[A1:0] and A12 status at T13.

Figure 53 — WRITE (BL8) to READ (BC4/BL8) OTF

EX1019 at 74;
EX1052 (Brogioli) at 116:2-117:9

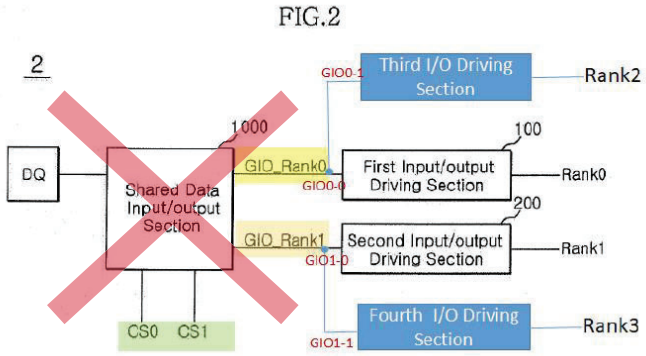
problem with shared TSVs. For instance, if the read and write latencies were increased by the same amount, read and write data would still appear at the same relative timing and collisions would still occur, as illustrated below.

'060: Paper 28 (Sur-Reply) at 6-8:
'160: Paper 28 (Sur-Reply) at 6-8:

'060: Paper 24 (Reply) at 4-5
'160: Paper 24 (Reply) at 4-5

Appx10713

Petition did NOT propose a combination with Kim's circuit operating *the same way* as in Kim's original 2-rank figure (Netlist's Strawman #4)



EX2023, p.70

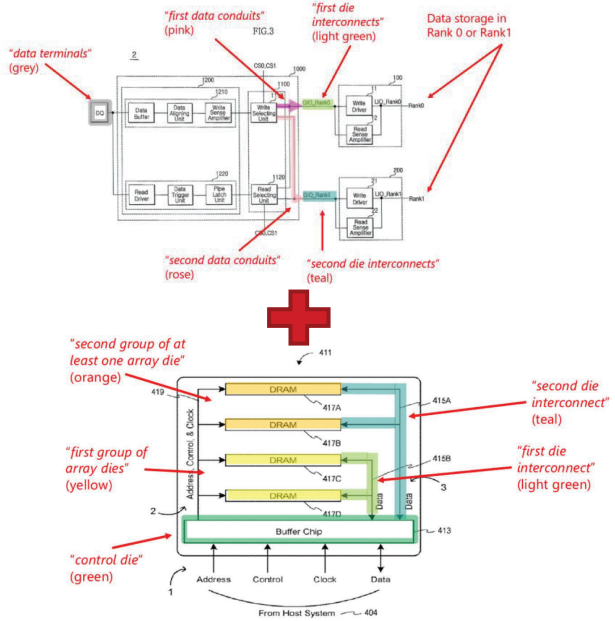
Q. In your four-rank version of Figure 2 of Kim shown at the top of page 70 of your expert declaration marked as Exhibit 2023, you assume that Kim's circuits would operate the same way as in Kim's original Figure 2, correct?

A. That is correct, as I recall figure -- the -- the earlier figures of Kim.



EX1052 (Brogioli) at 320:5-:11

Kim + Rajan



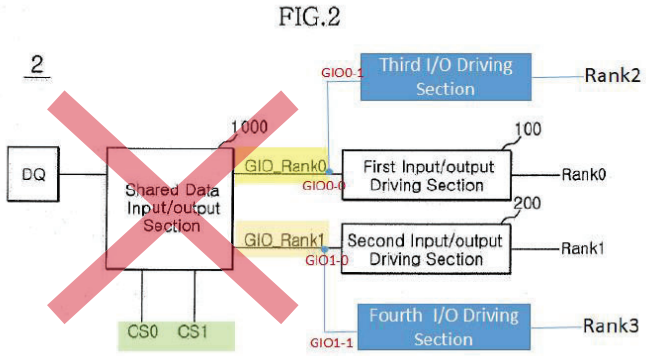
the extra address bits may be decoded by the buffer chip to individually select the DRAM chips, utilizing separate chip select signals (not shown) to each of the DRAM chips in the stack.

EX1015 (Rajan) at 6:34-:38

EX2025 (Wolfe) at 46:3-:12, 55:18-57:9, 79:13-80:11, 86:12-:17, 201:3-202:5, 204:2-205:8
'060: Paper 24 (Reply) at 11; Paper 20 (POR) at 32-38
'160: Paper 24 (Reply) at 12; Paper 20 (POR) at 32-38

Appx10724

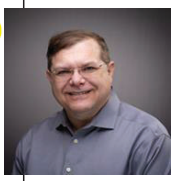
Petition did NOT propose a combination with Kim's circuit operating *the same way* as in Kim's original 2-rank figure (Netlist's Strawman #4)



EX2023, p.70

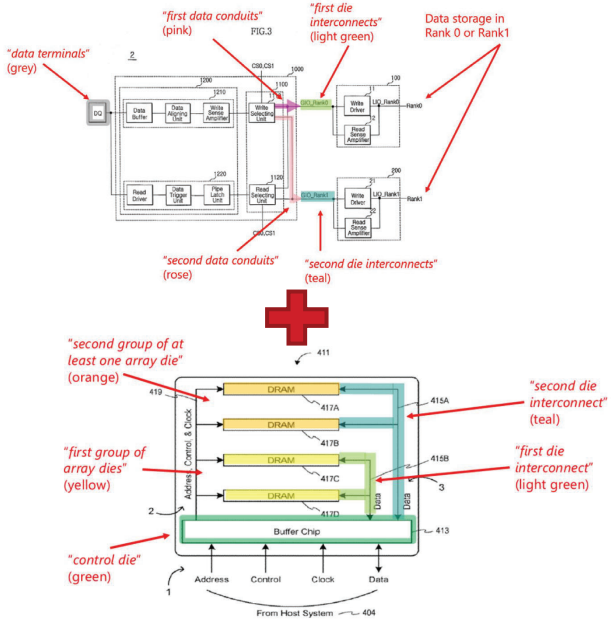
Q. So like the Rank0 is also going to be added to the GIO_Rank0 line, basically just a duplicate --

A. One option would be to add -- to connect Rank2 to the GIO_Rank0 line and to connect Rank3 to the GIO_Rank1 line, and create the appropriate control signals so that only one rank would drive each of those I/O lines at a time or only one rank would respond to write operations on them. There are lots of different ways to make rank multiplication work.



EX2025 (Wolfe) at 46:3-:12

Kim + Rajan



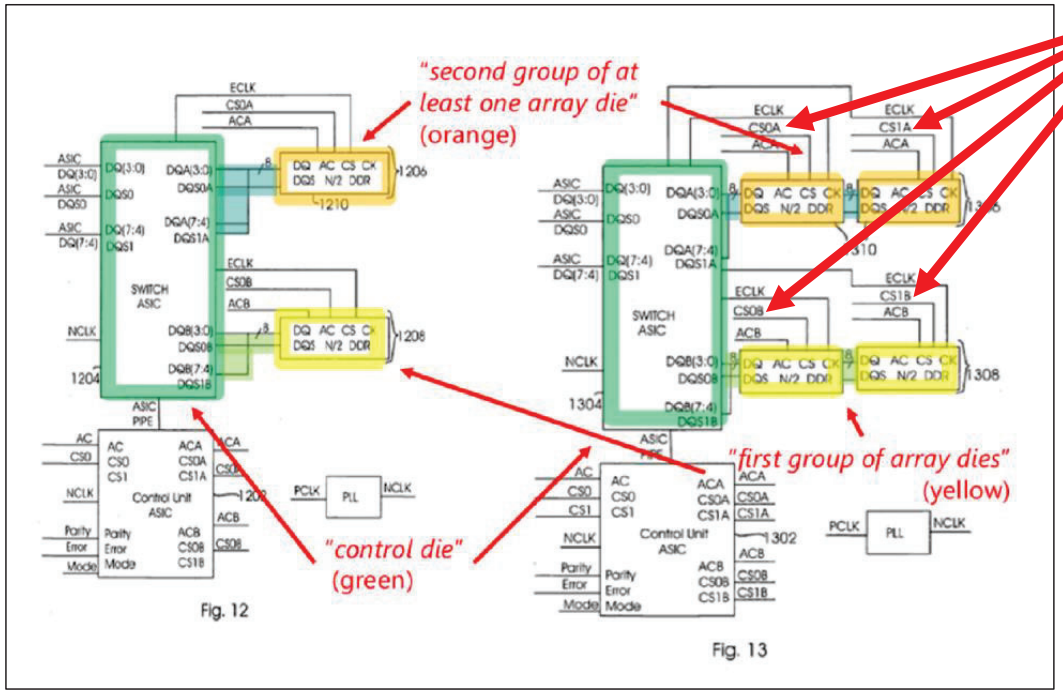
the extra address bits may be decoded by the buffer chip to individually select the DRAM chips, utilizing separate chip select signals (not shown) to each of the DRAM chips in the stack.

EX1015 (Rajan) at 6:34-:38

EX2025 (Wolfe) at 46:3-:12, 55:18-57:9, 79:13-80:11, 86:12-:17, 201:3-202:5, 204:2-205:8
'060: Paper 24 (Reply) at 11; Paper 20 (POR) at 32-38
'160: Paper 24 (Reply) at 12; Paper 20 (POR) at 32-38

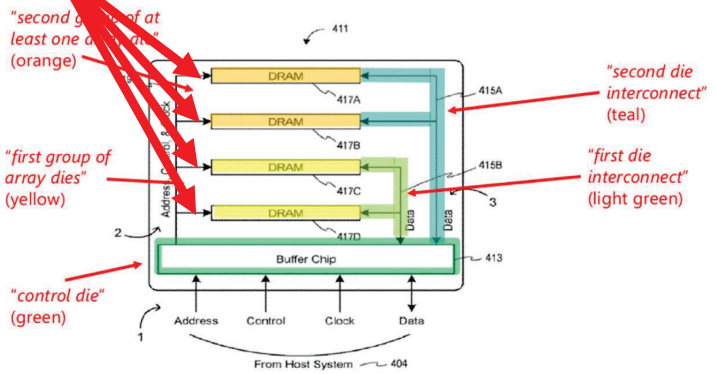
Appx10725

Rajan's separate chip selects and shared data buses (on the right) were well known for rank multiplication (without data collisions) as shown by Ellsberry (on the left)



EX1026 (Ellsberry) at Figs. 12-13

Four separate chip selects (CS0A, CS1A, CS0B, CS1B), one for each rank (four ranks total)



the extra address bits may be decoded by the buffer chip to individually select the DRAM chips, utilizing separate chip select signals (not shown) to each of the DRAM chips in the stack.

EX1015 (Rajan) at 6:34-:38 & Fig. 4

'060: Paper 1 (Pet.) at 27, 29-30; Paper 24 (Reply) at 8-9
'160: Paper 1 (Pet.) at 29, 21-32; Paper 24 (Reply) at 9-10

Appx10726

Number of transistors in Wyman (on the right) corresponds to "driver size" in the '060/'160 Patents

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.

'060: EX1001 at 17:36-:48
'160: EX1001 at 17:24-:36

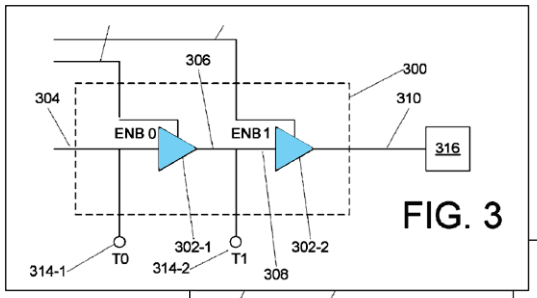


FIG. 3

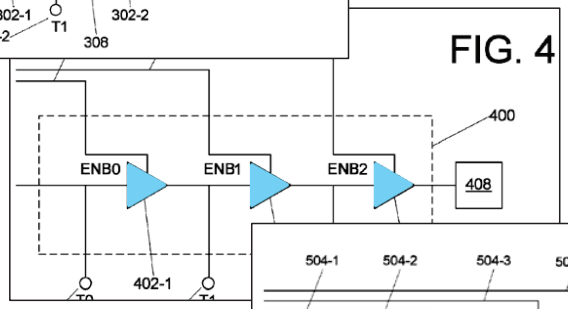


FIG. 4

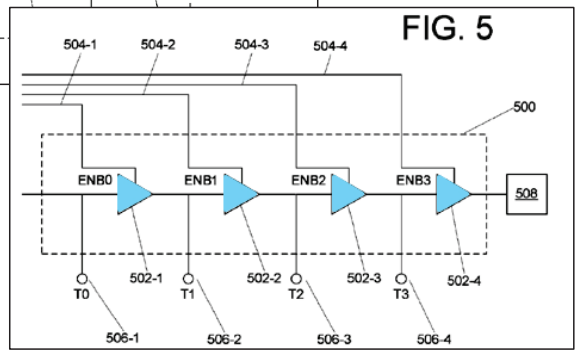


FIG. 5

EX1017 (Wyman) at Figs. 3-5

'060: Paper 24 (Reply) at 18-20
'160: Paper 24 (Reply) at 14-16

Appx10734

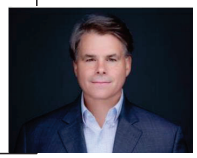
Number of transistors in Wyman (on the right) corresponds to "driver size" in the '060/'160 Patents

Q. So in the situation where only driver 502-1 is enabled, the signal is driven only by driver 502-1 to tap T1, which is also labeled "506-2"; is that correct?

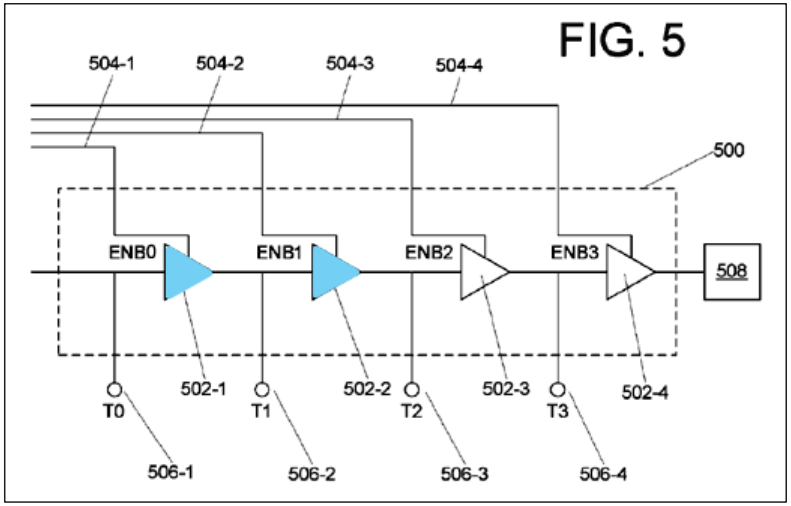
A. That appears to be at least what Figure 5 is illustrating.

Q. And if Figure 5 of Wyman marked as Exhibit 1017, if only drivers 502-1 and 502-2 are enabled, the input signal is driven by both driver 502-1 to tap T1 marked as "506-2," and driver 502-2 will drive to tap T2 marked as "506-3," correct?

A. I think that sounds right, in -- Wyman talked about different configurations or -- different configurations of the different drivers to divide up the driver level.



EX1052 (Brogioli) at 328:10-:24



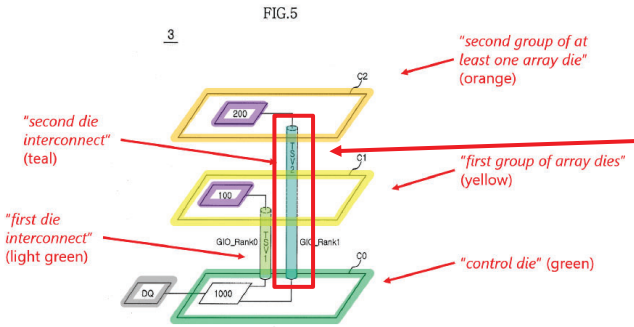
EX1017 (Wyman) at Fig. 5

'060: Paper 24 (Reply) at 18-20
'160: Paper 24 (Reply) at 14-16

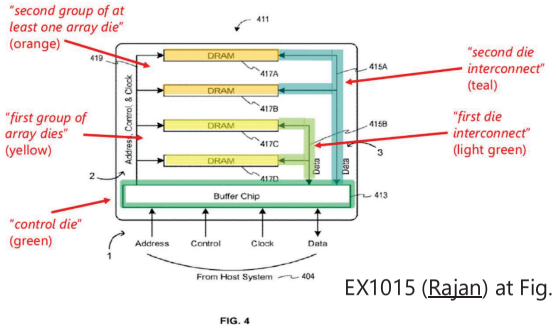
Appx10735

Using Wyman's drivers of different sizes for Kim's TSVs of different lengths was obvious

Kim + Rajan



EX1014 (Kim) at Fig. 5



EX1015 (Rajan) at Fig. 4

Wyman: additional drive for longer die interconnects

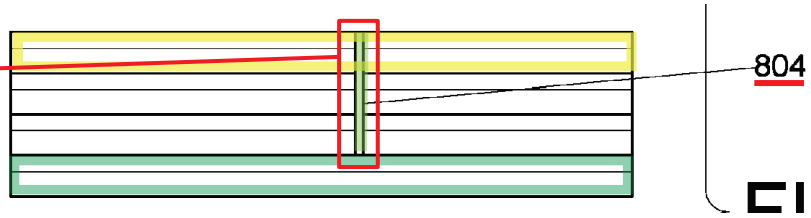


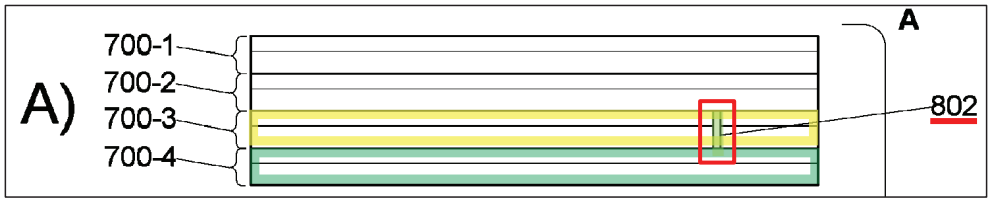
FIG. **8**B illustrates, in simplified form, a longer through-chip via connection (**804**) extending directly between the mother chip (**700-4**) and daughter **3** (**700-1**) in a straight line. The increased resistance, capacitance and impedance of such a connection (**804**) might require additional drive than that referred to in connection with FIG. **8**A. In other words, the

EX1017 (Wyman) at 6:21-:26 & Fig. 8B

Appx10736

'060: Paper 24 (Reply) at 18-20
'160: Paper 24 (Reply) at 14-16

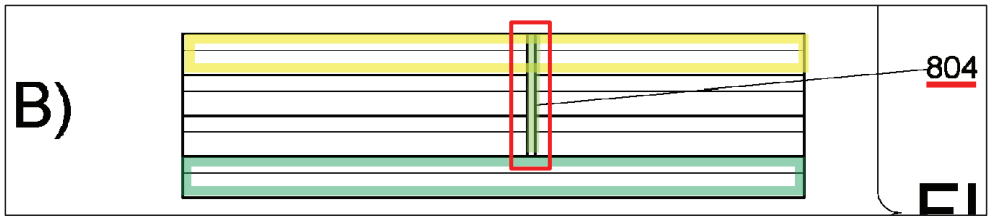
Wyman, like the '060/'160 Patents, teaches that the difference in load for TSVs of different lengths is NOT negligible



EX1017 (Wyman) at Fig. 8A

FIG. **8**A illustrates, in simplified form, a through-chip via connection (**802**) extending between the mother chip (**700-4**) and daughter **1** (**700-3**). Presuming that this connection would require minimal drive, if a driver portion **500** of FIG. **5** was

EX1017 (Wyman) at 6:15-:18



EX1017 (Wyman) at Fig. 8B

FIG. **8**B illustrates, in simplified form, a longer through-chip via connection (**804**) extending directly between the mother chip (**700-4**) and daughter **3** (**700-1**) in a straight line. The increased resistance, capacitance and impedance of such a connection (**804**) might require additional drive than that referred to in connection with FIG. **8**A. In other words, the

EX1017 (Wyman) at 6:21-:26

'060: EX1001 at 4:42-:45
'160: EX1001 at 4:45-:48

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

'060: Paper 24 (Reply) at 18-20; Paper 1 (Pet.) at 20-21
'160: Paper 24 (Reply) at 14-16; Paper 1 (Pet.) at 22-23

Appx10737

Rajan discloses a JEDEC (DRAM) interface, even for non-DRAM dies

In similar argument as to Riho, Petitioner's contention is that it would have been obvious "to create a package with an *interface* that complied with the well-known JEDEC standards." Pet. 93 (emphasis added); *see also* Pet. 94 (asserting that "emulating a standard JEDEC interface was well-known"). Thus, on this record, we do not agree with Patent Owner that Petitioner's reliance on JEDEC compliance undermines Petitioner's contentions.

'060: Paper 13 (ID) at 24;
see also '160: Paper 13 (ID) at 25

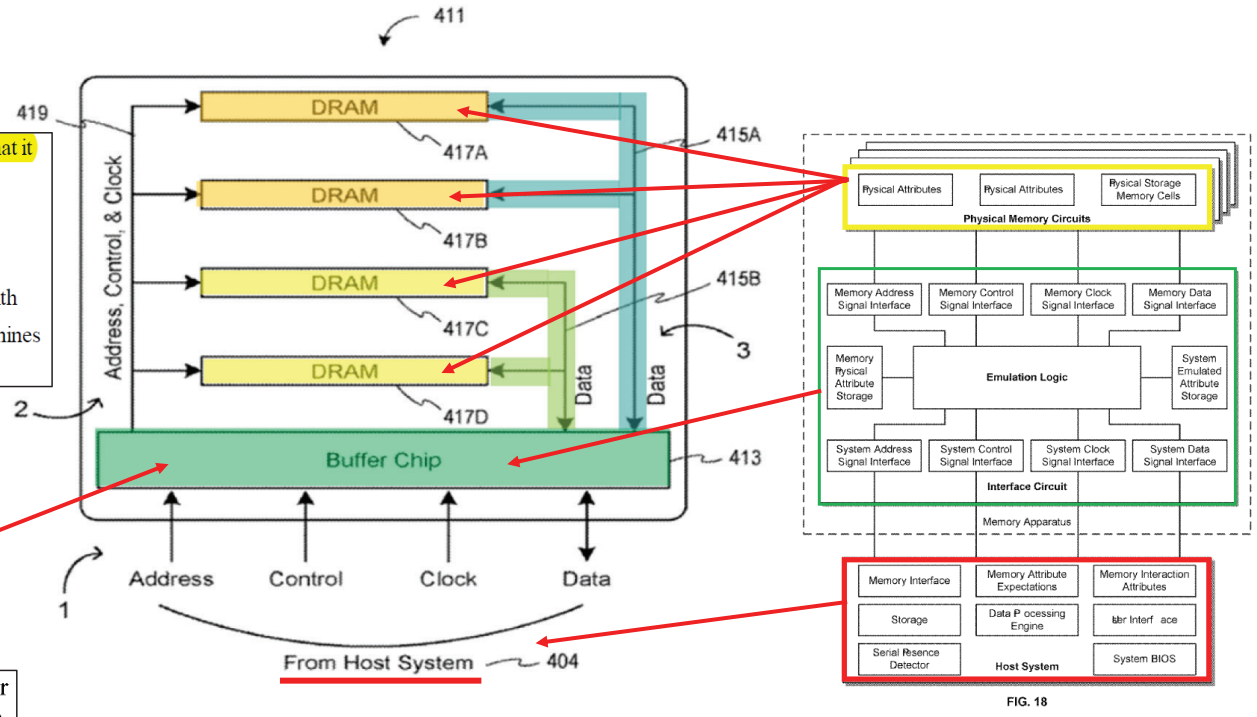


FIG. 4

FIG. 18

The interface circuit may take the form of or incorporate, or be incorporated into, a register, an AMB, a buffer, or the like, and may comply with Joint Electron Device Engineering Council (JEDEC) standards, and may have forwarding, storing, and/or buffering capabilities.

EX1015 (Rajan) at 4:20-:24; *see also id.* at 4:48-:50, 14:6-:11, 14:39-:62 & Fig. 18

'060: Paper 24 (Reply) at 13-15
'160: Paper 24 (Reply) at 16-18

Appx10747

Paper No. 1

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.,

Petitioner,

v.

NETLIST, INC.,

Patent Owner

Patent No. 8,787,060

Issued: July 22, 2014

Filed: November 3, 2011

Inventor: Hyun Lee

Title: Method and Apparatus for Optimizing Driver Load in a Memory Package

Inter Partes Review No. IPR2022-01428

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,787,060

CLAIM LISTING

Ref. #	Listing of Challenged Claims
1.a	1. A memory package, comprising:
1.b	a plurality of input/output terminals via which the memory package communicates data and control/address signals with one or more external devices;
1.c	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;
1.d.1	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,
1.d.2	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
1.e.1	a control die comprising
1.e.2	at least a first data conduit between the first die interconnect and a first terminal of the plurality of input/output terminals,
1.e.3	and at least a second data conduit between the second die interconnect and the first terminal, the first terminal being a data terminal,
1.e.4	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	The memory package of claim 1, wherein the control signals include data path control signals for controlling the first and second data conduits.
3	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.a	The memory package of claim 3, wherein the control signals include command/address signals and
4.b	wherein the control die is configured to provide the command/address signals to the plurality of stacked array dies.

Ref. #	Listing of Challenged Claims
5	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.
6.a	The memory package of claim 1, wherein the control die further comprises chip-select conduits, the memory package further comprising:
6.b	third die interconnects coupled between respective chip-select conduits and respective ones of the plurality of stacked array dies.
7.a	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 between a first load on the first data conduit and a second load on the second data conduit,
7.b	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.
8.a	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,
8.b.1	wherein the control die is configurable to operate in any one of a first mode and a second mode, and wherein:
8.b.2	in the first mode, the control die receives the data path control signals from the one or more external devices; and
8.b.3	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.
9	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.

Ref. #	Listing of Challenged Claims
10	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.
11.a	A memory package, comprising:
11.b	a plurality of input/output terminals via which the memory package communicates data and control/address signals with one or more external devices;
11.c	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;
11.d.1	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,
11.d.2	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
11.e.1	a control die comprising
11.e.2	at least a first data conduit between the first die interconnect and a first terminal of the plurality of input/output terminals,
11.e.3	at least a second data conduit between the second die interconnect and the first terminal, and
11.e.4	chip select conduits for providing chip select signals to respective array dies;
11.e.5	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	The memory package of claim 11, wherein the chip select conduits pass through the control die.
13	The memory package of claim 11, wherein the chip select conduits include drivers to drive the chip select signals to the respective array dies.

Ref. #	Listing of Challenged Claims
14	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.a	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
15.b	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	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	The memory package of claim 11, wherein the control circuit controls the respective states of the fist 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	The memory package of claim 16, wherein the control/address signals comprise: at least one command signal, at least one address signal, and at least one data path control signal.
19.a	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 second terminals of the plurality of terminals, and
19.b	wherein the control circuit controls the respective states of the first data conduit in response to the data path control signals.
20.a.1	A method for optimizing load in a memory package comprising
20.b.1	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,
20.b.2	at least a first die interconnect and a second die interconnect,
20.b.3	a control die, and

C. Background Technology

1. JEDEC Standards

As recognized by the provisional application, EX1005, ¶¶[0004-05], a POSITA would have been familiar with the JEDEC standards for memory devices and memory modules. EX1003, ¶¶132-134; EX1023, p.9 & Fig.16; EX1022, pp.332-35, 318-20. For example, a POSITA would have understood that DDR SDRAM memory devices were already standardized by JEDEC. EX1015, 8:6-15; EX1016, ¶[0027]; EX1003, ¶132. Moreover, a POSITA would have understood that JEDEC had standardized memory modules and corresponding interface circuits such as "a register, an AMB, a buffer, or the like." EX1015, 4:20-24; EX1003, ¶132. Examples of relevant standards include the JEDEC standard for DDR3 memory devices, EX1019-21, and the JEDEC specification for GDDR4 memory devices, EX1035.

2. Electrical Load and Drivers

A POSITA would have also been familiar with basic concepts like electrical loads of circuitry used in stacked memory devices and corresponding driving requirements to drive the load of such circuitry. EX1001, 2:8-15 (acknowledging known tradeoff between driver size, load, space, and power); EX1003, ¶135. For example, the greater the "fan-out," the greater the capacitive loading, and thus the greater the propagation delay, as shown below. EX1030, pp.132-138, Fig.3.55 (below); EX1003, ¶136.

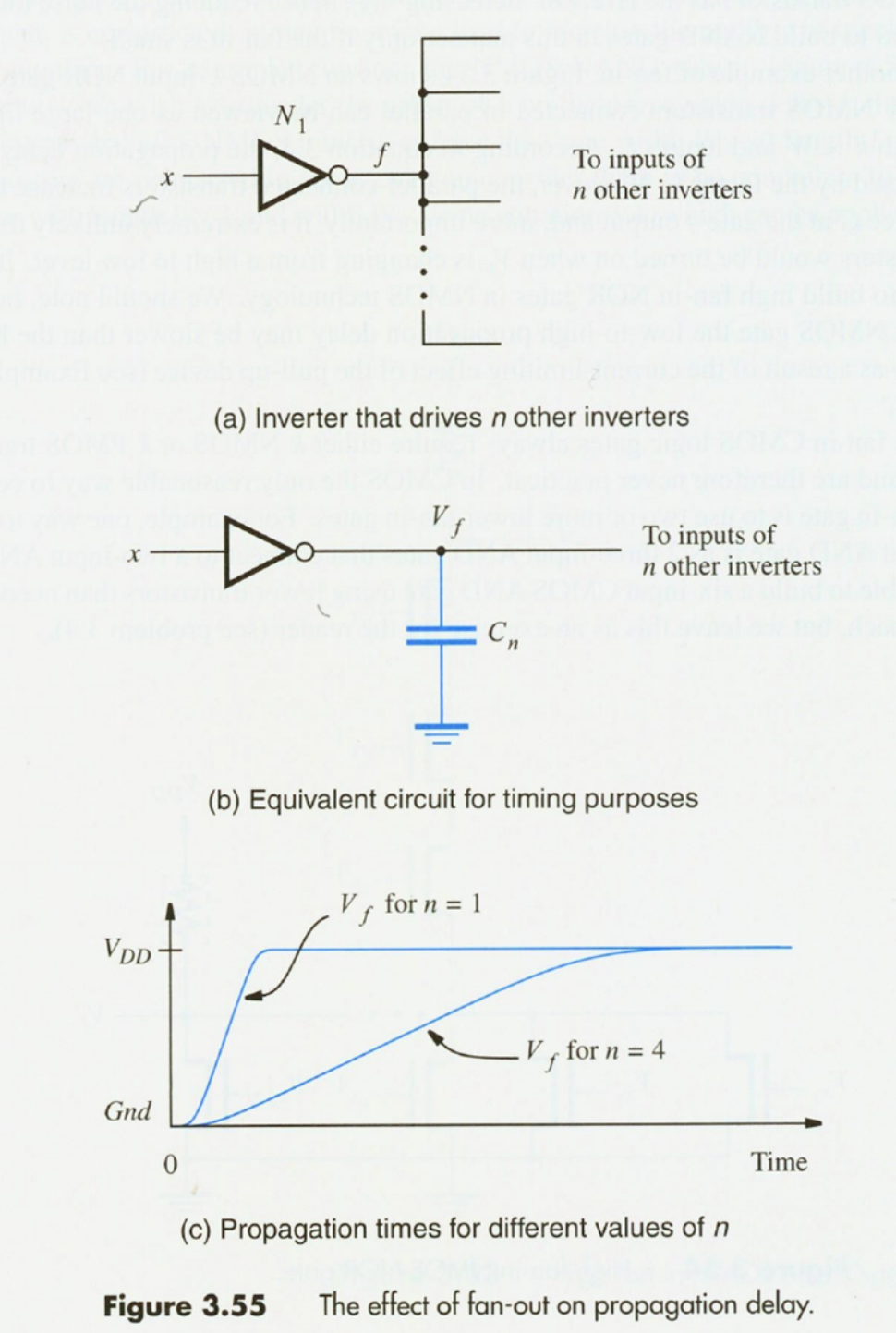


Figure 3.55 The effect of fan-out on propagation delay.

A POSITA would have understood that buffers (such as tri-state buffers) were often used to improve performance when driving a large capacitive load, and

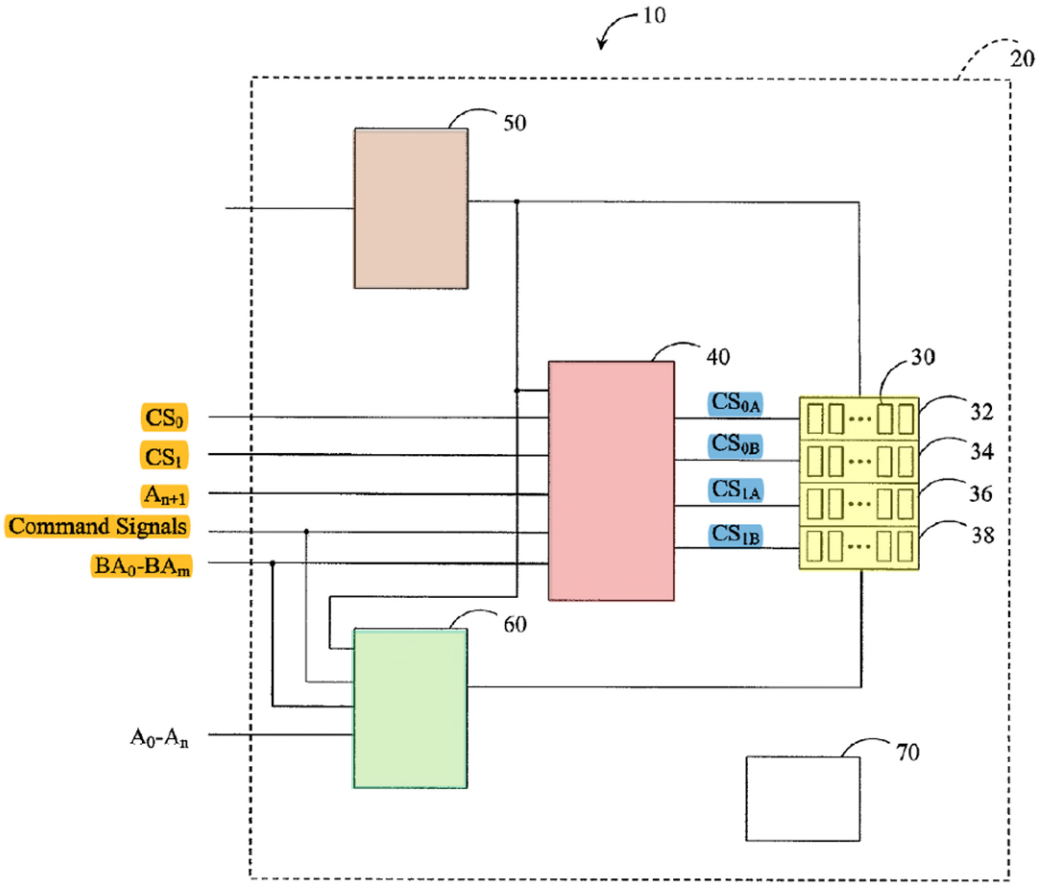
that "[b]uffers can be created with different amounts of *drive capability*." EX1030, pp.135-138; EX1003, ¶137.

3. Rank Multiplication

As admitted by the 060 Patent, a POSITA also would have been familiar with "rank multiplication"—e.g., where a memory module can appear to the computer system as if it has two ranks of higher density (and more expensive) memory devices complying with the JEDEC standard, when it physically has four ranks of lower density (and less expensive) memory devices—which can result in substantial cost savings, providing a motivation to implement the technique in a wide range of memory packages and memory modules. EX1001, 22:37-39 ("Embodiments of rank multiplication are described in greater detail in U.S. Pat. Nos. 7,289,386 ["the 386 Patent," EX1028, 4:47-5:10, 32:58-33:16] and 7,532,537 [EX1031] …."); EX1041, 2:15-19; EX1015, 6:30-7:67; EX1003, ¶138.

As illustrated in Figure 1A (below) of the prior-art 386 Patent, rank multiplication involves mapping chip select signals (e.g., CS_0, CS_1, orange) and address signals (e.g., A_{n+1}, orange) received from a system controller complying with the JEDEC standard into ***more*** chip select signals (CS_{0A}, CS_{0B}, CS_{1A}, CS_{1B}, blue) to control the higher number of ranks of memory devices (yellow) on the module. EX1028, 2:34-38, 5:11-50, 6:63-8:45, Fig.1A; EX1003, ¶¶139-140.

Figure 1A:



More ranks would result in more load, however, so the prior-art 386 Patent explains that an isolation device (e.g., 120 below) should be used to reduce the electrical load created by the additional ranks of memory devices, and to avoid certain data collisions. EX1028, 24:41-56, 26:12-19, Fig.6D (below); *see also* EX1031, 6:48-62; EX1003, ¶140.

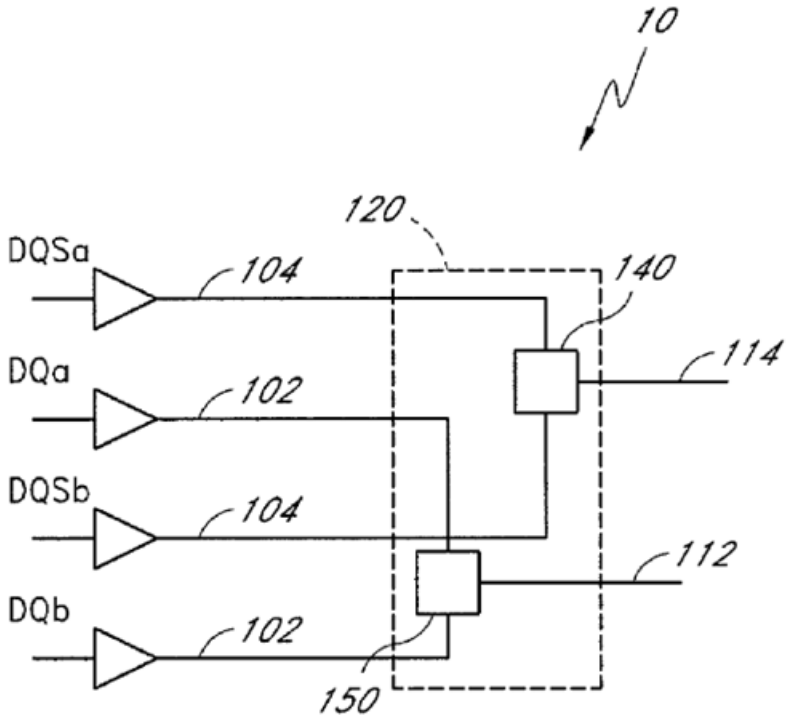


FIG. 6D

While the prior-art 386 Patent discloses rank multiplication implemented in a separate controller (the register), *see* EX1028, 5:42-48, POSITAs would have known that it could ***also*** be implemented within a single memory package on the control die, or in both locations, *see* EX1015, 6:30-7:67; EX1041, 2:15-19; EX1001, 18:38-19:7, Fig.7 (below, showing prior-art module implementing rank multiplication in both the module controller 712 and the memory package control chip 722); EX1003, ¶141.

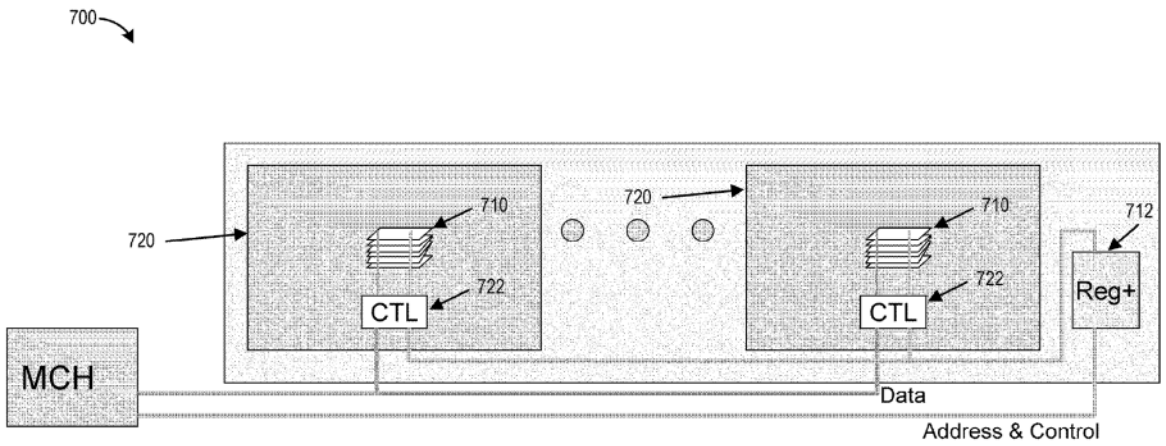


FIG. 7

D. The 060 Patent

1. Technical Overview

The 060 Patent "relates to memory devices and memory modules …. specifically … [ways] for reducing the load of drivers of memory packages included on the memory modules." EX1001, 1:18-22.

The 060 Patent admits as prior art the memory package designs depicted in Figures 1A and 1B (annotated below),[1] which include, for example, "array dies 160 [(yellow)] and a control die 170 [(green)] that…[includes] a driver 184 that drives data signals to each of the array dies 160 along a corresponding die

[1] Unless stated otherwise, all emphasis in quotes and color annotations in figures have been added.

interconnect 182 [(light green)]." EX1001, 1:30-33, 1:63-2:4, Fig.1B; EX1003, ¶61.

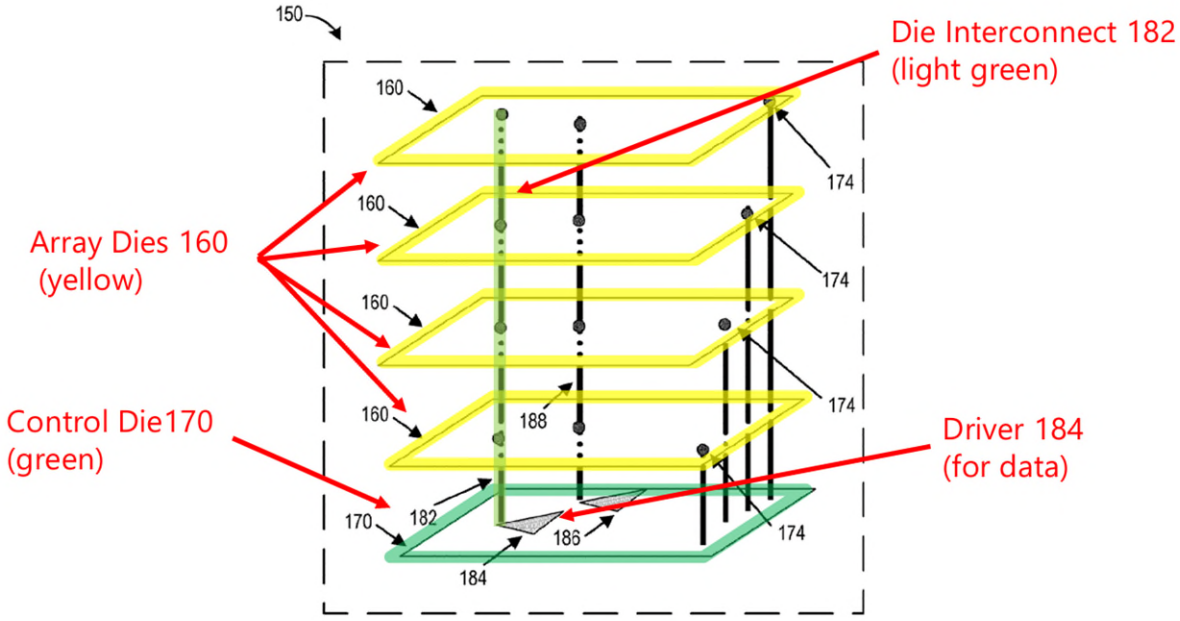
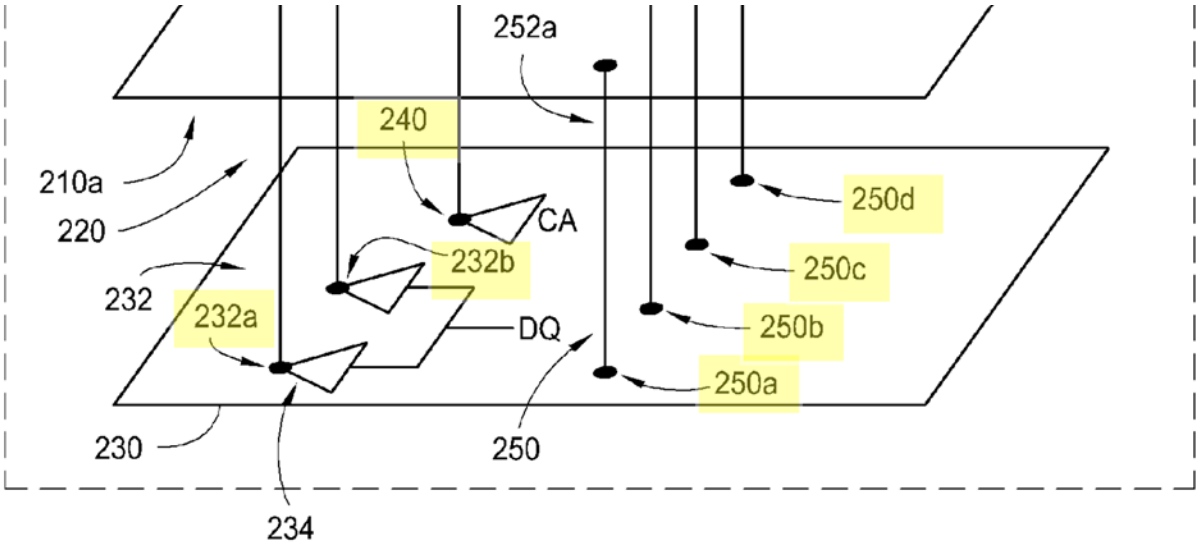


FIG. 1B

The 060 Patent teaches that each die interconnect has a corresponding "conduit" (e.g., 232a below), configured to transmit a signal to the die interconnect, and which may include a driver. EX1001, Fig.2, 6:48-62, 9:28-60, 21:18-20; EX1003, ¶62.



As illustrated by Figure 2 (below), the 060 Patent purports to improve upon the prior art (e.g., Figure 1B) by replacing ***one*** data driver (e.g., driver 184 in Figure 1B) with ***two*** or more drivers 234 coupled to the same DQ data terminal, each driving the data signal through a respective die interconnect (220a, 220b) in electrical communication with a ***subset*** of the array dies, thereby reducing the overall load of each data conduit. EX1001, 1:19-21, 1:63-2:4, 2:8-15, 7:9-8:18, Fig.2; EX1003, ¶¶61-63.

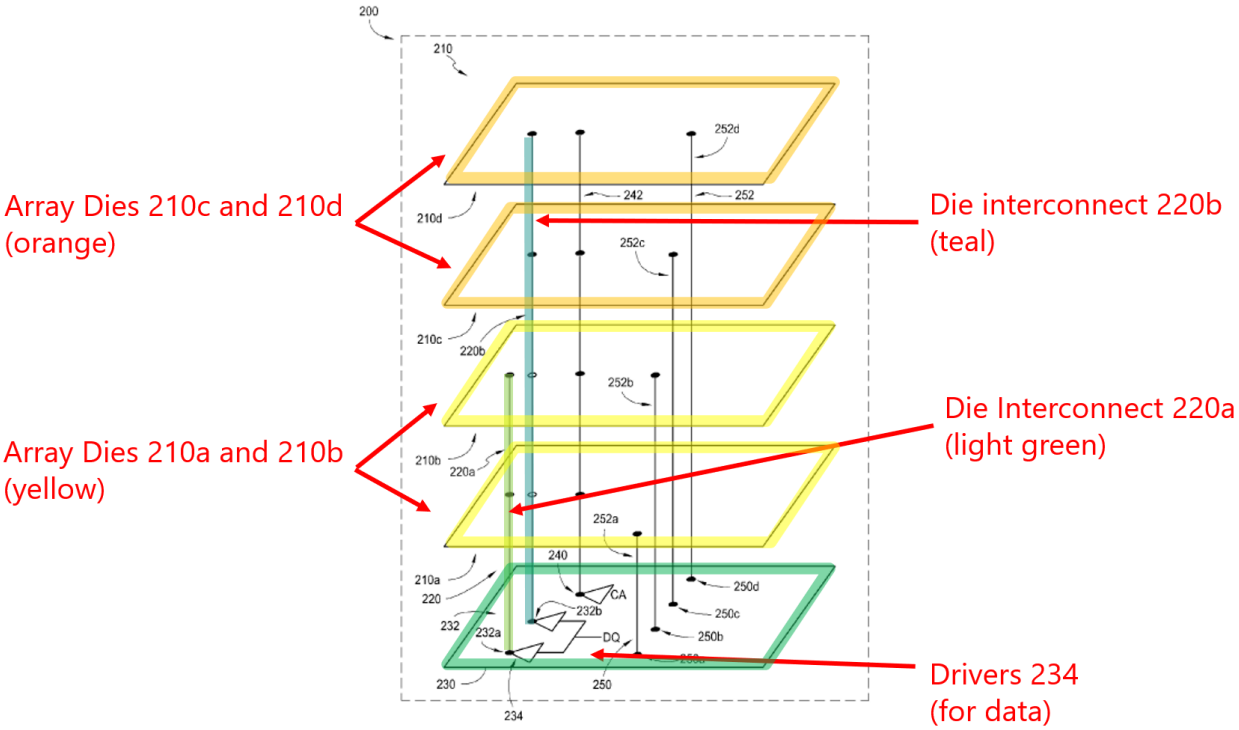


FIG. 2

In Figure 2, for a given data bit (e.g., D0), die interconnect 220a (light green) may be in electrical communication with data ports from array dies 210a and 210b (yellow) (illustrated by darkened circles) and not in electrical communication with data ports from array dies 210c or 210d (orange), while die interconnect 220b (teal) may be in electrical communication with data ports from array dies 210c and 210d and not 210a and 210b, even though die interconnect 220b may pass through these array dies (illustrated by unfilled circles), "e.g., through through-holes or vias of array dies 210a and 210b." EX1001, 5:63-6:1, 6:9-15; EX1003, ¶¶64-65.

optimizing output drive capacity "according to a change in the time constant caused by parasitic capacitance and parasitic resistance." EX1018, ¶¶[0003, 0010, 0097], Fig.7A (below); EX1003, ¶¶129-131.

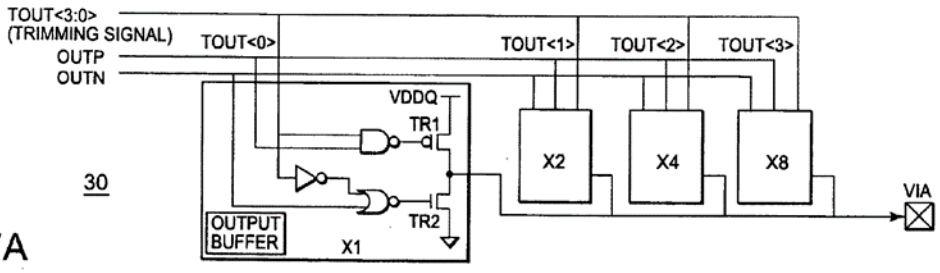


FIG. 7A

VI. CLAIM CONSTRUCTION

In related litigation, Patent Owner has interpreted some claim terms broadly for purposes of infringement, *see* EX1044, pp.63-68; EX1046, pp.63-69, even though a narrower interpretation may be more reasonable, but Petitioner contends that no express constructions are needed for this proceeding because the claims are obvious under either interpretation. EX1003, ¶112-113.

VII. ARGUMENT

A. Ground 1 (claims 1-6, 8-14, 16-19, 29-34)

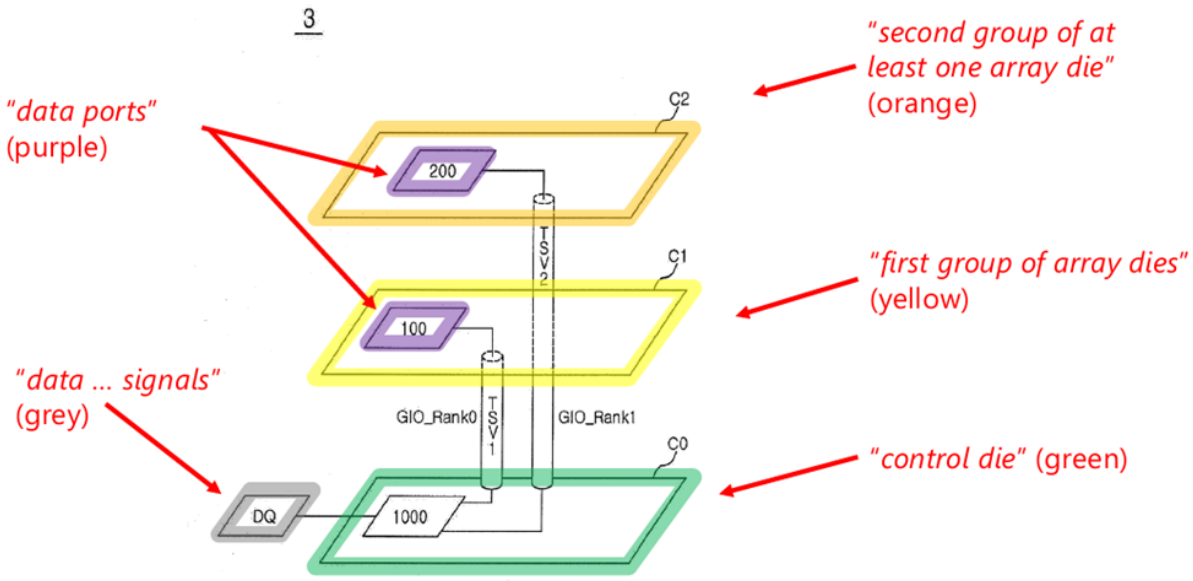
1. Ground 1 combination: Kim (EX1014) and Rajan (EX1015)

Ground 1 combines Kim (EX1014) with Rajan (EX1015), both of which, like the 060 Patent, disclose the same basic memory package structure including groups of stacked memory chips (yellow, orange) and a shared interface circuit (green). EX1014, ¶¶[0012, 50], Fig.5 (first below); EX1015, 1:51-60, 2:6-7, 4:48-

50, Fig.4 (second below); EX1001, 1:18-21, 5:23-26, Fig.2 (above p.14); EX1003, ¶¶150, 154-155.

Kim:

FIG.5



Rajan:

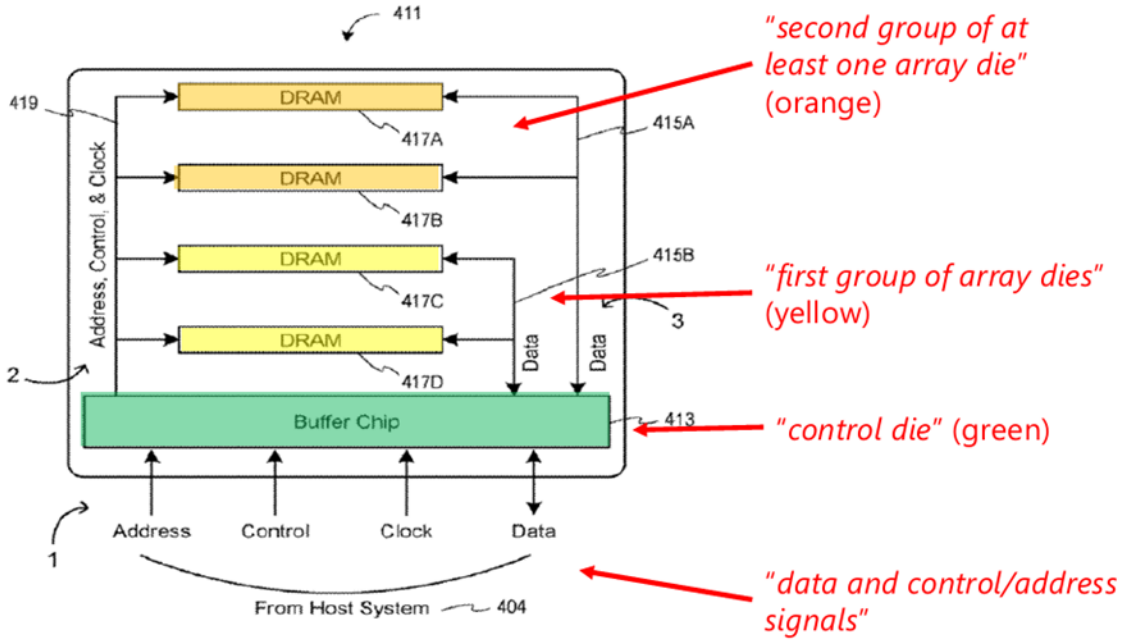


FIG. 4

A POSITA would have been motivated to combine Kim and Rajan for several reasons. First, as explained and shown above, they are analogous art with similar structures, so a POSITA would be motivated to look at the details taught by Rajan when implementing Kim, including details about the interface chip (green) and stacking the dies (yellow, orange). EX1003, ¶¶154-155.

Second, a POSITA would be motivated to create a package with an interface that complied with the well-known JEDEC standards, as taught by Rajan. EX1003, ¶¶152-153. Specifically, a POSITA would have found it obvious to use Rajan's terminals (e.g., in Figure 4, above) and corresponding data and control/address signals, in Kim's stacked memory package (e.g., in Figure 5, above), and been motivated to implement the functionality of a JEDEC compatible external data interface in Kim's control chip, as suggested by Rajan (including Figure 18, below), to make Kim's memory package compatible with the JEDEC standard. EX1014, ¶[0038], Fig.5; EX1015, 3:52-54, 8:8-11, Figs. 4, 18 (below); EX1003, ¶¶151-153. Indeed, the JEDEC standards were influential and well-known, as discussed above (p.6).

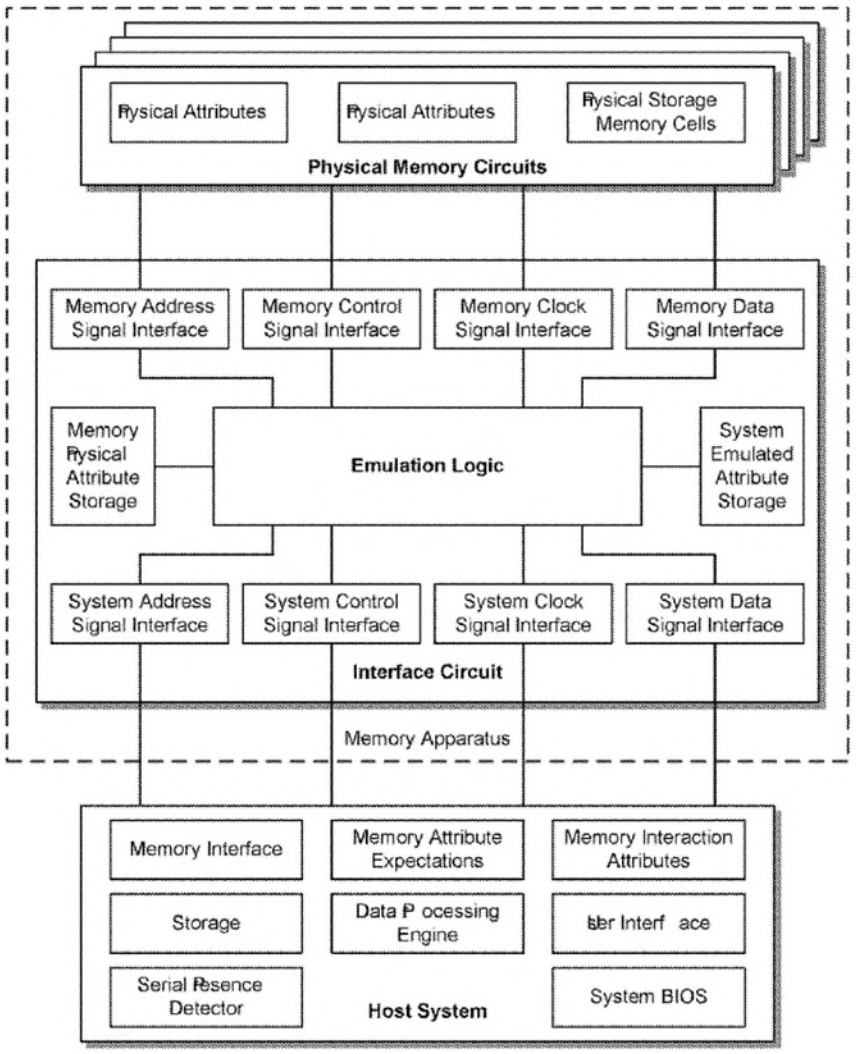


FIG. 18

Such a combination — following Rajan's suggestion to implement address, control, and data terminals following the JEDEC standards (including the DDR3 standard for stacked memory devices shown below, EX1019, p.12, Fig.3) — would have been well within a POSITA's level of skill since, as demonstrated by the 060 Patent's admitted prior art, emulating a standard JEDEC interface was a known option (as discussed above, pp.8-11). EX1003, ¶¶155-157; EX1001, 22:37-39; EX1015, 5:36-43, 4:20-24, Fig.4.

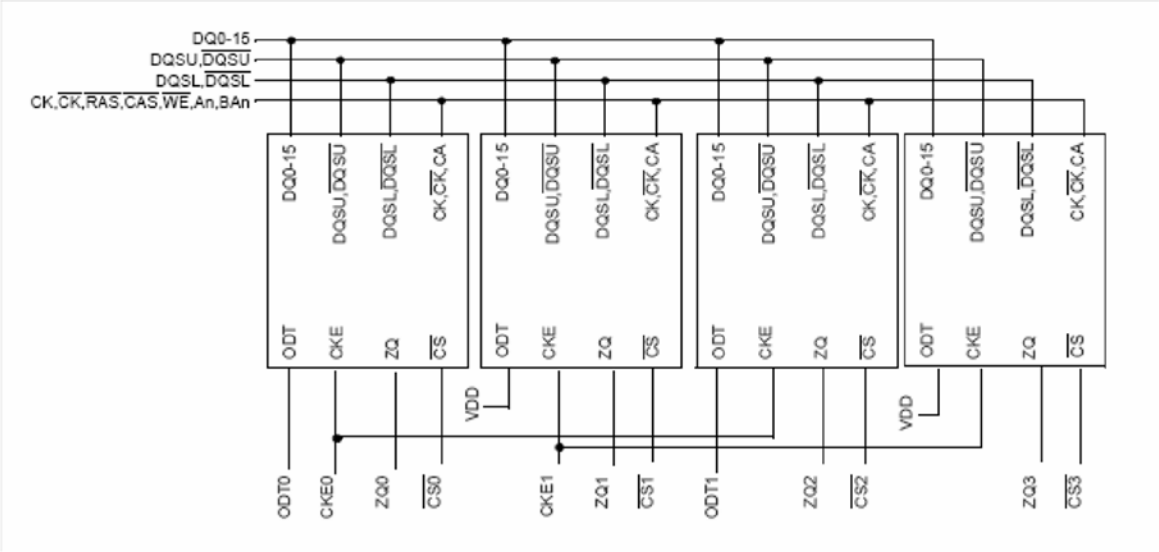
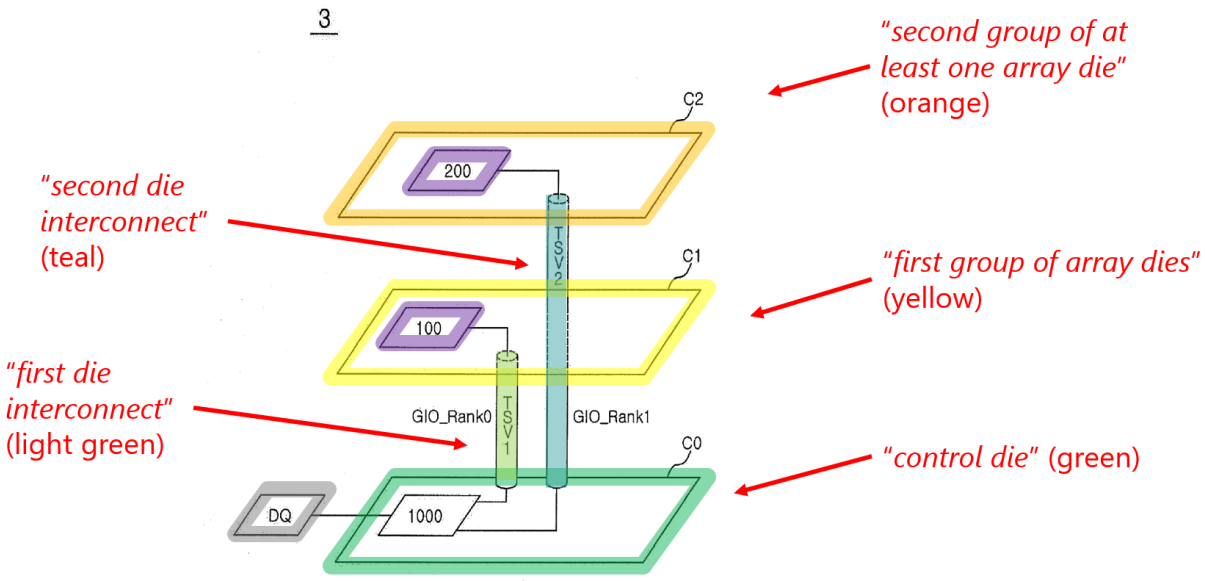


Figure 3 — Qual-stacked / Quad-die DDR3 SDRAM x16 rank association

Third, a POSITA would have been motivated to look at Rajan for the details about adding more memory chips in the stack (resulting, e.g., in four chips, yellow and orange, in two groups, as shown below), as suggested by Kim's disclosure that "any number of … chips may be used." EX1003, ¶¶158-160; EX1014, ¶¶[0048, 50].

Kim:

FIG.5



Rajan:

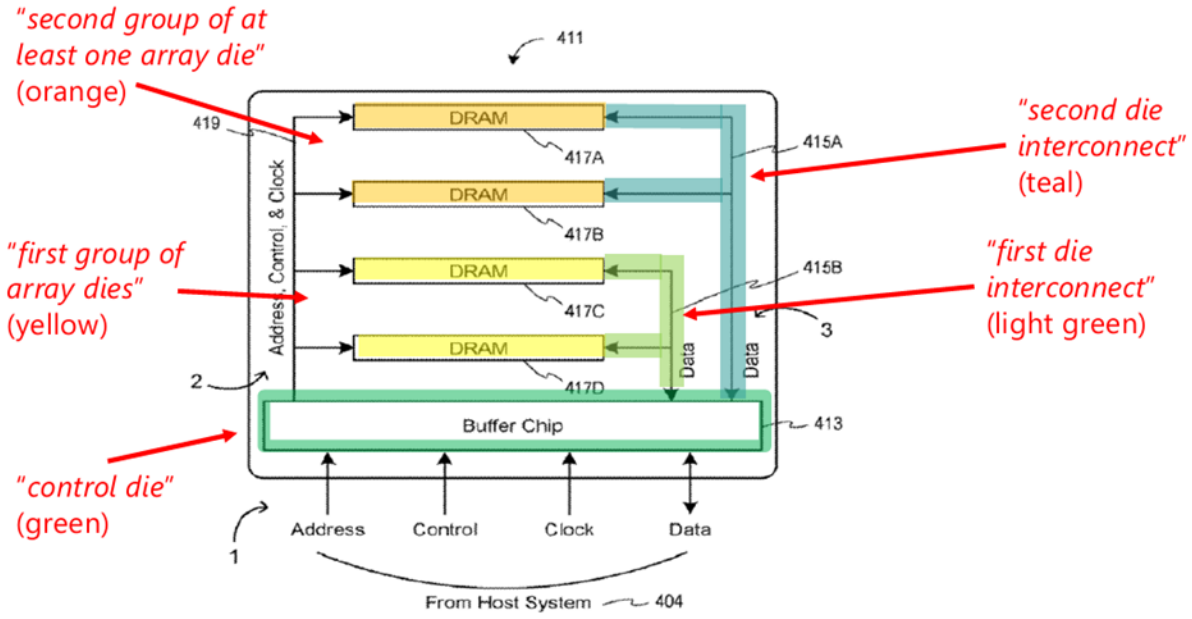


FIG. 4

A POSITA would have been motivated to implement a shared data bus for multiple memory chips as taught by Rajan (e.g., 415A and 415B, above) using, e.g., Kim's TSV interconnects (light green and teal, above), in part because a POSITA would have understood that there were a finite number of known ways to connect additional dies, including having a subset of dies sharing a through-silicon via (TSV). EX1003, ¶¶161-163; EX1015, 5:36-43, Figs. 2-6. Sharing a TSV among two or more dies was a known option, as confirmed by other references at the time (below). *See, e.g.*, EX1025, Fig.5 (below).

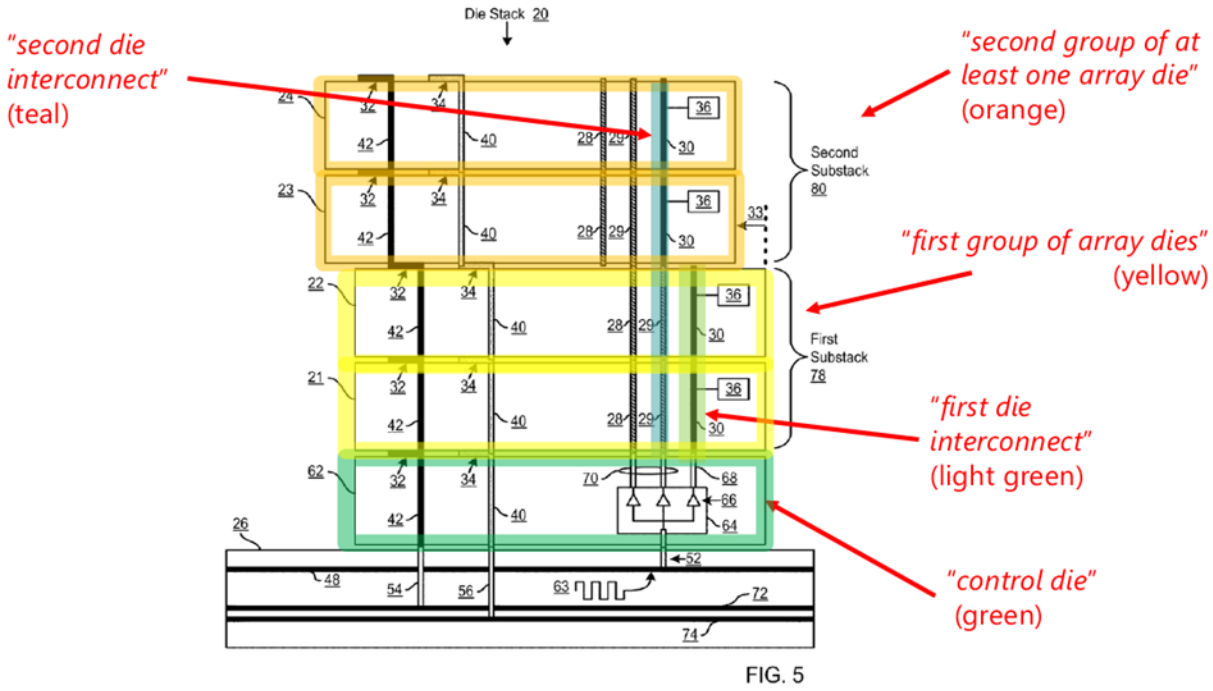


FIG. 5

Moreover, a POSITA would have been motivated to connect additional memory dies to Kim's TSV1 and TSV2 because it would not require creating new TSVs (which would add space and circuitry) and would allow emulating the JEDEC

standard required by external devices. EX1003, ¶164; EX1015, 3:27-30, 3:52-61. Furthermore, it was common at the time to have multiple memory dies sharing a single data bus. EX1003, ¶¶165-166. For example, when implementing "rank multiplication" (discussed above, pp.8-11), it was common to use a "fork-in-the-road" arrangement with two data buses for two ranks of memory devices (shown below left, yellow and orange), with the option of adding two additional ranks of memory devices (resulting in four ranks) to the existing data busses (shown below right, two yellow, two orange), meaning two memory dies (e.g., two orange) would share a given data bus. EX1003, ¶165; EX1026, Figs.12-13 (below); EX1027, pp. 77-82 (Final Written Decision analyzing EX1026); *see also supra* pp.8-11; EX1001, 18:65-19:7, 19:41-47, 22:34-57.

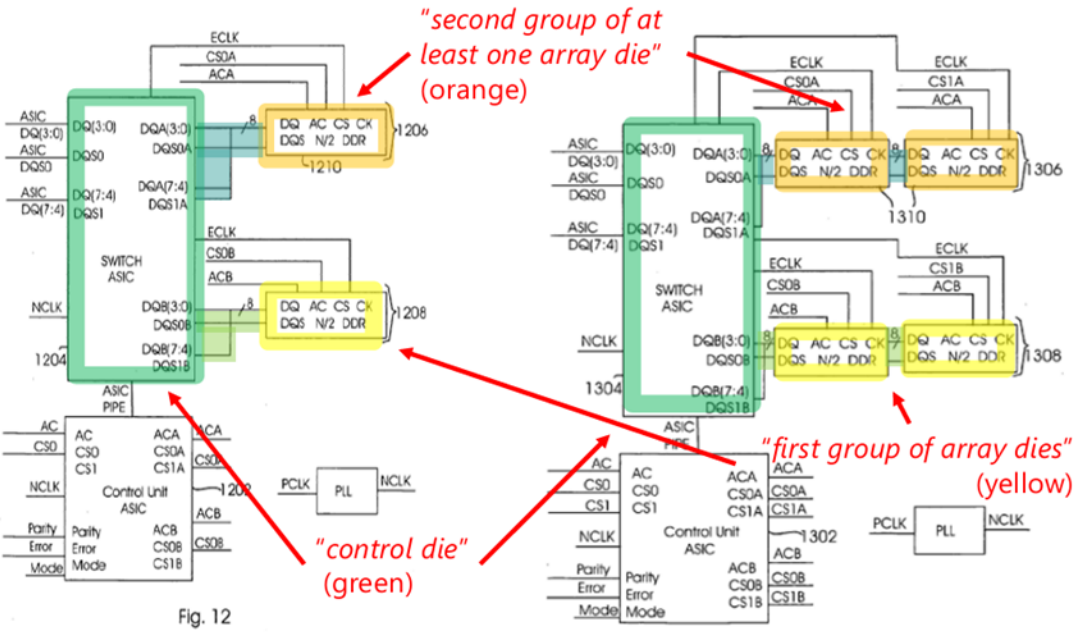


Fig. 12

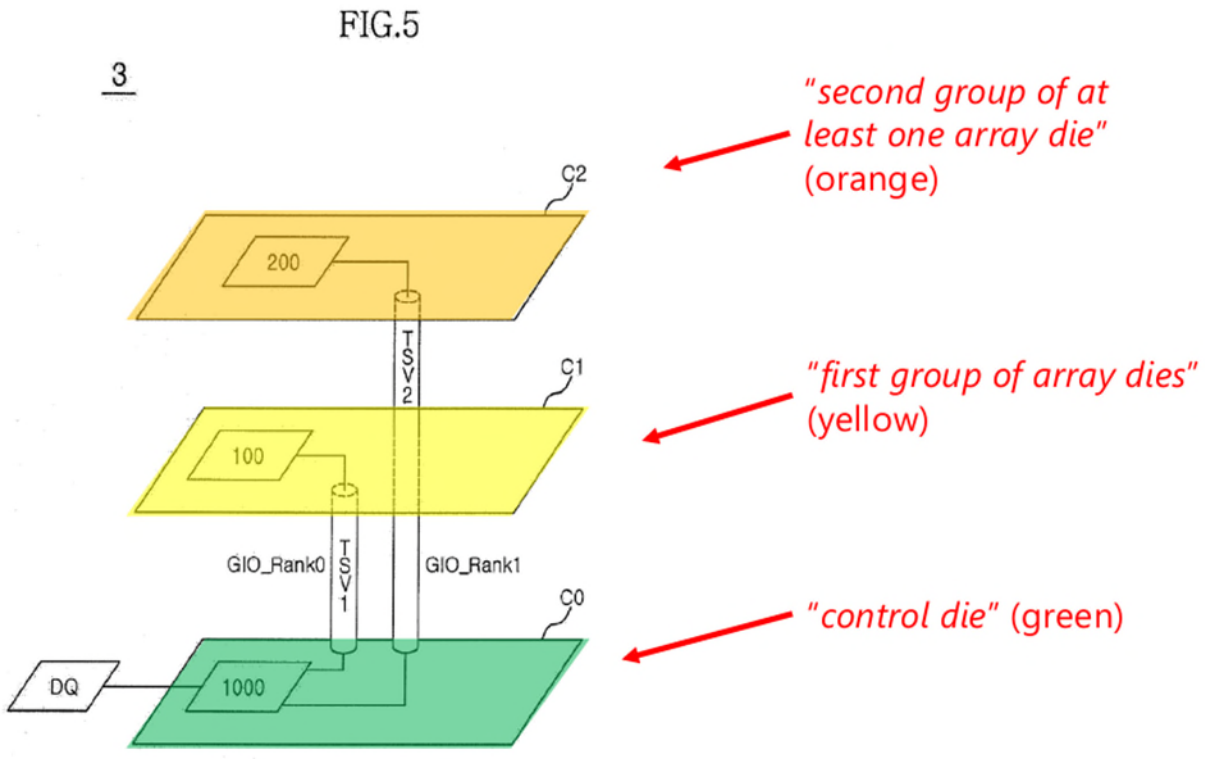
Fig. 13

Figure 4 of Rajan teaches a similar arrangement as shown above (p.27), and a POSITA would have been motivated to implement that arrangement in Kim to preserve Kim's data collision avoidance feature without creating additional TSVs. EX1003, ¶167; EX1014, ¶¶[0042-44], Figs.4A-4B.

2. Independent Claim 1

a) ***[1.a] Preamble***

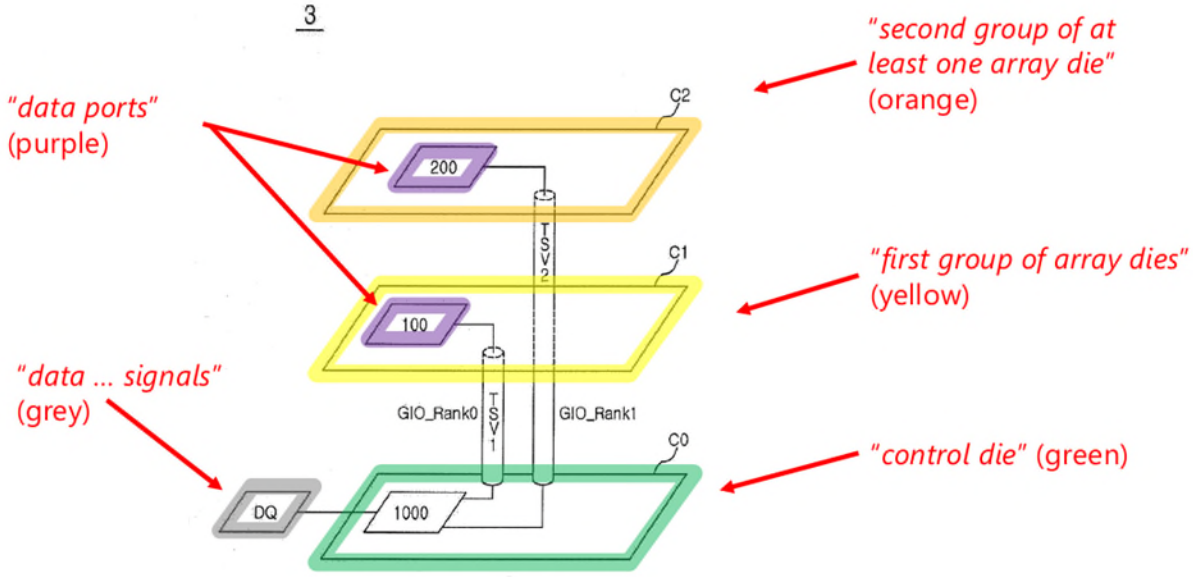
To the extent the preamble is limiting, Ground 1 teaches "*[a] memory package* [e.g., Kim's main chip C0 and stacked memory chips C1 and C2 in Figure 5 (below) which are "package[d] in a single package," EX1014, ¶¶[0046-47]], *comprising*." EX1003, ¶¶184-188.

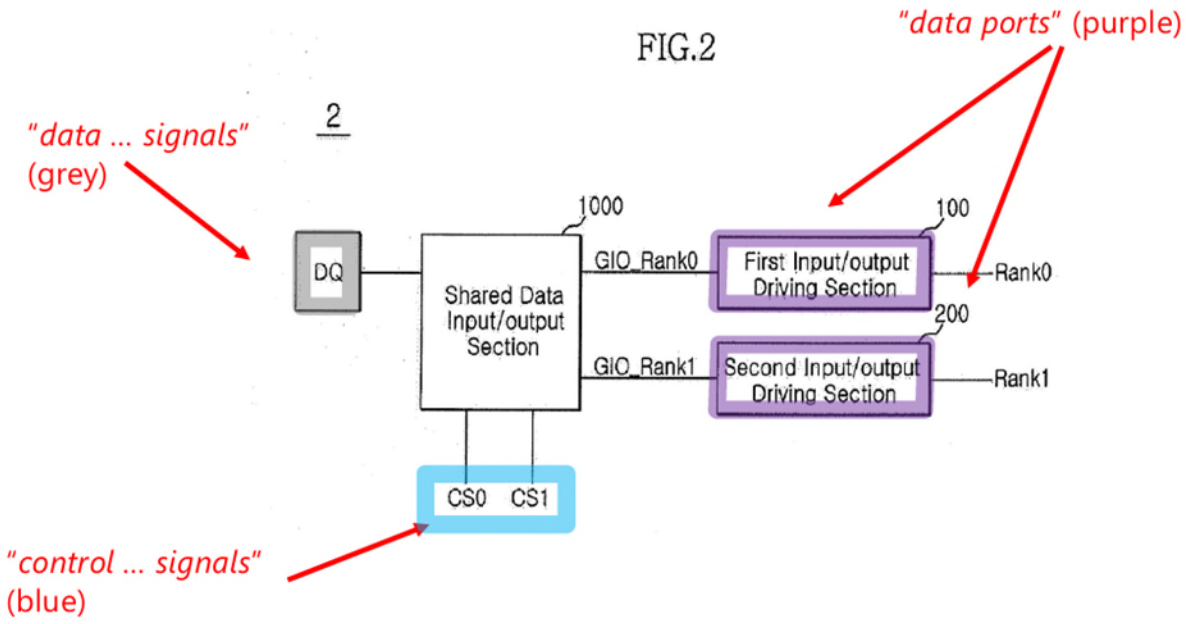


b) ***[1.b] Input/Output Terminals***

Ground 1 teaches "*a plurality of input/output terminals via which the memory package communicates data* [e.g., through Kim's data pad DQ (grey) in Figure 5 (below)] *and control/address signals* [e.g., read/write command signals including chip-select control signals CS0 and CS1 in Figure 2 (below) and address signals A0-A15 per the JEDEC standard, e.g., EX1019, pp.6-13, 18, 33] *with one or more external devices*." EX1003, ¶¶189-210.

FIG.5





Kim describes the input/output section 1000 communicating "*data…signals*" through a data pad "DQ" (grey above) during read and write operations. EX1014, Abstract, ¶[0028]; EX1003, ¶¶193-195.

In Ground 1 (pp.21-30), it would be obvious for Kim's signals to comply with JEDEC. Thus, Kim's chip selection signals CS0 and CS1 would be "*control…signals*" that are part of the command signals that control the memory devices. EX1014, ¶¶[0006, 0032], Fig.2 (above); EX1019, p.13 (chip select signals are "considered part of the command code"). Under the JEDEC standards for stacked memory devices, and as admitted by the 060 Patent, chip-select signals are received from an "*external device[]*" at terminals on the memory package. *See, e.g.*, EX1019, pp.6-14 (describing pinout for stacked memories including terminals for data (DQ) and chip-select (CS) signals); EX1001, 1:30-56 (admitting prior art

included memory packages with chip-select ports "configured to receive corresponding chip select signals to enable or select the array dies for data transfer."); EX1023, p.9, Fig.16; EX1022, pp.318-20, 332-35; EX1003, ¶196.

Under the JEDEC standards, Kim's "write and read command signals," EX1014, ¶[0038], would also include "*address signals*" received via terminals from an external controller ("*external device[]*"). *See, e.g.*, EX1019, pp.6-14 (package pinout for stacked DDR3 memories including terminals for address signals A0-A15), p.18 (explaining that read/write commands use row and column address signals), p.33 (Command Truth Table with Activate command with row address, and Read/Write command with column address); EX1001, 1:30-56 (admitting prior art included "command and/or address signals"); EX1003, ¶¶197-198.

Furthermore, a POSITA would have been motivated to implement Kim's control/address signals in light of Rajan (*see* pp.21-30), including Rajan's interfaces for sending/receiving address, control, and data signals to/from the host system in accordance with the JEDEC standards discussed above. EX1015, 2:6-7, 4:20-24, 5:36-43, 14:11-18, Figs.4 (below), 18 (interface circuit); EX1003, ¶¶201-202.

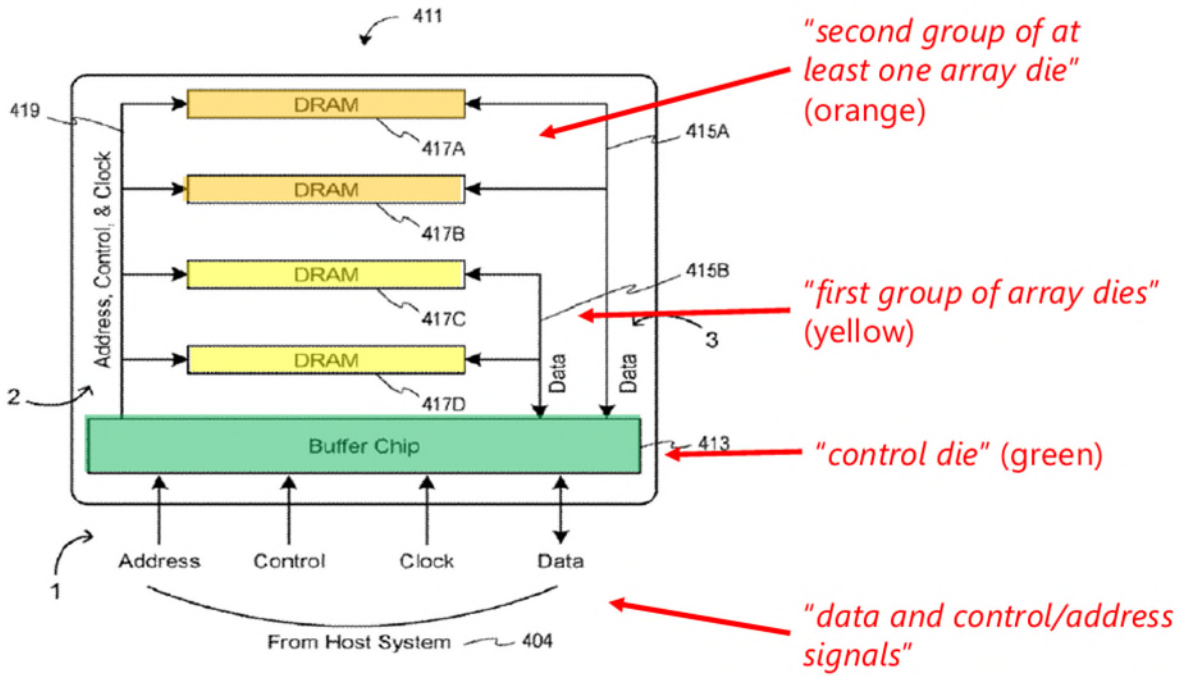
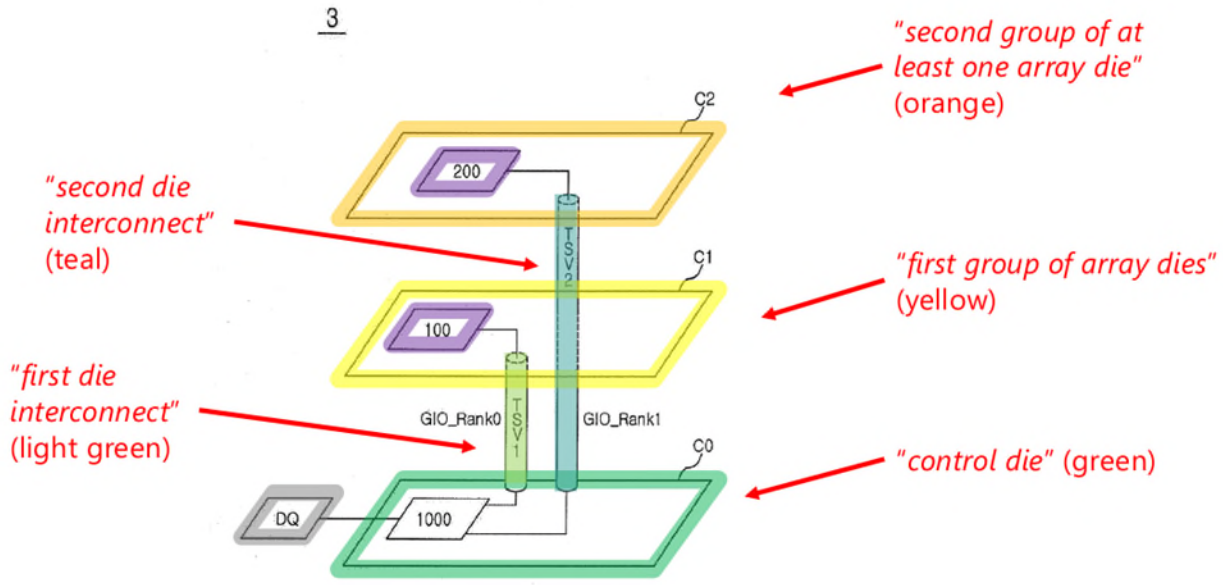


FIG. 4

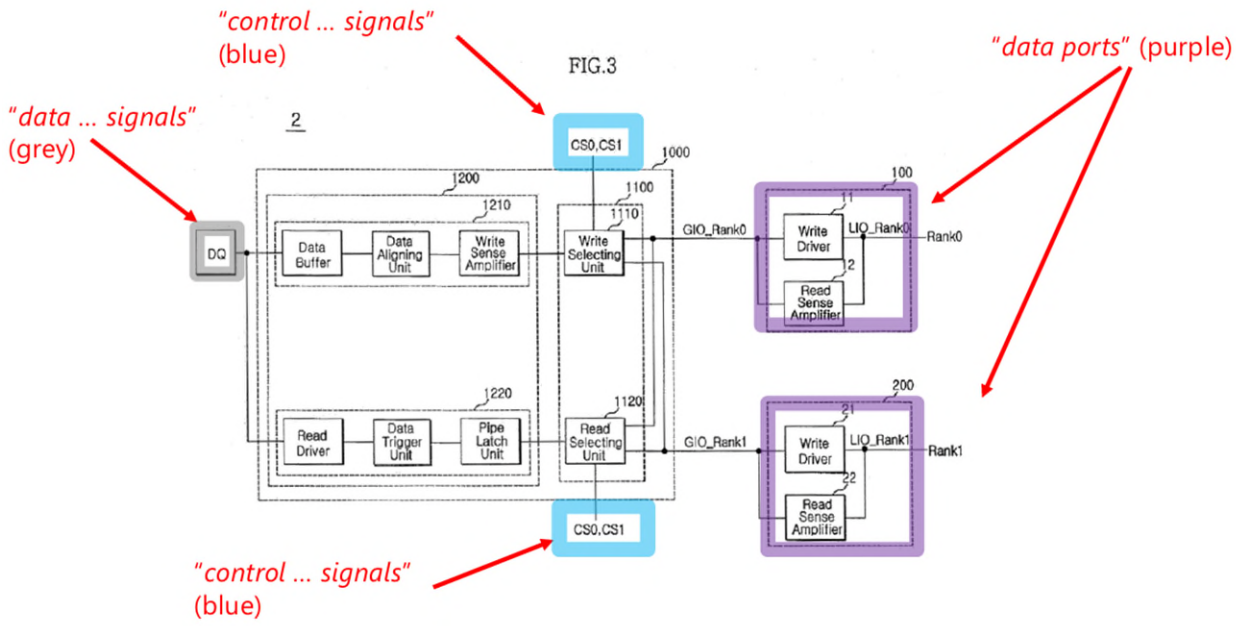
c) ***[1.c] Stacked Array Dies***

Ground 1 teaches "*a plurality of stacked array dies including a first group of array dies* [e.g., including Kim's slave chip C1 (yellow), coupled to TSV1] *and a second group of at least one array die* [e.g., including Kim's slave chip C2 (orange), coupled to TSV2], *each array die having data ports* [e.g., input/output driving sections 100 and 200 (purple)]." EX1014, ¶[0048] ("any number of … slave chips may be used"), Fig.5 (below); EX1003, ¶¶211-229.

FIG.5



A POSITA would have understood from Kim's explanation how the first and second input/output driving sections 100 and 200 store and drive data that they are "*data ports*" through which data is transmitted for storage and through which the stored data is retrieved from the slave chip. EX1014, ¶[0026], Figs.5, 3 (below); EX1003, ¶¶215-216.



A POSITA would further have understood that Kim's "memory bank" for data storage in slave chips C1 and C2 would include a memory array (consistent with textbooks, below, and the JEDEC standards), rendering obvious the claimed "*array die[s]*." *Id.*; EX1019, p.13; EX1023, pp.1-2, Fig.2 (below); EX1022, pp.316-320 (same); EX1003, ¶¶217-218.

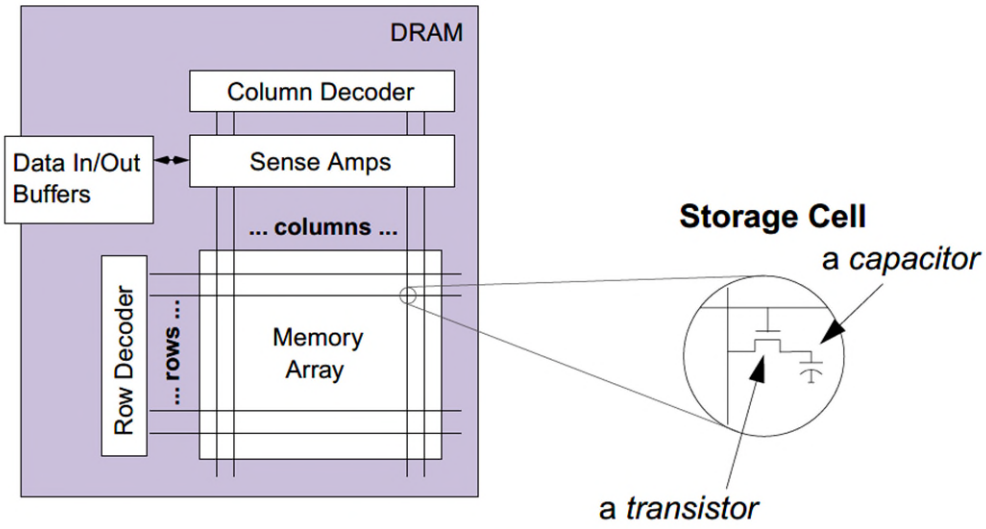


Figure 2: Basic Organization of DRAM Internals
The DRAM memory array is a grid of storage cells, where one bit of data is stored at each intersection of a *row* and a *column*.

Insofar as one might argue that "*first group of array dies*" requires multiple slave chips coupled to TSV1, Kim discloses that "any number of … slave chips may be used," EX1014, ¶¶[0048, 0050], consistent with other contemporaneous references (e.g., below) showing multiple groups of stacked dies, where each die in a group is connected to the same data interconnect, EX1024, 9:14-18, Fig.5 (below); EX1003, ¶¶219-220.

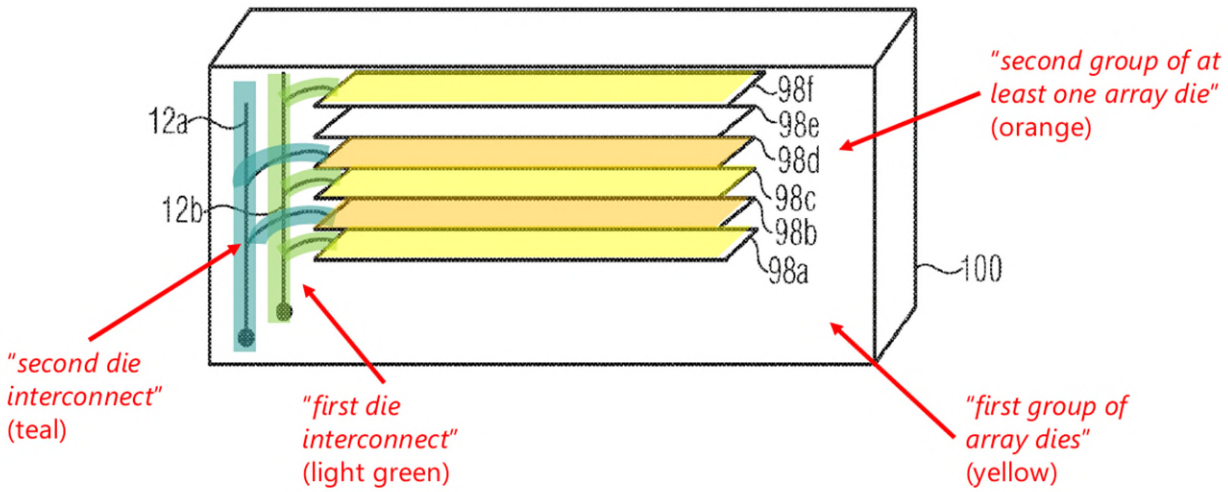


FIG 5

Additionally, Rajan discloses multiple stacked memory chips connected to the same data bus (below), thus further rendering obvious in the combination for Ground 1 that both the "*first*" and "*second*" "*group of array die[s]*" each include multiple "*array dies*" as explained in more detail above (pp.26-30). EX1003, ¶¶221-227.

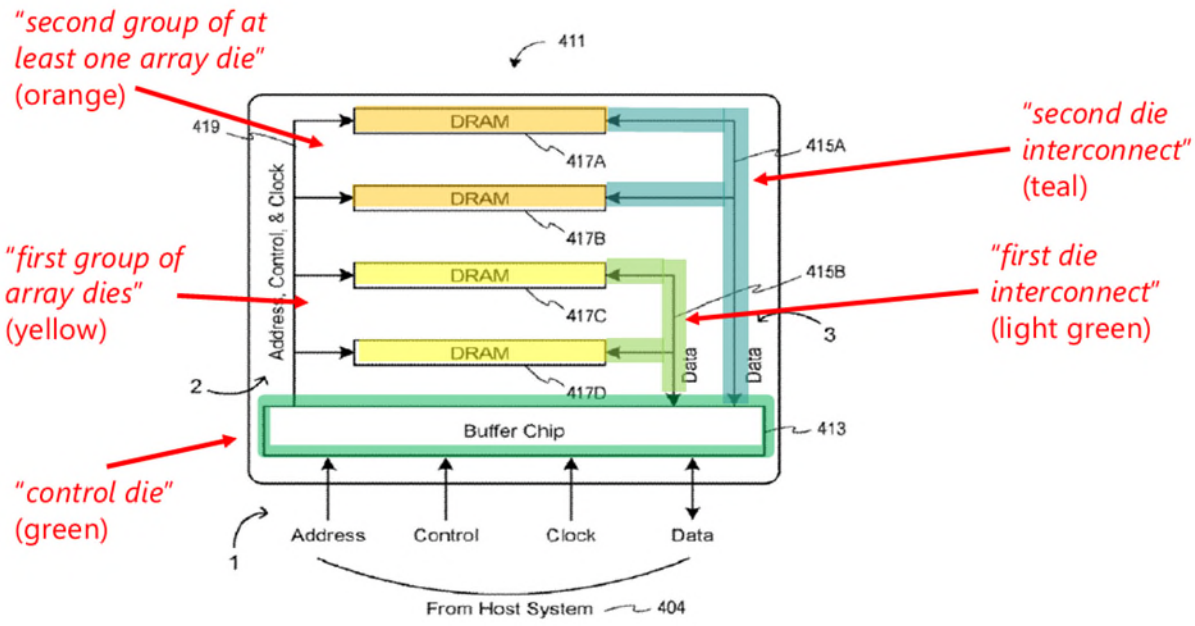
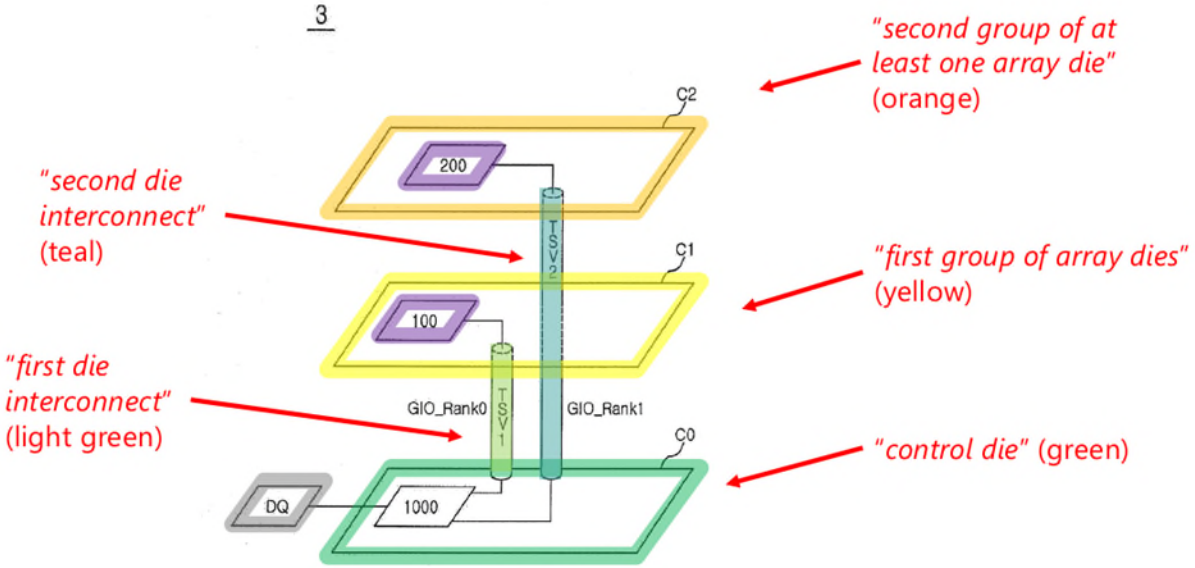


FIG. 4

d) ***[1.d.1]-[1.d.2] Die Interconnects in Electrical Communication with Array Dies***

Ground 1 teaches "*at least a first die interconnect* [e.g., Kim's through-silicon via TSV1] *and a second die interconnect* [e.g., Kim's TSV2], *the first die interconnect in electrical communication with the first group of array dies* [e.g., including Kim's slave chip C1] *and not in electrical communication with the second group of at least one array die* [e.g., including Kim's slave chip C2]," and "*the second die interconnect* [TSV2 is] *in electrical communication with the second group of at least one array die* [including C2] *and not in electrical communication with the first group of array dies* [including C1]; *and*." EX1014, ¶[0049], Fig.5 (below); EX1003, ¶¶230-239.

FIG.5

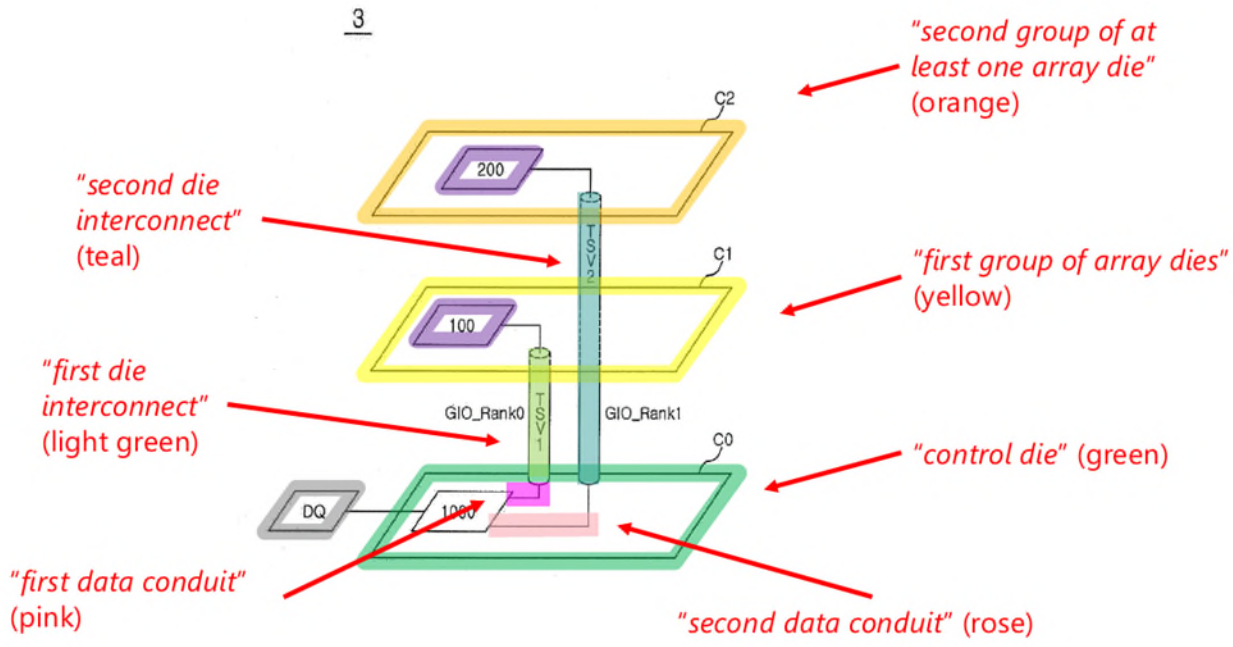


As shown above, TSV1 does not extend up to C2 and thus is not "*in electrical communication*" with C2, while TSV2 passes through C1 without being "*in electrical communication*" with C1 (e.g., there is no data port (purple) or anything else connected to TSV2 in C1). *Id.* Furthermore, a POSITA would have understood from Kim's disclosure that data collisions do not occur because, as illustrated in Figure 5, the data input/output lines through TSV1 and TSV2 are in electrical communication with one and not the other of chip C1 implementing Rank0, and C2 implementing Rank1. EX1014, Abstract, ¶[0045]; EX1003, ¶¶233-237.

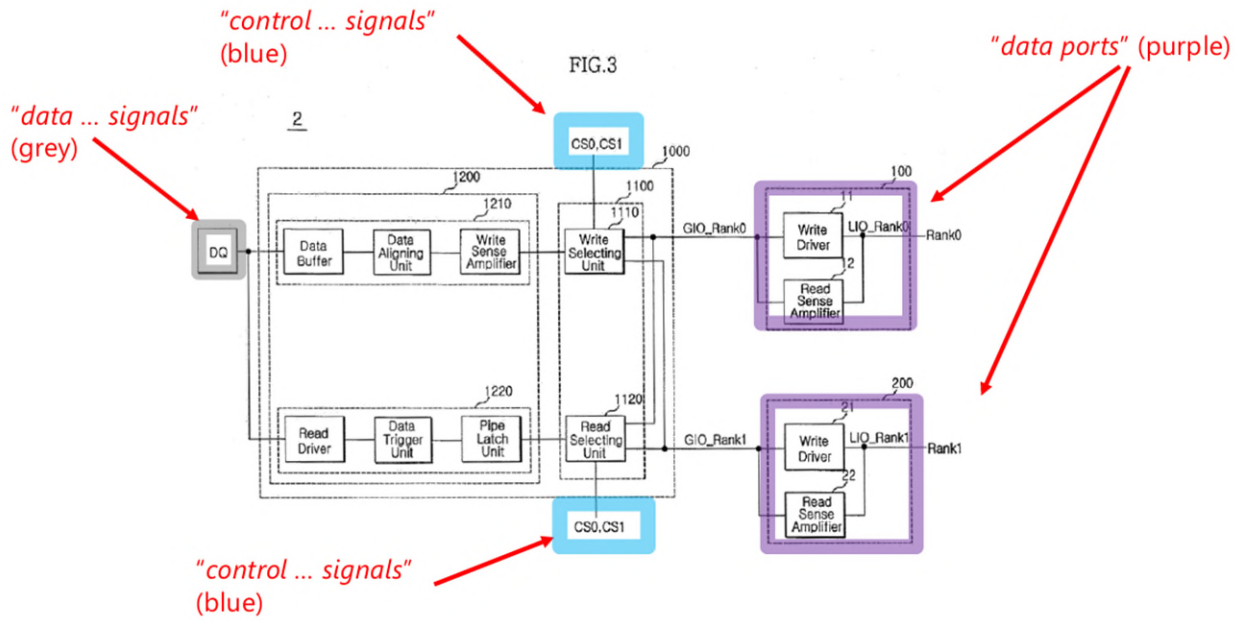
e) ***[1.e.1] Control Die***

Ground 1 teaches "*a control die* [e.g., including Kim's main chip C0 (green)] *comprising.*" EX1014, ¶¶[0038, 45], Figs. 3, 5 (below); EX1003, ¶¶240-247.

FIG.5



Kim teaches that the main chip C0 ("*control die*") controls the operation of the memory apparatus, including read/write operations in response to their respective commands from an external device, and that rank selecting unit 1100 (below) in the "*control die*" distinguishes these operations for the first rank (in the "*first group of array dies*") and the second rank (in the "*second group of … array die[s]*"). EX1014, ¶¶[0038, 45], Fig.3 (below); EX1003, ¶¶243-244.

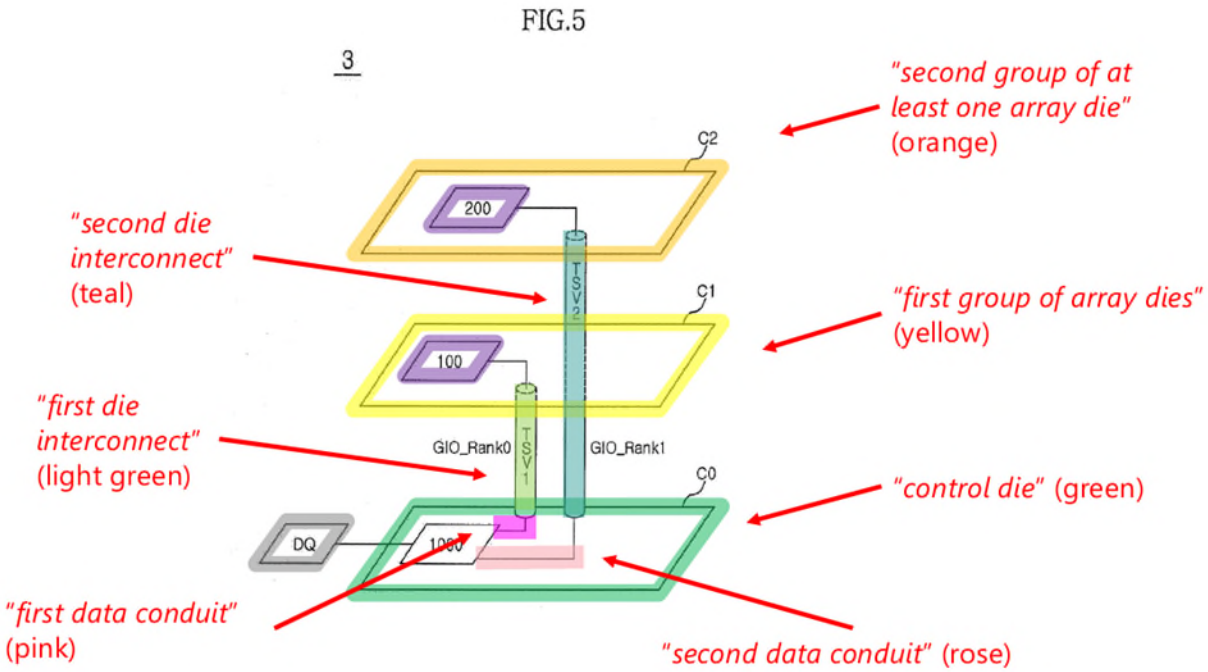


Moreover, the Kim-Rajan combination for Ground 1 (*see* pp.21-30) teaches that the "*control die*" can emulate one or more characteristics (e.g., using rank multiplication, *see* pp.8-11) different from those of the physical memory devices. EX1003, ¶245; EX1015, 6:30-7:67, Fig.18.

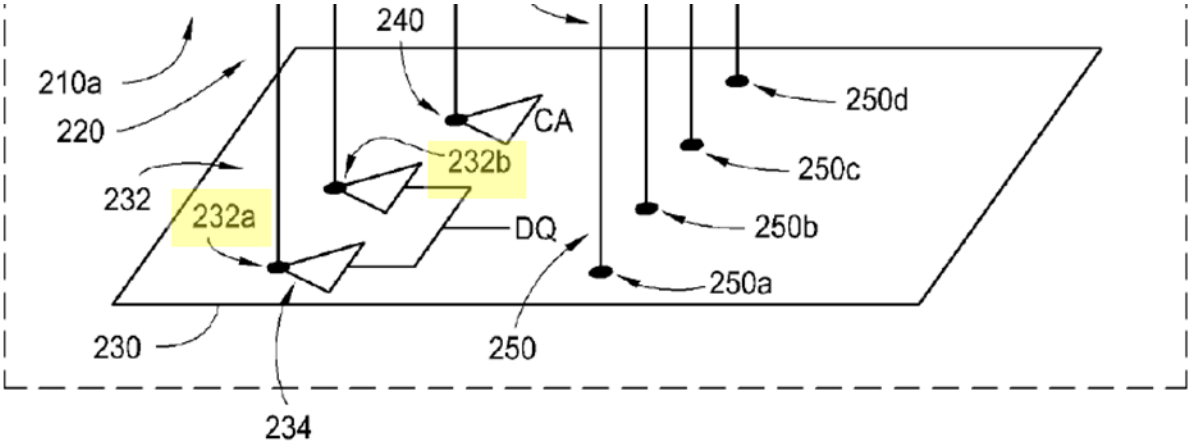
f) ***[1.e.2]-[1.e.3] First and Second Data Conduit***

Kim's Figure 5 (below) shows that Ground 1 teaches "*at least a first data conduit* [including a signal line (pink) for communicating data through data input/output section 1000] *between the first die interconnect* [TSV1 (light green)] *and a first terminal* [e.g., data pad DQ (grey)] *of the plurality of input/output terminals, and at least a second data conduit* [including a signal line (rose) for communicating data through data input/output section 1000] *between the second die interconnect* [TSV2 (teal)] *and the first terminal, the first terminal being a data*

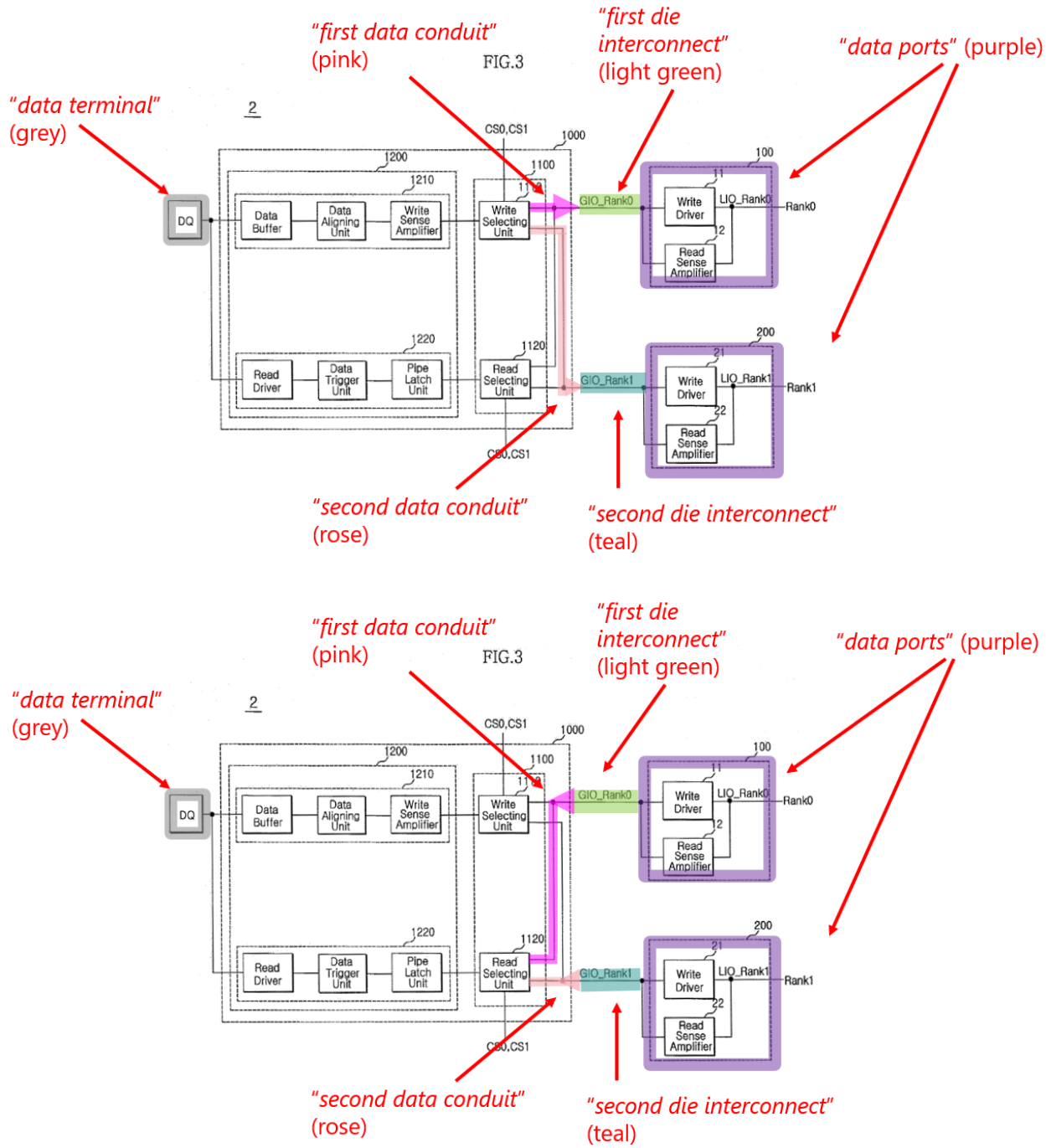
terminal [e.g., data pad DQ (grey)]." EX1014, ¶[0049], Fig.5 (below); EX1003, ¶¶248-264.



The "*data conduit[s]*" in Kim (above) are similar to the "*data conduit[s]*" 232a and 232b in the 060 Patent (below) that transmit a data signal to a respective die interconnect. EX1001, 6:48-62, 9:28-60, Fig.2 (below in part).



A POSITA would have understood from Kim's description of rank selection by chip selection signals (CS0, CS1) for read/write operations that the "*data conduit*" connecting the write selecting unit 1110 (below) and the read selecting unit 1120 (below) in data input/output section 1000 (above and below) to a TSV to Rank0 is between "*the first die interconnect*," e.g., TSV1, and a "*data terminal*," e.g., data pad DQ (grey, above and below), and the "*data conduit*" connecting the data input/output section 1000 (above and below) to a TSV to Rank1 is between "*the second die interconnect*," e.g., TSV2, and the "*data terminal*," e.g., data pad DQ (grey, above and below). EX1014, ¶¶[0029-30, 0035-38], Fig.3 (first below, showing a first "*data conduit*" (pink) active for write operation to Rank0, or a second "*data conduit*" (rose) active for a write operation to Rank1), and Fig.3 (second below, showing a first "*data conduit*" (pink) active for a read operation from Rank0, or a second "*data conduit*" (rose) active for a read operation from Rank1); EX1003, ¶¶252-261.

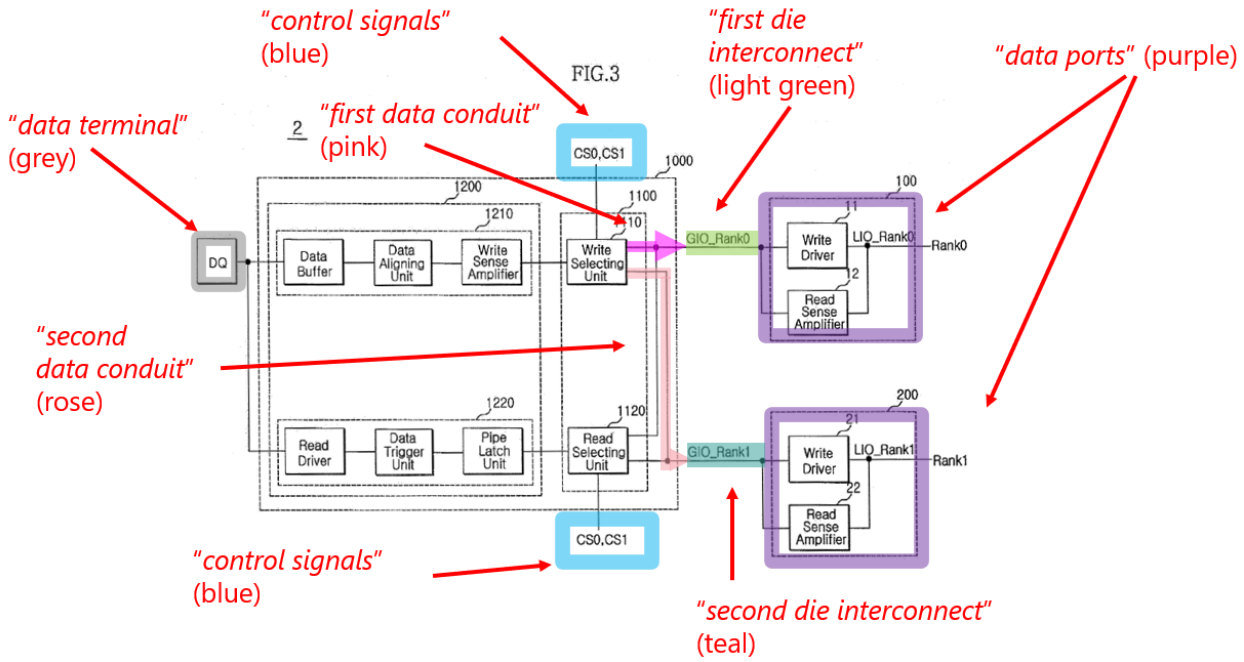


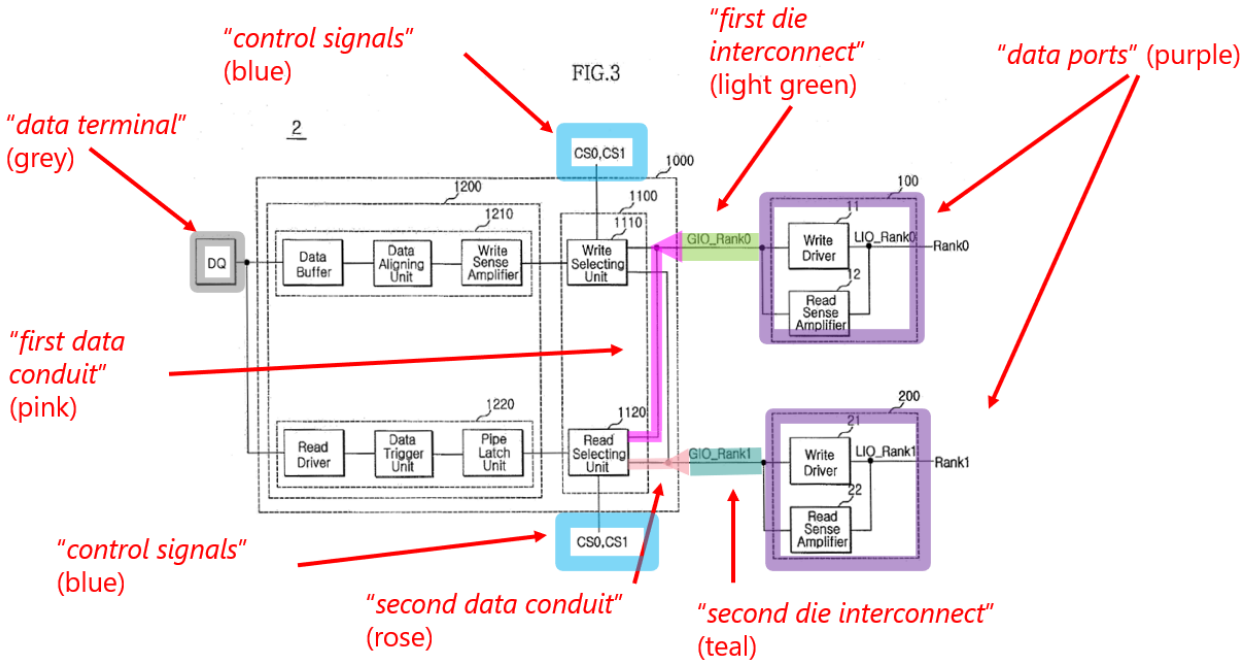
"first data conduit" (pink)
FIG.3
"first die interconnect" (light green)
"data ports" (purple)
"data terminal" (grey)
"second data conduit" (rose)
"second die interconnect" (teal)

As explained above for claim limitation [1.b] (pp.31-34), Kim discloses that data pad DQ (grey) is a "*data terminal*." EX1003, ¶262.

g) ***[1.e.4] Control Circuit***

Kim's Figure 3 (below, annotated for write (upper) and read (lower) operations) shows that Ground 1 teaches "*the control die* [from [1.e.1], pp.41-42] *further comprising a control circuit* [e.g., in rank selecting unit 1100, including write selecting unit 1110 and read selecting unit 1120] *to control respective states* [e.g., to be active or not active, and if active to transmit data in the read direction or in the write direction] *of the first data conduit* [(pink) to TSV1] *and the second data conduit* [(rose) to TSV2] *in response to control signals* [including chip selection signals CS0 and CS1 and read/write command signals pursuant to the JEDEC standard, EX1019, pp.6-14, 18, 33; EX1023, p.9, Fig.16; EX1022, pp.318-20, 332-35,] *received via one or more second terminals of the plurality of terminals* [from an external device, as discussed for [1.b], pp.31-34]." EX1014, ¶¶[0032, 0035-36, 0038], Fig.3 (below); EX1003, ¶¶265-272.





3. Claim 2

Grounds 1 teaches "*claim 1, wherein the control signals* [from the external device] *include data path control signals* [e.g., chip-select and read/write command signals pursuant to the JEDEC standard] *for controlling the first* [(pink)]

A POSITA would have understood that Kim's chip selection signals and read/write command signals, received by Kim's rank selecting unit 1100 (*see* claim limitation [1.e.4], pp.46-47), are "*data path control signals*" since they determine direction of data transfer (read or write) and select the TSV (TSV1 or TSV2) through which data is transmitted from or to Rank0 or Rank1, respectively. EX1014, ¶¶[0031-38], Fig.3; EX1003, ¶277 (citing ¶¶265-272). Furthermore, Kim teaches that such "*data path control signals*" are needed to avoid collisions. EX1014, ¶[0031]; EX1003, ¶277. Moreover, in the combination of Ground 1, either the "*control circuit*" in the memory package and/or the register on the memory module external to the memory package can implement "rank multiplication," as discussed above (pp.8-11, 28-30), thus affecting the "*data path control signals*." EX1003, ¶278.

4. Claim 3

Ground 1 teaches "*claim 1, wherein the control circuit* [from [1.e.4], pp.46-47, e.g., in Kim's rank selecting unit 1100 (in Fig.3, below)] *is configured to generate data path control signals* [as discussed above for claim 2 (pp.47-49), including "internal write and read signals…[that are] generated," EX1014, ¶[0038]] *for controlling the first and second data conduits* [to or from TSV1 or TSV2] *in response to the received control signals* [including chip selection and JEDEC-standard read/write command signals, from an external device, which may

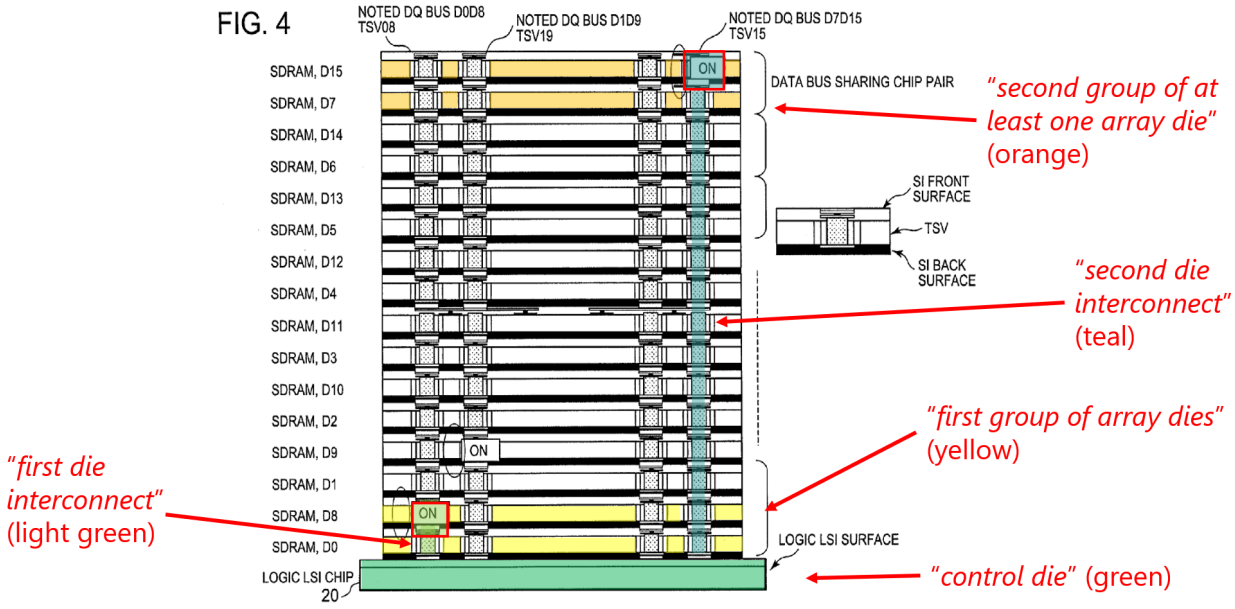
states of the first data conduit and the second data conduit," as explained for claim element [1.e.4] (pp.46-47) and claim 3 (pp.49-51). EX1003, ¶¶677-681.

B. Ground 2 (Ground 1 + claim 7)

1. Ground 2 combination: Ground 1 and Riho (EX1016)

Ground 2 is based on Ground 1 (Kim in view of Rajan) in further view of Riho (EX1016). EX1003, ¶¶169-173. Ground 2 thus renders obvious the same claims as Ground 1, plus claim 7 as discussed below.

As explained above (pp.21-30), a POSITA would have been motivated to look at analogous art, and Riho is analogous since it is also directed to the operation of stacked memory devices and discloses the same basic structure (below), including stacked memory chips (yellow, orange) and a shared interface circuit (green), called the "logic LSI chip" in Riho. EX1016, ¶¶[0002, 0062], Fig.4 (below); EX1003, ¶171.



Riho teaches using ***pairs*** of SDRAM memory chips, where each pair shares a TSV for data signals (DQ), such as pair D0/D8 (yellow) coupled to DQ bus D0D8 including TSV08, and pair D7/D15 (orange) coupled to DQ bus D7D15 including TSV15, with only ***one*** of the two SDRAMs in the pair connected to the TSV at a time, thus "reduc[ing] by half the load." EX1016, ¶¶[0062, 0103]; EX1003, ¶172.

Kim teaches that "any number" of chips may be stacked, EX1014, ¶¶[0048, 0050], and thus a POSITA would have been motivated with a reasonable expectation of success (when adding chips to Kim) to use Riho's technique described above to further reduce the load (allowing for higher-frequency operations) and to compensate for phase variations, as taught by Riho, without the need to create additional TSVs. EX1003, ¶¶172-173; EX1016, ¶¶[0119-20].

2. Claim 7

a) ***[7.a] First and Second Number of Array Dies***

Ground 2 teaches "*[t]he memory package of claim 1, wherein a first number of* [two] *array dies in the first group of array dies* [sharing a data signal TSV] *and a second number of at least one array die* [two dies] *in the second group of at least one array die* [sharing another data signal TSV] *are selected in consideration of a load of the first die interconnect and a load of the second die interconnect* [including the load of each of those data signal TSVs] *so as to reduce a difference between a first load on the first data conduit and a second load on the second data conduit* [so the "skew between data signals DQ and data strobe signals DQS/B can be minimized" allowing a "single synchronization signal with high accuracy," EX1016, ¶¶[0050, 0053-54]]." EX1003, ¶¶336-343; *see also supra* pp.6-8 (more load creates more delay); EX1019, p.13 ("Data Strobe" for reads and writes).

b) ***[7.b] First and Second Load***

Ground 2 teaches "*the first load including a load of the first die interconnect* [on one data signal TSV], *and a load of the first group of array dies* [including a pair of SDRAM chips], *and the second load including a load of the second die interconnect* [on another data signal TSV] *and a load of the second group of at least one array die* [including another pair of SDRAM chips]." EX1016, ¶¶[0013, 0132] ("load of the interconnections" and "load…from the…SDRAM"); EX1003, ¶¶344-350; *see also supra* pp.6-8 (load was well known).

C. Ground 3 (Ground 1 + claims 15, 20-28)

1. Ground 3 combination: Ground 1 and Wyman (EX1017)

Ground 3 is based on Ground 1 (Kim in view of Rajan) in further view of Wyman (EX1017). EX1003, ¶¶174-181. Ground 3 thus renders obvious the same claims as Ground 1, plus claims 15 and 20-28 as discussed below.

Wyman is analogous art to Kim and the 060 Patent: its goal is to improve efficiency of signaling between stacked chips. EX1017, 1:45-48; EX1001, 1:19-21, 5:23-26; EX1014, ¶¶[0003, 0046]; EX1003, ¶¶178, 180. Wyman discloses that shorter paths (e.g., 802 below) have less load and thus require less drive, while longer paths (e.g., 804 below) have more load and thus require a larger drive, but it would be "wasteful" and "overkill" to use the "full capacity" of a driver for either path. EX1017, 1:22-24, 1:45-48, 6:15-50, 7:13-18, Fig.8 (below); *see also supra* pp.6-8 and EX1030, pp.135-138 (disclosing different-sized transistors to achieve different driver strengths).

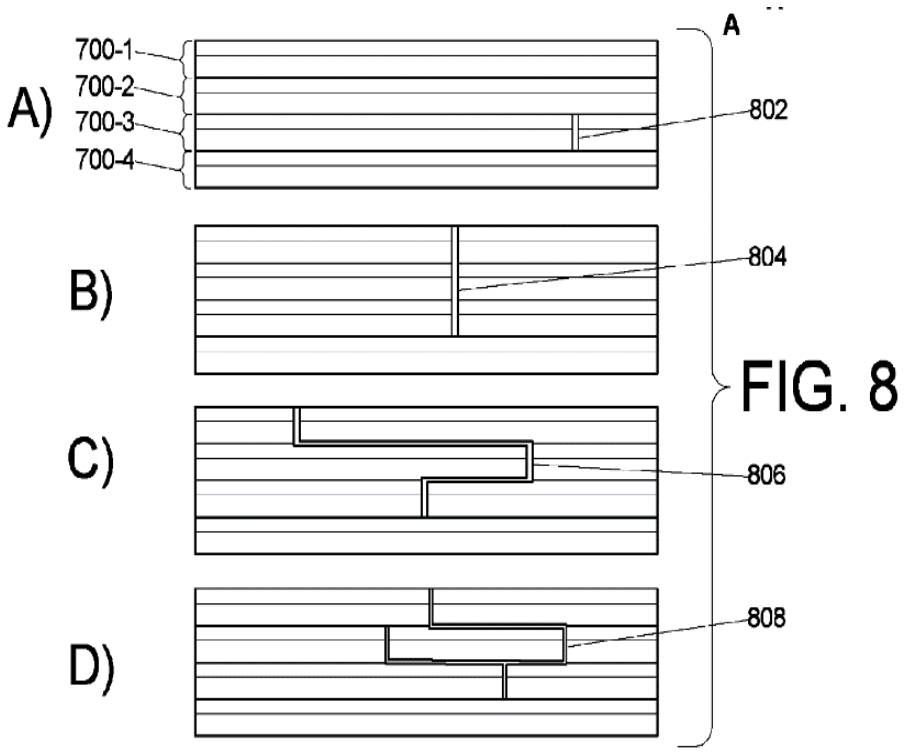
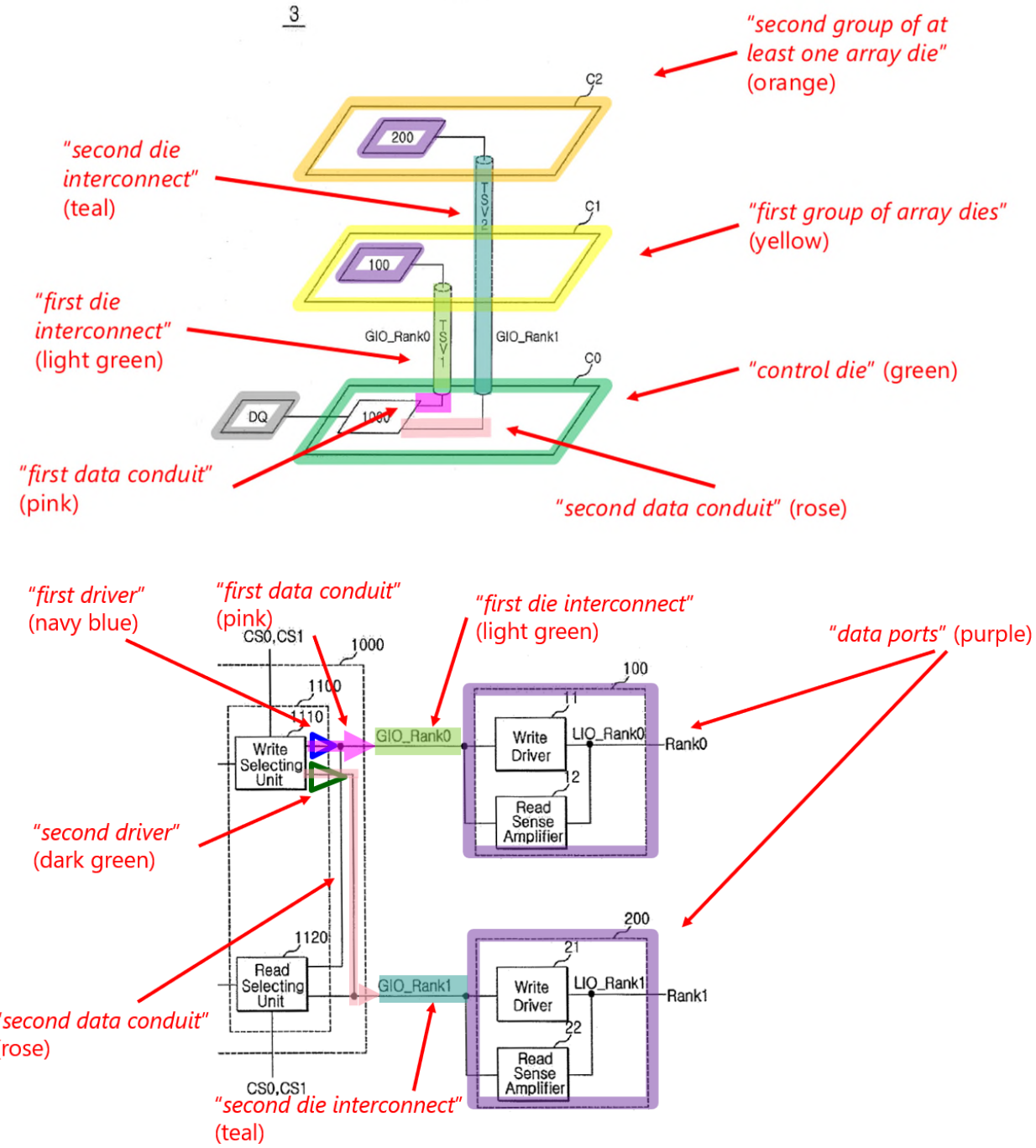


FIG. 8

A POSITA would have been motivated to implement Wyman's teachings in Kim to improve power efficiency by using a smaller driver (navy blue) for the shorter TSV1 (light green) and a larger driver (dark green) for the longer TSV2 (teal), as shown below, without using the full strength of a driver which Wyman teaches would be "wasteful and inefficient." *Id.*; EX1003, ¶¶177-180.

FIG.5

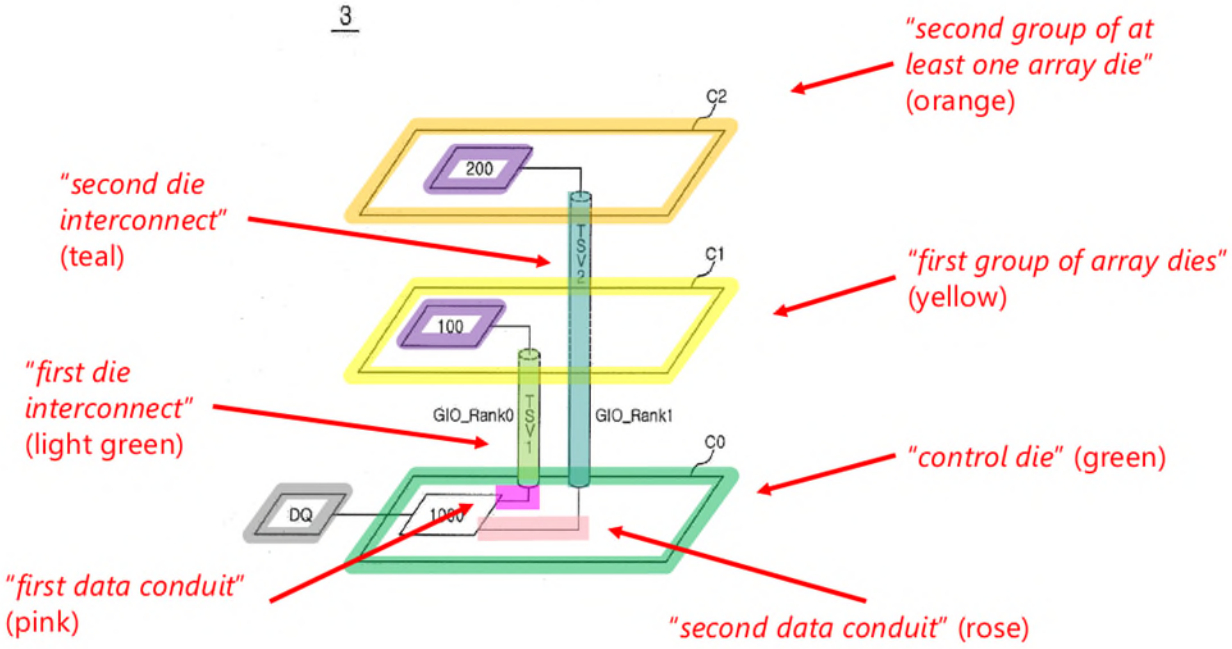

3
"second group of at least one array die" (orange)
C2
200
"second die interconnect" (teal)
TSV2
C1
"first group of array dies" (yellow)
100
"first die interconnect" (light green)
GIO_Rank0
TSV1
GIO_Rank1
C0
"control die" (green)
DQ
1000
"first data conduit" (pink)
"second data conduit" (rose)
"first driver" (navy blue)
"first data conduit" (pink)
"first die interconnect" (light green)
"data ports" (purple)
CS0,CS1
1000
1100
1110
Write Selecting Unit
GIO_Rank0
100
11
Write Driver
LIO_Rank0
Rank0
12
Read Sense Amplifier
"second driver" (dark green)
1120
Read Selecting Unit
200
21
Write Driver
LIO_Rank1
Rank1
22
Read Sense Amplifier
"second data conduit" (rose)
GIO_Rank1
CS0,CS1
"second die interconnect" (teal)

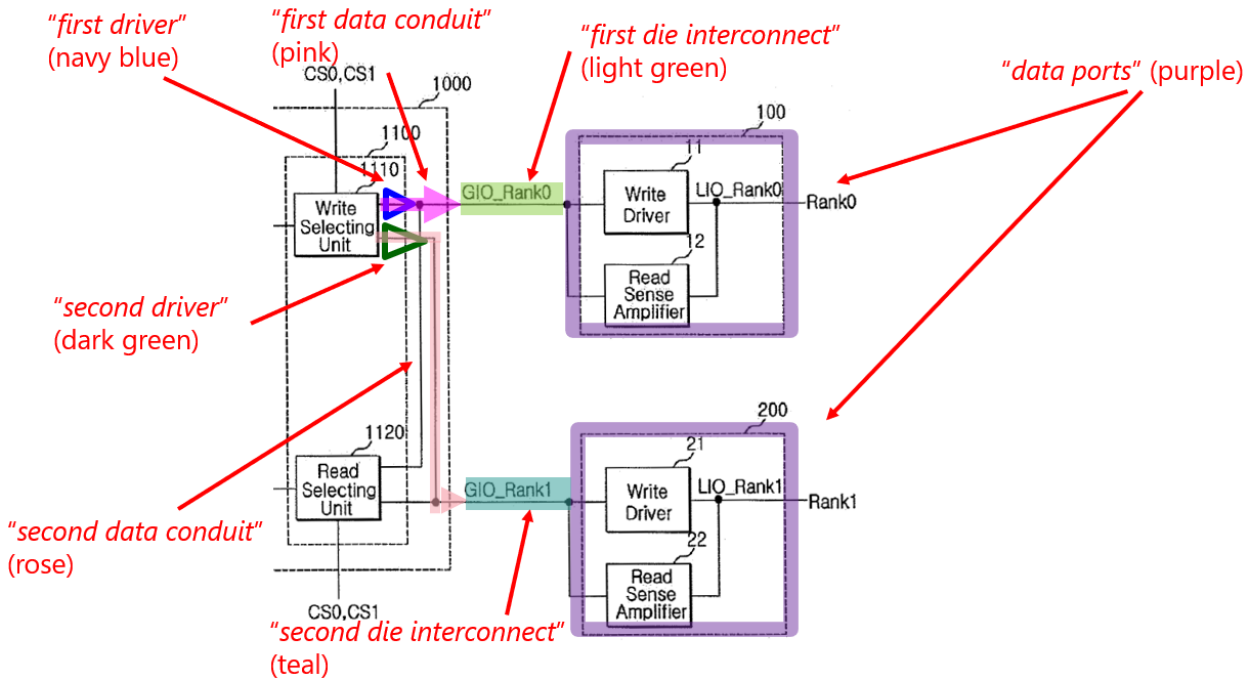
2. Claim 15

a) *[15.a] First and Second Data Conduit*

Ground 3 teaches "*[t]he memory package of claim 11, wherein: the first data conduit* [(pink) to TSV1, *see* claim [1.e.2] (pp.42-45)] *comprises at least a first driver having a first driver size* [smaller navy blue triangle, below]*, and the second data conduit* [(rose) to TSV2, *see* claim [1.e.3] (pp.42-45)] *comprises at least a second driver having a second driver size* [larger dark green triangle, below]*, and*." EX1014, Figs. 3, 5 (below); EX1003, ¶¶456-464.

FIG.5





Given that Kim expressly illustrates in Figure 5 above that TSV1 is shorter than TSV2, a POSITA would have understood from Wyman the benefit of using a different driver size for each data conduit that is optimized for efficiently driving data signals through corresponding TSVs to avoid "wasteful and inefficient" use of power, as explained more fully above (pp.84-86). EX1003, ¶¶456-464.

b) ***[15.b] Driver Size***

Ground 3 teaches "*wherein the first driver size* [smaller navy blue triangle, above] *and the second driver size* [larger dark green triangle, above] *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* [from [11.c] (p.69)] *without significant signal degradation*." EX1003, ¶¶465-471. As shown above, Kim's TSV1 and TSV2 are each coupled to circuitry in some, but not all, of the slave

chips in the stack, requiring less driver strength than if ***all*** the slave chips were connected, and Wyman teaches to choose a driver size that will "adequately drive" the signal but is still less than the "total drive," which would be "overkill" and inefficient. *Id.*; EX1017, 1:45-47, 6:57-60; *supra* pp.84-86.

3. Claims 20-28

The limitations of claims 20-28 are substantially identical to earlier limitations, as shown in the following table, and thus they are obvious in light of Ground 3 for at least the same reasons discussed above:

This limitation…	…is substantially similar to this limitation…	…and thus obvious for at least the same reasons above and as discussed in EX1003:
[20.a]	[1.a][6]	¶¶499-502 (¶¶184-188)
[20.b.1]	[1.c]	¶¶503-506 (¶¶211-229)
[20.b.2]	[1.d.1]-[1.d.2]	¶¶507-510 (¶¶230-239)
[20.b.3]	[1.e.1]	¶¶511-514 (¶¶240-247)
[20.b.4]	[1.b]	¶¶515-518 (¶¶189-210)
[20.c]	[20.a]	¶¶519-522 (¶¶499-502)
[20.d.1]	[1.b], [1.e.2]-[1.e.3]	¶¶523-526 (¶¶189-210, 248-264)
[20.d.2]	[1.b], [1.e.4]	¶¶527-530 (¶¶189-210, 265-272)

[6] The specific steps for "*optimizing [the] load*" are explained in the subsequent limitations.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.
Petitioner,

v.

NETLIST, INC.
Patent Owner.

Case IPR2022-01428
Patent No. 8,787,060

NETLIST, INC.'S

PATENT OWNER RESPONSE

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

111237898

TABLE OF CONTENTS

Page

I. Introduction 1

II. Overview of the '060 Patent 2

III. Skill Level of a POSITA 6

IV. Asserted Prior Art 7

A. Kim 7

B. Riho 10

V. Claim Construction 15

A. "array die" 16

B. "chip select signals" 18

VI. The Petition Improperly Relies Solely On Expert Testimony For Many Claim Elements 19

VII. Kim Grounds Are Not Obvious 20

A. Kim, alone or in combination, does not disclose or render obvious designs with elements 1.d.1 and 1.d.2 20

1. Modifying Kim by having chips of the same rank share the same TSV would result in an inoperable package 23

2. Modifying Kim by grouping one chip from each of two ranks and having them share one TSV would not yield an operable memory system 28

3. Modifying Kim with four separate TSVs for four different ranks would not result in the claimed inventions 30

TABLE OF CONTENTS
(Cont'd)

Page

4. Modifying Kim by having two separate ranks of chips share a TSV would re-introduce data collisions that Kim tries to avoid 32

5. A POSITA would not have combined Kim with Rajan in the manner proposed 38

B. Petitioner has not shown why a POSITA would have used non-DRAM stacks with a JEDEC-compliant DRAM interface 44

C. Petitioner has not shown that the Kim combinations would render obvious "chip-selection" signals or conduits 48

D. Claim 7 (Ground 2) 51

E. Ground 3 54

VIII. Riho Grounds Are Not Obvious 56

A. Riho, alone or in combination, does not disclose or render obvious elements 1.d.1 and 1.d.2 56

1. Petitioner's position in this proceeding is inconsistent with its infringement position while Netlist has applied a consistent interpretation consistent with the specification 56

2. Riho's TSVs are each in communication with all dies in its stack 63

3. Dr. Wolfe's new theory at the deposition is incorrect 68

4. Rajan would not cure Riho's deficiency 73

B. Riho, alone or in combination, does not disclose or render obvious elements 1.e.2-1.e.3 75

TABLE OF CONTENTS
(Cont'd)

Page

C. Petitioner has not shown why a POSITA would have used non-DRAM stacks with a JEDEC-compliant DRAM interface 79

D. Riho's chip select signal does not satisfy petitioner's proposed construction (claims 6, 11, 20, 29 and dependent claims) 80

E. Claim 7 82

F. Ground 5 83

IX. Conclusion 83

TABLE OF AUTHORITIES

Page(s)

Cases

Arendi SARL v. Apple Inc.,
832 F.3d 1355 (Fed. Cir. 2016) 48

Cisco Sys., Inc. v. C-Cation Techs, LLC,
IPR2014-00454, Paper 12 (PTAB, Aug. 29, 2014).....19

DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.,
567 F.3d 1314 (Fed. Cir. 2009) 7

Fidelity Nat'l Info. Serv.s, Inc. v. DataTreasury Corp,
IPR2014-00491, Paper 9 (P.T.A.B. Aug. 13, 2014).....19

Microsoft Corp. v. Enfish, LLC,
662 Fed. App'x 981 (Fed. Cir. 2016) 51

Other Authorities

37 C.F.R. § 42.1(b) 15

83 Fed. Reg. at 51353 15

EXHIBIT LIST

Exhibit No.	Description
2001	Samsung's P.R. 4-2(a) Preliminary Claim Construction and Extrinsic Evidence in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. August 5, 2022)
2002	Exhibit B (Samsung's proposed construction) to Joint Claim Construction Statement in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. August 19, 2022)
2003	Samsung's Responsive Claim Construction Brief in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. September 16, 2022)
2004	Claim Construction Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. December 14, 2022)
2005	Netlist's Objections to the Claim Construction Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. December 28, 2022)
2006	Samsung's Responses to Netlist's Objections to the Claim Construction Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. January 11, 2023)
2007	*Cho, J. H. et al.*, "A 1.2V 64Gb 341 GB/s HBM2 Stacked DRAM with Spiral Point-to-Point TSV Structure and Improved Bank Group Data Control," ISSCC 2018
2008	Netlist's Opposition to Samsung's Motion for Summary Judgement of Non-infringement in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 281, 2:21-cv-00463-JRG (E.D. Tex. December 28, 2022)
2009	Deposition Transcript of SungJoo Park in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. November 23, 2022) (excerpt)

Exhibit No.	Description
2010	Netlist's Unopposed Motion to Withdraw its Objections to the Claim Construction Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 192, 2:21-cv-00463-JRG (E.D. Tex. Feb. 2, 2023)
2011	Order Granting Netlist's Unopposed Motion to Withdraw its Objections to the Claim Construction Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 194, 2:21-cv-00463-JRG (E.D. Tex. Feb. 3, 2023)
2012	2023-02-22 Netlist Email to Samsung
2013	Samsung's Fourth and Amended Initial and Additional Disclosures in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. December 6, 2022) (excerpt)
2014	Robins Report TOC in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex. December 27, 2022)
2015	Reserved
2016	Reserved
2017	PC2100 and PC1600 DDR SDRAM Registered DIMM Design Specification by JEDEC, v.1.3 (January 2002)
2018	Order Amending Docket Control Order in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 236, 2:21-cv-00463-JRG (E.D. Tex. Feb. 10, 2023)
2019	Netlist's Motion for Partial Summary Judgement on Samsung's Invalidity Defenses in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 236, 2:21-cv-00463-JRG (E.D. Tex. Feb. 3, 2023)
2020	Samsung's Opposition to Netlist's Motion for Partial Summary Judgement on Samsung's Invalidity Defenses in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 277, 2:21-cv-00463-JRG (E.D. Tex. Feb. 17, 2023)
2021	2023-02-17 Netlist Email to Samsung re IPR2022-01428 Sotera Stipulation

Exhibit No.	Description
2022	2023-02-17 Netlist Email to Samsung re IPR2022-01427 Sotera Stipulation
2023	Declaration by Dr. Michael C. Brogioli
2024	2023-06-26 2023-06-27 Deposition Transcript of Dr. Harold Stone in *Netlist, Inc. v. Micron Technology, Inc. et al.*, 2:22-cv-203-JRG-RSP (E.D. Tex. Jun. 26, 2023)
2025	Combined 2023-06-27 and 2023-06-28 Deposition Transcripts of Dr. Andrew Wolfe with Errata
2026	Deposition Transcript of Dr. Andrew Wolfe in IPR2022-00615
2027	2023-04-14 Trial Transcript (excerpt) in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex.)
2028	2023-04-18 Trial Transcript (excerpt) in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex.)
2029	2023-04-19 Trial Transcript (excerpt) in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex.)
2030	Robins demonstrative excerpt in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex., April 19, 2023)
2031	Milton demonstrative excerpt in *Netlist, Inc. v. Samsung Electronics Co., Ltd.*, 2:21-cv-00463-JRG (E.D. Tex., April 14, 2023)
2032	*Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 236, 2:21-cv-00463-JRG (E.D. Tex.), Trial Exhibit PX1778
2033	*Netlist, Inc. v. Samsung Electronics Co., Ltd.*, Dkt. 236, 2:21-cv-00463-JRG (E.D. Tex.), Trial Exhibit PX0446

Exhibit No.	Description
2034	Shiah, Tien (Samsung Semiconductor Inc., Memory Product Marketing), "Pivotal Memory Technologies Enabling New Generation of AI Workloads," downloaded from https://developer.download.nvidia.com/video/gputechconf/gtc/2019/presentation/s91028-2-pivotal-memory-technologies-enabling-new-generation-of-ai-workloads-presented-by-samsung.pdf
2035	News release by Samsung ("Samsung Begins Mass Producing World's Fastest DRAM – Based on Newest High Bandwidth Memory (HBM) Interface, downloaded from https://hothardware.com/news/samsung-begins-mass-production-of-the-worlds-fastest-dram-hbm2-launched
2036	Choi, J.S., "Next Big Thing: DDR4 3DS," JEDEC Server Forum 2014, available from https://www.jedec.org/sites/default/files/files/JS_Choi__Server_Forum_2014(1).pdf
2037	SK hynix Newsroom, "The Story of SK hynix_s HBM Development," downloaded from https://news.skhynix.com/the-story-of-sk-hynixs-hbm-development/
2038	2011-02-13 Netlist Presentation ("HyperCloud and N3DS")
2039	Building an SRAM Substitute, downloaded from https://www.bigmessowires.com/2015/09/04/building-an-sram-substitute/ (2015)
2040	"How Long Does NAND Flash Memory Last?", downloaded from https://www.networkcomputing.com/data-centers/how-long-does-nand-flash-memory-last#
2041	"Micron boosts NAND flash endurance six-fold," downloaded from https://www.networkcomputing.com/data-centers/how-long-does-nand-flash-memory-last#
2042	U.S. Patent No. 8,301,833 to Netlist

Exhibit No.	Description
2043	Micron NAND Flash Memory Datasheet for MT29F2G08AAWP/MT29F2G16AABWP, MT29F4G08AAWP/MT29F4G16AABWP, MT29F8G08AAWP (from IPR2022-00996, EX1049)
2044	STMicroelectronics SRAM datasheet (M48Z512A, M48Z512AY, M48Z512AV), available from https://www.st.com/resource/en/datasheet/cd00001187.pdf
2045	Everspin Technologies, MR4A08B MRAM compatible with SRAM (2018)
2046	Image of a JEDEC presentation, available from https://pc.watch.impress.co.jp/img/pcw/docs/488/696/html/28.jpg.html https://pc.watch.impress.co.jp/img/pcw/docs/488/696/html/28.jpg.html
2047	Neil H.E. Weste et al., "CMOS VLSI Design A Circuits and System Perspective" (2011)
2048	Kim, Ji-Young, et al. "An Energy-Efficient Design of TSV I/O for HBM With a Data Rate up to 10 Gb/s." IEEE Journal of Solid-State Circuits (2023)
2049	Standard Cells Databook from https://github.com/mflowgen/freepdk-45nm/blob/master/stdcells-databook.pdf
2050	UMBC Lecture Notes on Capacitance and Resistance Model, available from https://redirect.cs.umbc.edu/courses/undergraduate/CMPE315/Fall21/cpatel2/lectures/chap2_lect06_rc_model.pdf
2051	Sohn, K. et al., "A 1.2V 20nm 307 GB/s HBM DRAM With At-Speed Wafer-Level IO Test Scheme and Adaptive Refresh Considering Temperature Distribution," IEEE Journal of Solid-State Circuits, 52(1) (Jan. 2017)
2052	Flash 101
2053	"SRAM" notes, available from http://ece-research.unm.edu/jimp/vhdl_fpgas/slides/SRAM.pdf

Exhibit No.	Description
2054	Lee, D. et al., "Improving 3D-Stacked Memory Bandwidth at Low Cost," ACM Transactions on Architecture and Code Optimization, 12(4), Article 63 (2015), https://users.ece.cmu.edu/~omutlu/pub/smla_high-bandwidth-3d-stacked-memory_taco16.pdf
2055	Samsung, "High Bandwidth Memory is Catching on for Data-intensive Computing," printed from https://semiconductor.samsung.com/news-events/tech-blog/high-bandwidth-memory-is-catching-on-for-data-intensive-computing/

I. Introduction

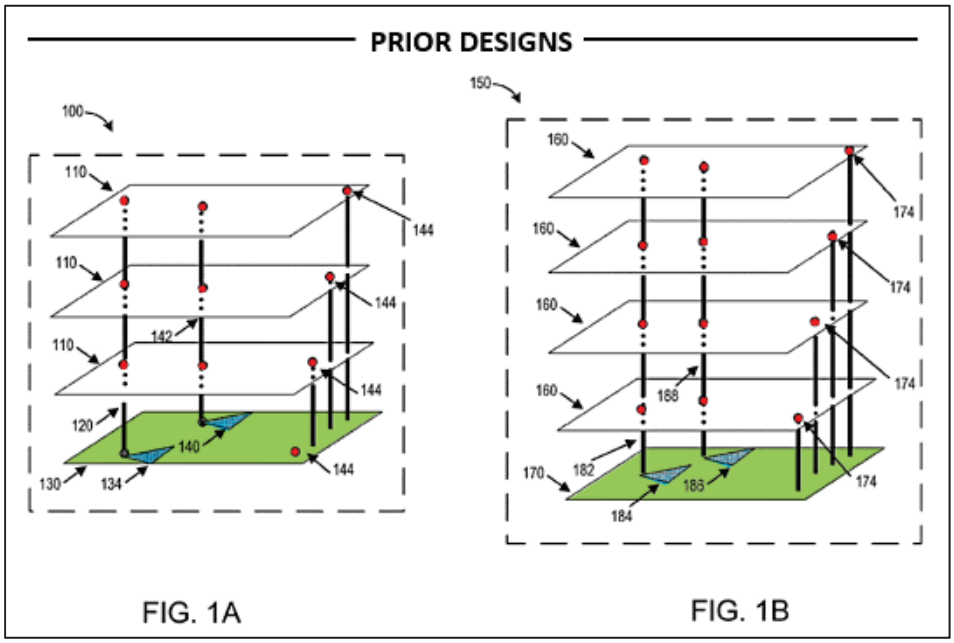
At the time of the invention, then-existing 3D-stacked memory package designs used multi-drop configurations, in which each TSV is in electrical communication with every memory die. EX1001, Figs. 1A-1B, 1:30-33. Samsung's 3DS-DDR4 packages used multi-drop configurations. EX2023 ("Brogioli"), ¶299; EX2036, 7-8. The first generation high bandwidth memory (HBM), first launched in 2014 by SK hynix, likewise used such multi-drop configurations. *See, e.g.*, EX2037, 3; EX2007, 1 (noting previous generation HBM had "multi-drop TSV"); Brogioli, ¶299.

The history of commercial 3DS and HBM in particular is objective evidence that the patented design is not obvious. Petitioner does not explain why the largest and most advanced memory manufacturers in the world (including, Samsung, Micron and SK hynix) did not include the claimed features when they developed earlier generations of 3D-stacked products. Samsung launched its first (infringing) HBM products in 2016, after receiving from Netlist a presentation on the '060 patent in April 2015. EX2027, 220:17-222:1; EX2032 [PX1778]-0001, -0004; EX2034 at 8 ("Samsung launched HBM2 in Q1 2016."). SK did not publish a paper announcing the need to reduce "heavy TSV loading" associated with multi-drop configurations until 2018. EX2007. That was years after Netlist first shared the inventive concept with the industry in 2011. *E.g.*, EX2038; Brogioli, ¶52.

As explained in further detail below, the asserted art did not teach or suggest elements 1.d.1-1.d.2 (or limitations 11.d.1-2, 20.e.1-2 or 29f) as related to electrical communication of a die interconnect with some, but not all, of the array dies. This is consistent with the objective evidence of the development history of the HBM and 3DS DRAM stacks. For these and further reasons below, the Board should uphold the challenged claims.

II. Overview of the '060 Patent

According to the '060 patent, prior art 3D-stacked memory package designs applied a multi-drop configuration, whereby each data die interconnect is in electrical communication with every array die 110. *See* EX1001, Figs. 1A-1B (annotated).



Dr. Hyun Lee, the inventor of the '060 patent, realized, however, that multi-

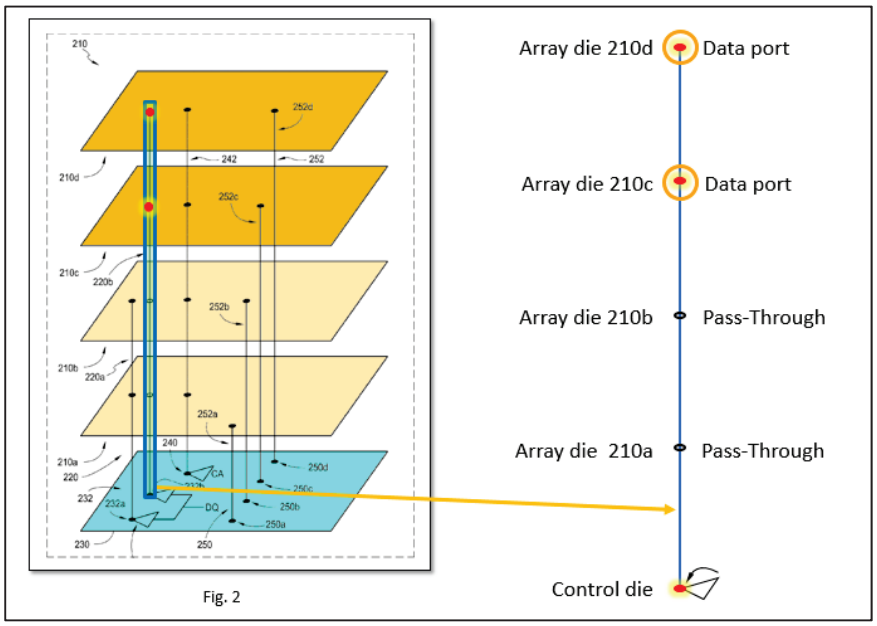
drop configurations often required a large driver to drive the signals along a die interconnect, and "generally a larger driver not only consumes more space on the control die, but also consumes more power." EX1001, 2:8-15; 17:43-49.

To overcome the above problem and to reduce the power consumption as well as the driver size, the '060 patent teaches that

> 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.

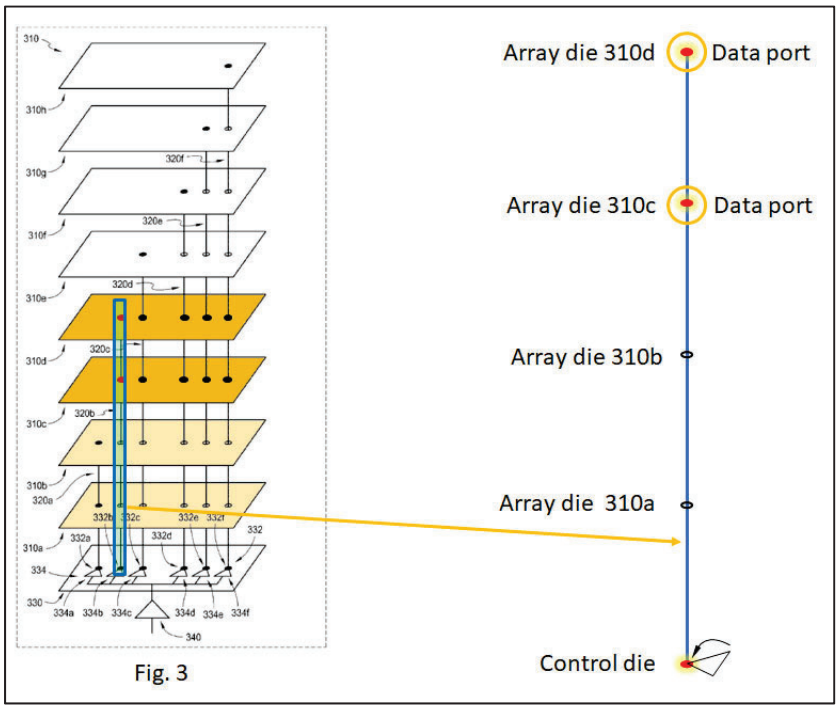
Id., 4:22-29.

As shown in Figures 2 and 3, instead of electrically communicating with every array die, each die interconnect is in electrical communication with some, but not all, of the array dies. For instance, Figure 2 shows that die interconnect 220a is in electrical communication with dies 210a and 210b, but not with dies 210c and 210d. In contrast, die interconnect 220b is in electrical communication with dies 210c and 210d but passes through dies 210a and 201b. *See id*., 5:46-6:35.



(EX1001, Fig. 2 annotated)

Likewise in Figure 3, the die interconnect 320a can be in electrical communication with dies 310a and 310b but not with dies 310c-h while the die interconnect 320b can be in electrical communication with dies 310c and 310d but not with other dies. *Id.*, 10:41-55.



(EX1001, Fig. 3 annotated)

The array dies in the memory package are stacked on top of a control die. *E.g.*, *id.*, Figs. 2-3, 6:36-67, 10:35-67. The control dies include data conduits with drivers that are configured to transmit data signals along corresponding data die interconnects. *Id.* The specification teaches, for example, "[t]he 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 interconnects 320 is in electrical communication via a port of the die interconnect 320." *Id.*, 11:12-16. By reducing the number of array dies in electrical communication with the die interconnect, the load and the required driver size are reduced. *Id.*, 11:16-31, 17:50-54. Thus, if a data die

interconnect is in electrical communication with only dies a and b, but not with dies c and d, then the size of the driver may be reduced to account for that reduced load to drive the signals to dies a/b. *Id*.

III. Skill Level of a POSITA

For purposes of this proceeding, Patent Owner applies the education and work-experience level specified on Petition, page 5. Notably, Petitioner does not allege that a POSITA would have been familiar with designs and operations of non-DRAM devices or non-JEDEC specifications. Pet. 5. Petitioner's expert testified unequivocally, however, that he "d[id]n't believe an SDRAM die is ever different than a DRAM circuit," and that he focused on obviousness combination that do not involve DRAM or SDRAM dies. EX2025 ("Wolfe"), 94:12-18. That is, under Petitioner's unpatentability theories, POSITAs would be engaged in the project of making changes to memory devices that they have no experience with. Brogioli, ¶30.

Petitioner's allegation that a POSITA "would … have been familiar with the structure and operation of circuitry used in stacked memory devices …." is also of no help because although they might have some general knowledge or familiarity with low-level circuits, they would not have had enough knowledge or experience to understand the interrelations and interoperations among all the different circuitry, or how to apply those general knowledge to substantially modify a

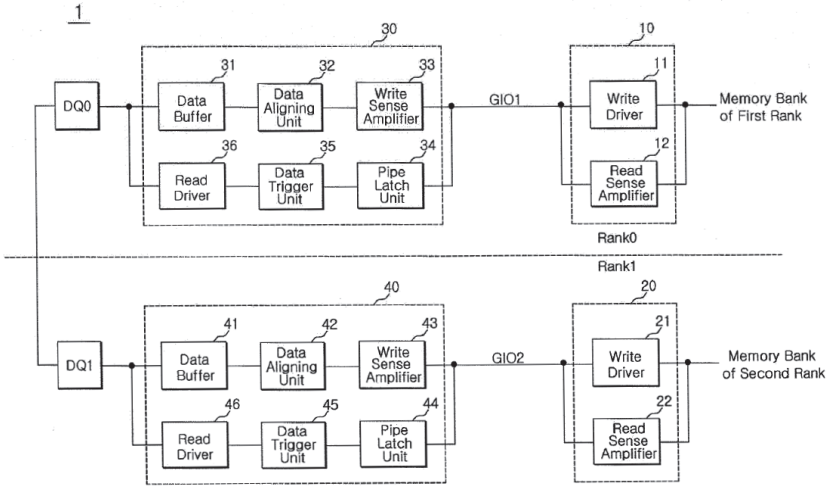
memory system, including accounting for the very different timing and electrical requirements. Brogioli, ¶32.

IV. Asserted Prior Art

A. Kim

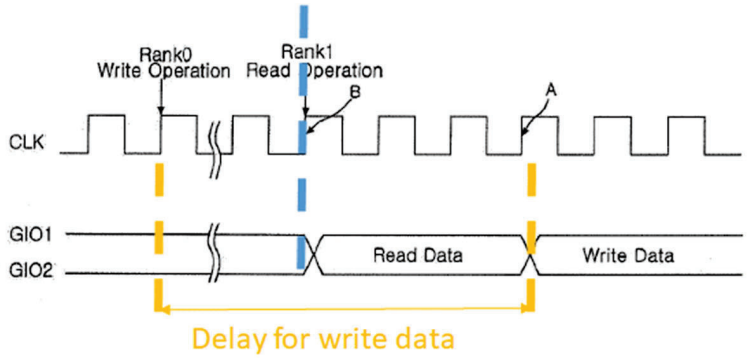
According to Kim, in a "typical semiconductor memory apparatus," each rank0/1 of memory device is connected individually to a separate input/output driver 10/20, global input/output line GIO1/2, data input/output section 30/40 and data pad DQ0/1. EX1014, [0006]-[0007], [0019], Fig. 1. As shown, the data pads, DQ0 and DQ1, for the two ranks are connected to each other. *Id.*

FIG.1
(PRIOR ART)



Kim notes that in such a configuration, "in the case where the write operation of the first rank Rank0 is performed and then the read operation of the second rank Rank1 is performed, a data collision is unavoidable." *Id.*, [0042]. Kim discloses that this problem would exist "[a]summing that the first and second

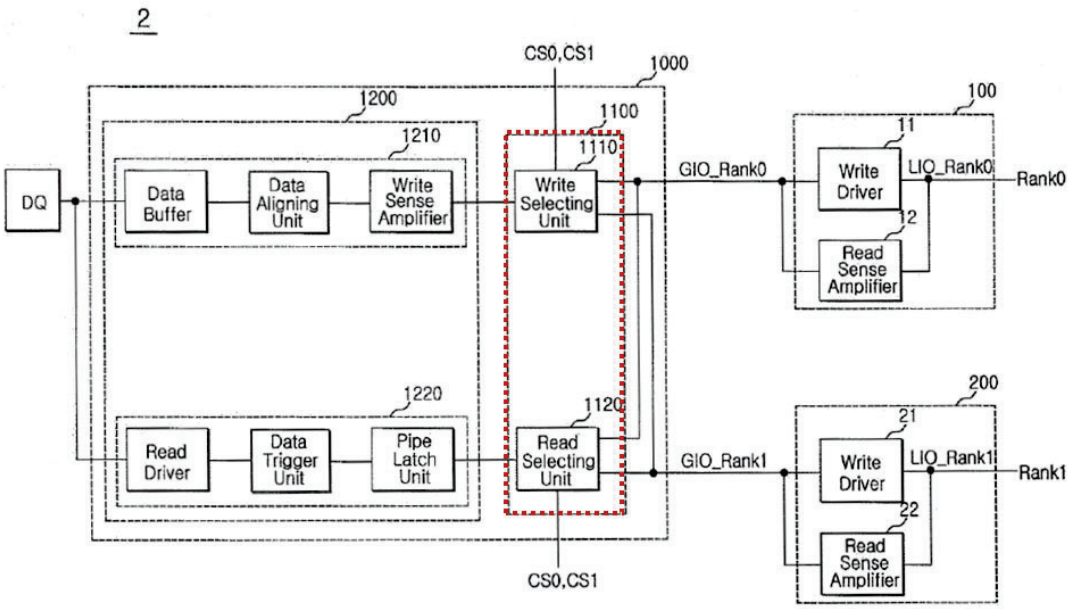
ranks Rank0 and Rank1 share a data input/output circuit," and "first and second data input/output lines operate in the same manner." *Id.* This is because "it takes time for [write] data inputted through the data pad DQ to pass through the data input/output section" such that "a time A at which the [write] data is actually loaded on the data input/output line [GIO1 and GIO2] is after a certain time interval is lapsed." *Id.* In contrast, in a read operation, the read data "is immediately loaded [at a time B] on the data input/output line GIO1 or GIO2." *Id.*, [0043]. Thus, because the write data appears on the data input/output line at time A, which is after time B at which the read data appears, "the data to be outputted to the data pad DQ is changed from the read data to the write data, resulting in a data collision." *Id.*; *see* annotated Fig. 4A below; Brogioli, ¶¶66-69.



To address the above problem, Kim proposes "a data input/output circuit compris[ing] a rank selecting section configured to be selectively connected to one of a first and second ranks in response to a chip selection signal, and output data to a connected rank or receive data from the connected rank; and a data input/output

section configured to output the data transmitted from the rank selecting section through a data pad to an [sic] device during a read operation, and output the data inputted to the data pad to the rank selecting section during a write operation." Kim, [0013]. A schematic illustration of such a circuit is shown below in Figure 3.

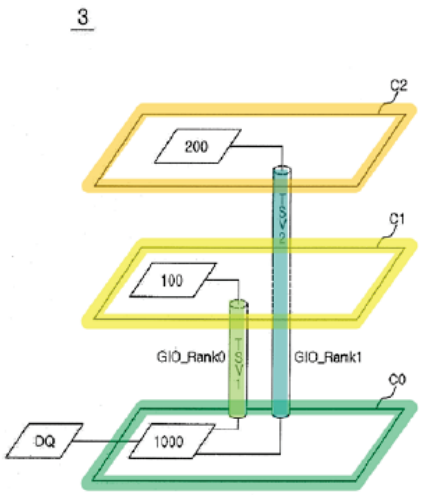
FIG.3



Kim asserts that the above configuration could avoid data collision in the cross-rank read-after-write operation because the rank selecting unit 1100, in response to CS0 and CS1, selects only one of the data input/output lines GIO_Rank0 and GIO_Rank1 to place data. *See id.*, [0044]-[0045], Fig. 4B. This allows write data to appear only on the GIO line associated with the rank selected for a write operation, instead of also appearing on the GIO line associated with the

non-selected rank. *Id.*; Brogioli, ¶¶77-79. Figure 5 relied on by Petitioner is an instantiation of the above concept with vertical integration. *See* Kim, [0046].

FIG.5



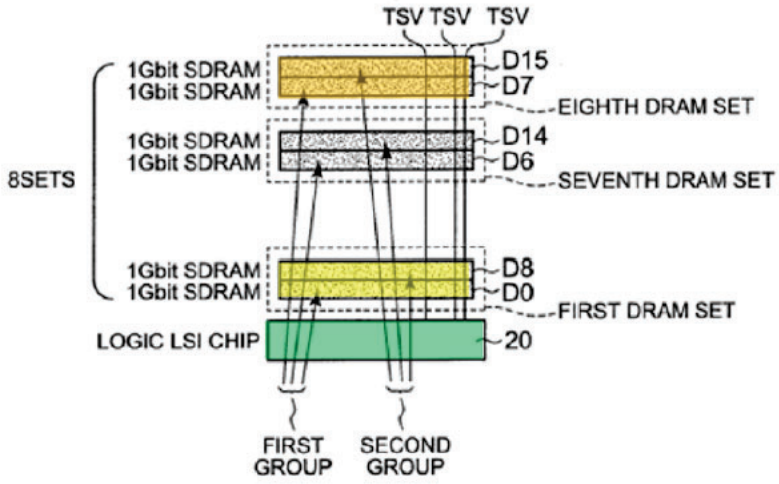
In this configuration, the data input/output section and rank selecting unit are implemented on the main chip C0 as element 1000; and slave chips C1 and C2 are connected respectively to TSV1 and TSV2 so that data associated with each chip travels along the dedicated TSV only. Kim, [0049].

B. Riho

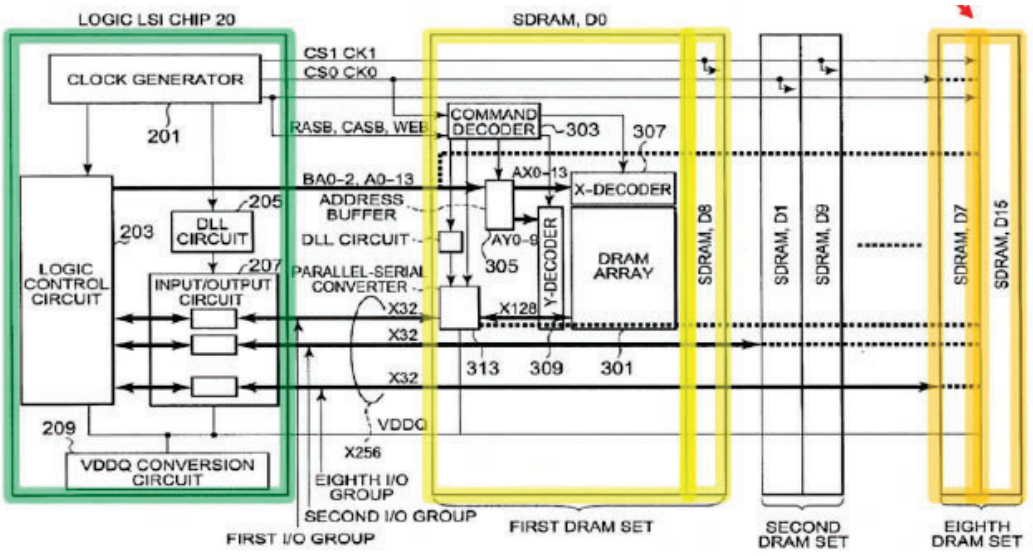
Riho discloses a memory package having sixteen SDRAM chips D0-D15 placed on top of a logic chip. Riho divides the 16 SDRAM chips into two "chip selection groups": D0-D7 and D8-D15. EX1016, [0028] and [0033]. In each access cycle, only the eight chips from a same chip selection group are accessed. *Id.*, [0033]. "Since the first and second groups are mutually exclusively controlled

by the control chip 20 at the first and second access cycles, each of the TSVs corresponding to one I/O bit is shared by the first and second groups." *Id.*

FIG. 1



In particular, one chip from a first group is paired with another chip from the second group to form a "set." *Id.*, [0029], Fig. 1. The eight sets are: D0 and D8, D1 and D9, D2 and D10, D3 and D11, D4 and D12, D5 and D13, D6 and D14, and D7 and D15. *Id.* "The first to eighth DRAM sets are accessed in parallel by the control chip (master chip) 20 …." *Id.*



In another embodiment (Figure 2, reproduced above as annotated by Petitioner), "[e]ach input/output circuit 207 sends and receives 32-bit width data signals DQ between itself and the SDRAMs D0 to D15, respectively, and thus sends and receives parallel data signal DQ of 256-bit width in total." Riho, [0040]. "The data signal DQ is transmitted or received as an I/O data signal." *Id.* Each DRAM set (*e.g.*, D0 and D8) is allocated with a common 32-bit I/O group. *Id.*, [0045]. "These eight I/O groups are accessed in parallel by the control chip …." *Id.*, [0040]. In this embodiment, "the DRAM sets defined by the I/O groups determine the data transfer rate" whereas "the chip selection groups determine the memory capacity value." *Id.*

The vertical structure of Figure 2 is shown in Figure 4. *Id.*, [0062]. In the Figure 2 embodiment, SDRAM D0 and the input/output circuit 207 of the logic

LSI chip 20 are connected to each other through 32 data signal DQ TSVs. *Id.*, [0045], Fig. 2. The "data signal DQ TSVs for the SDRAM D0 are also used by the SDRAM D8 …." *Id.* "The data signal DQ TSVs between the SDRAM D0 and the logic LSI chip 20 pass through the SDRAMs D8, D1, D9 … D14, D7 to extend to the uppermost SDRAM D15, thereby forming the continuous type TSVs." *Id.* The portion of the TSVs between SDRAM D1 (2nd set) through SDRAM D15 (8th set) is unused. *Id.*

Other DRAM groups utilize similar configurations. *Id.* For instance, the 32 "data signal DQ TSVs for the SDRAM D7 are provided between the input/output circuit 207 of the logic LSI chip 20 and the SDRAM D7" and "are commonly used by the SDRAMs D7 and D15." *Id.*, [0046]. "In this manner, all the data signals DQ TSVs form the continuous type TSVs connected between the logic LSI chip 20 and the uppermost SDRAM D15 and are substantially equal in length to each other," as illustrated in Figure 4. *Id.*, [0045], [0062].

FIG. 4

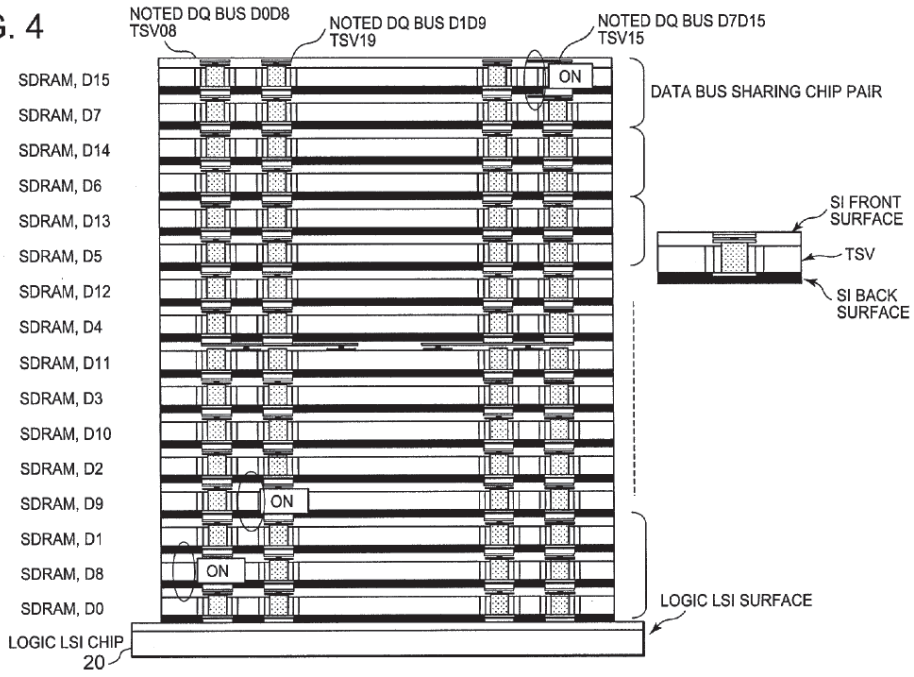
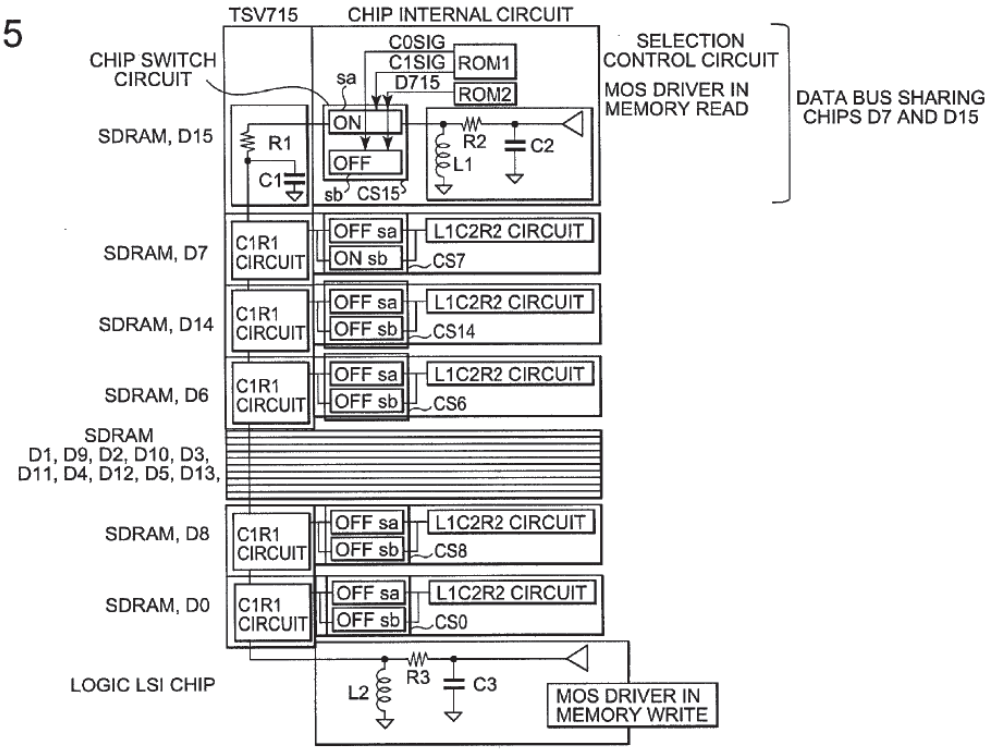


Figure 5 provides "an equivalent circuit" for what is depicted in Figure 4. *Id.*, [0069]. As seen in Figure 5, TSV715 is electrically coupled to each of the switching circuits CS0 through CS15 on D0 through D15. *Id.* Each switch circuit includes two switch elements sa and sb "for determining an on/off state with respect to the data signal DQ TSV715" under the control of ROM1 and ROM2, which "set indication signals collectively for all the SDRAMs." *Id.*, [0070]. For TSV715, sa of D15 or sb of D7 is on[1] while all other switch elements in the stack

[1] *See id.*, [0091] (switch "is turned on when the group indication signal C and the set indication signal D both take a logic 1," otherwise the switch is off); [0093]-[0094] (when the second chip selection group indication signal C1SIG is logic 1 and the set indication D715 is logic 1, SDRAM D15 is used; and when the first chip selection group indication signal C0SIG is 1 and the set indication D715 is 1, SDRAM D7 is on).

are off. *Id.*, [0071]-[0075], Fig. 5; *see also id.*, [0076] (D0-D6 and D8-D14 having the same equivalence circuit as D7 and D15); Figs. 6A-6C.

FIG. 5



V. Claim Construction

The Board should adopt the same constructions as the District Court because both forums apply the same *Philips* claim construction standard and the parties applied those constructions at the jury trial. A consistent construction promotes the policy goal of a "***just***, speedy and inexpensive resolution of every proceeding," 37 C.F.R. § 42.1(b), by applying the same construction for infringement and invalidity. *See, e.g.,* 83 Fed. Reg. at 51353 ("The PTAB will consider prior claim constructions from district courts or the ITC and give them appropriate weight.").

A. "array die"

In the District Court, Petitioner persuaded the Court to construe "array die" as "array die that is different from a DRAM circuit." EX2004-p.32. There, citing to arguments made during prosecution to distinguish Rajan137, "Samsung contend[ed] Netlist disclaimed array dies that are 'DRAM circuits' during prosecution." *Id.*, 31. The referenced prosecution statements state:

> … Rajan[137] does not disclose "a plurality of stacked array dies." Rajan[137] merely stacks DRAM circuits 206A-D, which are different from array dies. As a result, Rajan[137]'s buffer chip 202 also operates very differently from the control die in claim 1.

EX1002, 465.

Samsung argued that the above statements amounted to "a clear and unmistakable disclaimer":

> During prosecution, the examiner rejected pending claim 1 (among other claims) based on prior art reference Rajan[137]. See Ex. B at 4-9 (2013-10-11 Non-Final Office Action). The applicant argued Rajan[137]'s DRAM circuits are "different from array dies" recited in claim 1. See Ex. C at 10 (2014-01-14 Amendment) ("Rajan[137] merely stacks DRAM circuits 206A-D, which are different from array dies . . . ."). ***The applicant's argument was a clear and unmistakable disclaimer assertion that an array die does not include DRAM circuits***.

EX2003-p.19 (citation omitted).

The District Court agreed that Netlist disclaimed the structural features of the Rajan137 DRAM circuits 206A-D:

> Netlist *structurally* distinguished "stacked DRAM circuits" from "stacked array dies" to obtain the patent. Based on that distinction, the Court construes "array die" as "array die that is different from a DRAM circuit."

EX2004-p.32 (emphasis in original). Netlist agreed to be bound by this construction and proceeded with the jury trial on that construction. EX2010-EX2011 (Netlist's motion to withdraw objections to Markman order granted).

Rajan137's (EX1011) DRAM circuits 206A-D are identical to Rajan's (EX1015) DRAM circuits 417A-D. *Compare* EX1011, Figs. 2A-2E, [0050]-[0052] *with* EX1015, Figs. 2-6, 5:14-6:2.

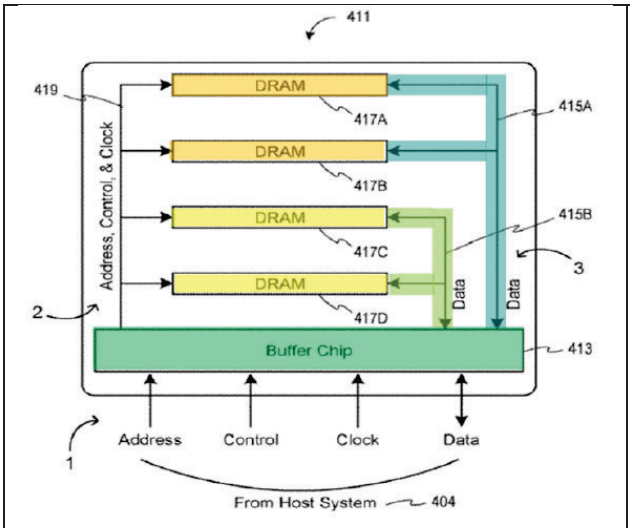


FIG. 4

(EX1015, Fig. 4)

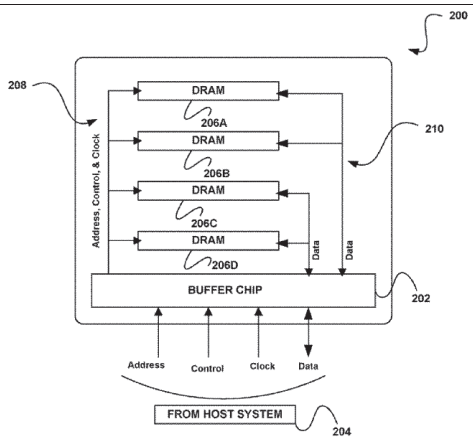


FIGURE 2C

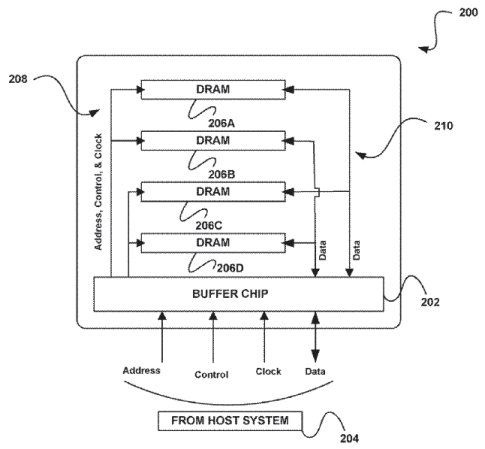


FIGURE 2B

(EX1011, Figs. 2B-2C)

B. "chip select signals"

In the District Court, Petitioner successfully argued that the plain and ordinary meaning of "chip select signals" and the associated "chip select conduits" "excludes situations in which a chip select signal could enable multiple array dies at once." EX2004-pp.33-34. Petitioner argued that such a construction is needed to give meaning to the term "select":

> Netlist seeks to rewrite the claim and ignore the word "select" under the guise of claim construction. It does so by expanding the coverage to a situation in which a chip select signal could enable multiple and even all array dies at once. But this is no selection at all… Instead, as Netlist's own evidence shows, and as the plain language of "chip select" suggests, a chip select signal enables or selects a "corresponding" or "respective" array die. *See* '060 patent at 1:49-56 …; *id.* at 25:16-25 …; *id.* at 24:31-36 …; *id*. at 25:28-30[.] … Netlist's

> cited evidence is consistent with the rest of the description in the '060 and '160 patents, which routinely describes chip select signals in this manner rather than as enabling or selecting "one or more" array dies. *See*, *e.g.*, '060 patent at 10:57-60 …; *id.* at 15:54-57[.]

EX2003-pp.21-23 (quotations omitted). The District Court agreed with Petitioner. EX2004-pp.33-34. Patent Owner also agreed to be bound by this construction. EX2010—EX2011. The construction should be applied by the Board.

VI. The Petition Improperly Relies Solely On Expert Testimony For Many Claim Elements

Throughout the Petition, Petitioner tries to evade the word limit by including no substantive analysis in the Petition itself, but attempting to incorporate paragraphs and paragraphs of its expert's declaration. *See, e.g.*, Ground 1: elements 1.c., 1.d.1-1.d.2 (and similarly, 11.d.1-11.d.2, 20.e.1-20.e.2, 29.f), 11.e.5, 29.b, 31-32; Ground 2: elements 7.a and 7.b; Ground 4: elements 1.d.1-1.d.2, 1.e.2-1.e.3, 7.a, 7.b 8.a, 10, 11.e.5, 29.b, 1-35; and Ground 5 elements 15.a-b.

For instance, for elements 1.d.1-1.d.2, Petitioner devotes two paragraphs in Ground 1 and a single paragraph in Ground 4, but cites respectively to eleven (11) and eighteen (18) of its expert declaration. Pet. 39-40, 100-101. This is improper, and such arguments should be deemed as absent from the Petition itself. *See, e.g., Cisco Sys., Inc. v. C-Cation Techs, LLC*, IPR2014-00454, Paper 12 (PTAB, Aug. 29, 2014); *Fidelity Nat'l Info. Serv.s, Inc. v. DataTreasury Corp*, IPR2014-00491,

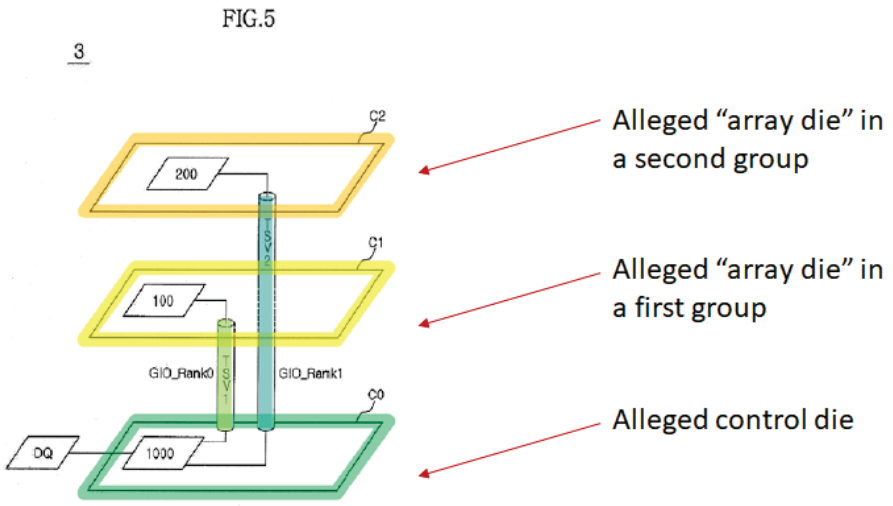
Paper 9 at 8-10 (P.T.A.B. Aug. 13, 2014) (treating Petitioner's citations to the 287-page expert declaration in lieu of citations to the record as impermissible incorporation-by-reference arguments).[2]

VII. Kim Grounds Are Not Obvious

A. Kim, alone or in combination, does not disclose or render obvious designs with elements 1.d.1 and 1.d.2

Each of the challenged claims require "[a] first die interconnect in electrical communication with [a] first group of array die***s*** and not in electrical communication with [a] second group of at least one array die," and "a 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 die***s***." *See* EX1001, 23:65-24:5, 25:9-15, 26:15-22, 26:65-27:5. That is, at least one of the die interconnects must be in electrical communication with more than one die. *Id.* And there must be at least one other die that is not in electrical communication with that die interconnect. *Id.* In Kim, however, each TSV is in communication with only a single chip. *See* Kim, Fig. 5; Brogioli, ¶84.

[2] This lack of analysis is especially problematic when considering that the Petition relies extensively on block diagrams from the figures of the asserted references, which lack crucial detail such as even basic circuitry information.



(Kim, Fig. 5 annotated)

Petitioner points generally to the prior art to argue that "[a] POSITA would have been motivated to implement a shared bus for multiple memory chips …." Pet. 27-28. Petitioner argues that "[s]haring a TSV among two or more dies was a known option," and thus presumably obvious. Pet. 28. Petitioner does not explain, however, how TSVs are shared in the prior art or how Kim's TSVs are to be shared. Pet. 26-30, 37-40.

Petitioner looks to Rajan, Figure 4 for die grouping.[3] In Rajan, Figure 4, Rajan's DRAM circuits 417A and 417B share one data bus and DRAM circuits 417C and 417D share another data bus. EX1015, Fig. 4 (reproduced below as

[3] Petitioner also cites Ruckerbauer, Foster and Ellsberry for die grouping. Petitioner does not assert that these references disclose any of the claim limitations and they do not cure Kim's deficiencies. Brogioli, ¶¶152-165.

annotated on Pet. 27).

Rajan:

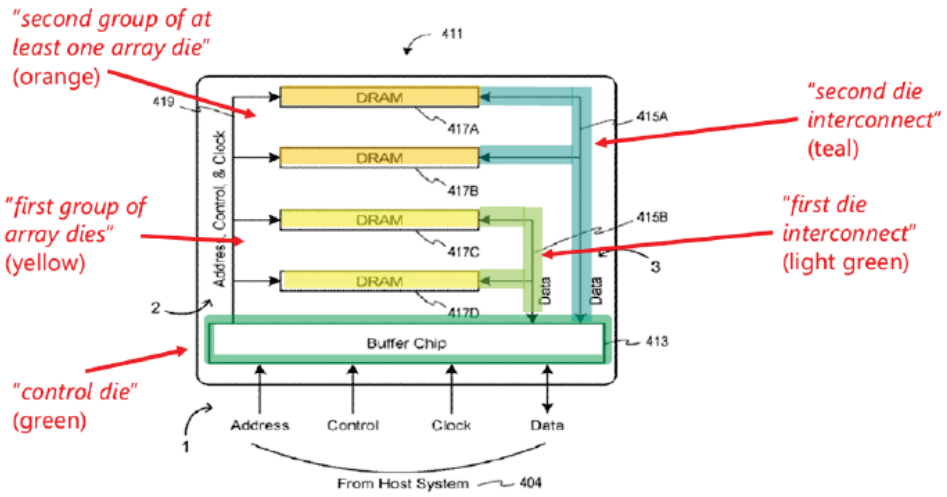


FIG. 4

Rajan, however, is silent on what is the relationship among the DRAM circuits sharing a data bus. *See* EX1015, 2:6-7, 5:36-43. But regardless, as Dr. Brogioli explains, the four possibilities proposed by Petitioner or its expert would either yield inoperable configurations or would not arrive at the claimed inventions. Brogioli, ¶86. These four possibilities are: (1) four chips belong to two ranks and two chips from the same rank share a common TSV; (2) four chips belong to two ranks and a chip from one rank shares a TSV with a chip from another rank; (3) four chips belong to four ranks and each of the four ranks has its own TSV; and (4) four chips belong to four ranks and every two ranks share a TSV. *Id.* The sections below discuss each possibility in turn and explain why they

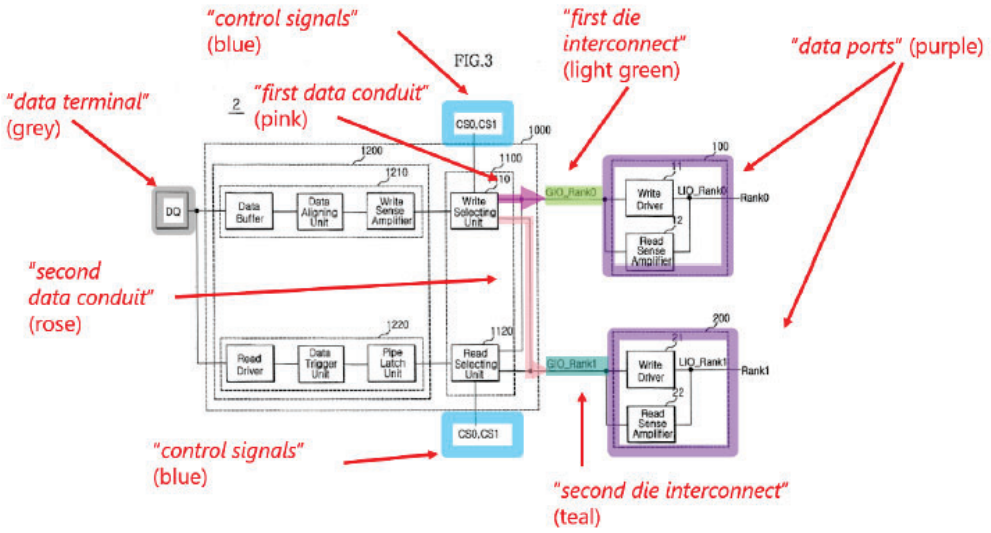
do not render the challenged claims obvious.

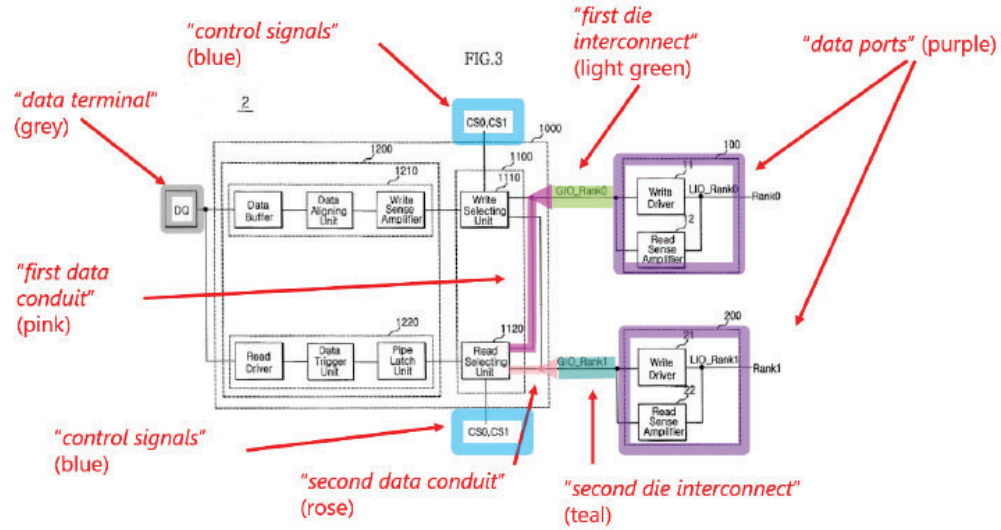
1. Modifying Kim by having chips of the same rank share the same TSV would result in an inoperable package

For elements 1.e., Petitioner states:

> A POSITA would have understood from Kim's description of rank selection by chip selection signals (CS0, CS1) for read/write operations that the "*data conduit*" connecting the write selecting unit 1110 … and the read selecting unit 1120 … in data input/output section 1000 … to a TSV to Rank0 is between "*the first die interconnect*," e.g., TSV1, and a "*data terminal*," e.g., data pad DQ (grey…), and the "*data conduit*" connecting the data input/output section 1000 … to a TSV to Rank1 is between "*the second die interconnect*," e.g., TSV2, and the "*data terminal*," e.g., data pad DQ (grey…).

Pet. 44; *see also* Pet. 46 (arguing that Kim discloses a control circuit that controls respective states of the first data conduit (to TSV1, pink) and the second data conduit (to TSV2, rose) in response to chip selection signals CS0 and CS1 and JEDEC read/write command signals).





That is, Petitioner argues that TSV1 and TSV2 should respectively connect to Rank0 and Rank1 chips. Brogioli, ¶88.

Samsung's and Micron's experts have conceded, however, that dies from the same rank should not share the same data wire (such as a TSV) because then the dies would impermissibly output data concurrently, resulting in data collisions, electrical problems and damages to circuits. *See* quotes below.

For example, Dr. Wolfe testified that "ordinarily we would refer to the complete set of dies that can operate concurrently, that can output at the exact same time in a single data word as a rank …." EX2025, 39:12-16. And he testified "it would generally be difficult to have two die[s] both output simultaneously to the same TSV and have it operate correctly." *Id.*, 40:1-4. That is because "a wire with TSV is just a specific kind of a wire," such that connecting two chips belonging to the same rank to the same TSV could cause "a data collision," "an electrical problem," and "that's just not the way a wire [such as a TSV] in a memory circuit normally operates." *Id.*, 39:18-40:12; *see also id.*, 29:13-23 (in Riho, D14 and D15 belonging to the same rank would not be connected to the same TSV, because the TSV is "only one wire," "can only be high or low," and can't carry values at the same time from two different DRAM arrays"); 29:24-31:1 (connecting two chips in the same rank to the same TSV could cause data collision, could led to, *inter alia*, "speed-related problems," "current-related problems" and "damages to circuits), 91:14-22 (agreeing that if Rajan's DRAM circuits 417A and 417B are in the same rank, then "you wouldn't put them in the same rank" because the two would then output data simultaneously in a read operation).

Micron's expert in the district court, Dr. Harold Stone, testified to the same effect when discussing Figure 2 of the 060 patent. EX2024 ("Stone"), 122:13-123:13. Dr. Stone testified that the assumption that 210a and 210b, which share a

same die interconnect, would belong to the same physical rank "becomes unreasonable" and "that you have a physical data on two die layers –two array die layers, and they're all going to drive the same die interconnect at the same time, you know, I could say, Hey, yeah, that looks very strange. I don't think you can build that." *Id*.

Dr. Brogioli is in accord, explaining that

> … normally, in a JEDEC-compliant DIMM, when a read or write command is sent, devices in the rank selected by the chip select signal would output or receive data bits simultaneously in parallel with each bit accessed in an access cycle carried on a separate data line to a separate data terminal, as illustrated below and as would understood by a person of ordinary skill in the art.

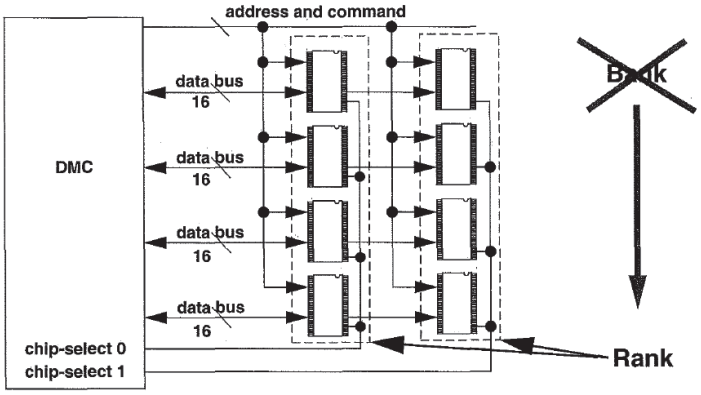
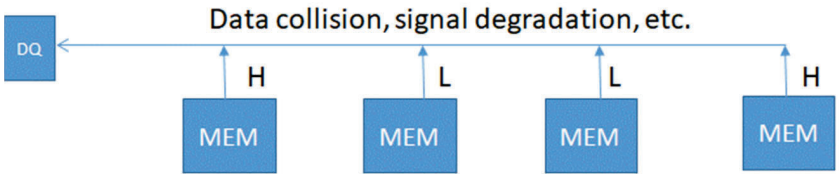


FIGURE 10.5: Memory system with 2 ranks of DRAM devices.

(EX1022, 413)

In contrast, in Kim's configuration, the chips in the same rank are connected to the same wire (*e.g.*, TSV) and the same DQ, as illustrated below.



> Therefore, when these chips in the same rank respond to the same read command, they would output data concurrently to the same wire, resulting in data collision when the devices output data bits of different values to the same wire concurrently. A write command, on the other hand, would result in the same data being written to different chips in the same rank instead of having different bits written to different chips as intended. As such, coupling chips belonging to the same physical rank to the same TSV would result in an inoperable configuration.

Brogioli, ¶¶94-96; *see also id.*, ¶¶97, 99-101 (providing examples on how ranks in a DIMM operate, and how prior art like Riho have memory devices in the same rank output or receive data via different TSVs in parallel in response to a read or write command).

In summary, Drs. Wolfe, Stone and Brogioli all agree that a POSITA would not have modified Kim in such a way that devices belonging to the same rank share the same TSV because that would cause data collisions, electrical problems and even damages to circuits. Brogioli, ¶¶108, 97. Consequently, a POSITA would not have modified Kim by having two chips from the same rank share a TSV.

2. Modifying Kim by grouping one chip from each of two ranks and having them share one TSV would not yield an operable memory system

At his deposition, Dr. Wolfe denied that he modified Kim in reliance on an assumption that the dies sharing the same wire and data bus belong to the same rank. EX2025, 90:22-91:22. Instead, he testified that for all claims, he relied on a modification in which the two dies sharing the same data bus each come from one rank of the memory devices, as illustrated below. *See* EX2025, 91:23-92:14.

Rajan:

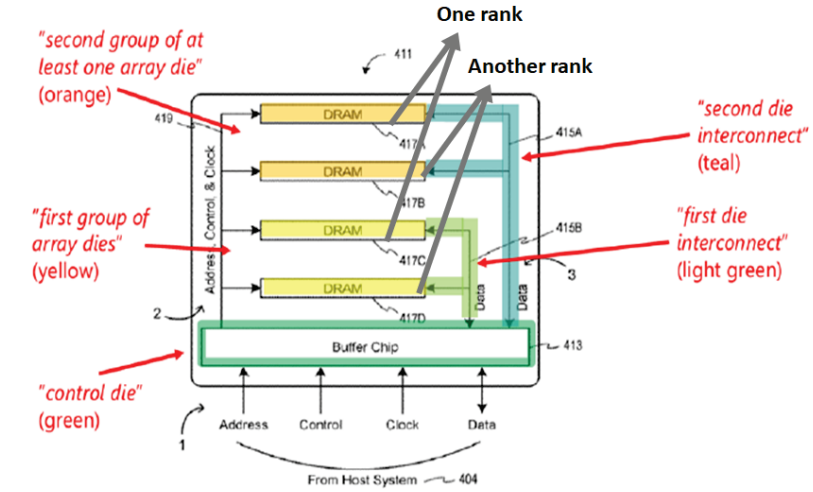
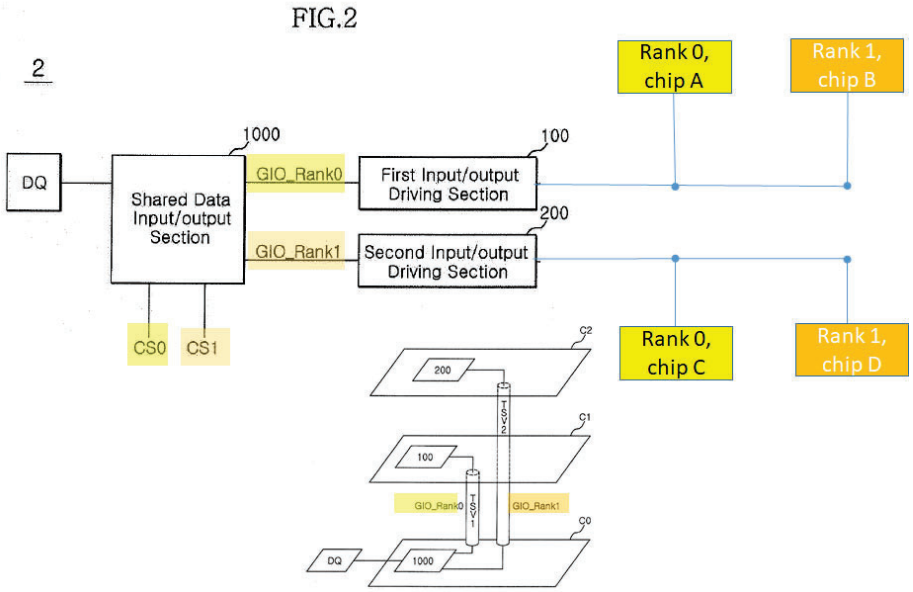


FIG. 4

(Further annotation to Samsung's annotation on Pet. 27 for illustration)

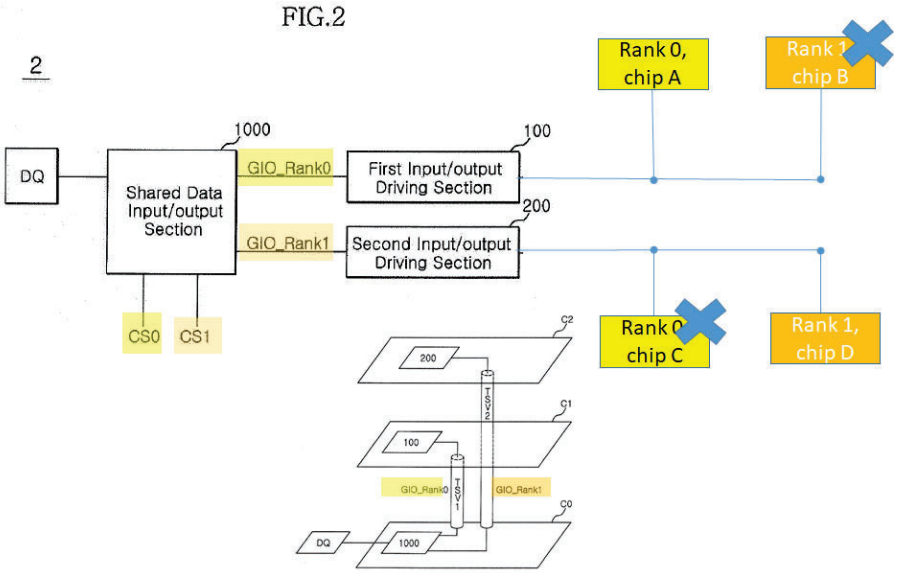
As Dr. Brogioli explains, applying that mapping to Kim would yield an arrangement like below (Brogioli, ¶110), in which TSV1 corresponding to GIO_Rank0 would be shared by a rank0 memory (Chip A) and a rank1 memory (Chip B) and TSV2 corresponding to GIO_Rank1 would also be shared by a rank0

memory (Chip C) and a rank1 memory (Chip D).



In Kim, "[t]he rank selecting unit 1100 distinguishes the read operation and the write operation for the first and second ranks Rank0 and Rank1 in response to the chip selection signals CS0 and CS1, and is selectively connected to the first and second data input/output lines GIO_Rank0 and GIO_Rank1." Kim, [0045]. That is, one of GIO_Rank0 and GIO_Rank1 is selected in response to CS0 or CS1. Brogioli, ¶112; *see also* Kim, [0028]-[0030].

As such, "(1) in response to a CS0 signal, only GIO_Rank0 line would be selected, making the Rank0 chip C connected to GIO_Rank1 inaccessible; and (2) in response to a CS1 signal, only GIO_Rank1 line would be selected, making the Rank1 chip B connected to GIO_Rank0 inaccessible." Brogioli, ¶113; *see* illustration below.



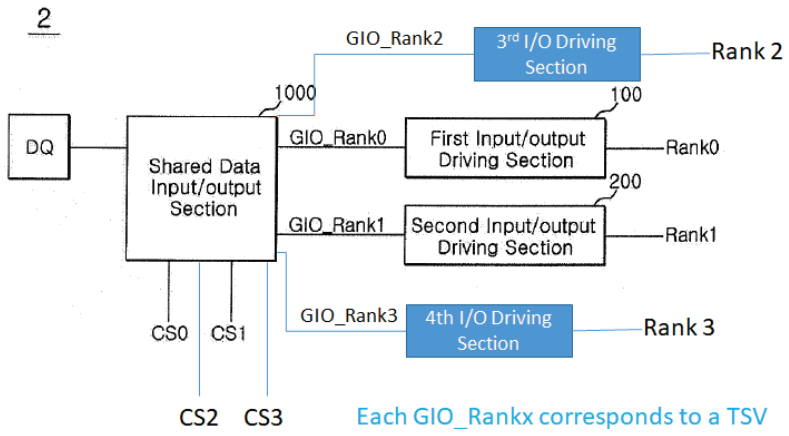
A POSITA would not have made this arrangement because this would result in half of the memory devices inaccessible, resulting in wasting half of the memory capacity, and worse still, loss or corruption of half of the data when half of the memory devices in a rank cannot be accessed in a read or write cycle. Brogioli, ¶114. For example, in a Rank0 write operation, chip C on the GIO_Rank1 line would receive data of indeterminate values or end up storing read data from Chip D instead if a cross-rank read-after-write operation occurs (*see* section VII.A.4). *Id*. Simply put, a POSITA would not have had a reason to modify Kim in such a way that the same TSV is shared by two chips, one from each of the ranks, as Dr. Wolfe apparently suggested at his deposition. *Id.*, ¶115.

3. Modifying Kim with four separate TSVs for four different ranks would not result in the claimed inventions

Dr. Wolfe also suggests that a POSITA would have modified Kim to have

four separate ranks for "many of the claims." EX2025, 92:2-8. Dr. Wolfe stated that there are two ways to go from two dies in a stack to four dies, one of which involves "hav[ing] four GIOs and four chip selects and expand[ing] the read and write selecting units to each select 1 of 4 …." EX2025, 45:5-13; *see also id.*, 44:13-45:4 (expanding two ranks to three ranks "is just generally difficult and complicated").

Dr. Brogioli explains that in this option where there are four separate GIOs, one for each rank, each GIO would be associated with one chip, instead of multiple chips for at least one of the TSVs. *See* illustration below; Brogioli, ¶117.



This modification therefore would not lead to the claimed inventions, which require at least one die interconnect be "in electrical communication with [a] first group of array dies" (plural). EX1001, 23:66-67; Brogioli, ¶118.

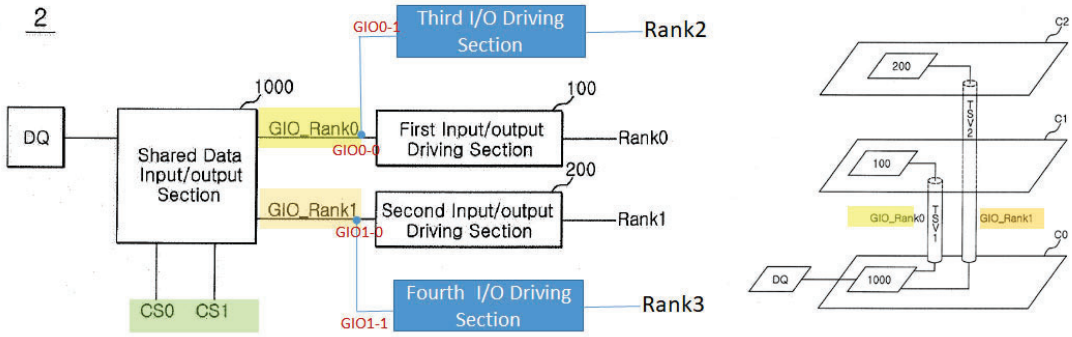
4. Modifying Kim by having two separate ranks of chips share a TSV would re-introduce data collisions that Kim tries to avoid

Dr. Wolfe also suggested an option of "hav[ing] 4 ranks sharing 2 ports," (EX2025, 45:5-13), for which he further testified that:

> One option would be to add -- to connect Rank2 to the GIO_Rank0 line and to connect Rank3 to the GIO_Rank1 line, and create the appropriate control signals so that only one rank would drive each of those I/O lines at a time or only one rank would respond to write operations on them.
> ….

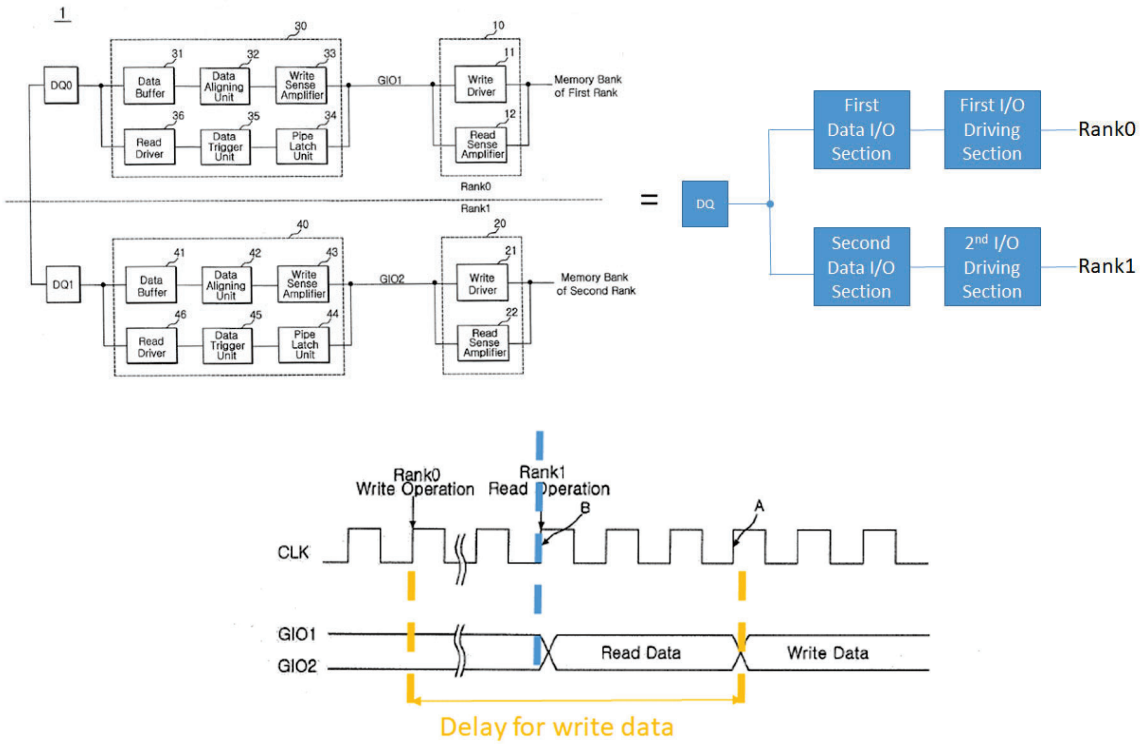
EX2025, 46:3-12. A POSITA, given Kim's express concern regarding data collisions in cross-rank read-after-write operations would not have made the proposed modification, even if they could "create the [unspecified] appropriate control signals." Brogioli, ¶¶121, 136-137. Dr. Brogioli illustrates Dr. Wolfe's proposal as below. *Id.*, ¶120. This configuration applies regardless of whether Rank2 and Rank0 (and Rank3 and Rank1) are completely independent or whether they belong to the same logic rank but different physical ranks (*e.g.*, rank multiplication). *Id.*

FIG.2



Kim notes an issue of "unavoidable" data collisions in cross-rank read-after-write operations[4] in systems where more than one rank of memory devices are coupled to a data pad in a memory access but "cannot share a data input/output circuit" that selects a path associated with a particular rank. Kim, [0013], [0042]-[0045]. This results in write data associated with an earlier issued Rank0 write command appearing on GIO2 later in time than read data associated with a later-issued Rank1 read command, causing the Rank1 read data to be changed to the later appearing Rank0 write data and an unavoidable data collision. *Id.*, [0042]-[0043]; Brogioli, ¶¶122-124.

[4] For example, "where the write operation of the first rank Rank0 is performed and then the read operation of the second rank Rank1 is performed." Kim, [0042]-[0043].

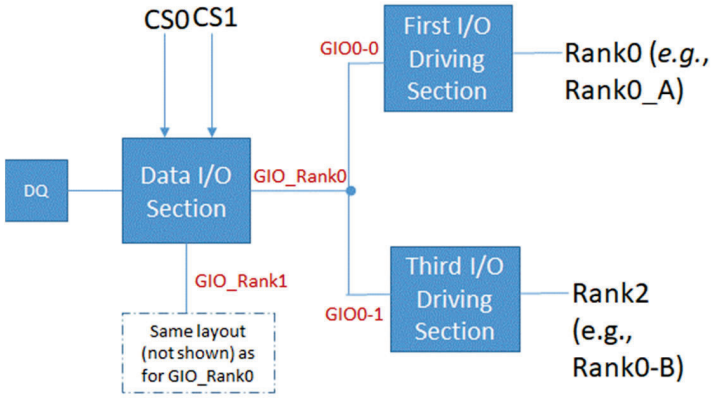


(Kim Figs. 1, 4A & high-level block diagram illustration of Fig. 1)

Kim solved the "unavoidable" data collision problem by providing a shared data input/output section 1000 with a rank selecting unit 1100 to the two memory ranks in its disclosure. *See* Kim, [0045], [0021]; Brogioli, ¶125. As shown in Figure 3, "[t]he rank selecting unit 1100 distinguishes the read operation and the write operation for the first and second ranks Rank0 and Rank1 in response to the chip selection signals CS0 and CS1, and is selectively connected to the first and second data input/output lines GIO_Rank0 and GIO_Rank1." *Id.*, [0045]; Brogioli, ¶126. In this manner, "even when the read and write operations of the first and second ranks Rank0 and Rank1 are repeatedly performed, a data collision

does not occur …." *Id.*

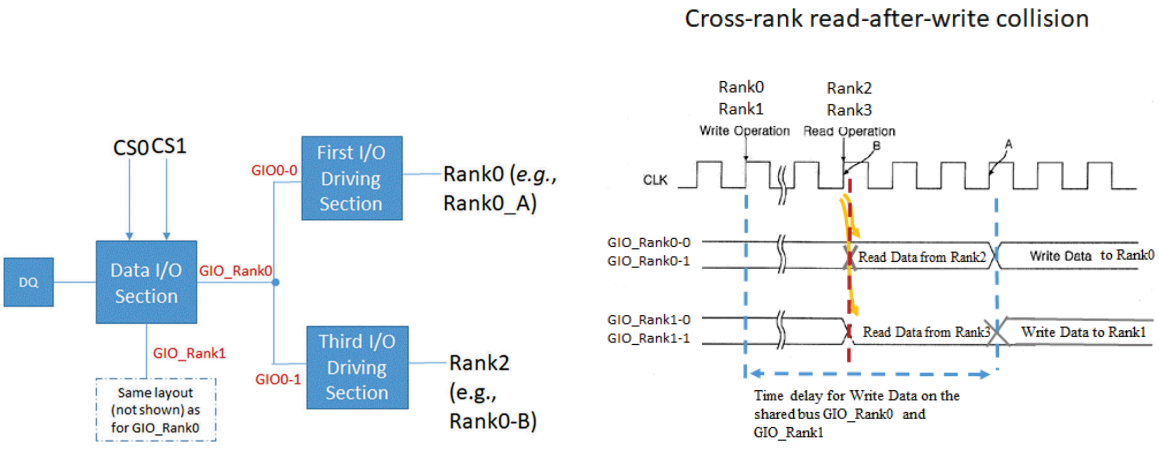
But in Dr. Wolfe's modification, Rank0 and Rank2 memories would share the GIO_Rank0 line without an intervening rank selection unit that selects between GIO0_0 and GIO0_1. *See* illustration below; Brogioli, ¶¶127-128. Likewise, Rank1 and Rank3 memories would share the GIO_Rank1 line without an intervening rank selection unit that selects between GIO1_0 and GIO1_1. *Id.*



In the absence of such additional rank selecting units, data placed on the shared GIO_Rank0 path would appear on both GIO0-0 and GIO0-1, and data placed on the shard GIO_Rank1 path would appear on both GIO1-0 and GIO1-1. Brogioli, ¶129. This would reintroduce the data collision problem experienced by a "typical semiconductor memory apparatus." *Id.*

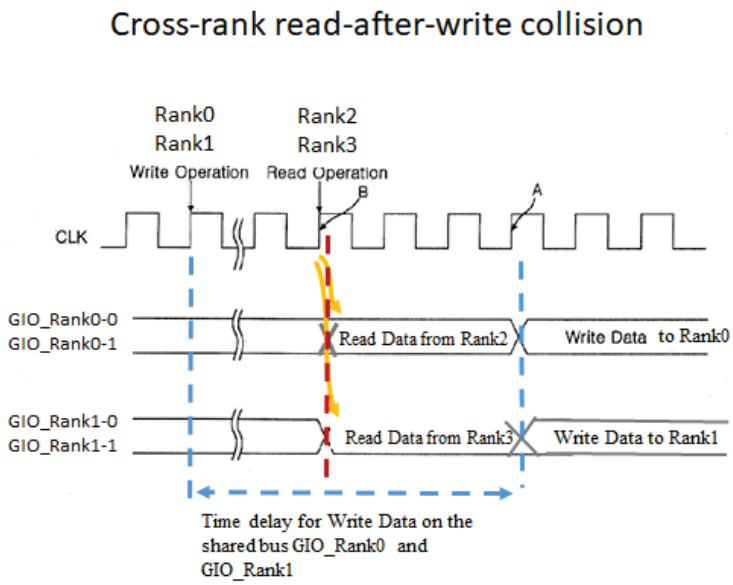
For instance, as in a "typical semiconductor memory apparatus," when a write operation for Rank0 occurs, it takes time for the write data to propagate through the data I/O section and shared GIO_rank0 path, before being loaded on

both GIO0-0 and GIO0-1 lines. *See* Kim, [0042]; Brogioli, ¶130.



The write data for Rank0 would appear on both GIO0-0 and GIO0-1 because, in Dr. Wolfe's proposal, nothing directs write data appearing on the shared GIO_Rank0 line selectively to only one of the GIO0-0 and GIO0-1 to prevent write data intended for Rank0 to also appear on GIO0-1 line associated with Rank2. Brogioli, ¶131.

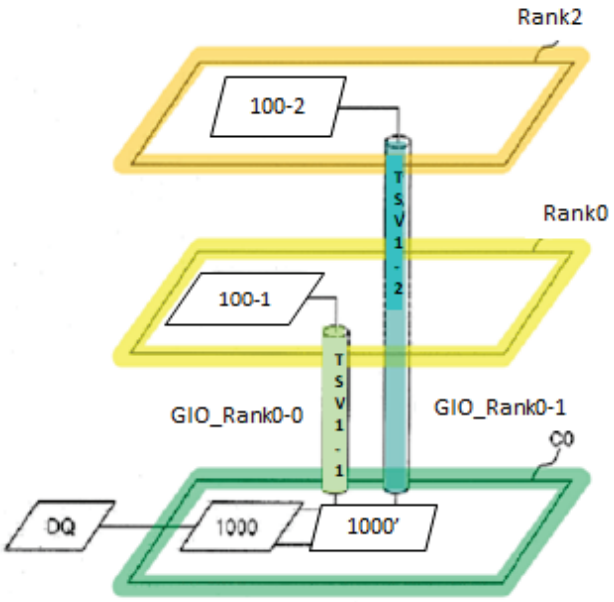
In that situation, if a Rank2 read operation follows a Rank0 write operation, then the read data from Rank2 would be immediately loaded onto GIO0-1 line in Kim's memory system. Brogioli, ¶¶132-133; Kim, [0042]-[0043]. The Rank2 read data would appear on GIO0-1 earlier in time than the time at which Rank0 write data is loaded on GIO0-1, resulting in the Rank2 read data being changed to the Rank0 write data and hence a data collision. *Id.*; *see* illustration below.



Consequently, in light of the timing constraint inherent in Kim and Kim's explanation for the cause of the unavoidable data collisions, a POSITA would understand that the two ranks cannot share the same GIO line and thus would not have modified Kim in the manner suggested by Petitioner because this would re-introduce the unavoidable data collision that Kim tried to solve. Brogioli, ¶134; *DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.*, 567 F.3d 1314, 1326-28 (Fed. Cir. 2009) (not obvious when prior art warned that rigidity increases the likelihood that a screw would fall within the human body and become inoperative for its intended purpose, and hence taught away from the proposed combination for a rigid screw).

To be clear, Netlist is not proposing bodily incorporation, but merely responds to Dr. Wolfe's proposed modification that does not specify any additional

changes beyond having multiple chips share a common TSV. As Dr. Brogioli explains, adding rank-selecting units to select between GIO0-0 and GIO0-1 and between GIO1-0 and GIO1-1 would result in a four-rank-four-TSV configuration that does not meet the requirement for 1.d.1. Brogioli, ¶135.



1000' = rank selection between GIO0-0 and GIO0-1 paths and paths associated with Ranks1 and Rank3 not shown

5. A POSITA would not have combined Kim with Rajan in the manner proposed

At his deposition, Dr. Wolfe admits that expanding Kim from two ranks to four ranks would be "complicated," albeit allegedly "routine." EX2025, 56:14-24. But neither Petitioner nor Dr. Wolfe provides any analysis on what those admittedly "complicated" changes would be, including "all of the timing and all of

the driver control to make it all work." *Id.*, 57:1-4.[5]

As noted above, timing—especially the stated "unavoidable" "data collision" in the cross-rank read-after-write operation with a shared data bus—was a critical problem for Kim. Kim, [0042]-[0043]; Brogioli, ¶139. Petitioner does not provide any solution to overcome Kim's timing restrictions for its particular memory system. *See also* Brogioli, ¶137 (increased read latency would require chip redesign beyond a POSITA's skill level and would introduce undesirable decrease in data throughput and speed). Nor has it provided any evidence that a POSITA, who is only expressly alleged to be familiar with JEDEC DRAM devices, would have had the right knowledge to solve Kim's timing problem (other than following Kim's teaching). *Id.*, ¶¶29-32, 138-142 (proposed modification involves "a multi-dimensional problem which could not be easily solved"). As such, Petitioner has not shown that a POSITA would have modified Kim in light of Rajan for a shared bus approach. *See also id.*, ¶143 (unlike proposed modifications for Kim, Rajan does not share interconnects and therefore does not have the same data collision issues).

In its reason for modification, Petitioner also argues that a POSITA would

[5] Without showing how a POSITA could allegedly overcome the complicated timing issues, Petitioner's obviousness analysis is legally deficient because that difficulty would have demotivated a POSITA from the modification.

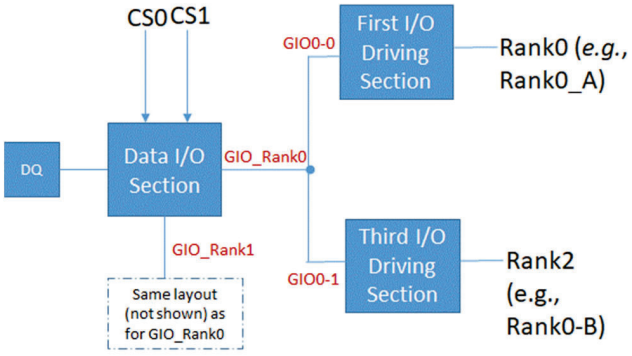
have looked to Rajan for an allegedly JEDEC-compliant interface. Pet. 24-25. But that modification deals with the interface between the memory controller and the control circuit and not the interface between the control circuit and the memory devices. *Id.*; Brogioli, ¶145; EX2025, 94:25-95:12. Rajan's DRAM circuits need to be "different" from Kim's (or Riho's) chips for the Petition to succeed, based on the claim construction the District Court adopted at Petitioner's urge. EX2004-p.32. Despite that difference, Petitioner does not provide any evidence that Rajan's JEDEC-interface would have worked with the "different" memory devices.

Petitioner also does not argue any modification to Kim memory devices, including its read/write latency. Pet. 26-30, 34-40; EX2025, 95:7-12 (Dr. Wolfe testifying "Kim … is clear that the external interface is distinct from the internal chip level interface").[6]

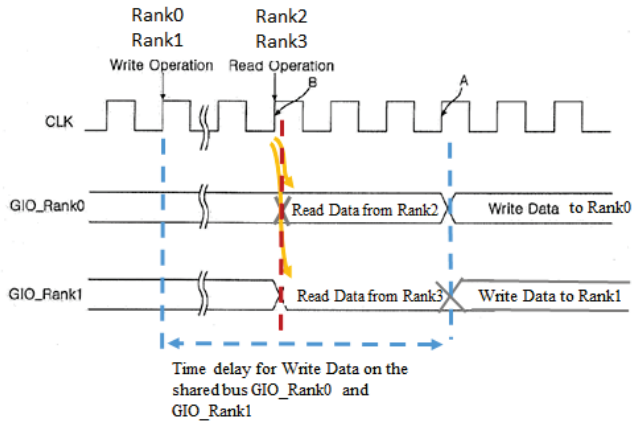
As such, even with the modification, when Kim's memory device receives a read command, it will immediately output read data on the corresponding local data line, that is, GIO0-0 for rank0 and GIO0-1 for rank2. Brogioli, ¶146; discussion in section VII.A.4 above; illustration below. Write data, on the other

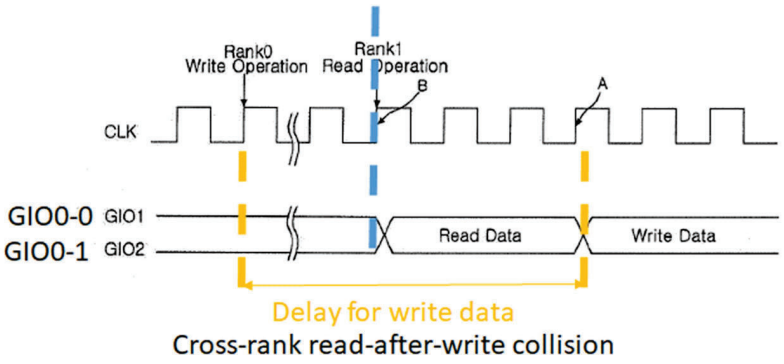
[6] In addition, Dr. Wolfe makes clear that for the IPR, he is relying on non-DRAM dies. EX2025, 94:12-18. Thus, JEDEC-specification for DRAMs would not apply to those devices.

hand, takes much longer to appear on the GIO0-0 and GIO0-1 line because "it takes a time for data inputted through the data pad DQ to pass through the data input/output section." *See id.*; Kim, Figs. 4A-4B (delay between read command and first piece of read data much smaller than delay between write command and first piece of write data); [0042].



Cross-rank read-after-write collision



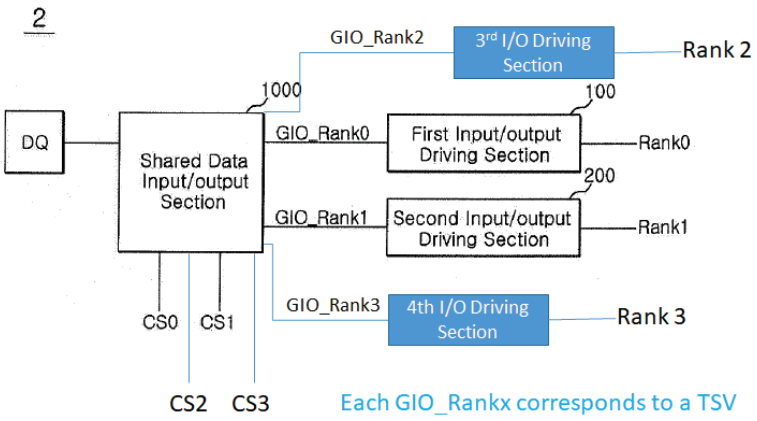


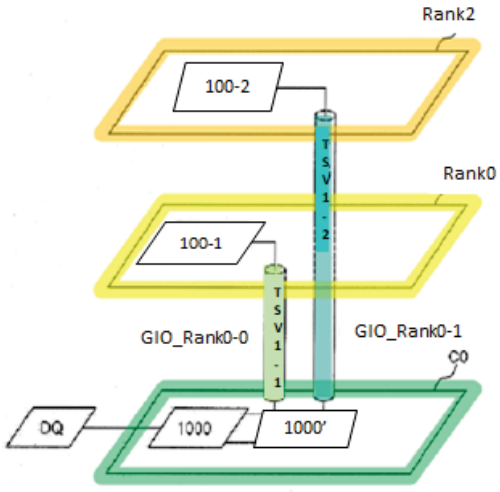
Kim's "timing difference between read and write operations would not change even if the interface circuit was modified to comply with JEDEC specification, because that compliance … is with respect to timing from the memory controller's perspective, and not with respect to when the memory circuit itself would output data in a read operation or receive data in a write operation." Brogioli, ¶147.

Given the timing difference between Kim's read and write operations, a POSITA would understand that in Kim's memory system, data collision for cross-rank read-after-write operations would indeed be "unavoidable" for Rank0 and Rank2 (or Rank1 and Rank3) that share a GIO line, but without any intervening rank selecting unit. Kim, [0042]-[0043]; Brogioli, ¶148.

As Dr. Brogioli explains, a POSITA would understand that in Kim's memory system, to prevent the "unavoidable" cross-rank read-after-write collisions, Rank2 and Rank0 (and Rank3 and Rank1) must not share a path after the shared input/output section 1000. *Cf.*, configurations for Rank0 and Rank1

depicted in Figures 2, 3 and 5; Brogioli, ¶149. But in those configurations, each rank would have its own non-shared GIO lines, and no two ranks would share a same TSV. *Id.*; *see* illustrations below. That is, modifying Kim's TSV structures to meet the requirement that at least one die interconnect be in connection with a group of array die*s* (plural) would result in inoperable apparatuses for Kim's intended purposes. Brogioli, ¶¶149-150.





1000' = rank selection between GIO0-0 and GIO0-1
The above does not show Ranks1 and 3 chips or rank selection units for GIO1-0 and GIO1-1

For these reasons, Petitioner has not shown that modifying Kim in the manner suggested in the Petition or newly at the deposition by Dr. Wolfe would have yielded the recited configuration.

B. Petitioner has not shown why a POSITA would have used non-DRAM stacks with a JEDEC-compliant DRAM interface

As noted in the claim construction section, the District Court adopted Samsung's construction that "array die" should be construed as "array die that is different from a DRAM circuit." EX2004-pp.31-32.

Despite the structural construction that Petitioner advocated and won in the District Court, the Petition contains no analysis on how or why Kim's memory chips are different from Rajan137's DRAM circuits. Dr. Wolfe testified that he "do[es]n't believe an SDRAM die is ever different than a DRAM circuit," which is why he focused on the fact that prior art was allegedly "not restricted to DRAM dies or SDRAM dies." *Id.*, 94:16-18.[7]

[7] Petitioner consistently treats Kim's memory devices as DRAMs, however. Pet. 36 (referencing textbooks reproduced on Pet. 37 and JEDEC standards); Pet. 37 (showing "DRAM" memory arrays as reproduced from EX1022-EX1023); EX1019 ("DDR3 SDRAM" JEDEC specification, cited at Pet. 36).

Q. When would a SDRAM die be different than a DRAM circuit?

A. I don't believe an SDRAM die is ever different than a DRAM circuit. I think all SDRAM dies within this claim construction would be DRAM circuits.

That's why I focused on the fact that both Rajan and Riho and Kim and Wyman are not restricted to DRAM dies or SDRAM dies.

But Petitioner relies on DRAM references' disclosure of "memory array" to argue why Kim allegedly discloses "array dies." Pet. 36-37 (citing EX1019, EX1022-EX1023, all related to DRAMs). For instance, for claims 29-34 that recite "a plurality of DRAM packages" that comprise "array dies," Dr. Wolfe maps "a plurality of DRAM packages" to "buffered stacked DRAM circuits" of Rajan. EX1003, ¶¶607, 1106 (citing EX1015, Rajan). There is no analysis, however, why the "buffer DRAM circuits" would have array dies that are different from Rajan's DRAM circuits. Pet. 77-78, 134-135. The challenge therefore fails.

Furthermore, if, in Petitioner's combination, the stacked dies are non-DRAM dies, neither the Petition nor the declaration explains why a POSITA would be motivated to have the package emulate a DRAM interface or how the resulting structure would work given Petitioner's theory that the interface circuit should provide a JEDEC-compliant interface to the host, citing to SDRAM standards. Pet. 31-33. While Petitioner asserts "emulating a standard JEDEC

interface was a known option," (Pet. 25), the associated discussion on Pet. 8-11 relates to rank multiplication where one or more DRAM dies are presented as a single DRAM die to the host. *See* Pet. 8-11; Brogioli, ¶172.

Petitioner has not presented any evidence that with non-DRAM circuits such as NANDs or SRAMs, a POSITA would have or could have emulated the memory package as JEDEC-compliant DRAM devices. Pet. 8-11, 24-33; Brogioli, ¶¶173-177. For instance, NAND devices are orders of magnitude slower than DRAMs, must erase a whole block of memory instead of being able to write individual bits as in DRAMs, and generally last on the order of 100,000 write cycles, which would be unacceptable to DRAM applications that would involve frequent data update. Brogioli, ¶¶174-175 (Dr. Brogioli explaining the many differences between NAND and SDRAMs; and opines that a POSITA would not have been able to make the combination as Petitioner now asserts); EX2040-EX2041.

As another example, SRAMs are much faster, of lower capacity, have more complex cell structure, and far more expensive than DRAMs. Brogioli, ¶¶176-177. Petitioner has not presented any evidence that a POSITA would have used the much faster and expensive SRAMs, only to have it slowed down at the buffer circuit to mimic a much slower and cheaper DRAM chip. *Id.* (explaining that a POSITA would not have done so and the many differences between SRAMs and SDRAMs); *cf.*, EX2039 (2015 post-invention blog, suggesting using cheaper and

more available DRAMs to emulate SRAMs, not the other way around; and there is compatibility issues).

Indeed, while Riho suggests that its disclosures might be applied to "SRAMs or nonvolatile memories," (EX1016, [0135]), it never suggests that the controller chip would have a JEDEC-compliant DRAM interface when the controlled dies in the stack are SRAMs or non-volatile dies. *See* EX1016 *generally*; Brogioli, ¶179. Even in the system-in-package example with mixed DRAM and NAND chips, Riho only suggests one stack comprising "DRAM chips as slave chips and a master chip" and a second stack "comprising a plurality of NAND flash memory chips as slave chips and a master chip." EX1016, [0136]; Brogioli, ¶179.

While Riho also suggests a "single master chip" that controls both the DRAM chips and NAND flash chips, that could be a processor controlling exchange of information between the DRAM and the NAND, as in non-volatile memory DIMMs described in Netlist's 918 patent. Brogioli, ¶180; EX2043 [US8,301,833], Abstract., Figs. 1-3. What Riho does not suggest is having the master present a JEDEC-compliant interface to the host while having non-DRAM chips operate behind the emulated interface. Brogioli, ¶180. Nor does Kim, Rajan, Ma or any other reference make that suggestion. *See* EX1014, EX1015, EX1039, *generally*.

As such, Petitioner has not shown that a POSITA would have been

motivated to use non-DRAM dies in combination with an interface circuit that emulates a JEDEC-compliant interface to the host. Given the materiality of the limit and the lack of any evidence in the Petition that this was done, Petitioner's obviousness theory fails. *Arendi SARL v. Apple Inc.*, 832 F.3d 1355, 1362-66 (Fed. Cir. 2016).

Netlist maintains that the structures used in Samsung, SK Hynix and Micron HBM2, 2E and/or 3 designs can satisfy the definition of array dies because of their unique differences from Rajan. Petitioner, however, presents no evidence that the primary references have designs with those or other unique differences.

C. Petitioner has not shown that the Kim combinations would render obvious "chip-selection" signals or conduits

Claim 6 requires, *inter alia*, "third die interconnects coupled between respective chip-select conduits and respective ones of the plurality of stacked array dies." EX1001, cl. 6. The Petition claims that the third die interconnects are allegedly used to "couple buffered chip-select signals CS0, CS1 from main chip C0 to slave chips C1 and C2, respective, to perform read/write operations." Pet. 59. But Dr. Wolfe claims he is not relying on a combination in which the alleged "array dies" are DRAMs. EX2025, 94:12-18. In light of this, there is no substantial evidence that alleged "array dies" have any need to receive chip select signals, because the only evidence the Petition cites on this issue relates to

SDRAMs designed to receive chip-select signals. *See* Pet. 61-62 (citing to SDRAM specifications and references to argue that "chip-select bus … is essential in a JEDEC style memory system").

For instance, the Petition does not present any evidence that memory systems other than JEDEC-compliant SDRAMs would use chip-selection signals to select chips in the stack or process chip-selection signals on a device level (such that there would be a need to send CS signals to the device). Brogioli, ¶¶169-171, 193; *cf.*, EX1015, 6:34-48 (Rajan discloses "utilizing separate chip select signals … to each of the DRAM chips in the stack" "to individually select the DRAM chips"); EX2043 (NAND datasheet does not show chip-select signals); EX2044 (SRAM datasheet does not show chip-select signals).[8]

As such, Kim's allegedly non-DRAM devices (as opposed to the rank-selecting circuit in the control die) need not receive chip-select signals under Petitioner's theory. Brogioli, ¶194. For instance, Kim's control chip can use CS0/CS1 signals transmitted on the chip-select conduits to select the appropriate GIOs for data transmission (*see* Kim, [0044]-[0045]) without needing another wire

[8] There is also no evidence that the concept of "rank"/"rank multiplication," repeatedly cited as motivation for modification, would apply to non-DRAM memory systems. *See*, *e.g.*, EX1022, 413 ("The word *rank* is now used to denote a set of ***DRAM*** devices that operate in lockstep to a given command in a memory system.").

to transmit chip-select signals to the devices themselves. Brogioli, ¶194. A POSITA would not construct a TSV to transport CS signals to memory devices that are not designed to receive or decode CS signals. Brogioli, ¶195. As Dr. Brogioli explains, TSVs that penetrate through multiple dies were difficult and expensive to form at the time of the invention. *Id.*; EX2046 (a JEDEC presentation that indicated commercially viable speed bin yield would not be viable until 2016). Moreover, one failed TSV would cause the entire die stack to be discarded. Brogioli, ¶196; EX2007 ("in case of an 8Hi-stack HBM, one failed TSV connection causes 9 dies to be discarded"). Thus, a POSITA would not have incurred the cost or the risk to build TSVs that are not needed for the memory chips' operation. Brogioli, ¶¶196-197. For this same reason, there would no motivation for a POSITA to "provid[e] chip select signals to respective array dies," as recited in claims 11, 20, and 29 and their dependent claims. *See also* Pet.71, 73 (alleging chip-select conduits obvious for "transmit[ting] chip-select signals up to the 'array dies'" (or driving CS to respective array dies), without showing that array dies needed CS for operation).

To the extent the Petitioner changes course again and argues that the combination involves SDRAM dies, Petitioner has not explained why it would be obvious for the CS signals to be sent to only one die, instead of all dies in the same rank. *See* EX1022, p.319 ("there is a chip-select network that connects from the

memory controller to every DRAM in a *rank* ….”). As noted in the claim construction section, Samsung advocated and won a claim construction that the term “chip select signal” would “exclude[] situations in which a chip select signal could enable multiple array dies at once.” EX2004 at 33-34. Such a construction would be incompatible with SDRAM ranks that each include multiple devices, because CS is sent to all of them so the devices in the same rank receive or output data concurrently in response to a read/write command. Brogioli, ¶198; Riho (CS0CK1 and CS1CK1 sent to all dies in a rank simultaneously).

D. Claim 7 (Ground 2)

Ground 2 combines Kim and Rajan with Riho. Petitioner argues that a POSITA would have combined the references because Riho is analogous art, (Pet. 81), but the general similarity does not explain why a POSITA would have combined Riho with Kim/Rajan. *See Microsoft Corp. v. Enfish, LLC*, 662 Fed. App’x 981, 990 (Fed. Cir. 2016) (generic assertion that the references were “directed to the same art or same techniques, *viz*, … they ‘discuss[ed] information search and retrieval using computer databases’ or ‘describe[d] storing data in databases, allowing users to query the data, and using indexes to facilitate those queries’ was “inadequate to show persuasively that a relevant skilled artisan would have been motivated to combine the references”).

Indeed, unlike Kim where different ranks of memories cannot share a same

TSVs, in Riho, pairs of dies from two different ranks are connected to the same TSV. *See* Kim, Fig. 5, [0042]-[0045], discussion in Section VII.A; Riho, Fig. 4 and [0028]-[0029], [0062] (D0-D7 belong one to group selected by CS0/CK0 and D8-D15 belong to a second group selected by CS1/CK1, with D_n and D_{n+8} forming a set that shares a TSV); Brogioli, ¶202. Petitioner ignores this material difference and does not explain how it proposes to modify Kim/Rajan in light of Riho, asserting only that a POSITA would have used Riho's technique to further reduce load. Pet. 82.

Petitioner also asserts that a POSITA would have combined Kim and Riho to select two dies in communication with each respective TSV "so the 'skew between data signals DQ and data strobe signals DQS/B can be minimized.'" Pet. 83. But in Riho, the skew between DQ and DQS/B is minimized by making the corresponding TSVs "substantially equal in length to each other." EX1016, [0050]; *see also* [0053] (discussing the distance of the signal lines between the controller chip and the controlled chip); Brogioli, ¶203. In Kim, the TSVs are of different lengths and Petitioner does not suggest any change to Kim's TSV lengths or the way TSV intersects (or does not intersect) different dies in the stack. *See* Kim, Fig. 5. Thus, combining Kim and Riho would not achieve the alleged objective of minimizing skew in DQ and DQS. This represents another reason why a POSITA would not have combined Kim and Riho under Petitioner's

rationale. Brogioli, ¶204.

Combining Riho and Kim would also not result in the claimed inventions. Brogioli, ¶¶205-209. The '060 patent teaches that because loads of TSVs could be significant, the number of array dies in a group might need to be adjusted to account for the TSV load, in order to achieve load balancing on different data conduits. *See*, *e.g.*, '060 patent, Table 1, 14:3-11.

Riho does not do that under Petitioner's own theory that "only one of the two SDRAMs in the pair [is] connected to the TSV at a time." Pet. 82. Under that theory (which is incorrect), regardless of the TSV load and the number of chips in a group, the load attributable to the memory chips would be the load of one chip. Brogioli, ¶206. There would thus be no need to consider the TSV loads or the number of array dies per group in load balancing because the end result would be the same (*id.*); and the Petition does not show Riho considers such factors when selecting the number of chips per set (Pet. 83). *Id.*

Further in Kim, TSV1 and TSV2 have different lengths, which Petitioner suggests elsewhere would lead to different loads of the TSVs. *See* Pet. 84-89 (Petitioner argues for using different driver sizes to allegedly account for different loads associated with TSVs of different lengths). The Petition, however, simply announces that two chips per group would satisfy 7[a], without explaining why this resulted from "consideration of" the TSV loads. Pet. 82-83. There is thus a

complete lack of pertinent analysis of element [7a] for the Kim/Rajan/Riho combination. Brogioli, ¶207.

Element 7b requires that the first/second load on the first/second data conduit includes "a load of the first/second group" of memory devices. In the context of Riho, because each "group"/set includes two SDRAM chips, to satisfy element 7.b. the load on the conduit need to include a load associated with two chips. *See* EX1001, cl. 7; Brogioli, ¶208. But under Petitioner's theory, the load would only include load of a single chip and the combination therefore does not satisfy 7b. Brogioli, ¶208.

E. Ground 3

Petitioner adds Wyman to Ground 3. Wyman provides a single driver circuit with "a set of selectable drivers each having an individual drive capability." Wyman, 1:36-37. Thus, "depending on the necessary signal strength, one need only select the corresponding tap which supplies no less than the amount necessary for the particular element of device." *Id.*, 2:66-3:1. For example, a driver circuit shown in Figure 5 with T0-T3 taps is used to drive loads associated with through-chip via connections shown in Figures 8A-8D. *Id.*, 6:15-50; Brogioli, ¶211. Each tap provides drive strength sufficient to drive a range of load, *e.g.*, T1 tap could drive loads between those shown in Figures 8B and 8C. *Id.*, 6:38-43; Brogioli, ¶211.

Given Wyman's disclosures, a POSITA would not have even used different drivers to drive TSV1 and TSV2 as asserted. *Contra* Pet. 85. First, because Kim's TSV1 and TSV2 are used at different times, Wyman's use of a single driver circuitry to drive different connections would suggest using the allegedly more "efficien[t]" Wyman-style driver circuitry for both TSVs, instead of a separate driver for each TSV. Brogioli, ¶¶213-214; *see also*, 6:44-50 (using T3 tap to drive fan-out connections to two daughter chips).

Second, Wyman's driver circuitry, once fabricated, would be regarded as a single driver of a given size regardless of the number of driver components used. EX2024, 141:18-145:14, 145:24-146:7, 149:24-150:8; Brogioli, ¶215. Thus, even if different taps of a Wyman-style driver are used to drive TSV1 and TSV2, that would still involve the use of a driver of the same size. Brogioli, ¶215.

Third, Wyman and Kim do not even suggest that the two TSVs in Kim with different lengths would result in sufficiently different load to warrant using different taps. Brogioli, ¶216. For instance, load contribution from TSV "may be negligible compared to load contribution from the array dies." *See* EX1001, 4:36-38; Brogioli, ¶216. If so, the same tap may be used. Brogioli, ¶¶213, 217.

Thus, incorporating Wyman's teaching would result in an implementation that uses a single Wyman-style driver, and not first and second drivers or drivers of different sizes. Brogioli, ¶214.

VIII. Riho Grounds Are Not Obvious

A. Riho, alone or in combination, does not disclose or render obvious elements 1.d.1 and 1.d.2

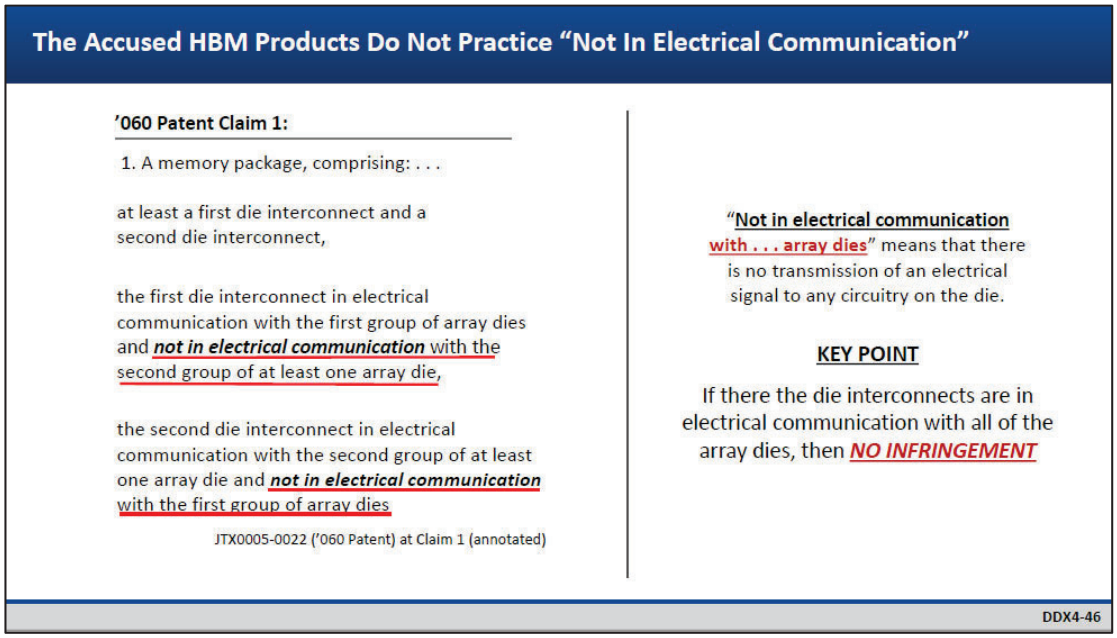
In this proceeding, Petitioner argues that Riho's TSVs, which are each connected to a switch circuit on every SDRAM in the stack, are nevertheless in electrical communication with only some of the SDRAMs. Pet. 100-101; *see* Riho, Fig. 5 (representative TSV715 coupled to switch circuits that connect to driver circuits); Fig. 6 (each SDRAM die coupled to each TSV). As explained below, this position is inconsistent with the positions taken by all parties in the district court proceedings. That is, under either side's application of the term "(not) in electrical communication with … array dies," Riho does not disclose or suggests elements 1.d.1-1.d.2.

1. Petitioner's position in this proceeding is inconsistent with its infringement position while Netlist has applied a consistent interpretation consistent with the specification

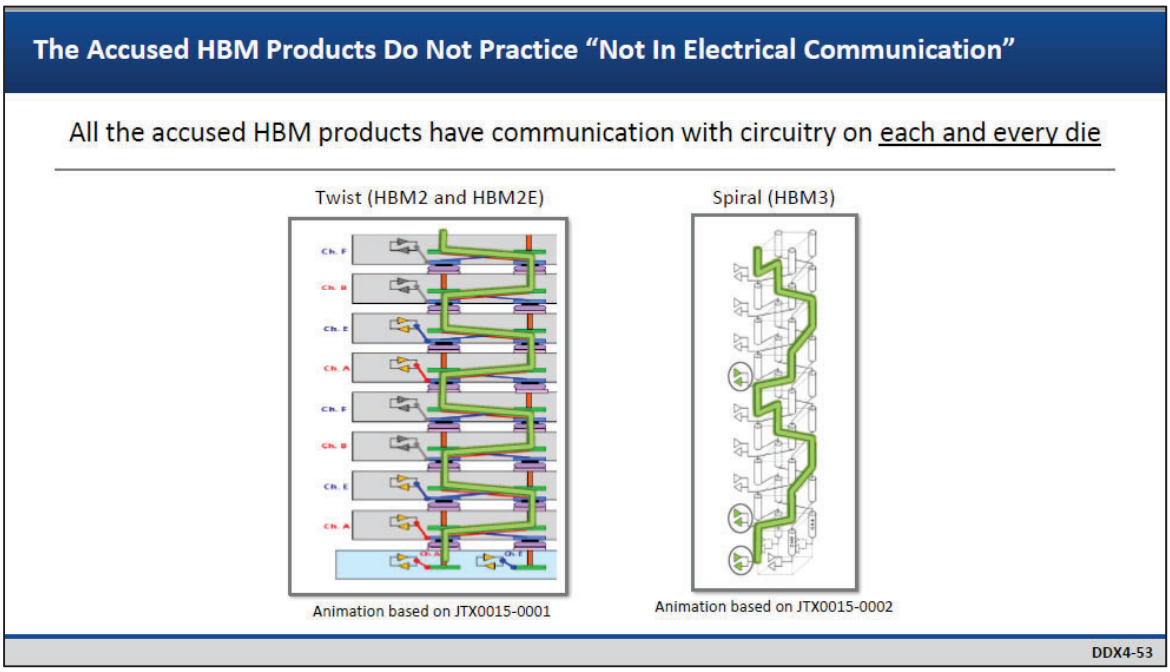
In the district court action, Samsung's expert for the '060 patent, Dr. Robins, testified that

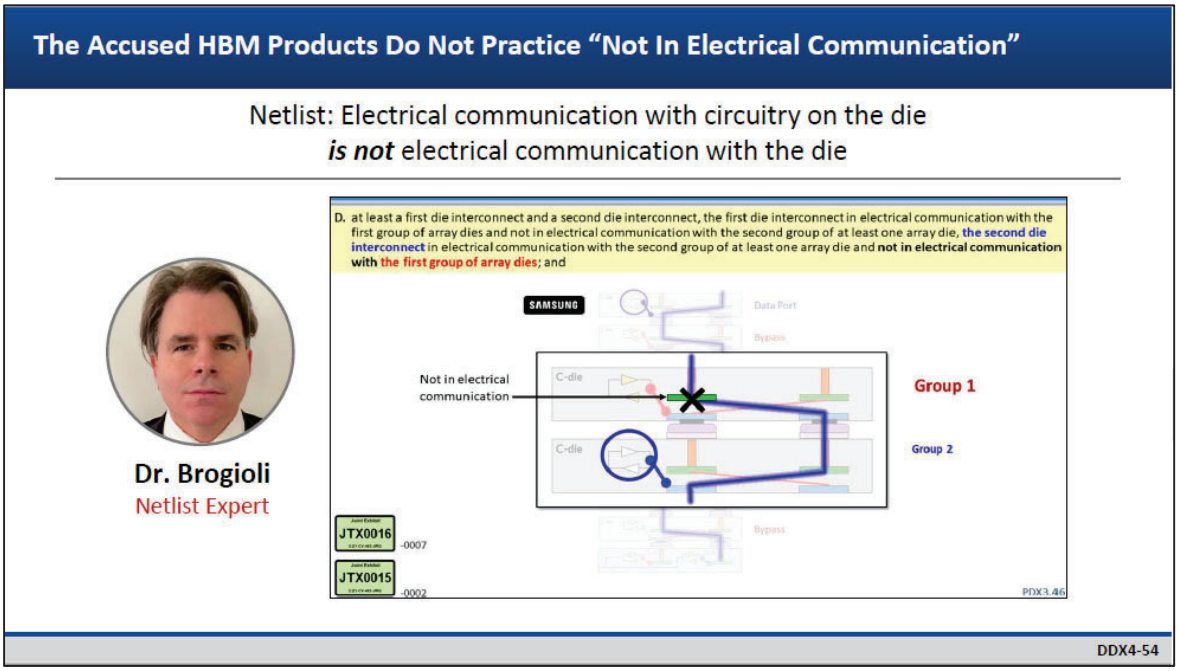
> the notion of electrical communication is a broad concept involving the transmission of electricity. So not in electrical communication with array dies means that there is no transmission of an electrical signal to any circuitry on the die.

EX2029, 1057:3-7; EX2030, DDX4-46. Accordingly, testified Dr. Robins, "if the die interconnects connect to all the array dies, then there can't be infringement and there's no infringement." EX2029, 1057:8-12.



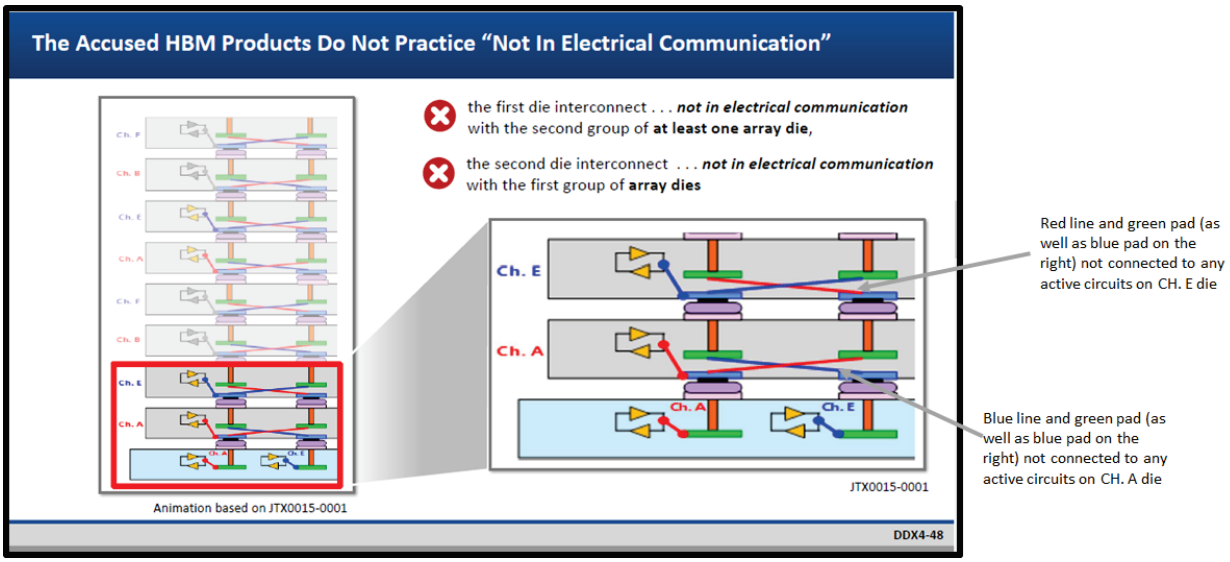
In the accused Samsung products, the die interconnects are connected to data ports on some of the array dies, but not on the others, as depicted below. EX2030, DDX4-53, DDX4-54; Brogioli, ¶236.





Samsung argued that even when the die interconnect was coupled only to dummy pads on an array die, it was nevertheless in electrical communication with that die because the die interconnect was "connected to various circuitry on the die,

including these horizontal paths … in blue and red, and also including the pads … in blue rectangle and green rectangle and so forth." EX2029, 1058:7-1059:14; 1063:21-1065:5. Samsung does not dispute that the blue and red lines or the pads in its products are not connected to any gates, ports or other active circuits in the particular array die. Brogioli, ¶237.



Micron similarly views "electrical communication" as encompassing any electrical wave exchanged between a die interconnect and an array die. For example, Micron's expert, Dr. Stone, testified that the term "electrical communication" simply means "electrical connection" to him. EX2024, 95:10-13. According to Dr. Stone, whether the die interconnect is connected to a stub (no active circuits) or a port array (active circuits), "wave propagates on the die along the stub or the port array," and "those waves are going back and forth and back and forth" to cause noises. *Id.*, 94:8-95:4. Thus, according to Dr. Stone, there would

be electrical communication even if the die interconnect is only connected with a stub,[9] "because the wave travels on that." *Id.*, 94:8-95:9; *see also id.*, 96:11-97:8.

The specification, of course, differentiates between electrical communication and electrical connection. For instance, it teaches that data ports on array dies "enable electrical communication and data transfer between the corresponding memory circuitry of the array dies … and a communication pathway (e.g., a die interconnect)." EX1001, 5:41-45. On the other hand, in some cases where there is no electrical communication between a die interconnect and certain array dies, "electrical connection leading from the TSV of the array dies that are not configured to be in electrical communication with the die interconnect … may be stubs." *Id.*, 8:57-60. The "stubs are not configured to provide electrical communication with the memory cells of the array die," even though there may be electrical connection between the TSV and the stubs. *Id.*, 8:60-62. Hence, it is incorrect that any electrical connection between a die interconnect and a portion of an array die would lead to electrical communication. Brogioli, ¶232.

While Netlist disagrees with Petitioner's overbroad interpretation of

[9] Dr. Wolfe states that "[a] stub wire would be a little piece of wire that's not connected to anything or not functionally connected." EX2025, 123:24-124:2. Dr. Wolfe also states that a POSITA would not view a TSV to be in electrical communication with a stub "because there's nothing being communicated with." *Id.*, 181:11-17.

"electrical communication," it agrees with Samsung's and Micron's district court experts that connection between a die interconnect and an active circuitry (such as a port) on the die would result in electrical communication even if the active circuitry is disabled. EX2024, 109:1-113:22; Brogioli, ¶¶233-234, 240. For instance, in reference to Figure 1A of the 060 patent, Dr. Stone testifies, *inter alia*:

> In other words, let the top layer have a driver that sends a wave down the die interconnect.
>
> Let the second layer be on a die interconnect in which the driver is disabled. That second layer still has capacitance between the die interconnect and the driver on that array die. And therefore, you have to charge or discharge it. And therefore, it's in electrical communication.

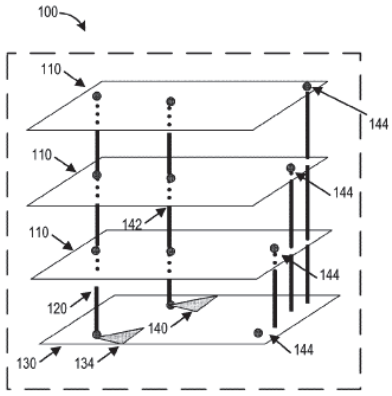
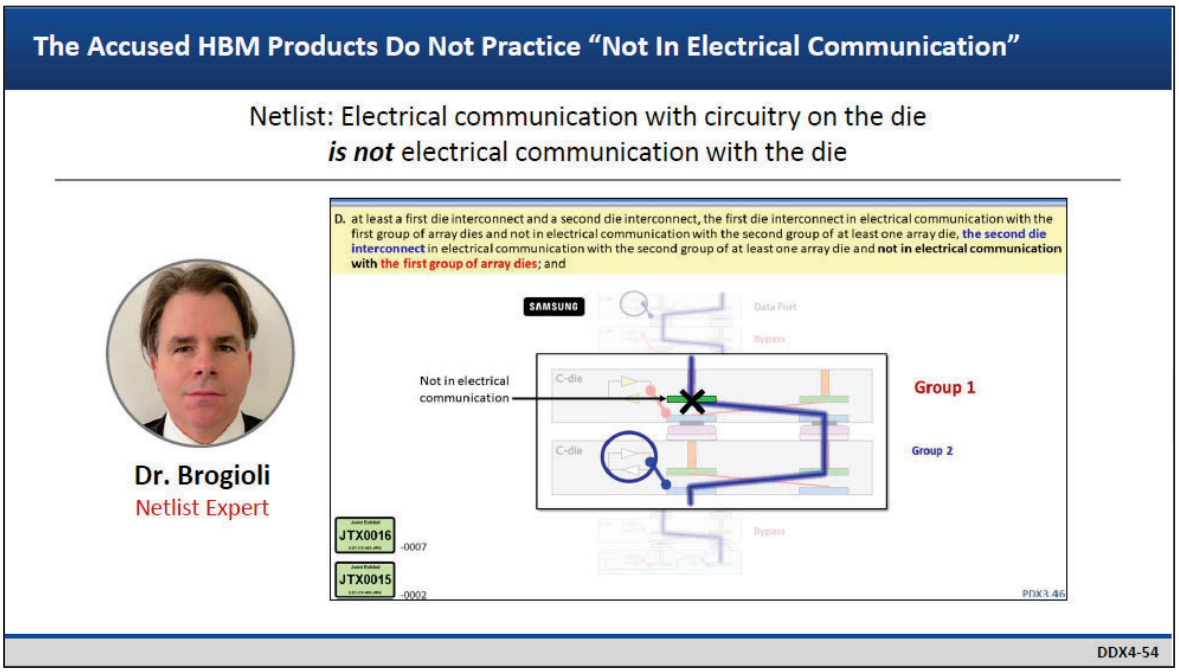


FIG. 1A

EX2024, 112:11-19.

This is consistent with the specification's disclosure that when a die interconnect is connected to a data port on an array die, there is electrical communication between the die interconnect and the array die. *See* '060 patent, Abstract, 8:22-32, 8:37-47, 2:19-31, 2:38-49, 5:41-45, 5:54-6:18, 6:27-35, 7:11-26, 8:3-8, 10:22-34, 11:12-16, 11:31-36, 12:45-51, 15:9-13, 17:29-36; Brogioli, ¶¶231, 241.

Accordingly, Netlist's expert, Dr. Brogioli explained that a die interconnect in Samsung's products was in electrical communication with an array die if it was connected to a data port (*e.g.*, receivers/drivers) on that die, even if the data port was disabled or turned off. *See* EX2030, DDX4-54; Brogioli, ¶¶242-243. Dr. Brogioli also explained that a die interconnect was not in communication with an array die if the TSV was connected to only dummy pads on the die, because there was nothing on the other side of the pad in the particular array die for the die interconnect to communicate with. *Id.*



For instance, in Samsung's demonstrative DDX4-54 that reproduced a demonstrative by Dr. Brogioli, the blue die interconnect is in electrical communication with a "Group 2" die because it is connected to a data port on that

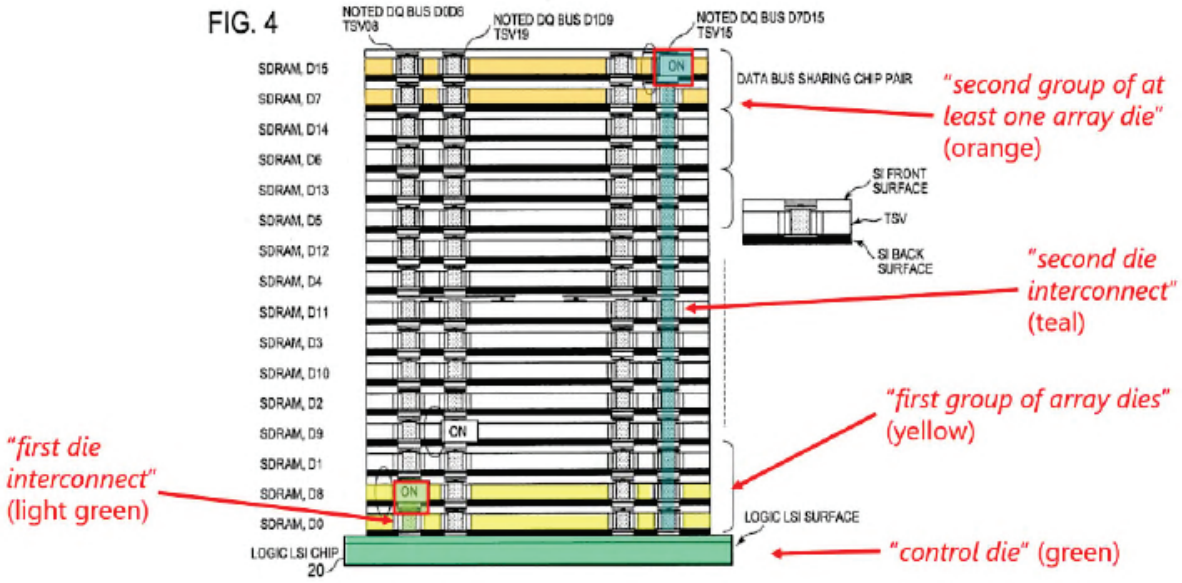
die (annotated with a blue circle) even though that port is turned off. Brogioli, ¶242; DDX4-54. On the other hand, the blue die interconnect is not in electrical communication with a "Group 1" die because it intercepts the die on a (green) pad having no data ports attached to it. *Id*.

This position is consistent with the objective of the '060 patent's inventions, which is to reduce the load that a driver in a corresponding data conduit must overcome. EX2024, 112: 6-19; Brogioli, ¶244. Dr. Borgioli opines that where a TSV is connected to a data port on a die, there would be a corresponding increase in load even when the data port is switched off, compared to when the TSV is not connected to data port on the array die. Brogioli, ¶244; EX2024, 112:11-19. Applying that understanding, Riho does not disclose or suggest a memory package with the recited characteristics of the '060 patent.

2. Riho's TSVs are each in communication with all dies in its stack

Riho's structure is as described in Section IV.B. As Dr. Brogioli explains, Riho's TSV structures adopt a multi-drop configuration in which each of the TSVs is electrically coupled to the switch circuits on every die, just like in the prior art mentioned by the '060 patent. *See* Brogioli, ¶245; Riho, Figs. 4-6; *see also* EX1018, Fig. 1 & abstract. For instance, as shown in Figure 4, each of the TSVs penetrates through each of the SDRAMs. *See*, *e.g.*, *id.*, Fig. 4 (annotation from

Pet. 101), [0045].



Per Figure 5, all sixteen (16) SDRAMs have the same equivalent circuits. EX1016, [0076], Fig. 5; *see also* [0080] (plan view shown in Figure 6 is the same for all SDRAM dies in Riho's stack). On each die, there are two switch circuits, sa and sb, for each data TSV. EX2025, 10:2-10, 14:7-13. The switch circuits are shown as having wires to a corresponding MOS driver circuit modelled as L1C2R2 circuit, which Dr. Brogioli explains would be understood as connected to memory arrays. Brogioli, ¶246. Dr. Brogioli also explains that Riho's switch circuits sa/sb would be understood as a data port or a portion thereof that can open or close. Brogioli, ¶247; *see also* EX2026, 169:19-22 ("port" "basically means a door").

FIG. 5

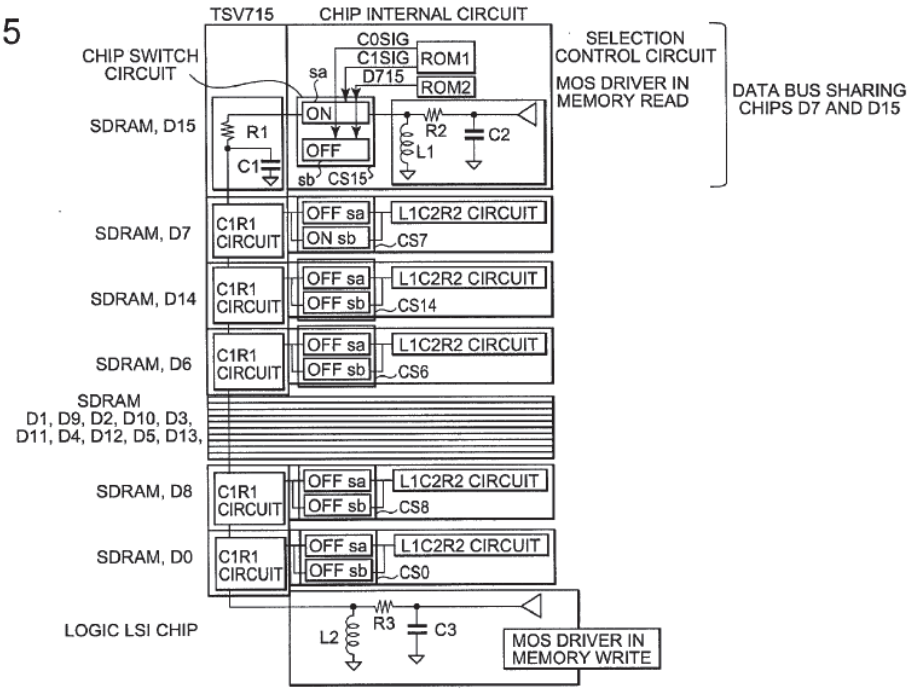


FIG. 6C

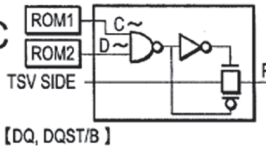
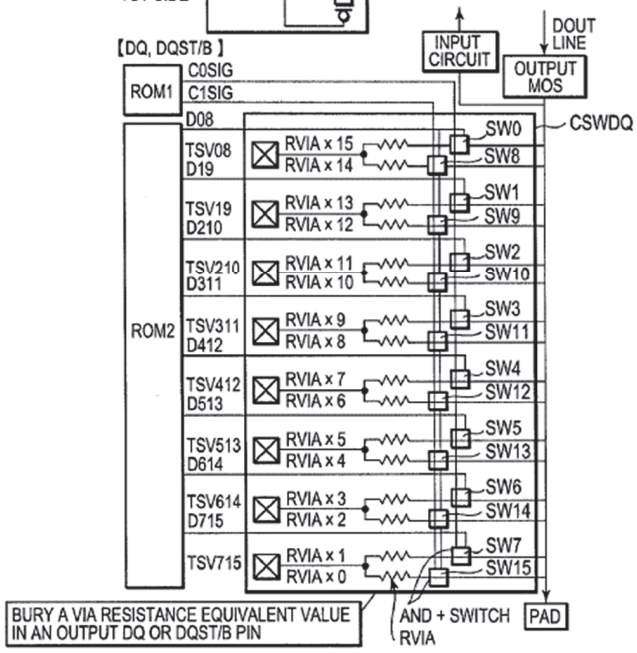


FIG. 6A



Dr. Brogioli explains that because each TSV in Riho's structure has a connection to a switch circuit that connects to a driving circuit and memory arrays in every die, each TSV would be regarded as in electrical communication with each and every die in the stack. Brogioli, ¶248; EX1016, [0086] ("An output MOS transistor and an input circuit shown in FIG. 6A are so-called I/O circuits … and, in the case of DRAM DQ, a memory cell is connected through the output MOS transistor and the input circuit"). There is thus no second group of at least one array die that is not in electrical communication with a first die interconnect (e.g., TSV715) or a first group of array dies that are not in electrical communication with a second die interconnect (e.g., TSV08). Brogioli, ¶248.

This conclusion is consistent with the view taken by both Drs. Robin and Stone who agree that, when a die interconnect can transmit a signal to a transistor or groups of transistors on an array die, there would be electrical communication. Brogioli, ¶249; EX2024, 109:1-113:22; Trial 1057:3-7; Robins DDX4-46; *see* discussion in Section VIII.A.1.

The conclusion would also be consistent with Dr. Wolfe's interpretation of Riho, which he alleges implements switch circuits sa/sb as transmission gates that couple or isolate the corresponding MOS driver circuit and the TSV. EX2025, 46:23-47:1; Brogioli, ¶¶250-251. EX1030, 138 depicts a transmission gate:

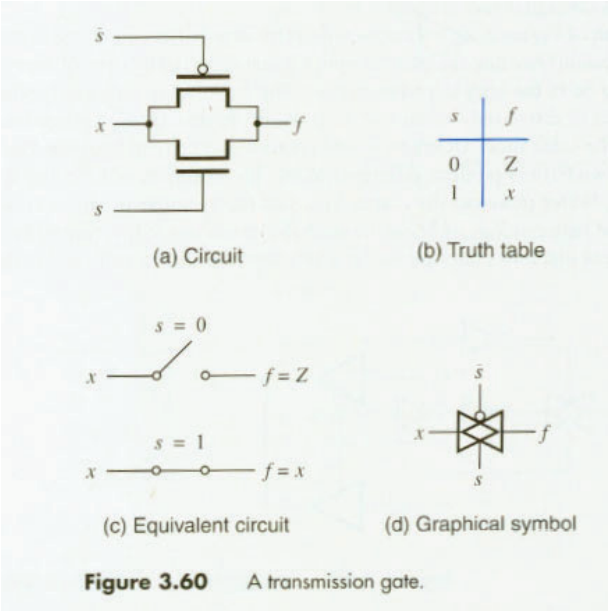
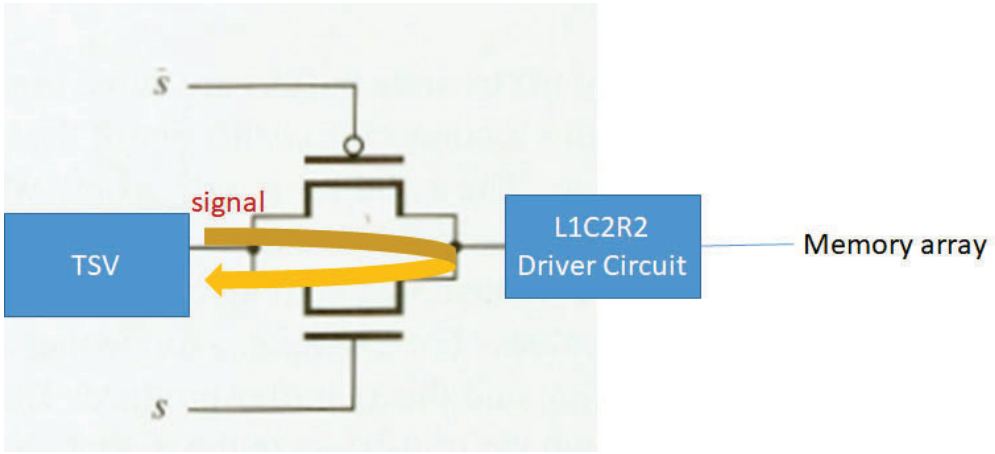


Figure 3.60 A transmission gate.

Dr. Brogioli explains that in Riho, the TSV would be connected to one side of the transmission gate and the driver circuit to the other side. Brogioli, ¶251. Even when the transmission gate is off, there would still be signals between the TSV and the input stage of the gate, *i.e.*, that there is electrical communication between the TSV and transistors on the array die. *Id.*; *see* illustration below. As such, there is electrical communication between the TSV and the array die. *Id.*



This would be consistent with the '060 patent disclosure where, for example, in Figs. 1A-1B, only one of the dies sharing a TSV would output data at a given time to avoid collision, with data ports of the other dies disabled. Brogioli, ¶¶35, 161. Nevertheless, every die is regarded as in electrical communication with the TSV. *Id*.

3. Dr. Wolfe's new theory at the deposition is incorrect

In his declaration, Dr. Wolfe asserts that "at least under Netlist's apparent interpretation, Riho's data signal DQ TSV08 is '*in electrical communication with the first group of array dies*' including SDRAMs D0 and D8, which are set into a conductive 'on state,' and that DQ TSV08 is '*not in electrical communication*' with any other SDRAMs, including SDRAMs D7 and D15 in the '*second group of at least one array die*,' which are in a non-conductive 'off state.'" EX1003, ¶732. Dr. Wolfe asserts that for the same reason for TSV08, "data signal DQ TSV715 … is in electrical communication with SDRAMs D7 and D15, but not with any other SDRAMs in the stack." *Id.*, ¶737. As explained above, that is not how Netlist applies the term.

At his deposition, Dr. Wolfe asserts that TSV715 is not in electrical communication with SDRAMs D0-D6 and D8-D14 because the switch circuits to those dies, when turned off, would isolate the load of driver from the TSV. EX2025, 9:11-14, 16:19-25, 17:8-18:2, 48:12-49:14 (Dr. Wolfe agrees that a TSV

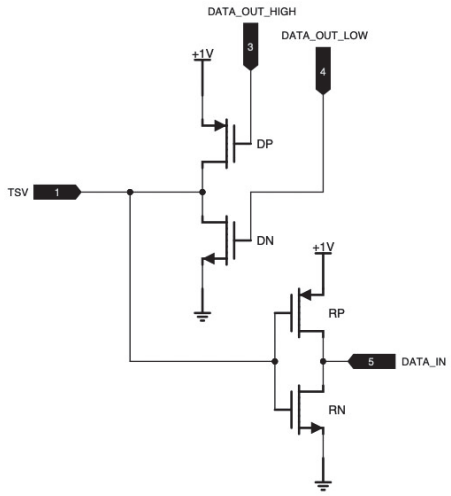
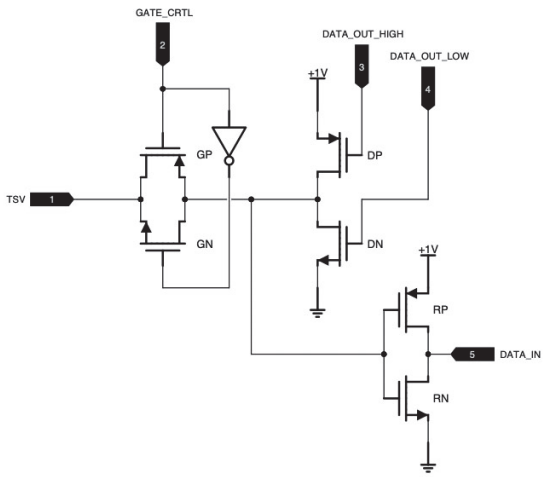
will experience some capacitive load from switch circuits sa and sb even if the switch circuits are turned off, but believes that the added load will be insignificant). While he accepts that a TSV would be in electrical communication with a driver circuit even if the driver circuit is off, he insists that there would be no electrical communication between a TSV and an array die if the switch circuits sa/sb (which he refers to as transmission gates) are turned off. *Id.*, 215:17-218:10, 224:8-16. Differentiating between transistors in a transmission gate and those in a driver circuit, Dr. Wolfe stated:

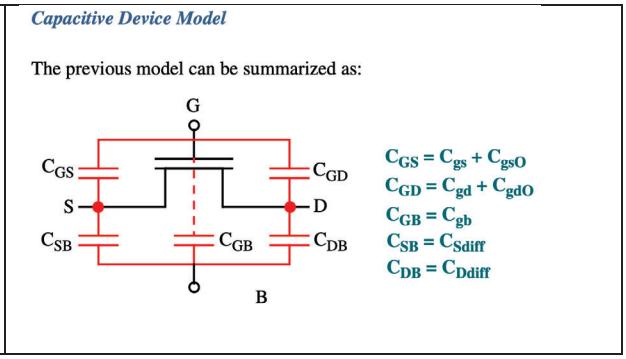
> There's a difference between the configuration of transistors in a transmission gate which is not processing the inputs and producing an output, but simply connecting the two sides or not connecting the two sides; and the input stage of a write driver which is designed to receive an input signal and then through logic produce an output signal rather than simply pass electricity through it.
>
> Those are two different things and a person of ordinary skill in the art would treat them differently. It would treat one as receiving a signal and another one as a pass-through switch.

EX2025, 219:8-20; *but see id.*, 107:12-21 (in Ruckerbauer, Fig. 5A, significant capacitive from buffer 10a likely even if the buffer is turned off) & 220:8-221:5 (allegedly no electrical communication between TSV 112a and buffer 10a if buffer 10a is turned off even though buffer 10a is designed to receive and process signals). But the claims are not to be interpreted by incorporating a subjective

intent of use. Dr. Wolfe also asserts that "the gates in the receiving circuit represent the bulk of the capacitance." EX2025, 154:3-13. None of these assertions appear anywhere in the Petition or in the referenced declaration. Pet. 100-102; EX1003, ¶¶727-744. The Board should decline to consider the new argument.

But even if considered, Dr. Wolfe's new theory is incorrect. Dr. Brogioli has provided an analysis that compares the capacitate load of on-chip drivers with and without a transmission gate, using the following circuit layout (Brogioli, ¶¶255-257):

without transmission gate	with transmission gate
	

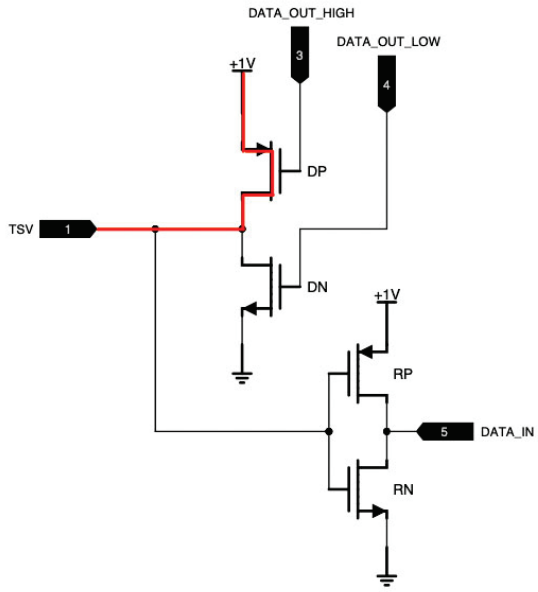
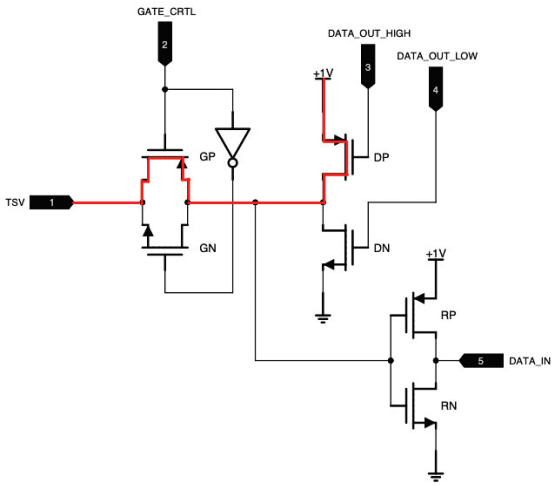


Dr. Brogioli first notes that contrary to Dr. Wolfe's assertion, "[t]he data receiver can use relatively small transistors because it only needs to be able to drive the capacitance of on-die logic" while "the output transistors need to drive the entire TSV." Brogioli, ¶258. Consequently, "a TSV driver [in an array die] must use 'huge' transistors compared to the input stage of the receiver" and "these 'huge' transistors [in the driver circuit] present significant input capacitance" on the TSV than the output capacitance associated with the input stage of a receiver circuit. *Id.*; EX2025, 24:7-13 (capacitive load on a bus is "predominantly the capacitance of the physical interconnect, the capacitance of any inputs connected to that bus, and the capacitance of the output stages of any disabled drivers that are currently connected to that bus"). Thus, to drive respective loads at the same rate, the receiver may be orders of magnitude smaller than the driver circuit, and correspondingly lower in capacitive contribution. Brogioli, ¶¶259-262.

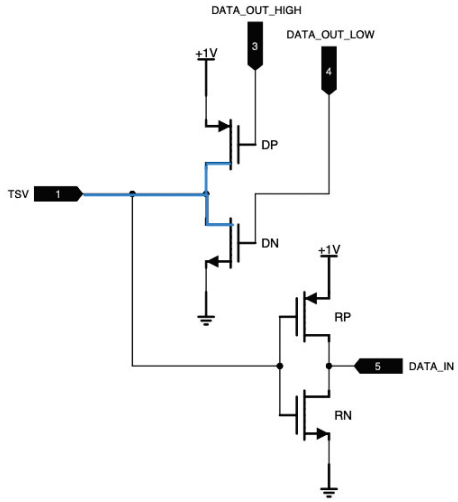
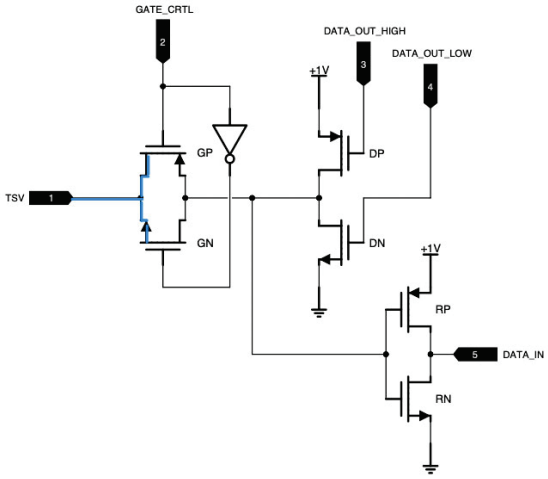
Second, in Riho, Figure 5, L1, R2, and C2 represent the internal parasitic loads of the components on die. EX1016, [0077]. On a typical design such as a static CMOS system, the capacitance of the transistors is the dominant form of the

parasitic capacitance on a chip. Brogioli, ¶264. The parasitic inductance and resistance will be dominated by the TSV interconnects and are relatively constant between the two designs shown on page 70. *Id.*

In the design without a transmission gate (left below), the driver drives the TSV through a single transistor. In the design with a transmission gate (below right), the driver drives the TSV through two transistors. To have equal drive strength in both designs, the transistors in the driver in the design on the right (with transmission gate) must be twice as large compared to the transistors in the design on the left (without a transmission gate). Brogioli, ¶265.

without transmission gate	with transmission gate
	

In both designs, the TSV sees the diffusion capacitance of two transistors on a disabled chip. Brogioli, ¶266. However, because the design that uses a transmission gate needs double-sized transistors, this gate will present double the capacitance. *Id.*

without transmission gate	with transmission gate
	

Therefore, a POSITA would recognize that adding a transmission gate in a practical system that maintains constant drive strength will materially increase the load presented to the TSV driver in the control die, not decrease it. Brogioli, ¶267. Accordingly, in Riho, each TSV is electrical communication with all dies. *Id.*

4. Rajan would not cure Riho's deficiency

Petitioner also asserts that Riho in light of Rajan renders obvious 1.d.1-1.d.2. Pet. 101. It is unclear what modified structure Petitioner proposes. Pet. 101. Dr. Wolfe's declaration that is improperly incorporated appears to suggest the same

structure as what Riho already discloses. EX1003, ¶739. As such, modified Riho still would not disclose the limitation as explained above in Sections VIII.A.2-VIII.A.3.

Petitioner's reasons to combine also fall short. Petitioner first argues that Rajan and Riho are allegedly analogous art, which speak to only their general similarity but not the specific reason to combine as required by the law. Petitioner has also suggested that a POSITA would look to Rajan for alleged details about the interface circuit and the external terminals to the host. Pet. 93. But that does not explain why designs related to interfacing with a host would have motivated a POSITA to modify the DRAM stack organization within Riho (instead of interacting with an external host). Brogioli, ¶270. Those are two different aspects of the design, as explained in the Kim grounds. *Id.*

Petitioner then argues that Rajan's supposed teaching of load reduction would have led to the combination. But in Rajan, the load refers to the use of the interface circuit to present a single load of the buffer chip to the host. Rajan, 5:63-6:2 ("… the host system device can be in electrical contact only with the load of the buffer chip and thereby be isolated from whatever physical memory circuits, data paths, busses, etc. exist on the (logical) other side of the buffer chip."); Brogioli, ¶271.

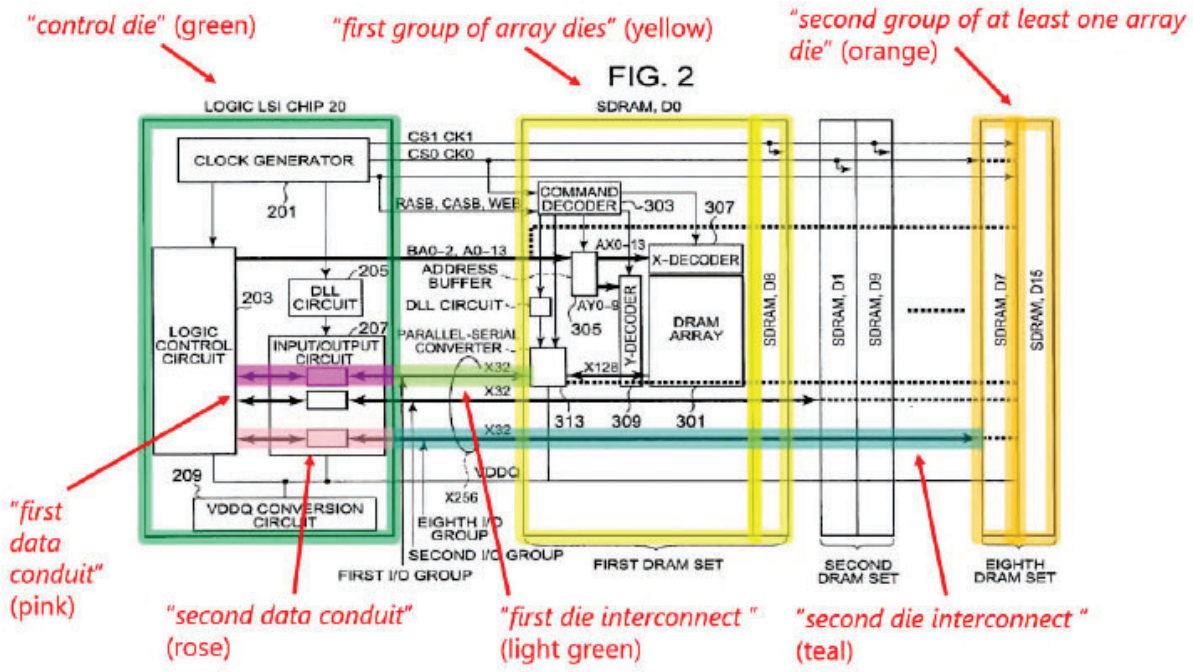
As Dr. Brogioli explains, this has nothing to do with the load on the data

conduits that is the concern in the '060 patent because the load in the '060 patent relates to the load internal in the memory package for data transmission between the control die and the array die(s). EX1001, 2:8-15; Brogioli, ¶¶271-272. Nor would Rajan's completely different load consideration leads to 1.d.1-1.d2 on the organization of the die interconnects in relation to the array dies. Brogioli, ¶¶271-272; EX2036, 7 & EX2046, 1 (even in 3DS DRAM packages having multi-drop configuration, the memory controller would only experience the load attributable to the master die). Hence, even if a POSITA would have implemented Rajan's interface circuit in Riho, that would not have yielded the 060 inventions. *Id*.

B. Riho, alone or in combination, does not disclose or render obvious elements 1.e.2-1.e.3

The claim require "a first data conduit between [a] first die interconnect and a first terminal …, and … a second data conduit between the second die interconnect and the first terminal …." Pet. 103. The plain language therefore requires that the first and second die interconnects be coupled to the same "first terminal." Petitioner maps the first die interconnect as TSV08 and the second die interconnect as TSV715. However, there is no evidence that TSV08 and TSV715 are connected to the same "first terminal" as required. That is, TSV08 is connected to data terminal DQ08 and TSV715 is connected to data terminal DQ715 given that all the bits are to be output in parallel. Brogioli, ¶273.

In particular, as Petitioner's own annotation shows, the outputs of TSV08 and TSV715 are in parallel. Pet. 104; *see also* Riho, [0033] ("Each SDRAM has an 8-bank structure and outputs a 32-bit data signal in parallel."); Brogioli, ¶274. While Riho has a parallel-to-serial converter circuit 313, the function of that circuit is to "send[] and receive[] x128-bit parallel data signals between itself and the DRAM array 301 and [to] further send[] and receive[] 32-bit (x32) parallel data signals between itself and the logic LSI chip 20." Riho, [0049]. "That is, the parallel-serial conversion circuit 313 has a function to convert a x128-bit data signal into a x32-bit data signals and to convert x32-bit data signals into a x128 data signal." *Id.* But there is no disclosure of further parallel-to-serial conversion in LSI chip 20 or multiplexing of *e.g.*, data from TSV715 and TSV08 so that those data would be output to the same "first [data] terminal." Brogioli, ¶274. What Riho does not disclose or suggest is that TSV08 and TSV715 are connected to the same sets of terminal(s). *See* Fig. 2 below as annotated by Petitioner (data from each TSV are output in parallel and not further serialized); Brogioli, ¶274.



To the extent that Petitioner suggests that in Rajan, Figure 4, the outputs from the two internal data busses are serialized and output through a single external data terminal (*see* Pet. 105), that is incorrect. Brogioli, ¶¶275-277. As Dr. Brogioli explains, a POSITA would understand the illustration in Rajan, Figure 4 as referring to that there are separate data, clock, control and address busses on the host-controller-chip interface, not that the data from the two internal data busses would then be merged. *Id*.

Rajan:

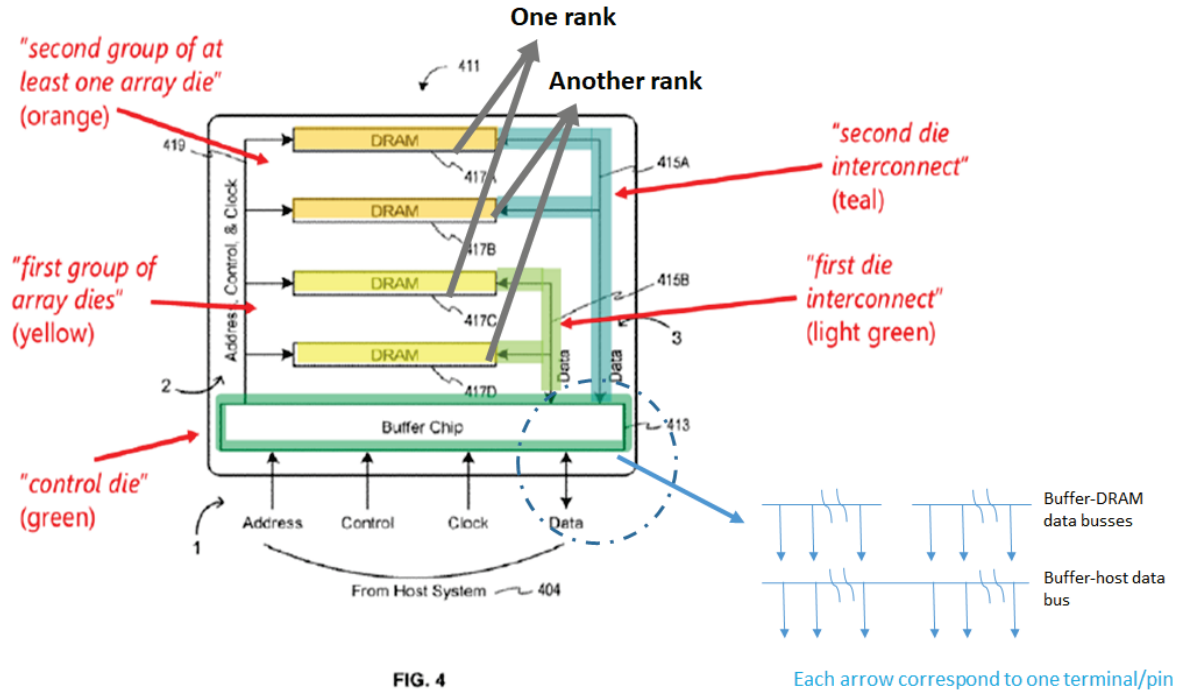


FIG. 4

That is, for example, if 417A and 417B share a data bus with internal pins X_0-X_n, and 417C and 417C share a data bus with internal pins X_{n+1} to X_{2n+}, the host-interface side would have the corresponding pins Y_0 through Y_{2n+1}. This is particularly the case given Dr. Wolfe's theory that 417A and 417C (and 417B and 417D) connecting to the two different internal data busses belong to the same rank. EX2025, 91:23-92:14. In such an arrangement, the two data busses would each transmit (or receive) a part of a data word and the supposedly JEDEC-compliant host-interface is required to transmit (or receive) all bits of the data word together in parallel, on different data pins. Brogioli, ¶276; *see* EX1022, 319 (bits from DRAMs in the same rank output on different terminals).

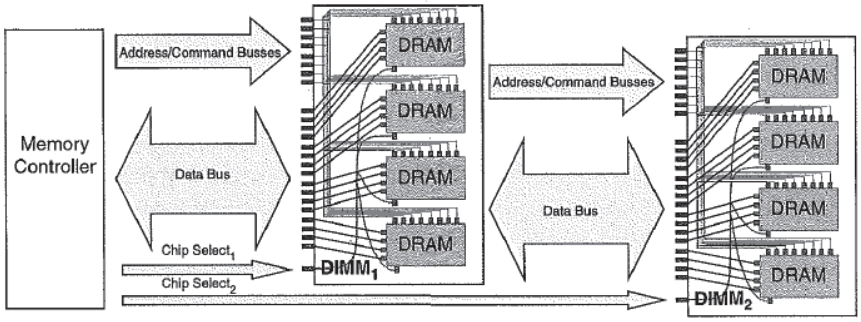


FIGURE 7.6: JEDEC-style memory bus organization. The figure shows a system of a memory controller and two memory modules with a 16-bit data bus and an 8-bit address and command bus.

Moreover, like Riho, nothing in Rajan suggests that data bits from the two different internal data busses (which are from two different dies) would be output from the same terminal. Brogioli, ¶277. Nor does Rajan suggest that data bits from TSV08 and TSV715 of Riho should be output on the same terminal. *Id.* There is thus a complete absence of analysis that "parallel" TSVs (wires) in Riho outputting parallel data would be connected to the same terminal, and the Petition therefore fails for this independent reason.

C. Petitioner has not shown why a POSITA would have used non-DRAM stacks with a JEDEC-compliant DRAM interface

As with Ground 1, Petitioner now asserts that for the "array die" element, it is relying on the fact that Riho is not restricted to DRAM dies or SDRAM dies. EX2025, 94:10-18. Of course, the Petition does not make that argument. *See* Pet. 91-108. Thus, the Board should not consider the belated argument. But even if considered, for the same reason as stated for the Kim grounds, the argument fails.

Brogioli, ¶¶172-180, 29-32.

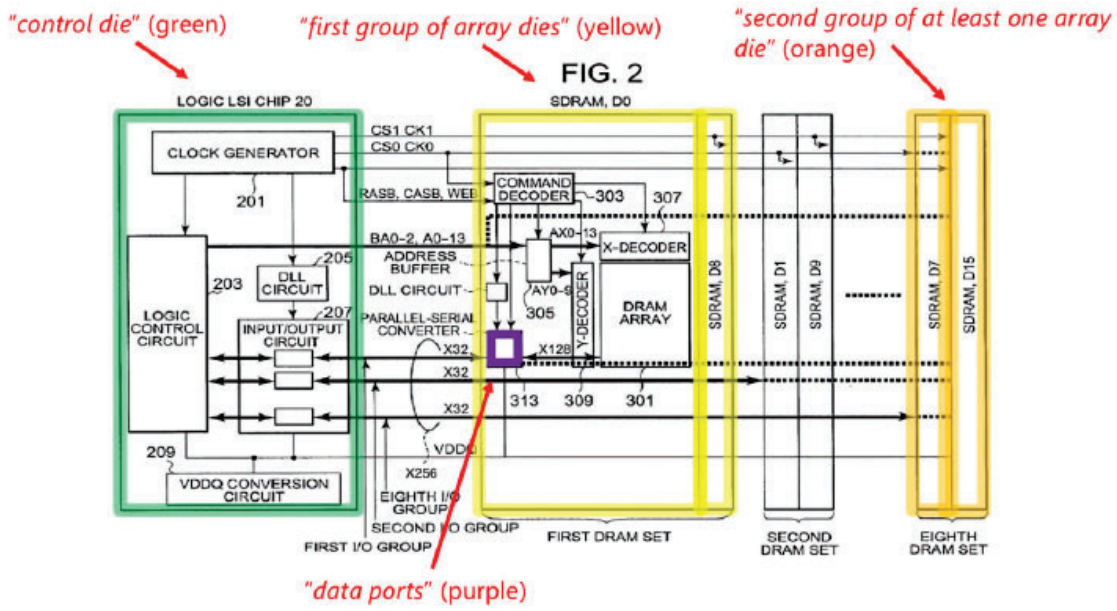
In the institution decision, the Board also suggested that Riho's SDRAMs differ from Rajan's because, for example, they have different capacity. ID at 16-17. Neither side is taking such a position because capacity was not the structural distinction that Netlist tried to draw between Rajan and the '060 inventions. EX1002, 465-66. Indeed, Dr. Wolfe stated categorically, "I don't believe an SDRAM die is ever different than a DRAM circuit." EX2025, 94:12-13. Hence, there is no evidence that "differen[ce]" can be satisfied by mere capacity differences.

D. Riho's chip select signal does not satisfy petitioner's proposed construction (claims 6, 11, 20, 29 and dependent claims)

For Ground 4 (as well as Ground 5), Petitioner again relies on JEDEC CS (or chip select) signals for the claimed "chip select signals." Pet. 117-118, 129-30 (citing to EX1019, a JEDEC specification on DDR3 SDRAMs). Petitioner suggests that, according to EX1019, each of a Quad-stack/Quad-die DDR3 Stack is associated with a separate CS signal. Pet. 117-118. But Petitioner's expert also testified that he was not relying on a combination that used SDRAM/DRAM chips, only non-DRAM ones. EX2025, 94:10-18. If that is the case, as with the Kim grounds (the discussion for which is incorporated), Petitioner has not shown that chip-select signals need be or would be sent to the respective memory dies because

there is no evidence that non-DRAM devices needed those signals for operation and it would increase manufacturing costs and failure risks by including unnecessary chip-select TSVs.

Moreover, Riho's express disclosure is at odds with Petitioner's advocated interpretation for "chip-select" signals as Riho's CS signals are sent to more than one die. Brogioli, ¶¶282-285.



For instance, CS0CK0 is sent to SDRAMs D0-D7 and CS1CK1 is sent to D8-D15. Riho [0037] (in Figure 2, the "clock generator 201 supplies a first clock signal CS0CK0 to the SDRAMs D0 to D7 … forming a first group (chip selection group) and further supplies a second clock signal CS1CK1 to the SDRAMs D8 to D15 … forming a second group (chip selection group)"); Pet. 128 (mapping chip select signals to CS0CK0 and CS1CK1); Brogioli, ¶283.

This is consistent with Petitioner's other evidence that teaches SDRAMs' chip-select signals are sent to all devices in a rank. EX1022 at 320 ("The chip select signal passes over a wire unique to each small set of DRAMs and enables or disables the DRAMs in that rank so that they, respectively, either handle the request currently on the bus or ignore the request currently on the bus."); EX2025, 35:3-6 ("In the Figure 5 embodiment of Rhio, if you issue a read operation, all of the DRAM die in a single rank will send data to the logic LSI chip simultaneously."); 36:2-7 (Dr. Wolfe acknowledging that it was his understanding that a rank is a collection of DRAMs that respond to a common memory command simultaneously"); Brogioli, ¶284.

Thus, the chip select signal that Petitioner relies on is sent to more than one die at a time, a situation that Petitioner successfully persuaded the District Court to exclude from the "plain and ordinary meaning" of "chip select signals." *See* EX2004 at 33-34; Brogioli, ¶285. Hence, Petitioner has not shown that CS0CK0 and CS1CK1 satisfy Petitioner's proposed construction of the term "chip select signal."

E. Claim 7

Petitioner's analysis for claim 7 for Ground 4 largely tracks that for Ground 2. *Compare* Pet. 83 *with* Pet. 119-120. For the same reasons as stated for Ground 2, Claim 7 is not obvious in Ground 4. Brogioli, ¶¶287, 205-208.

F. Ground 5

As explained above, in Riho, each TSV is in electrical communication with every die in the stack. As such, the drivers used to drive signals along the TSVs need to have at least the size of, and not less than the size of a driver "sufficient to drive a signal along a die interconnect in electrical communication with each of the plurality of array dies without significant signal degradation." *Contra* Pet. 141-42; Brogioli, ¶289. That is, given Riho's multi-drop configuration, Riho, alone or in view of Riho2, does not disclose or suggest claim 15[b] and other Ground 5 claims.

IX. Conclusion

Based on the foregoing, the Board should uphold the challenged claims.

Dated: July 27, 2023

Respectfully submitted,

/Hong Zhong/

H. Annita Zhong (Reg. No. 66,530)
Jason Sheasby (*pro hac vice*)
Irell & Manella LLP
1800 Avenue of the Stars, Suite 900
Los Angeles, CA 90067
Tel: (310) 277-1010
Fax: (310) 203-7199
Email: HZhong@irell.com
Email: JSheasby@irell.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.,

Petitioner,

v.

NETLIST, INC.,

Patent Owner

IPR2022-01428
Patent 8,787,060 B2

PETITIONER'S REPLY TO PATENT OWNER'S RESPONSE

(e.g., to select a "rank" of one or more memory chips to perform a read/write command). *See, e.g.*, EX1019, 13, 33; EX1022, 319-20, 413-14; EX1054, 5; EX2017, 13; EX1052, 299:24-300:6, 82:10-84:21, 88:22-89:24, 90:5-94:5, 335:13-336:19. Netlist attempts to change the court's construction to exclude all "situations in which a chip select signal could enable multiple array dies at once," POR 18-19, but the Board correctly determined that the court made no such blanket pronouncement, ID 11-12. As explained below (pp.15, 34), the prior art renders obvious "*chip select signals*" under the court's construction.

C. The Petition did not improperly incorporate by reference (POR 19-20)

Netlist argues the Petition devoted just a few paragraphs to elements 1.d.1-1.d.2, POR 19-20, but those paragraphs are replete with citations to supporting evidence and explanatory figures, Pet. 39-40, 100-102, which clearly show why 1.d.1-1.d.2 was obvious as the Board found, ID 13-19, 21-25.

D. Grounds 1-3 (with Kim+Rajan) (POR 7-10, 20-55)

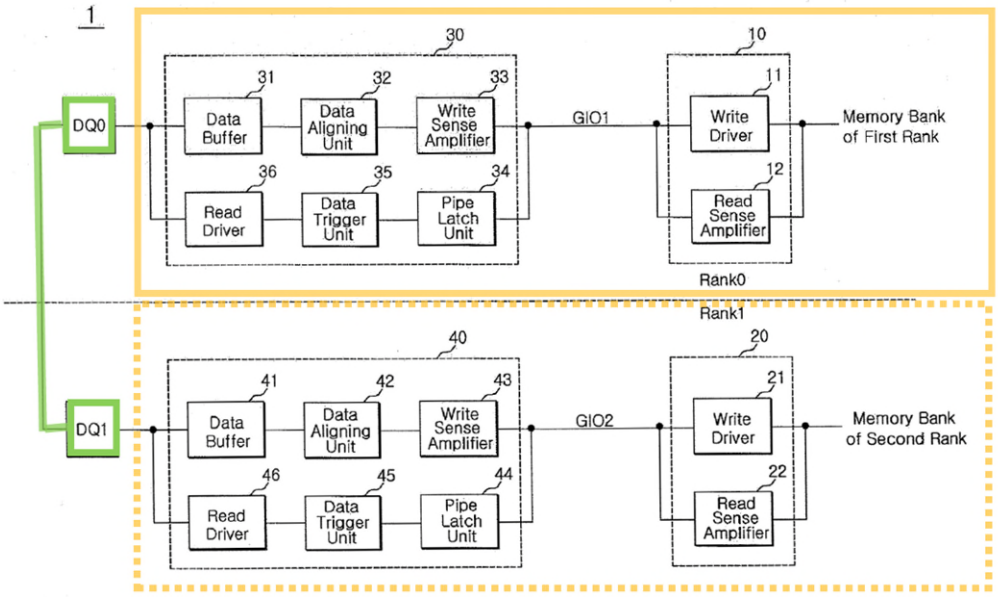
1. Avoiding collisions on shared data lines was well known (POR 7-10)

Netlist's entire argument against Grounds 1-3 is premised on the assumption that when the data pads (green below) of two memories are connected to each other, as shown in Kim's prior-art Figure 1, "data collision is ***unavoidable***"[1] when

[1] All emphases in quotes and color annotations in figures have been added.

a read by one device (solid orange below) follows a write by the other device (dashed orange below). POR 7-10; EX1014, [0042], Fig.1.

FIG.1
(PRIOR ART)



Netlist is incorrect. The "unavoidable collision" referenced by Kim above only applies to a ***unique*** hypothetical case "[a]ssuming…shar[ing] a data input/output circuit…operat[ing] in the ***same manner***" as in the prior art where the controller ***does not know*** the input/output circuit is being shared, EX1014, [0042], with the result that "command to data delays have not been properly selected" such that the write data collides with the read data, EX2025, 69:24-70:18. There were many known solutions to that problem, as discussed below, and thus a POSITA had a reasonable expectation of success in implementing the combinations for Grounds 1-3 ***without*** causing collisions, *see* Pet. 9-10, 21-30.

Indeed, both experts agreed that POSITAs were familiar with the problem of

collisions and knew of solutions. EX2025, 69:24-70:18 ("Data collision is always avoidable….[T]here are many, many known solutions to that problem."); EX1052, 315:20-:24 (agreeing "chips in different ranks could share the same data line ***without*** causing collisions"), 102:2-103:17, 106:19-107:12, 277:17-278:4, 278:19-279:3, 282:13-:24, 284:13-286:20, 314:13-19, 318:1-:6 (agreeing the problem of collisions was known and there was a motivation to avoid collisions).

For example, the JEDEC standard taught sharing the data pads (green boxes below) of stacked memories in different ranks, while using appropriately timed chip-select signals (red below) to avoid collisions when performing read and write operations on those shared data lines (green below):

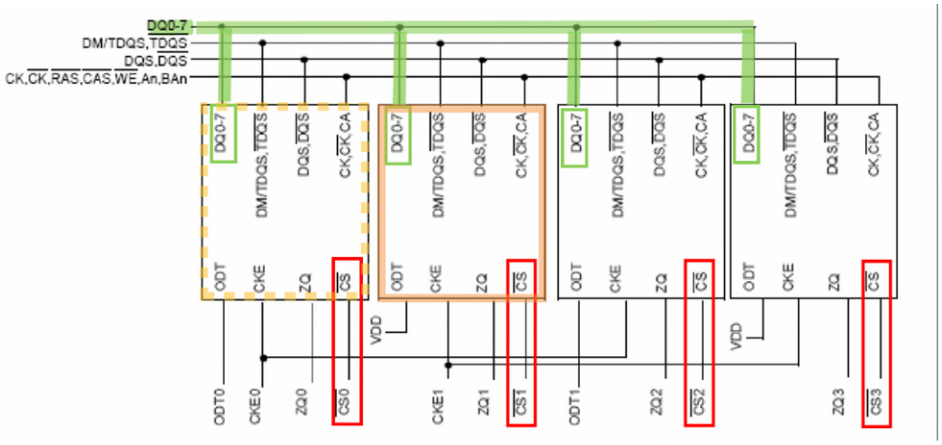


Figure 2 — Qual-stacked / Quad-die DDR3 SDRAM x8 rank association

Pet. 25-30 (reprinting EX1019, 12, above); EX1052, 301:8-304:1 (admitting the figure above was a "correctly operating device"), 279:4-:15 (agreeing that JEDEC provided "solutions to that problem [of collisions] that may have been acceptable"). To avoid collisions, the JEDEC standard provided a programmable

"CAS latency" value (e.g., 5, 6, 7, 8, or 9 clock cycles) to fine-tune the timing for read and write commands. EX1019, 23-24; EX1052, 109:18-111:16; EX1022, 343. The JEDEC standard also provided detailed timing diagrams for performing different sequences of commands without collisions (e.g., read followed by read, read followed by write, write followed by write, and write followed by read):

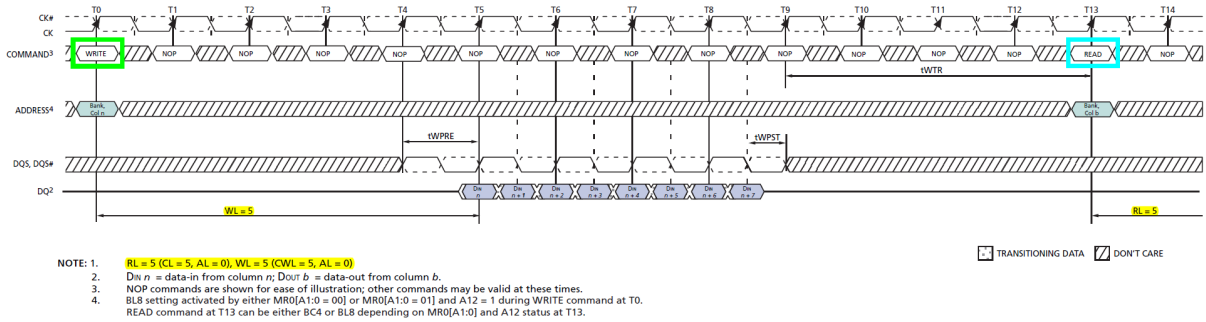


Figure 53 — WRITE (BL8) to READ (BC4/BL8) OTF

EX1019, 62-63, 73-74; EX1052, 113:2-118:10. Netlist's expert admitted a POSITA would have been aware of this JEDEC standard and its timing diagrams. EX1052, 76:14-77:20, 300:15-:24, 304:2-:15, 365:13-366:2.

The admitted prior art incorporated by reference in the 060 Patent also taught techniques for avoiding collisions. *See* EX1001, 4:7-:12, 19:44-:47, 22:37-:40 (incorporating the 436 patent, EX1057; the 386 and 912 patents that share the same disclosure, EX1028, EX1058; and the 537 patent, EX1031). For example, the 436 patent (filed in 2005) taught avoiding collisions by "inserting a delay between successive read commands" or by using "a serial presence detect (SPD) device to inform the memory controller of the physical boundaries of each memory

device." EX1052, 289:11-290:3. Similarly, the 912 patent (effectively filed in 2005) taught that collisions can be avoided by inserting "wait-time intervals or clock cycles." EX1058, 23:60-:67; EX1052, 279:16-280:6, 282:25-283:11, 291:3-292:12. Similarly, the 537 patent (filed in 2006) taught that adding ***one*** clock cycle to the CAS latency "provides sufficient time for read and write data transfers." EX1031, 21:28-:51; EX1065, 44-47.

Furthermore, the Rajan prior-art reference (included in all Grounds) provides much more detail than the 060 Patent (EX1001, 21:51-:56) with respect to timing solutions for avoiding collisions. EX1015, 8:59-12:13, figs.9-15; Pet. 79-80, 106-111, 136-137. Netlist's expert admitted that Rajan's buffer chip (shown in Figure 4, and which can include the interface circuit of Figure 18, *see* Pet. 16-18, 21-30) can change the timing of signals to emulate JEDEC and to avoid collisions:

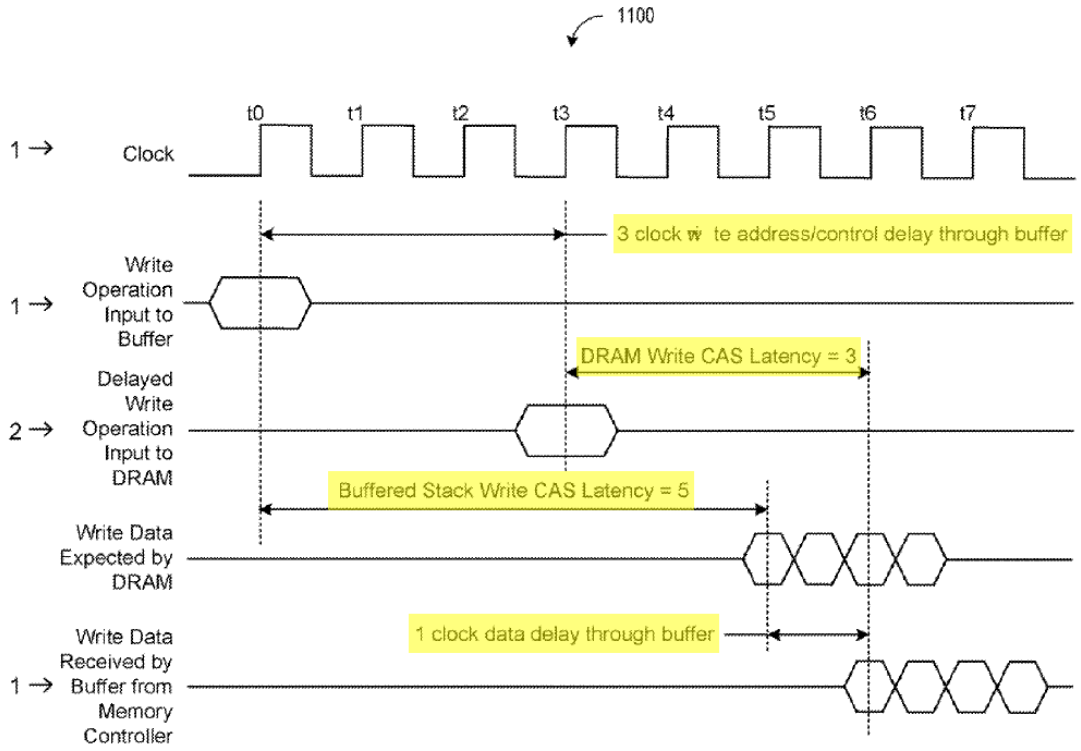


FIG. 11

EX1052, 190:6-192:8 (emulating CAS latency in Fig.9), 192:13-194:25 (emulating write CAS latency in Fig.11), 195:4-198:15 (delaying write operations in Fig.12, thus "making sure a piece of data shows up…at the exact point in time it's supposed to"), 198:20-200:9 (identifying potential bus conflicts and collisions caused by write operations in Fig.13), 200:10-201:11 (disclosing "methods of avoiding such collisions" in Fig.14); *see also* EX1022, 446-47 (similar).

Thus, a POSITA knew a number of different techniques to avoid collisions, contrary to Netlist's assumption that such collisions were always "unavoidable."

2. "*first group of array dies*" (POR 20-44)

Netlist argues that the claim language "*a first group of array dies*" requires

"***more than one*** die." POR 20-23. But the claim language does not exclude a "*group*" with only ***one*** die, which Netlist's expert never considered. EX1052, 130:10-131:3. In any event, the Petition explained that Kim discloses that "***any number*** of…slave chips may be used," EX1014, [0048], and Rajan teaches ***multiple*** dies in a group of stacked dies sharing the same data lines, *e.g*., EX1015, Fig.4, rendering obvious multiple dies in the "*first group*" of Kim. Pet. 34-40; *see also id.* 21-30; *In re Harza*, 274 F.2d 669, 671 (CCPA 1960) ("mere duplication of parts has no patentable significance unless a new and unexpected result is produced").

Netlist is incorrect that "Rajan…is silent on…the relationship among the DRAM circuits sharing a data bus." POR 22. To the contrary, Rajan clearly shows (below) which memory chips share which data bus, and teaches "utilizing separate chip select signals…to each of the DRAM chips in the stack" as found by the Board. EX1015, 6:34-:38, Fig.4 (below); ID 25-26; Pet. 26-29, 61-62.

Rajan:

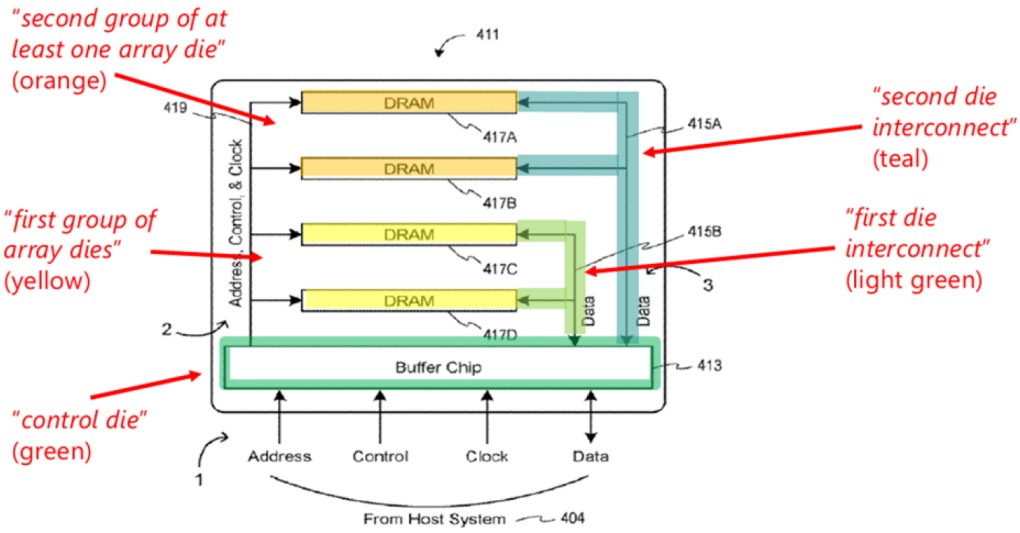


FIG. 4

The Petition further explained that this configuration of each memory in the stack having its own chip-select signal (thus being in separate ranks, *see* EX1019, 13; EX1022, 319-20, 413-14), and sharing two separate data busses, as taught by Rajan, was well-known and commonly used for rank multiplication, as evidenced for example by Ellsberry (below). Pet. 27, 29-30; EX1026, Figs.12-13 (below); EX1065, 37-53.

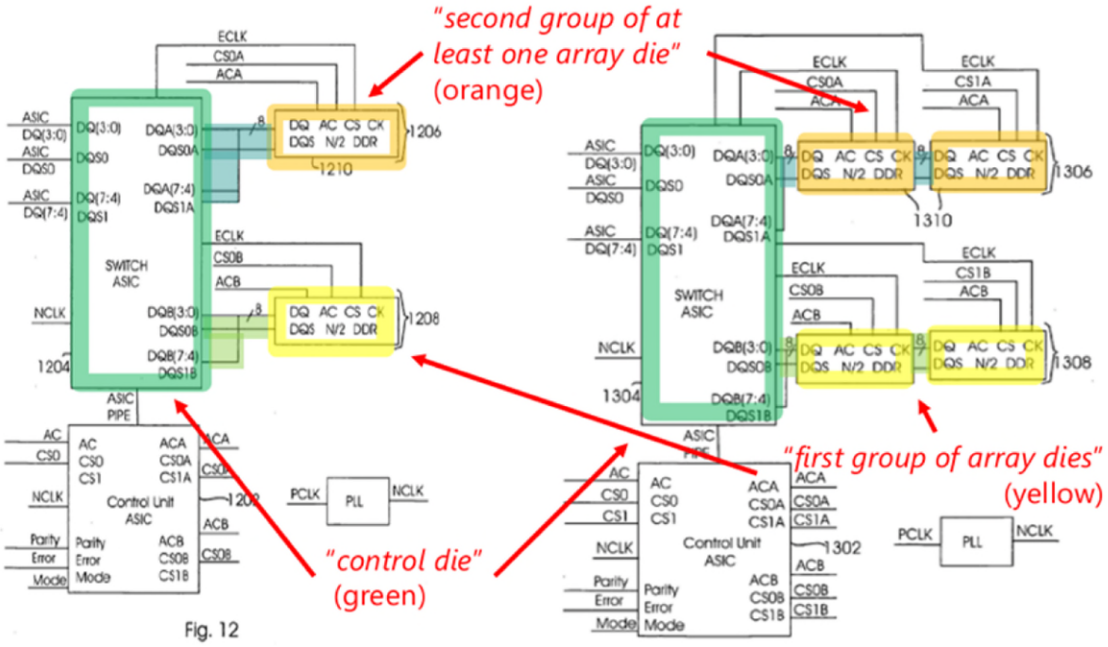


Fig. 12

Fig. 13

Instead of addressing the specific arrangements proposed by the Petition, Netlist focused on a series of strawman arguments, discussed below. EX1052, 319:2-:8, 324:2-:17, 325:7-:14, 325:24-326:14.

a) Slave chips in a stack can share a TSV for data transfers without creating collisions (POR 23-38)

Netlist's first strawman argument is that multiple chips "belonging to the ***same*** rank" cannot share a TSV for data transfers because that would lead to collisions. POR 23-27. But that is not what the Petition proposed. EX1052, 314:1-:24, 315:14-:19. As shown directly above (pp.8-10), the Petition arranged the four chips in ***different*** ranks, as was common to avoid collisions as discussed above (pp.3-5).

Netlist's second strawman argument is that it would not work to ***cross***-connect a chip in Rank0 to the TSV GIO_Rank1, and a chip in Rank1 to the TSV GIO_Rank0. POR 28-30. But Kim does not disclose cross-connecting ranks like that, *see* EX1052, 316:3-317:25, and that is not what the Petition proposed as shown above (pp.8-10).

Netlist's third strawman argument is that "four separate TSVs for four different ranks" would not result in the claimed inventions. POR 30-31. As discussed above (pp.7-8), that depends on the construction of "*group of array dies*" excluding a single die. In any event, the Petition proposed ***shared*** data busses as shown above (pp.8-10), not merely a ***separate*** TSV for each rank as imagined by Netlist, EX1052, 318:10-319:8.

Netlist's fourth strawman argument is to modify Kim's Figure 2 with "4 ranks sharing 2 ports," but ***without*** "the appropriate control signals" proposed by Dr. Wolfe. POR 32-38; EX2025, 46:3-:12, 79:13-80:11, 86:12-:17, 201:3-202:5, 204:2-205:8; EX1052, 319:9-320:11 (Netlist's expert admitting that he created the modified figure shown on POR 33 assuming no change to Kim's circuit). For example, even though Netlist's modified figure has ***four*** ranks, it shows only ***two*** chip select signals, EX1052, 320:12-322:7, rather than "***four*** chip selects" as actually proposed by Dr. Wolfe, EX2025, 55:18-57:9.

Based on yet another strawman figure, Netlist argues that two ranks cannot

share a data line because of "unavoidable" data collisions as allegedly shown in Kim's ***prior-art Figure 1***. POR 33-35; EX1052, 323:3-324:1. But the Petition did not rely on Kim's prior-art Figure 1, but rather the teaching of Kim, as shown for example in Figure 5, in combination with the teachings in Rajan (and Wyman). Pet. 21-30, 84-86.

Netlist also argues at length about allegedly "unavoidable" collisions when two ranks in Kim are connected to a "GIO_Rank1 line without an intervening rank selection unit between GIO1_0 and GIO1_1." POR 34-38. But as discussed above (pp.2-7), Kim mentions an "unavoidable" collision only with reference to a specific hypothetical scenario, and it was well known at the time how to avoid such data collisions (pp.2-7). Furthermore, Netlist's arguments ignore the teachings of Rajan and Ellsberry, as discussed above (pp.8-10). Indeed, "the [060] patent itself [e.g., EX1001, 21:51-:56] does not disclose the level of detail" found in the prior art for avoiding collisions. *Lockwood v. Am. Airlines, Inc.*, 107 F.3d 1565, 1570 (Fed. Cir. 1997).

b) Combining Kim and Rajan was obvious and would not result in data collisions (POR 38-44)

Netlist argues that combining Kim with Rajan would not solve the "unavoidable" data collision problem. POR 38-44. But all of Netlist's arguments depend on the same strawman arguments about Kim discussed above (pp.10-12).

Importantly, Netlist ignores that JEDEC provided for programmable "CAS

5. Ground 2 renders obvious claim 7 (POR 51-54)

For claim 7, Netlist argues that "general similarity does not explain why a POSITA would have combined Riho with Kim/Rajan." POR 51. But the Petition explained in detail the motivations for the combination, including reducing load and skew variations without the need to create additional TSVs. Pet. 81-83; EX1003, ¶¶171-73, 339-41, 347-48.

Netlist also repeats its argument that, in Kim, "different ranks of memories cannot share a same TSV[]," POR 51-52, which is incorrect for the reasons discussed above (pp.2-12). Furthermore, Netlist admits that Riho discloses "different ranks…connected to the same TSV," POR 52, thus motivating a POSITA to follow Riho's teaching and add more ranks as taught by Kim and Rajan without additional expensive TSVs. Pet. 26-30; EX2023, ¶195 ("TSVs…were…expensive to form").

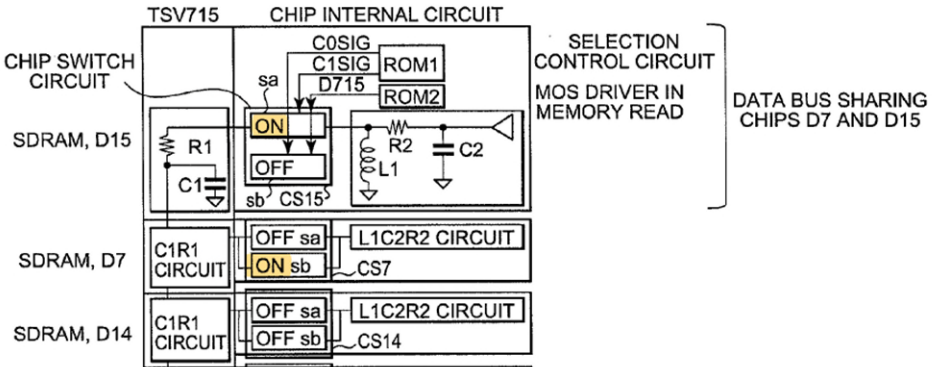
Netlist also sets up another strawman by arguing that "Petitioner does not suggest any change to Kim's TSV lengths." POR 52-53. On the contrary, the Petition explained that the combination would minimize the skew, as taught by Riho, i.e., making the TSVs "substantially equal in length to each other." EX1016, [0050]; Pet. 83 (citing EX1016, [0050]); EX1003, ¶¶340-41 ("equi-length interconnections…important…[for] synchronization").

Netlist further argues that, in Riho, the number of dies in each group does

not matter when "only one of the two SDRAMs…[is] connected to the TSV at a time." POR 53-54. But to maintain minimized skew with Riho's "equi-length" TSVs as discussed immediately above, there are an ***equal*** number of dies (e.g., two) in each group (with Riho also teaching the option of connecting only half of each group at a time to the TSV to further reduce the load by half). Pet. 82-83. Therefore, Riho teaches that the first/second number of dies in the first/second group is selected to be the ***same*** "*in consideration of a load of the…die interconnect*," i.e., to minimize skew. Pet. 82-83; EX2025, 197:12-199:2.

Netlist's argument that "the load on the conduit need to include a load associated with ***two*** chips," POR 54, is incorrect because a "*group of array dies*" can include just one die as discussed above (pp.7-8). In any event, Riho discloses (and the Petition identified) an embodiment connecting ***both*** SDRAMs D7 and D15 to TSV715 permanently according to ROM1-ROM2 (as shown in Figure 5 below), not just one of them at a time (as permitted by Figure 8). *See* Pet. 82-83 (annotating Figure 4, with "two" dies per group sharing a TSV); EX1016, [0062, 65, 71-74], Figs.4, 5 (below in part); EX1003, ¶¶339-41.

FIG. 5



6. Ground 3 with Wyman teaches using drivers of different sizes for TSV1 and TSV2 [15.a]-[15.b] (POR 54-55)

Netlist does not dispute that Wyman is analogous art to Kim and the 060 Patent, or that it teaches that shorter TSVs require a smaller driver than longer TSVs, and it would be wasteful and overkill to use the full strength of the driver for either. Pet. 20, 84-86, 88-89; POR 54-55. Instead, Netlist proposes alternative implementations for combining Kim with Wyman that allegedly would not satisfy the claims. Netlist misses the point: even if other implementations were known or even preferred, that does not make the proposed combination non-obvious. *See, e.g.*, *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380-81 (Fed. Cir. 2023) (combination need not be the "best option," just a "suitable option"); *Dome Pat. L.P. v. Lee*, 799 F.3d 1372, 1381 (Fed. Cir. 2015).

First, Netlist argues that a POSITA would have used a "***single*** driver for both TSV1 and TSV2," with one tap for the shorter TSV1, and another tap for the longer TSV2. POR 54-55. But as explained by Dr. Wolfe, "you can provide a

separate copy of [Wyman's] circuit 500 for each thing that needs to be driven" and "there's also the teaching in Wyman that you can make 500 at various different sizes" as shown, for example, in Figure 3 and 4. EX2025, 128:22-132:19; *see also* EX1017, Fig.1 (single output drive portion 100), 3:46-:49, 5:34-:44 (disabling unneeded downstream stages, instead of providing "taps" at each stage). Using separate drivers for separate data ports was a well-known technique to avoid collisions, especially for rank multiplication. *See, e.g.*, Pet. 29 (discussing "fork-in-the-road" arrangement); EX1026, fig.4; EX1062, Fig.4. Thus, a POSITA would be motivated to use separate drivers of different strengths as taught by Wyman for Kim's TSVs of different lengths, in particular for rank multiplication, as taught by Rajan.

In any event, Netlist's argument incorrectly assumes that the claimed "*first drivers*" and "*second drivers"* cannot share any circuitry. The claims require different "*first driver **size***" and "*second driver **size***" without any further restrictions. The 060 Patent teaches that the "size of the driver may be adjusted by the…***number of transistors included in the driver**.*" EX1001, 17:43-:45. That is similar to one embodiment of Wyman where each tap uses a different number of transistors. *See, e.g.*, EX1017, Fig.5; EX1052, 327:11-:17, 328:10-328:24, 329:16-330:5.

Next, Netlist argues that Wyman's driver circuitry would be regarded as a

fixed (physical) size, regardless of the number of components used, POR 55; EX1052, 330:6-:14, 331:19-332:22, contrary to the claim language and the specification discussed directly above teaching that "size" depends on the number of transistors used. In any event, as explained by Dr. Wolfe above (p.19), Wyman also teaches making the physical size of each driver different, as shown for example by Figures 3-6 of EX1017.

Finally, Netlist argues that load difference between Kim's shorter and longer TSVs may be too "negligible" to justify different driving strength, POR 55, which directly contradicts the express teaching of Wyman. *See, e.g.*, EX1017, Figs.8A-8B, 6:15-:30 ("increased resistance, capacitance and impedance [of longer TSV 804] might require additional drive"); EX1052, 334:4-:13 ("TSV"). The 060 Patent itself admits that, "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." EX1001, 4:42-:45.

E. Grounds 4-5 (with Riho+Rajan) (POR 56-83)

1. "*not in electrical communication with the…array die[s]*" [1.d.1]-[1.d.2] (POR 56-75)

For Grounds 4-5, Netlist's arguments all depend on the non-sensical claim construction position that there is still "*electrical communication*" even when an OFF switch electrically ***disconnects*** the SDRAM chip (and its load) from the TSV as taught by Riho (EX1016, [0065] ("TSV19 is ***not electrically connected*** to the

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD., MICRON TECHNOLOGY, INC., MICRON SEMICONDUCTOR PRODUCTS, INC., and MICRON TECHNOLOGY TEXAS LLC

Petitioners,

v.

NETLIST, INC.,

Patent Owner.

Case No. IPR2022-01428
Patent No. 8,787,060[1]

PATENT OWNER'S SUR-REPLY

[1] Micron Technology, Inc., Micron Semiconductor Products, Inc., and Micron Technology Texas, LLC, filed a motion and joinder and a petition in IPR2023-00882 and have been joined as petitioners in this proceeding.

The Board should decline to consider new arguments in the Reply or its liberal incorporation of Dr. Wolfe's testimony. But even if considered, the challenge fails on the merits.

I. Petitioner Advances Inconsistent Claim Constructions for Infringement and Validity

Having persuaded the district court to adopt its constructions for "array die" and "chip select," Samsung now runs away from them. *See* Reply 1-2; EX2004, 31-34. Samsung gained the benefit at trial of forcing Netlist to prove infringement under very narrow claim constructions. The same constructions should apply for validity.

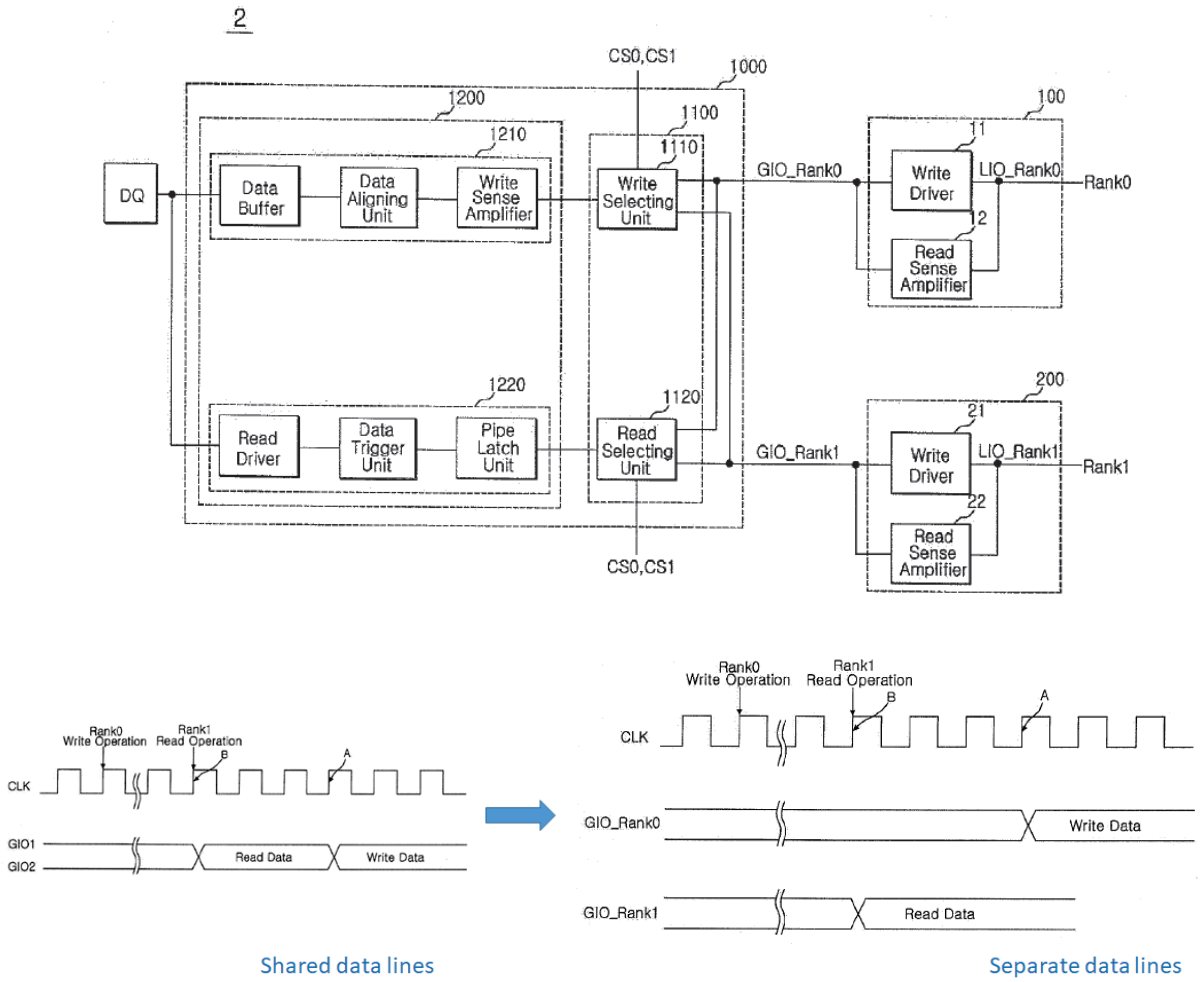
II. Petitioner Has Not Shown Unpatentability for Grounds 1-3

A. Contrary to Samsung's New Argument, JEDEC Timing Does Not Solve Kim's Data Collision Problem

The Petition had little detail on how modified Kim should be structured or operate. The POR explained why the four possible modifications would either be inoperable or differ from the inventions. POR 20-38. The Reply does little to defend three of the modifications and raises new arguments to save the fourth. Those new arguments should be disregarded. But even if considered, they do not provide substantial evidence of invalidity.

Samsung incorrectly argues that Kim's expressly disclosed "unavoidable" data collisions only occur when "first and second data input/output lines operate in the same manner." Reply 3. Kim discloses instead "[a]ssuming that …Rank0 and Rank1 share a data input/output circuit, first and second data input/output lines

operate in the same manner." EX1014, [0042]. That is, the first and second data I/O lines operate in the same manner because Rank0 and Rank1 share a data I/O circuit. *Id.* Hence, for the two I/O lines to operate differently, Rank0 and Rank1 must ***not*** share a data I/O circuit at a given time, just as Kim discloses. EX1014, [0044]-[0045], Fig. 3; EX2023, ¶¶77-79. Kim achieves this by interposing a rank selecting unit 1100 to select either GIO_Rank0 or GIO_Rank1 so that the two ranks do not share data lines at any given moment. *Id.*



Samsung argues that data collisions were known problems with known

solutions, new arguments that should be disregarded. Reply 3-8. Moreover, the known solution to Kim's problem was Kim's changed architecture. Notably, despite their knowledge of JEDEC specifications, Kim and his co-inventor at Hynix[2] (*see* EX1014, 1) concluded that certain data collisions were "unavoidable" in their system and added circuitry to decouple the GIO lines. EX1014, [0042]-[0045].

Samsung has not shown that other solutions existed. First, Samsung has not shown that the stacked dies in EX1019 actually share TSVs (*see* illustration below), asserting only that they share data pads. Reply 4.

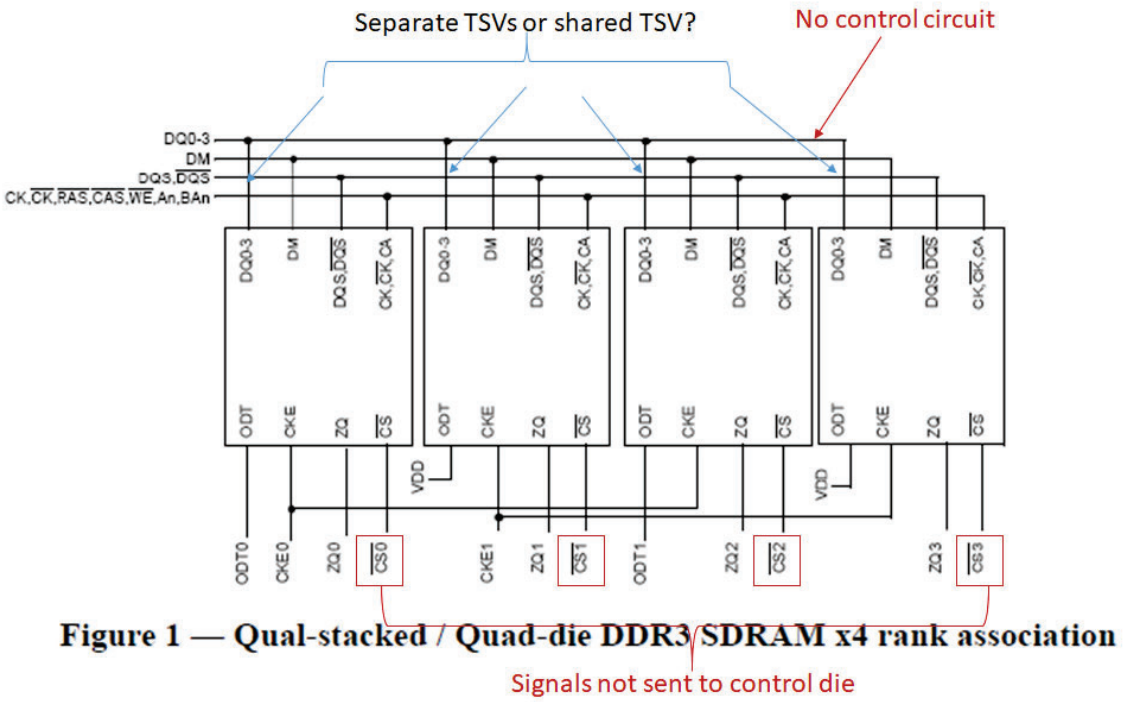
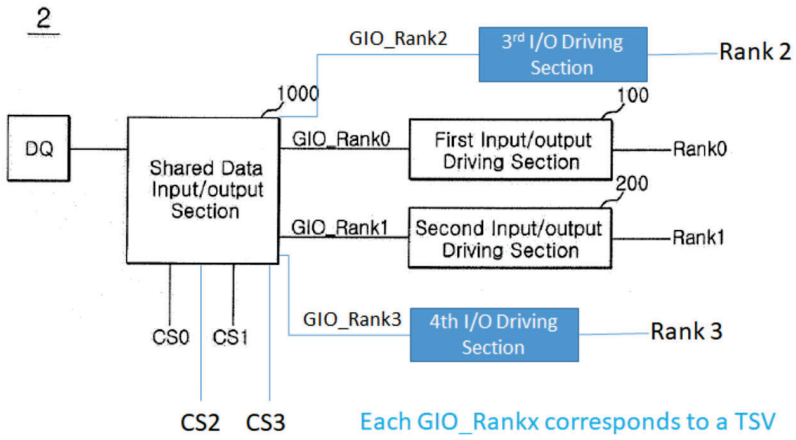


Figure 1 — Qual-stacked / Quad-die DDR3 SDRAM x4 rank association

TSVs can each be connected to one die but still share a pad, as illustrated

[2] Hynix, Micron and Samsung together control over 90% of the worldwide DRAM market.

below. POR 31.[3] Thus, sharing a pad does not mean sharing TSVs.



Second, if stacked dies in EX1019 do share TSVs without data collisions, then there would be no need for Kim's rank selection unit 1100 (mapped to the claimed control circuit), no first die interconnect versus second die interconnect, first group of array dies versus second group of at least one array die, or first data conduit versus second data conduit. EX2023, ¶137. Changes in one part of the system cause changes in other parts of the system, and not even Samsung asserts that EX1019 would lead to the claimed inventions. The Kim architecture that Samsung relies on—including the required control signal(s), data conduits, and control circuit—is specifically designed to address the otherwise "unavoidable" data collisions. If such data collisions did not exist in the first place, there would be no reason to add new circuit(s) to the "typical" arrangement. EX1014, [0042]-[0045].

[3] In this configuration, there are no "first group of array dies" in communication with any TSV, as the claims require. POR 31.

after-write operations. Moreoever, as shown above, given the constraints on CL and CWL, adding latency to read operations (i) would not change the relative timing between read and write data; (ii) would not delay read data sufficiently to avoid data collisions; and/or (iii) could even introduce other data collisions. *See* discussions on pages 6-8.

There is also no evidence that Kim's memory controller was unaware of the physical boundaries of Kim's ranks[5] such that the SPD approach could solve the data collision problem. Indeed, if SPD could have solved Kim's data collision problem, Hynix engineers would have known about it. *Contra* Reply 5-6.

Adding wait-time intervals (Reply 6) is no different than changing the CL and CWL values, and would run into the same problem as discussed earlier. Similarly, as explained above on pages 6-7, lengthening CAS latency by one clock cycle for both read and write (Reply 6) would not change the relative time of read or write data, and collisions remain. As Dr. Brogioli testified, handling data collisions involves "a lot of complexities" that is "beyond the scope of a person of skill." EX1052, 283:17-284:12.

Samsung also cites Rajan, which by definition does not have the array dies of the claims. *See* EX2004, 31-32. Samsung argues that Rajan teaches how to avoid

[5] *See* EX1014, [0006] ("chip selection command signals" used to control operations of different ranks).

collisions by delaying write operations, but none involved cross-rank read-after-write operations and certainly not for Kim's memory systems. Reply 6-7. Furthermore, Kim's data collisions are caused by read data for a later-issued read command appearing on the shared data line earlier than write data for an earlier-issued write command. EX1014, [0042], Fig. 4A. Delaying write commands will just further delay the write data, and does not solve Kim's data collision problem.

Likewise, having the buffer delay relaying commands or otherwise adjusting CAS latency (*see* EX1015, Figs. 13-15, 11:12-39) would just change the absolute time at which read data and write data appear, but not the relative timing of the data that underlies the collision. *See* pp. 6-8 above; EX1052, 346:17-347:4 ("just providing a [latency] value to the memory controller doesn't solve the – the scenario that Kim contemplates of the read-after-these-write operations").

EX1022 pages 446-447 also do not address Kim's data collisions. Reply 8. Figure 11.23 concerns operations "to the ***same*** rank," not different ranks. EX1014, [0042]. For Figure 11.25, Jacob does not even suggest there would be a suitable tRTRS for Kim's system, let alone how to determine it. EX1022, 448 (system-dependent tRTRS).

Samsung also cites as support Dr. Brogioli's unremarkable observation that ***"[i]n certain conte[x]ts***," chips in different ranks "***could***" share the same data line without causing collisions. Reply 4; EX1052, 315:20-24. Even Kim acknowledges

that data collisions do not occur in some operations, but observes that they are ***unavoidable*** in cross-rank read-after-write. EX1014, [0042]. The question is not whether there are operations in which data collisions do not occur, but how to address predicted data collisions for expected operations. In that regard, having a motivation to avoid data collisions (*see* Reply 4) says nothing about whether a known solution existed for Kim's system. Despite its knowledge in the memory area, SK Hynix (Kim) concluded that in Kim's system, the data paths for different ranks of memories should be decoupled, or there would be "unavoidable" data collisions in cross-rank read-after-write operations. *See* EX1014, [0042]-[0044]; *see also* EX2023, ¶137 (Dr. Brogioli explaining why latency would not solve Kim's problem and all the drawbacks in lengthening latency). Samsung's attorney argument that other solutions existed for Kim's system ignores Kim's express disclosure and should be accorded no weight.

B. When TSVs are Shared among Different Ranks, Kim's System Has Unavoidable Data Collisions in Cross-Rank Read-after-Write Operations

"In accordance with common English language, "a "first group of array dies" means a group of "two or more" array dies. *Apple Inc. v. MPH Techs. Oy*, 28 F.4th 254, 261 (Fed. Cir. 2022) (interpreting "information fields" (plural) as "two or more fields"). Nothing in the patent disclaims the plain and ordinary meaning of the plural form as two or more. In fact, the claims as a whole reinforce this meaning: When

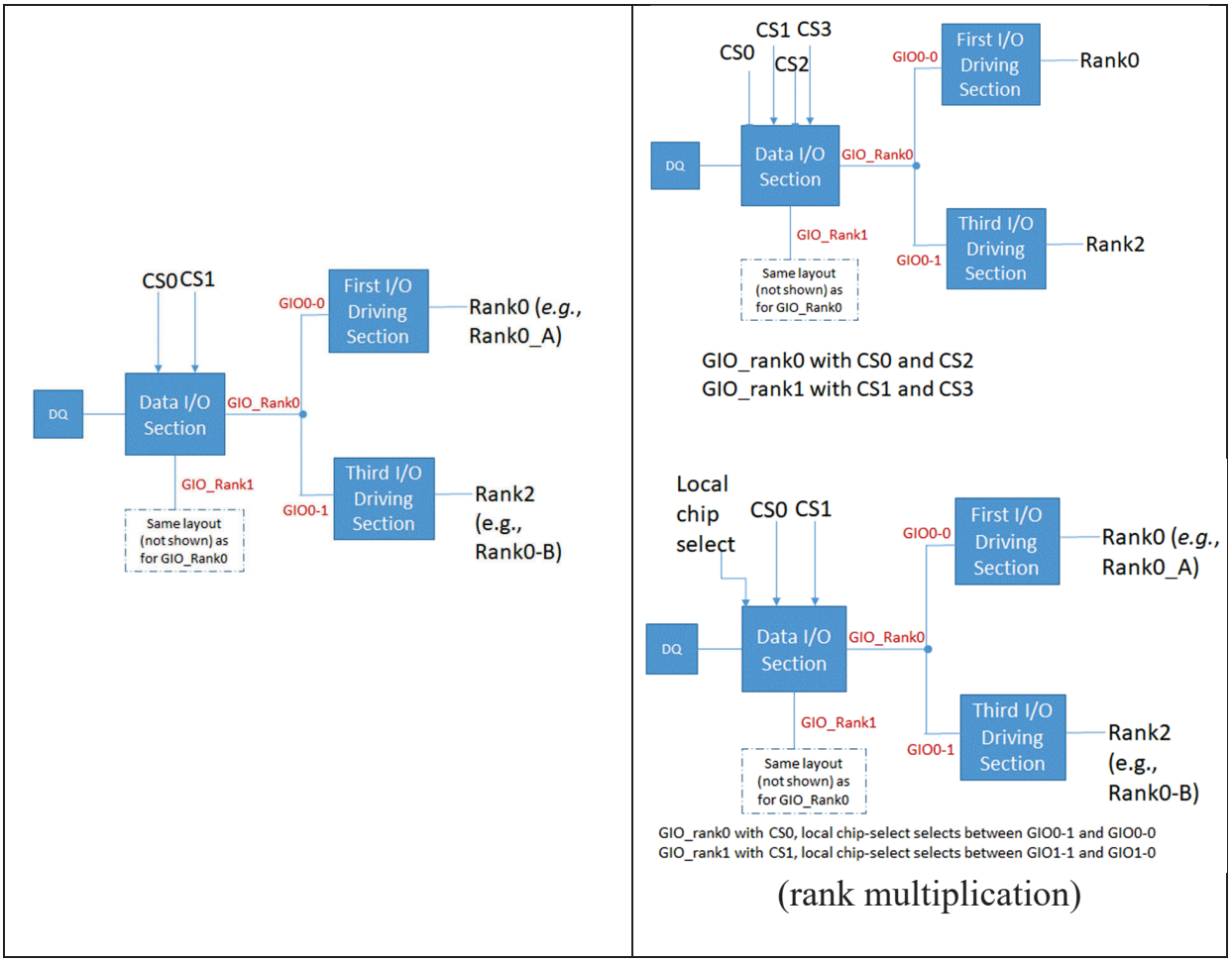
the inventor, Dr. Lee, intended to cover a group of one or more dies, he recites a "group of at least one die." '160, claim 1. Thus, the Board should reject Samsung's attempt to construe a "group of array dies" as a group of one or more dies. *Contra* Reply 7-8.

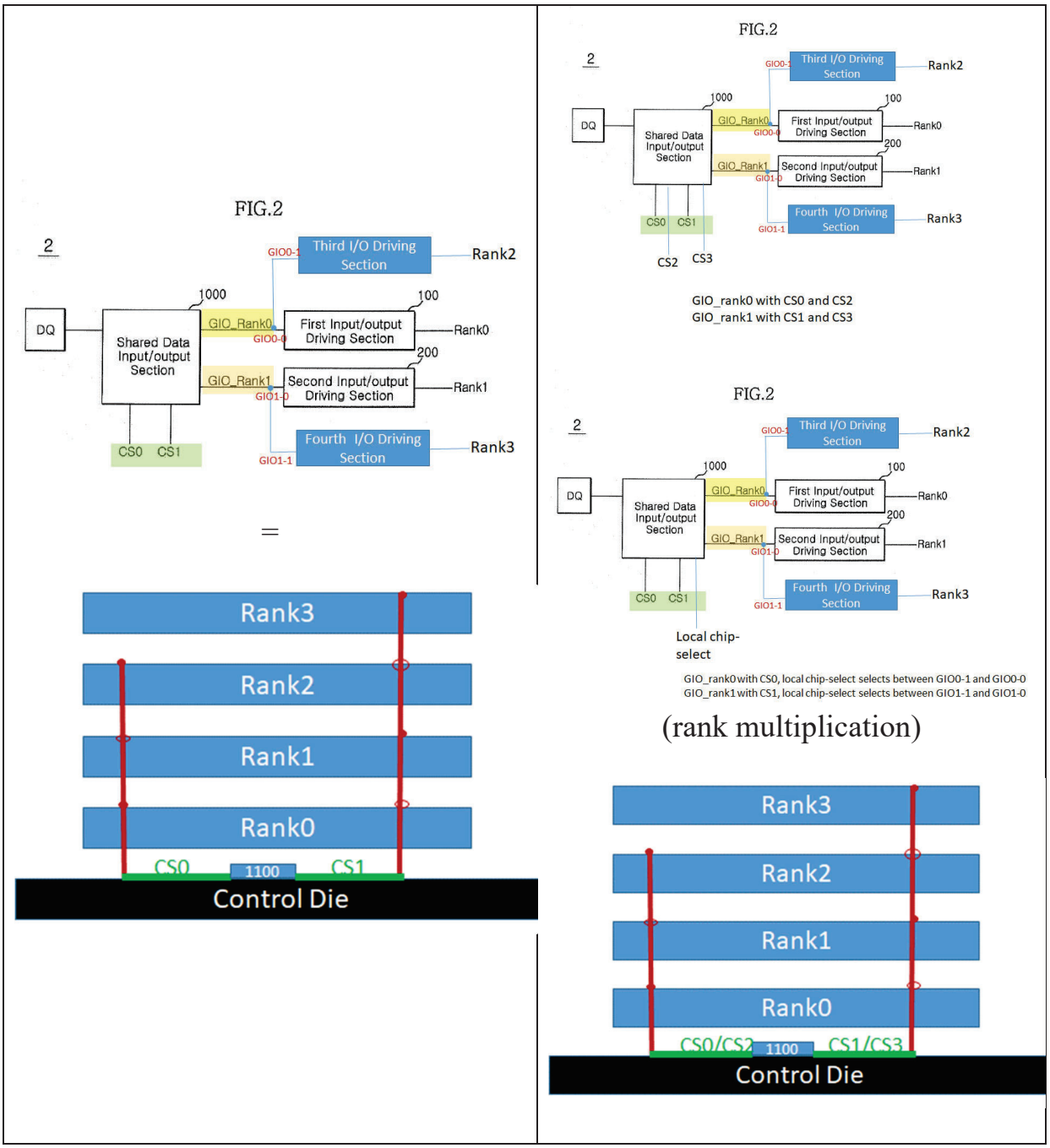
The difference in language matters, because requiring at least one die interconnect be in electrical communication with multiple "array dies" introduces the risk of data collisions over a shared data path. Hence, going from point-to-point to shared connections is not trivial as Samsung asserts, a complication that it clearly has not thought through. *Contra* Reply 8.

In the POR, Netlist explained in detail the problems associated with Samsung's proposed modifications. POR 21-22. The Reply barely defends three of the four configurations. Reply 10-11. Although the Reply characterizes them as "strawman argument[s]," each is based on Dr. Wolfe's testimony as explained in the POR. POR 23-31. For example, the Reply asserts that Samsung's modification would involve "four chips in ***different*** ranks." Reply 10 (emphasis in original). But Samsung acknowledges that it looked to Rajan Figure 4 for die grouping. Reply 8-9. Dr. Wolfe testified the four dies in Rajan Figure 4 were grouped into two ranks, "one rank would be 417A and 417C, and the other rank would be 417B and 417D." EX2025, 90:1-21. This is an example of Samsung's ever evolving theories.

Samsung complains that in Netlist's illustration of the fourth possible

configuration, Netlist included only two chip-select signals instead of four (a theory that is absent from the petition or Dr. Wolfe's declaration). Samsung does not and cannot show that using four instead of two chip selects would have made any difference in the analysis. Indeed, if the figures on POR 36 and 33 are replaced with the figures on the right, the analysis remains the same as illustrated below. *Cf.*, Rajan, 6:30-39 (using additional address bits to select chips individually); EX1014, [0006] (Kim already using chip-selects to control operations of different ranks).

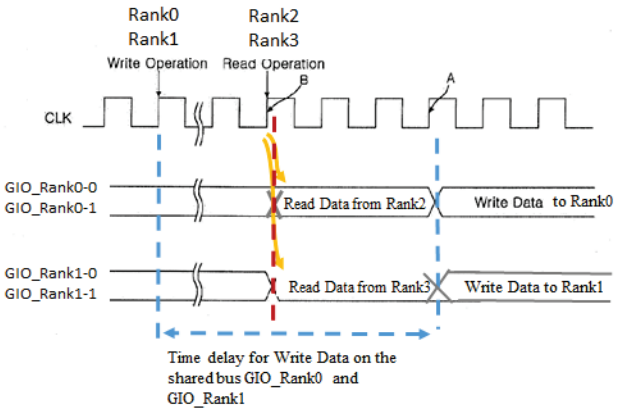




Instead of CS0, Chip select signals CS0 and CS2 would cause the data being transmitted onto GIO_Rank0 line as before, and respectively onto GIO0-0 and GIO0-1, also as before. Similarly, CS1 and CS3, instead of just CS1, would cause

the data being transmitted onto GIO_Rank1 line, and respectively onto GIO1-0 and GIO1-1, again as before. Thereafter the timing issue in the cross-rank read-after-write operations between Rank0 and Rank2 or between Rank1 and Rank3 would be the same as discussed in detail on POR 33-38. EX2023, ¶¶120-121, 128-133. This is because Kim's data collisions concern the relative timing of the read and write data on GIO-Rank0 via coupled GIO0-0 and GIO0-1 lines (or the relative timing on GIO-Rank1 via coupled GIO1-0 and GIO1-1 lines). *Id.* It does not matter whether Rank0 and Rank2 data are switched to the GIO_Rank0 by a single CS0 signal or by either CS0 or CS2. The same timing conflicts exist. The same is true for the other two ranks. *See* illustration below (same as POR 36).

Cross-rank read-after-write collision



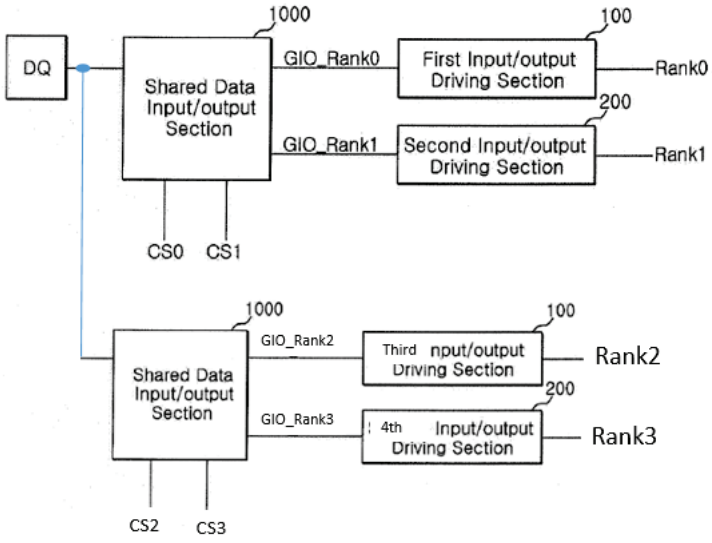
Dr. Wolfe testified that, even with 4 chip-selects, timing issues must be solved and that it is a "multistep" "nontrivial" process. EX2025, 55:18-58:10. Yet, neither Samsung nor Dr. Wolfe address whether any of the proposed modifications would

have timing issues, especially in light of Kim's stated problem. The POR explains why such timing issue exists in Kim. POR 32-38.

The Reply asserts that Netlist is relying on Kim's Figure 1. Reply 11-12. That is incorrect. The POR shows how Samsung's modification would look, and explains why the modification would reintroduce the data collision problem that Kim is trying to solve. Samsung has not shown why Dr. Brogioli's analysis in EX2023, ¶¶119-137 is incorrect, nor has Dr. Wolfe explain how the modification would actually work.

Samsung also references Rajan and Ellsberry's disclosures on rank multiplication, but does not explain how that would solve Kim's specific timing issue (or how modified Kim would operate). Reply 8-10, 12.[6] If Samsung's argument is that there would be two data buses, each having two separate paths controlled by a chip selecting unit, then—as illustrated below—that would just revert to the four-die-four-TSV configuration discussed in the POR.

[6] Ellsberry does not even involve stacked dies within a memory package.



Timing is system specific; and Samsung has no provided any evidence to support its new theory that a POSITA would or even could have modularly pieced together parts of different memory systems to arrive at the inventions. As Dr. Brogioli observed, if known solutions such as lengthening latency could have solve Kim's data collision problem, "there would be no need for the improvement that Kim describes through its Figures 2, 3 or 5 implementations." EX2023, ¶137. Samsung's ground 1 theory uses Kim's Figure 5 as a starting point.

C. Petitioner Has Not Shown How Rajan Could Have Solved Kim's Timing Problem

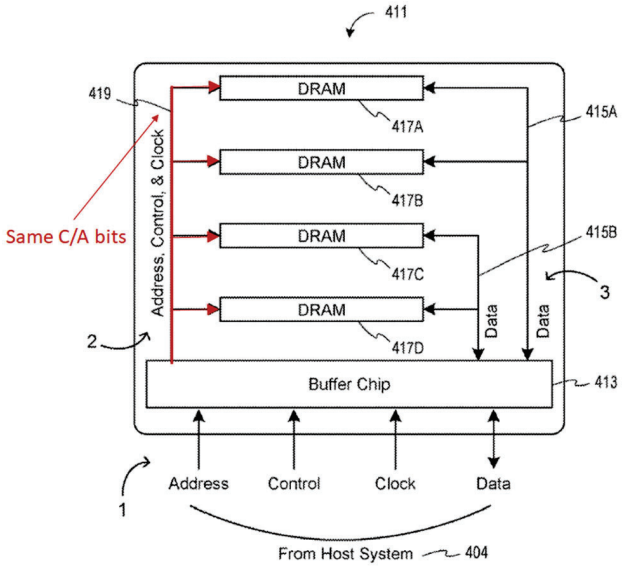
Netlist has explained above why the prior art solutions would not have addressed the challenges identified in Kim. *See* Section II.A. Samsung suggests that Kim's timing for read and write operations would change in the Kim-Rajan combination. Reply 13. That is a new theory: Netlist pointed out in the POR that Samsung does not suggest any modification to Kim's memory devices, and the

Reply does not contend that such a theory existed in the Petition. POR 40; Reply 13. Samsung criticizes Netlist for bodily incorporation, (Reply 13), but it is Samsung's burden to show what the proposed modification is and how that modification is supposed to work, including any changed parameters and operation. Samsung has not done any of that. Nor does the Reply provide any explanation what the timing parameters would be in the Kim-Rajan combination or why that would be the case given that the Kim's array dies need to be different from Rajan's DRAM circuits.

D. Samsung Has Not Shown that a POSITA Would Have Combined Kim and Rajan in the Manner Proposed

Samsung argues that because Netlist has prevailed on its infringement theory, it now does not have to show why SDRAMs in Kim and Riho are different from a DRAM circuit. Reply 13. But Netlist did the work and showed how the core dies in Samsung's HBM products differ from a DRAM circuit. EX2029, 1077:15-1122:3. In contrast, Samsung has not done any work to even show how DRAM dies it relies on in the prior art differ meaningfully from Rajan's DRAM circuits. Its expert testified: "I don't believe an SDRAM die is ever different than a DRAM circuit." EX2025, 94:10-18.

Samsung also argues that a POSITA would have used Rajan's interface die with non-DRAM dies. Reply 13-14. Contrary to Samsung's assertion, the '060



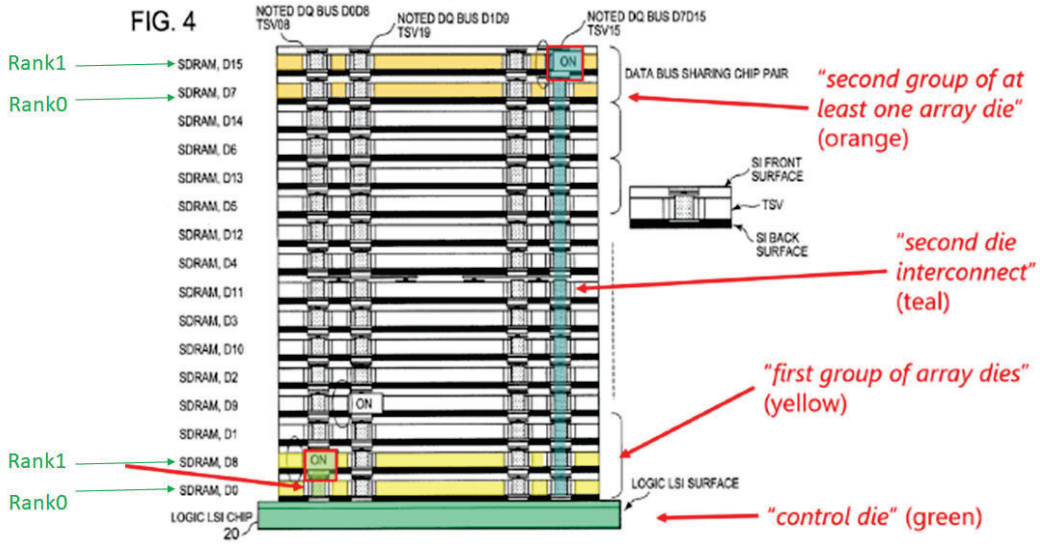
Moreover, in Rajan, Figure 4, the four dies are arranged in two ranks. EX2025, 90:1-21. Thus, each chip-select is sent to two dies at once, (EX1023, 2 (chip-select "enables or disables the DRAMs in that rank")), again not meeting the requirement of chip-select.

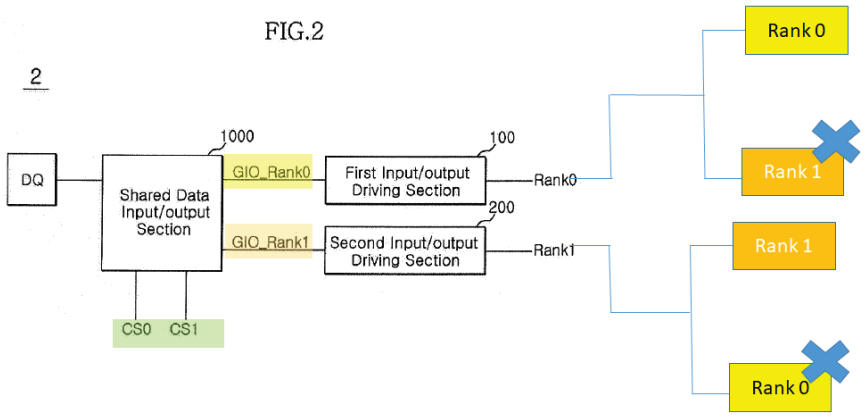
F. Claim 7 is Not Obvious

The POR explained why the combination of Kim-Rajan and Riho would not achieve the alleged reduction in DQ-DQS skew. POR 52-53. The Reply now argues that the combination intended to adjust the length of Kim's TSVs, even though the words "equi-length" or "substantially equal in length" never appeared in the petition. Pet. 81-83. Samsung also improperly incorporated Dr. Wolfe's testimony: the Petition cited 8 paragraphs from EX1003, spanning 4.5 pages. The Reply also cited pages of testimony. Nevertheless, neither EX1016, [0050] nor EX1003, ¶¶340-341 proposes modifying Kim's TSV lengths. The Petition only suggests using Riho's

die grouping for reducing load and compensating phase variations. Pet. 82.

Samsung also makes the new argument that a POSITA would have applied Riho's die grouping to Kim. Reply 16. Riho's die grouping would be the same as the second scenario discussed in the POR, and Samsung does not dispute its inoperability.[7] POR 28-30; Reply 11.





[7] Additionally, Riho's TSVs are in electrical communication with every die and output data in parallel with no two TSVs connected to the same data terminal as required by the claims. *See* discussions below.

Samsung improperly treats each part of the reference as a modular piece that it can reassemble at will. *ActiveVideo Networks, Inc. v. Verizon Commc'ns, Inc*., 694 F.3d 1312, 1327-28 (Fed. Cir. 2012) (rejecting expert's modular piece approach). As noted in the POR, Samsung has simply failed to consider how the combination would operate; and it still has not done so in the Reply. POR 51-52; Reply 16-17.

Furthermore, as explained in the POR, Riho does not teach selecting numbers of dies in a group in consideration of the load of TSVs or to minimize the differences in loads on different data conduits. EX1016, [0050]; EX2023, ¶¶203, 205-206; POR 53-54. Riho discloses only equalizing the TSV lengths to minimize skew, but nothing about load balance or recited die selection. *Id.* That each of Riho's TSVs can transmit data from two dies results from the fact that there are two chip-select groups, not from load considerations. EX1016, [0028]-[0029]. That a POSITA, after being presented with the solution, would understand that load balance could be achieved as claimed does not show that they would have had thought of doing so beforehand. *Contra* EX1003, ¶341. The record does not contain such evidence. The Board should reject Samsung's "impermissible *ex post* reasoning and hindsight bias" because it "use[s] the challenged patent as a roadmap to reconstruct the claimed invention using disparate elements from the prior art." *TQ Delta, LLC v. Cisco Sys., Inc.*, 942 F.3d 1352, 1361 (Fed. Cir. 2019).

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD, MICRON TECHNOLOGY, INC., MICRON SEMICONDUCTOR PRODUCTS, INC., and MICRON TECHNOLOGY TEXAS LLC,[1]
Petitioners,

v.

NETLIST, INC.,
Patent Owner.

Case No. IPR2022-01428
Patent No. 8,787,060

PATENT OWNER'S REQUEST FOR DIRECTOR REVIEW

[1] Micron Technology, Inc., Micron Semiconductor Products, Inc., and Micron Technology Texas LLC filed a motion for joinder and a petition in IPR2023-00882 and have been joined as petitioners to this proceeding.

TABLE OF CONTENTS

I. INTRODUCTION 1

II. BACKGROUND 2

III. LEGAL STANDARD 5

IV. ARGUMENT 6

A. The Board Abused its Discretion in Refusing to Order Petitioner to Provide its Expert's Inconsistent Deposition Testimony, Contradicting the Board's Regulations and Federal Circuit Law 6

B. The Board Was Not Legally Powerless and Should Have Authorized the Filing and Ordered Petitioner to Provide the Deposition Transcript for Use in this Proceeding 9

C. Director Review Is Necessary to Prevent Parties from Shirking Discovery Obligations with No Consequence 12

V. CONCLUSION 14

UNITED STATES PATENT AND TRADEMARK OFFICE

_____

BEFORE THE PATENT TRIAL AND APPEAL BOARD

_____

SAMSUNG ELECTRONICS CO., LTD., MICRON TECHNOLOGY, INC., MICRON SEMICONDUCTOR PRODUCTS, INC., and MICRON TECHNOLOGY TEXAS LLC,[1]

Petitioners,

v.

NETLIST, INC.,

Patent Owner.

_____

Case No. IPR2022-01428
Patent No. 8,787,060

_____

PATENT OWNER NETLIST, INC.'S NOTICE OF APPEAL

[1] Micron Technology, Inc., Micron Semiconductor Products, Inc., and Micron Technology Texas, LLC, filed a motion and joinder and a petition in IPR2023-00882 and have been joined as petitioners in this proceeding.

Pursuant to 35 U.S.C. §§ 141, 142, and 319, and in accordance with 37 C.F.R. §§ 90.2-90.3, Patent Owner Netlist, Inc. ("Netlist") appeals to the United States Court of Appeals for the Federal Circuit from the Final Written Decision of the Patent Trial and Appeal Board ("Board") entered on April 1, 2024 in IPR2022-01428 (Paper No. 48) ("Final Written Decision"), attached as Exhibit A; the Order denying Director Review of the Final Written Decision dated June 17, 2024 (Paper 52) ("Decision on Director Review Request"), attached as Exhibit B; and from all underlying findings, determinations, rulings, opinions, orders, and decisions regarding the *inter partes* review (Case IPR2022-01428) of U.S. Patent No. 8,787,060 (the "'060 Patent").

In accordance with 37 C.F.R. § 90.2(a)(3)(ii), Netlist states that the issues on appeal include, but are not limited to: the Board's determination that Claims 1 through 34 of the '060 Patent have been shown by a preponderance of the evidence to be unpatentable; the Board's construction of the challenged claims and application of its construction of the claims to the facts of record; the Board's procedural rulings, including its rulings regarding the adequate and timely preservation of certain of the parties' arguments and denial of submission of supplemental evidence that came into being after the sur-reply; the adequacy of the Board's consideration of the expert testimony, prior art, and other evidence in the record, including, but not limited to, evidence and testimony from related district court proceedings between Petitioner

and Patent Owner; the Board's factual findings, conclusions of law, or other determinations supporting or related to those issues (such as motivation to combine or reasonable expectation of success); the Board's compliance with the Administrative Procedure Act, including whether the Final Written Decision, denial of Director Review and the denial of Netlist's request for submitting supplemental evidence are arbitrary, capricious, an abuse of discretion, not in accordance with law, or in excess of the Board's jurisdiction, and any procedural irregularities associated with the review proceeding; as well as all other issues decided adversely to Netlist in any orders, decisions, rulings, and opinions.

This Notice of Appeal is being e-filed with the Clerk's Office for the United States Court of Appeals for the Federal Circuit, along with payment of the required docketing fees. In addition, a copy of this Notice of Appeal is being filed simultaneously with the Patent Trial and Appeal Board and filed by e-mail to the Director of the United States Patent and Trademark Office at efileSO@uspto.gov.

Dated: August 19, 2024

Respectfully submitted,

/Hong Zhong/

H. Annita Zhong (Reg. No. 66,530)
Jason Sheasby (*pro hac vice*)
IRELL & MANELLA LLP
1800 Avenue of the Stars, Suite 900
Los Angeles, CA 90067
Tel: (310) 277-1010
Fax: (310) 203-7199
HZhong@irell.com
JSheasby@irell.com

Blair A. Silver (Reg. No. 68,003)
IRELL & MANELLA LLP
750 17th Street NW, Ste. 850
Washington, DC 20006
Tel: (202) 777-6500
bsilver@irell.com

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. § 42.6, the undersigned certifies that on August 19, 2024, a copy of the foregoing **PATENT OWNER'S NOTICE OF APPEAL** was served by electronic mail, as agreed to by the parties, upon the following:

BAKER BOTTS L.L.P.
Eliot D. Williams, Reg. No. 50,822
Theodore W. Chandler, Reg. No. 50,319
Ferenc Pazmandi, Reg. No. 66,216
Michael E. Knierim, Reg. No. 78,268
Brianna L. Potter, Reg. No. 76,748
DLSamsungNetlistIPRs@BakerBotts.com

WINSTON & STRAWN LLP
Matthew A. Hopkins, Reg. No. 76,273
mhopkins@winston.com
Michael R. Rueckheim, *pro hac vice*
mrueckheim@winston.com
David Lin
Winston-IPR-Netlist@winston.com

I also certify that in addition to being filed electronically with the Board, a copy of this Notice of Appeal was filed on August 19, 2024, for delivery to the Director of the United States Patent and Trademark Office at the following e-mail address: efileSO@uspto.gov.

I further certify that a copy of the foregoing Notice of Appeal is being filed via CM/ECF on August 19, 2024, with the United States Court of Appeals for the Federal Circuit.

/Susan M. Langworthy/
Susan M. Langworthy

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.,
Petitioner,

v.

NETLIST, INC.,
Patent Owner

Patent No. 8,787,060
Issued: July 22, 2014
Filed: November 3, 2011
Inventor: Hyun Lee

Title: Method and Apparatus for Optimizing Driver Load in a Memory Package

---

Inter Partes Review No. IPR2022-01428

DECLARATION OF DR. ANDREW WOLFE REGARDING U.S. PATENT NO. 8,787,060

Samsung Electronics Co., Ltd.
Ex. 1003, p. i

144. Ground 2 involves the references in Ground 1 (Kim in view of Rajan) in combination with Riho.

145. Ground 3 involves the references in Ground 1 (Kim in view of Rajan) in combination with Wyman.

146. Ground 4 involves Riho in view of Rajan.

147. Ground 5 involves the references in Ground 4 (Riho in view of Rajan) in combination with Riho2.

V. COMPARISON OF THE PRIOR ART TO THE CLAIMS OF THE 060 PATENT

148. My analysis below is based on my understanding of the law and the other information discussed above, including the five Grounds discussed above.

A. Reasons to Combine for Grounds 1, 2, and 3

149. A Skilled Artisan would have been motivated to make the combinations in Grounds 1, 2 and 3 for at least the following reasons.

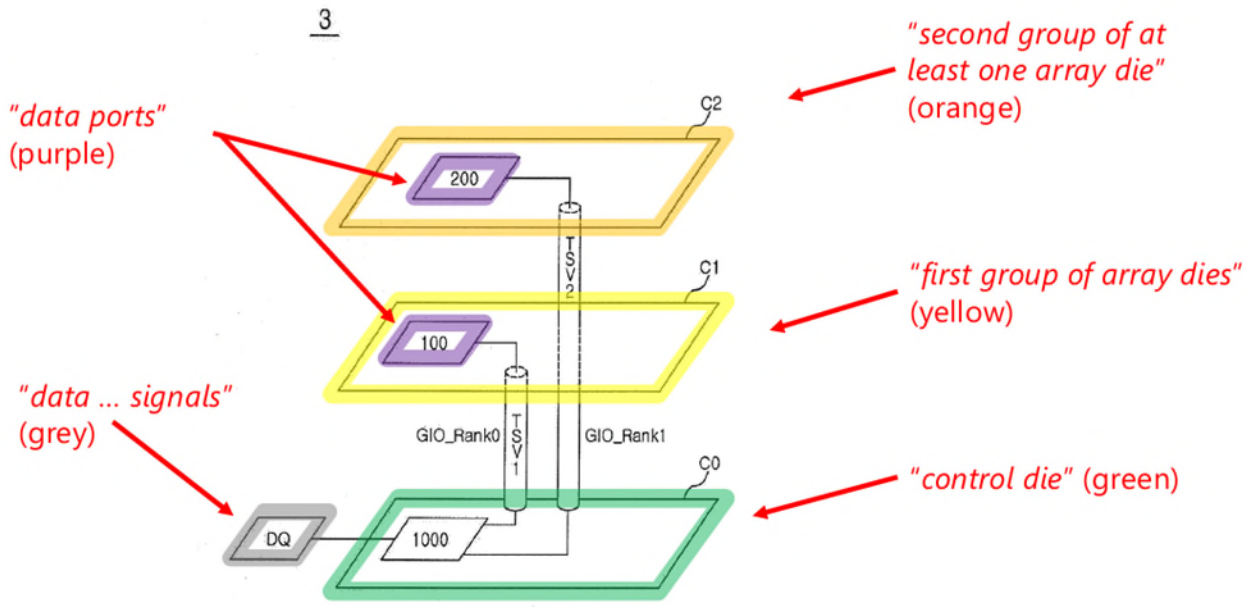
1. Ground 1: Kim in view of Rajan

150. As mentioned above, Ground 1 involves Kim in view of Rajan. *Supra* ¶143.

151. As discussed below in detail with reference to specific claim limitations, Kim discloses a memory package that includes stacked memory chips, like C1 (yellow) and C2 (orange), and a main chip C0 (green) including a shared data input/output section 1000. Ex. 1014 (Kim) at Fig. 5 (reproduced below,

annotated). Kim expressly discloses a data pad DQ (grey) to communicate data with external devices and further explains that "write and read command signals [are] applied from a device external to the semiconductor memory apparatus." *Id.* at [0038]. A Skilled Artisan would have understood from this disclosure that Kim's memory package includes "*input/output terminals"* to communicate "*data and control/address signals with one or more external devices*" as required by the 060 Patent claims.

FIG.5



152. Furthermore, the claimed terminals and corresponding data and control/address signals would have also been obvious over Kim in view of the teachings of Rajan. *See, e.g.*, Ex. 1015 (Rajan) at Fig. 4 (reproduced below, annotated, showing a memory package including stacked memory chips (yellow

and orange) and a buffer chip (green) which has input/output terminals for data and control/address signals communicated with a host system).

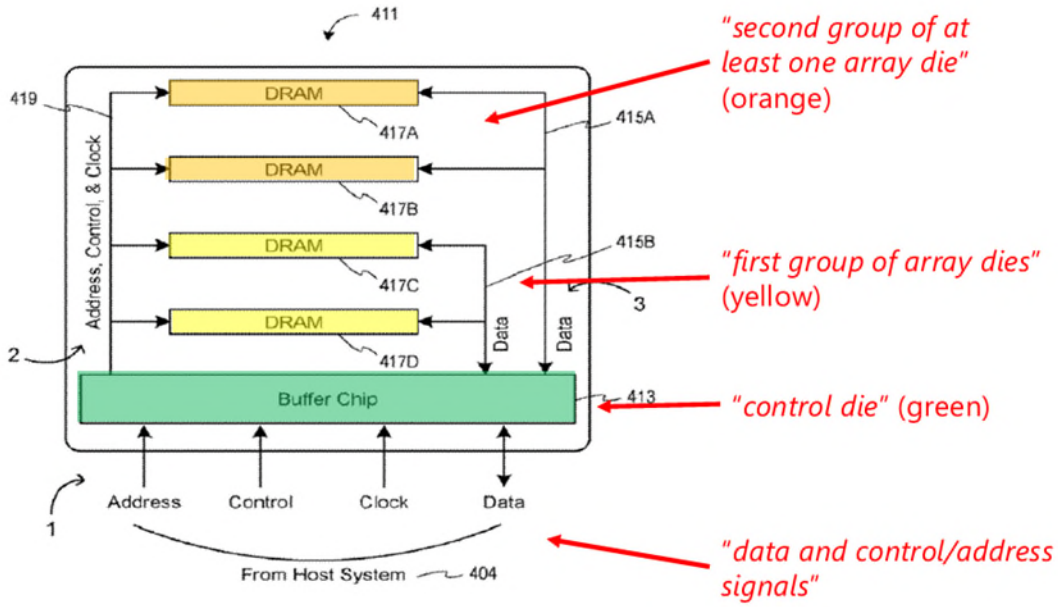


FIG. 4

153. For example, a Skilled Artisan would have been motivated to implement the functionality of a JEDEC compatible external interface as suggested by Rajan in the control chip of Kim to make Kim's stacked memory package compatible with the corresponding JEDEC standard. *See, e.g.*, Ex. 1015 (Rajan) at 8:8-11 ("Behavior of many DRAM circuits is specified by the JEDEC standards, and it may be desirable to exactly emulate a particular JEDEC standard DRAM."), 3:52-54 ("The interface circuit may present to the system device an emulated interface to e.g. DDR3 memory, while the physical memory chips are, in fact, DDR2 memory."); *see also id*. at Fig. 18 (reproduced below, disclosing that an

interface circuit includes emulation logic which interfaces, on one side with the host system and, on the other side, with the physical memory circuits).

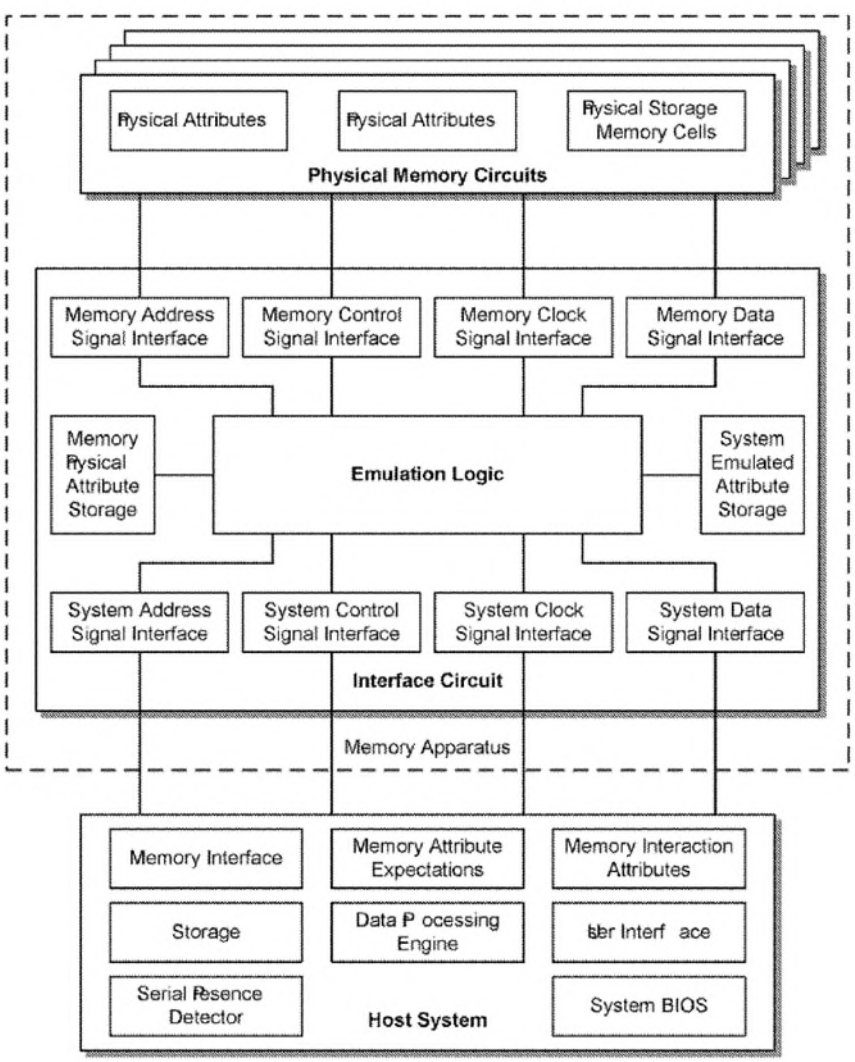


FIG. 18

154. Rajan is analogous art to Kim and the 060 Patent, as each is directed to the problem of effective operation of memory devices, including interfacing with stacked memory devices. Ex. 1015 (Rajan) at 1:51-60 ("What is needed, then, is an effective way to make use of low cost memory chips in manufacturing high capacity memory modules. … A memory subsystem is provided including an

interface circuit adapted for communication with a system and a majority of address or control signals of a first number of memory circuits."), 4:48-50 ("In one embodiment, the DRAM chips are stacked, and the buffer chip is placed electrically between them and the host system."); Ex. 1014 (Kim) at [0012] ("Various aspects of the present invention comprise a semiconductor memory apparatus capable of sharing a data input/output circuit."), [0050] ("It is to be understood that the shared data input/output section 1000 according to the embodiment of the present invention can be applied to a semiconductor memory apparatus in which a plurality of chips are stacked."); Ex. 1001 (060 Patent) at 1:18-21 ("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."), 5:23-26 ("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).").

155. Therefore, a Skilled Artisan would have been motivated to look at analogous art, like Rajan, for details about control and address signals for memory packages. Rajan discloses the same basic structure of a memory package like Kim and the 060 Patent: the package includes stacked memory chips and a shared interface circuit, called "buffer chip" in Rajan. Ex. 1015 (Rajan) at Fig. 4

(reproduced below, annotated), 2:6-7 ("FIG. 4 shows a buffered stack of DRAM circuits having one address, control, and clock bus and two data busses.").

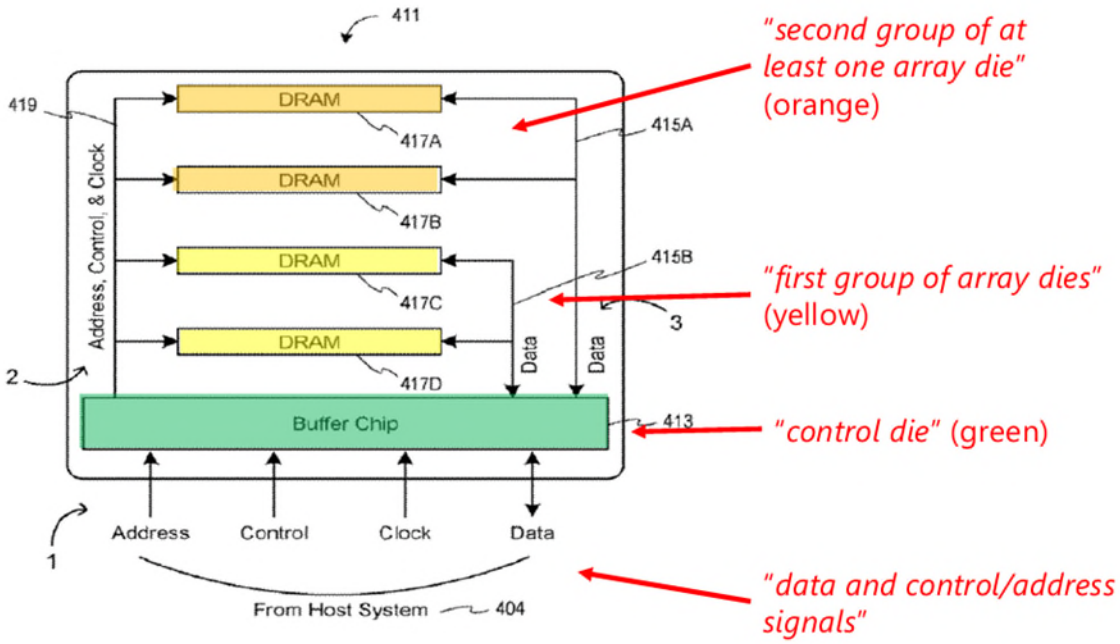


FIG. 4

156. Rajan describes that "the buffer chip 413 is interfaced to a host system 404," including the use of address, control, and data signals. Ex. 1015 (Rajan) at 5:36-43, Fig. 4 (reproduced above, annotated). Rajan also discloses that the interface may comply with JEDEC standards. *See id.* at 4:20-24 ("The interface circuit may take the form of or incorporate, or be incorporated into, a register, an AMB, a buffer, or the like, and may comply with Joint Electron Device Engineering Council (JEDEC) standards, and may have forwarding, storing, and/or buffering capabilities."). A Skilled Artisan would have understood that a JEDEC standard interface includes terminals for address, control, and data signals. *See,*

e.g., Ex. 1019 (JESD79-3C) at p.12 (Fig. 3 describing package pinout for stacked DDR3 memories, including terminals DQ0-15 for x16 data signals, and terminals for control/address signals, including RAS, CAS, WE, An, and BAn), Fig. 3 (reproduced below).

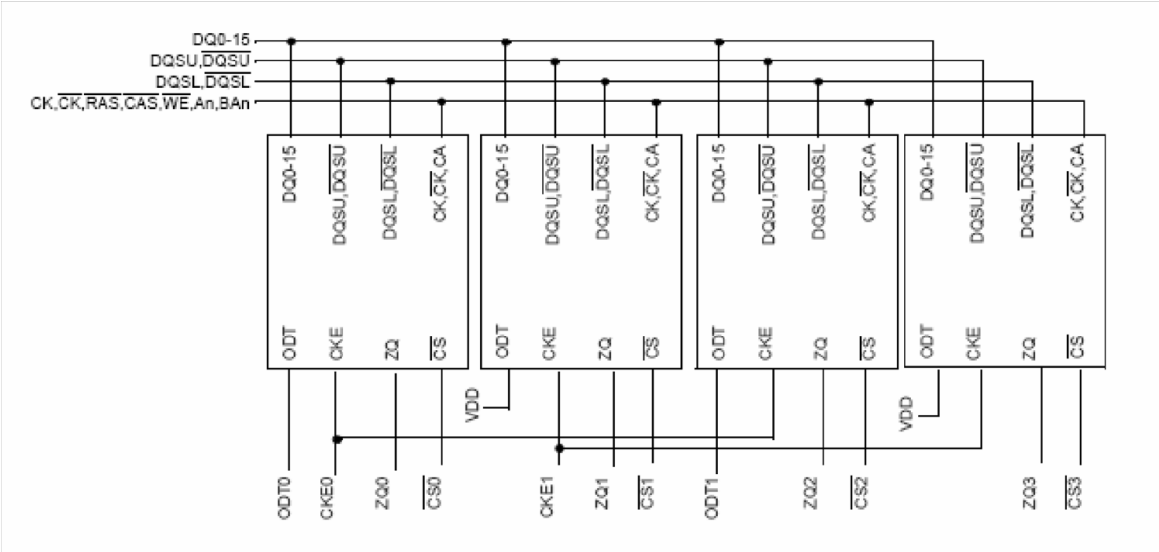


Figure 3 — Qual-stacked / Quad-die DDR3 SDRAM x16 rank association

157. A Skilled Artisan would have been motivated to follow Rajan's suggestion and implement address, control, and data terminals according to the JEDEC standards in Kim's memory package in order to use the package in JEDEC compliant systems. Making such a combination would have been well within the level of skill at the time, since emulating a standard JEDEC interface was known for years, as evidenced by Rajan and as also demonstrated by the prior art referenced in the 060 Patent. *See, e.g.*, Ex. 1001 (the 060 Patent) at 22:37-39 ("Embodiments of rank multiplication are described in greater detail in U.S. Pat.

Nos. 7,289,386 and 7,532,537 …."); Ex. 1028 (U.S. Pat. No. 7,289,386); Ex. 1031 (U.S. Pat. No. 7,532,537). As discussed above, the JEDEC standards were influential and well known. *See supra* ¶¶ 132-134; Ex. 1022, pp. 332-35, 318-20.

158. Furthermore, Kim explains that, "although only one main chip and two slave chips are shown [in Fig. 5], it is to be understood that any number of main and slave chips may be used." Ex. 1014 (Kim) at [0048]; *see also id*. at [0050] ("It is to be understood that the shared data input/output section 1000 according to the embodiment of the present invention can be applied to a semiconductor memory apparatus in which a plurality of chips are stacked. While the semiconductor memory apparatuses having two ranks are explained with reference to FIGS. 2 and 5, a person having ordinary skill in the art will appreciate that the technical concept of the present invention can be applied to a semiconductor memory apparatus which are divided into three or more ranks.").

159. A Skilled Artisan would have understood from this disclosure that Kim's invention can include more slave chips (and more ranks) and that some of those additional slave chips proposed by Kim can be in the same group as chips C1 or C2 and connected to TSV1 or TSV2, respectively. Furthermore, to the extent one might argue that Kim does not sufficiently disclose a group of multiple memory chips connected to the same TSV, it would have been obvious in light of Rajan, including Figure 4 of Rajan shown below.

160. As discussed above, both Rajan and Kim are analogous art to the 060 Patent, and a Skilled Artisan would have been motivated to look at analogous art, like Rajan, for details about adding more chips (and ranks) in the stack as proposed by Kim. Ex. 1015 at Fig. 4 (reproduced below, annotated).

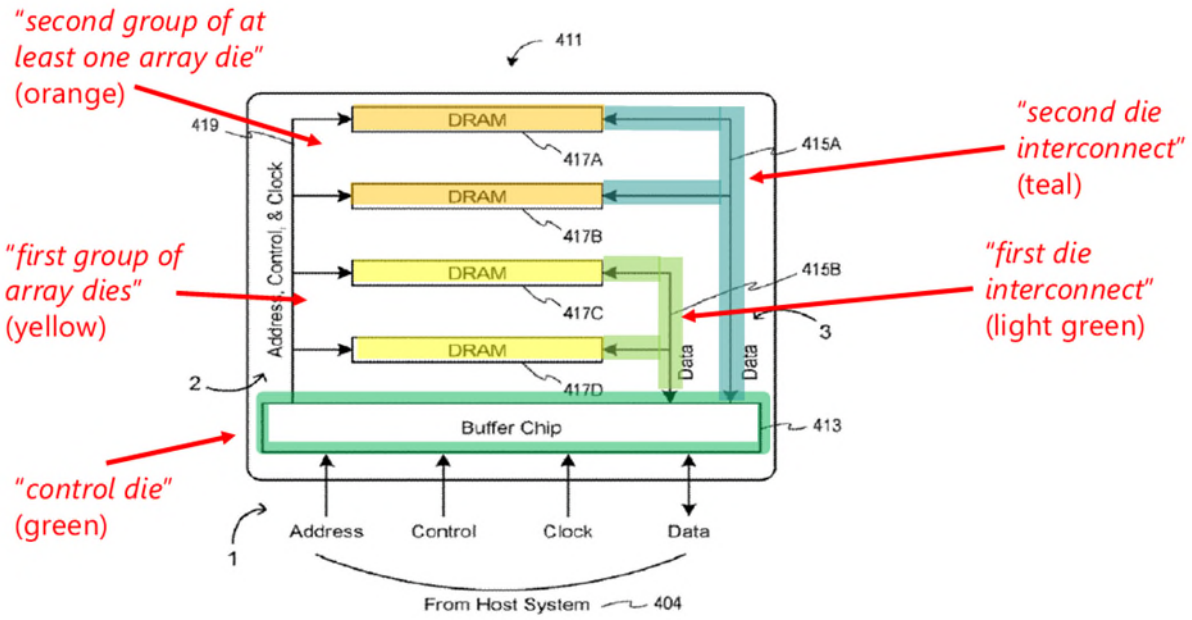


FIG. 4

161. For example, Rajan explains that "FIG. 4 illustrates another embodiment of a system 411 in which the buffer chip 413 is interfaced to a host system 404 and is coupled to the DRAM chips 417A-417D somewhat differently than in the system of FIG. 2 or 3. There is a shared address, control, and clock bus 419, and a plurality of shared data busses 415A and 415B. Each shared bus has two or more DRAM chips coupled to it." *Id.* at 5:36-43.

162. A Skilled Artisan would have also understood that creating such a shared data bus (415A, 415B above) as taught by Rajan was well within the level of skill at the time using, e.g., TSV interconnects like Kim. *See, e.g.*, Ex. 1025 (Foster) at Fig. 5 (reproduced below, annotated, showing two dies 21 and 22 (yellow) in a first substack being coupled to signal lines 68 by TSV 30, and two other dies 23 and 24 (orange) in a second substack being coupled to signal lines 70 using TSV 30 and pass-through vias PTV 29 which have no connection to circuitry in dies 21 and 22 in the first substack).

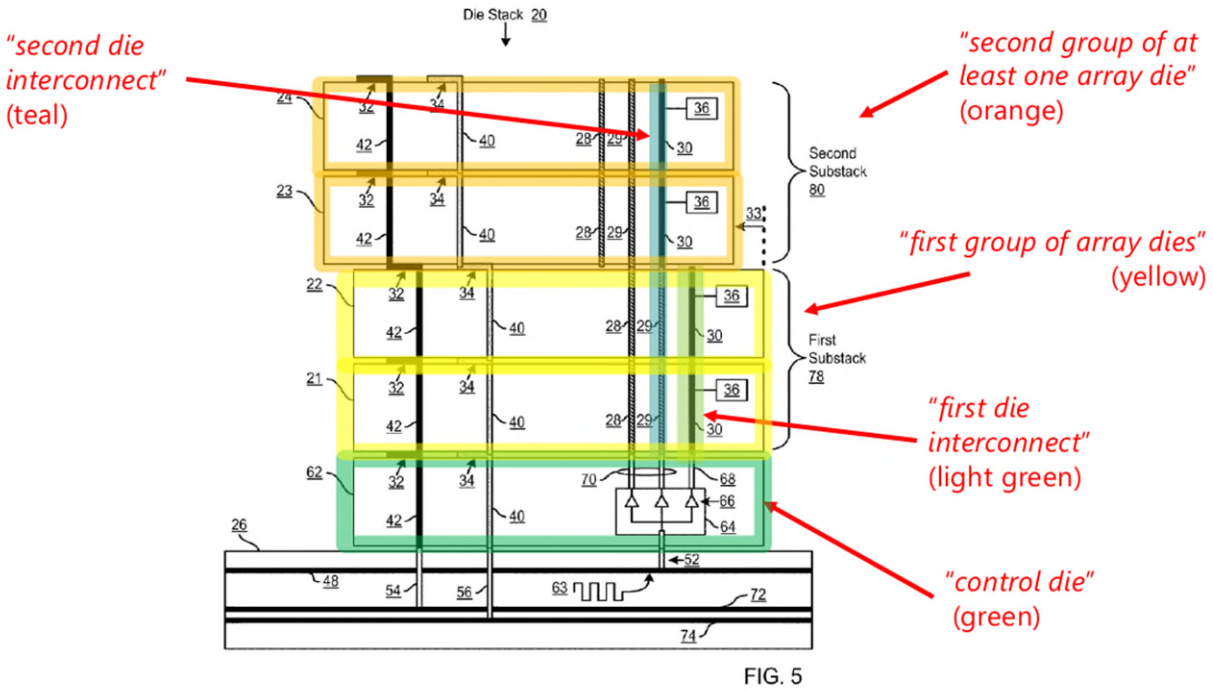


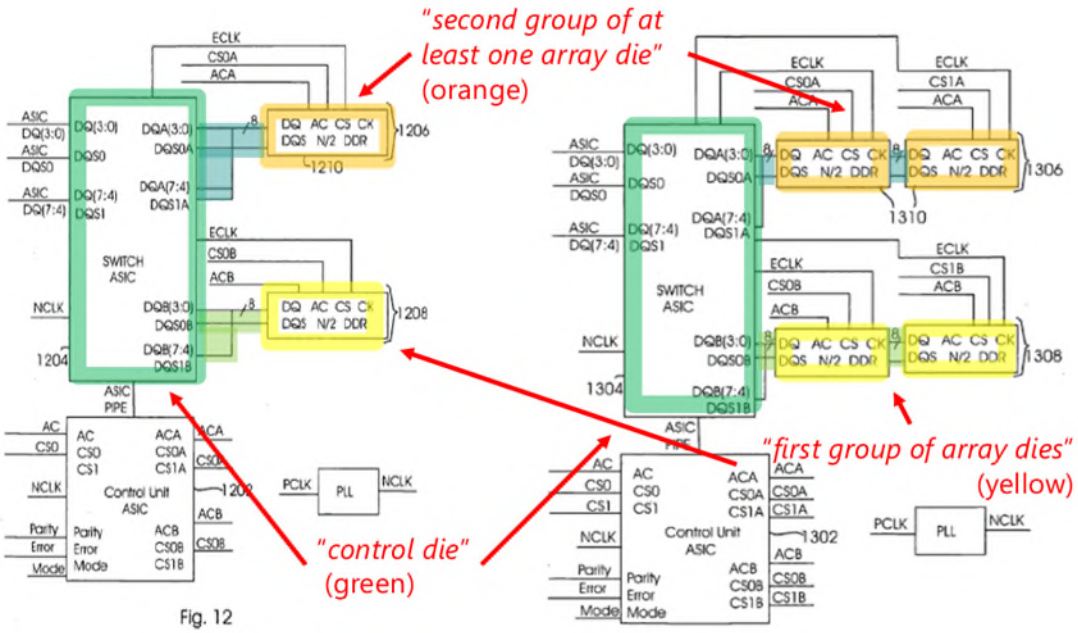
FIG. 5

163. Therefore, a Skilled Artisan would have understood that there were a finite number of known ways to connect the additional dies, including sharing a through-silicon via by a subset of the dies, so it would have been obvious to a

Skilled Artisan to follow this well know architecture when adding memory dies to the stack as suggested by Kim. *See*, *e.g.*, Ex. 1015 (Rajan) at Figs. 2-6 (describing a finite number of different interconnections of stacked memory chips).

164. Furthermore, a Skilled Artisan would have been motivated to connect the additional dies to the existing interconnects TSV1 or TSV2 in Kim's device, because such an arrangement could use Kim's chips without changing the number of interconnects TSV1 and TSV2, thus without creating new TSVs, which would require space and additional circuitry. Such a use of the interconnects TSV1 and TSV2 would also allow emulating certain characteristics required by external devices. Ex. 1015 (Rajan) at 3:27-30 ("For example, the interface circuit may tell the system device that the number of emulated memory circuits is different than the actual number of physical memory circuits."), 3:52-61 ("The interface circuit may present to the system device an emulated interface to e.g. DDR3 memory, while the physical memory chips are, in fact, DDR2 memory. The interface circuit may emulate an interface to one version of a protocol such as DDR2 with 5-5-5 latency timing, while the physical memory chips are built to another version of the protocol such as DDR2 with 3-3-3 latency timing. The interface circuit may emulate an interface to a memory having a first capacity that is different than the actual combined capacity of the physical memory chips.").

165. Indeed, a Skilled Artisan would have understood that Rajan's emulation techniques, including rank multiplication techniques, used "fork-in-the-road" arrangements where a two-rank arrangement has two data busses, and two more ranks can be added by sharing the existing two data busses of the two-rank arrangement. *See, e.g.*, Ex. 1026 (US 2006/0277355) at Figs. 12 and 13 (reproduced below, annotated, showing a switch ASIC connecting an 8-bit wide external DQ bus to two 8-bit wide internal busses DQA and DQB, and adding two more ranks in Figure 13 by sharing the DQA and DQB busses already used in the two-rank arrangement of Figure 12).



166. A Skilled Artisan's understanding that rank multiplication techniques used "fork-in-the-road" arrangements is further evidenced by the prior art

referenced in the 060 Patent. *See, e.g.*, Ex. 1001 (the 060 Patent) at 19:41-47, 22:34-57 (explaining that rank multiplication was well known prior art and incorporating by reference corresponding prior art patents disclosing fork-in-the-road arrangements, including Exhibits 1028 and 1031), 18:65-19:7 (describing prior-art memory package with a "programmable secondary 1-to-2 rank decoder" in the controller die of the memory package); *see also supra* at ¶¶138-141.

167. A Skilled Artisan would have been further motivated to use the interconnects TSV1 and TSV2 when adding more chips, because Kim explains that its structure of using two interconnects can avoid certain data collisions. Kim at Figs. 4A (explaining the problem of collision in the prior art), 4B (explaining the operation of the invention avoiding the collision), [0042-44] (same). Therefore, when adding more memory dies to Kim's stack, a Skilled Artisan would have been motivated to connect the additional dies to Kim's interconnects TSV1 and TSV2 as taught by Rajan to maintain Kim's data collision avoidance feature without creating additional TSVs.

168. For at least all these reasons, it would have been obvious to a Skilled Artisan to combine the teaching of Kim and Rajan to implement Kim's stacked memory package with the functionality of Rajan's interface circuit and inter-die connections.

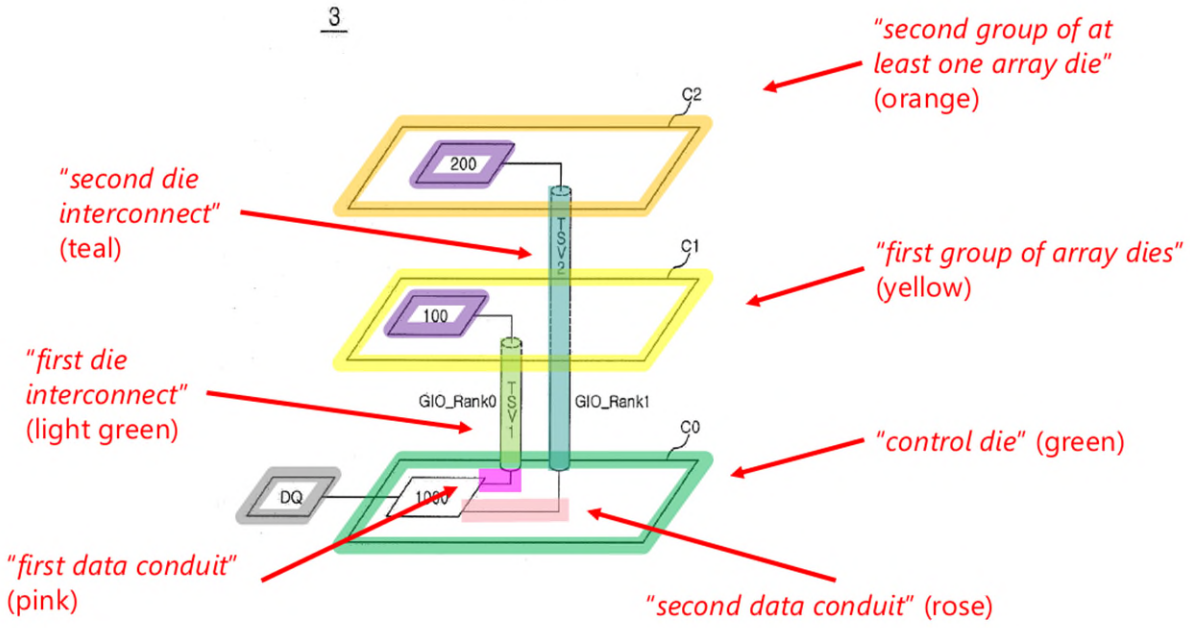
3. Ground 3: Ground 1 in view of Wyman

174. As mentioned above, Ground 3 involves Ground 1 (the combination of Kim and Rajan) in view of Wyman. *See supra* ¶145.

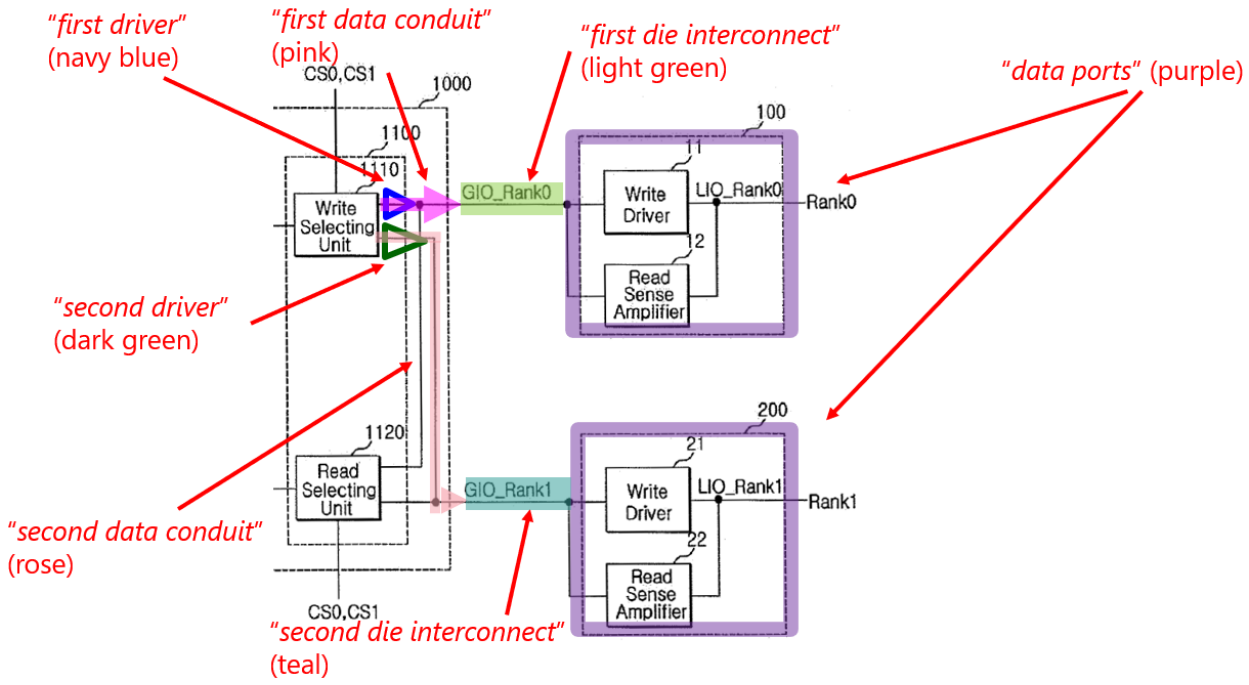
175. Wyman teaches to save power by using drivers that depend on the length (and load) of the connection in the stack. *See, e.g.*, Ex. 1017 (Wyman) at 1:22-24 ("In instances where chips are stacked, creating shorter runs, utilizing the full capacity of a driver would be both wasteful and inefficient since far less power is required.").

176. As expressly shown in annotated Fig. 5 of Kim (reproduced below), Kim discloses that TSV1 ("*first die interconnect*," light green) is shorter than TSV2 ("*second die interconnect*," teal). Ex. 1014 at Fig. 5 (reproduced below, annotated).

FIG.5



177. A Skilled Artisan would have understood that, due to the difference in length between TSV1 and TSV2, each of the "*first data conduit*" (pink) and the "*second data conduit*" (rose) can include a respective driver that is optimized for driving the data signal through the corresponding TSV in an efficient way. For example, Wyman explains that "[i]n instances where chips are stacked, creating shorter runs, utilizing the full capacity of a driver would be both wasteful and inefficient since far less power is required." Ex. 1017 (Wyman) at 1:22-24. Therefore, a Skilled Artisan would have been motivated to follow the teaching of Wyman in the system of Kim and, instead of using a full capacity driver to drive both TSV1 and TSV2, implement a respective driver for each of TSV1 and TSV2 that is optimized for the length of the corresponding TSV as illustrated in annotated Fig. 3 of Kim. Ex. 1014 (Kim) at Fig. 3 (reproduced below in part, annotated, showing a "*first driver*," navy blue, corresponding to TSV1 and a "*second driver*," dark green, corresponding to TSV2). Implementing Kim's system in this manner avoids the "wasteful and inefficient" use of power as taught by Wyman.



178. Wyman is analogous art to the 060 Patent and Kim, as each is directed to the problem of efficient signaling between chips in a stacked arrangement. Ex. 1017 (Wyman) at 1:45-48 ("The reduced path distances of stacked hybridized chips therefore require less drive to adequately drive a signal to an off-chip board and/or chip. These conditions warrant the use of restructuring the drive portion of existing chip architecture."); Ex. 1001 (the 060 Patent) at 1:19-21 ("More specifically, the present disclosure relates to systems and methods for reducing the load of drivers of memory packages included on the memory modules."), 5:23-26 ("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)."); Ex. 1014 (Kim) at [0003] ("Various embodiments of the present disclosure generally relate to a semiconductor memory apparatus, and more particularly, to data input/output of a

semiconductor memory apparatus."), [0046] ("In order to increase the integration of a semiconductor apparatus, a three-dimensional (3D) semiconductor apparatus comprising a plurality of stacked chips has been developed. The stacked chips provide a structure that enables the 3D semiconductor apparatus to be packages in a single package. Recently, a through-silicon via (TSV) type semiconductor apparatus has been developed in which silicon vias are formed to pass through the plurality of stacked chips so that all of the chips are electrically connected to one another.").

179. A Skilled Artisan would have understood, and been motivated, to turn to analogous art, like Wyman teaching about efficiently driving signals between stacked devices, in order to implement the stacked devices of Kim. For example, Wyman discloses that longer connections in a stacked arrangement require additional drive. *See, e.g.*, Ex. 1017 (Wyman) at Fig. 8 (reproduced below).

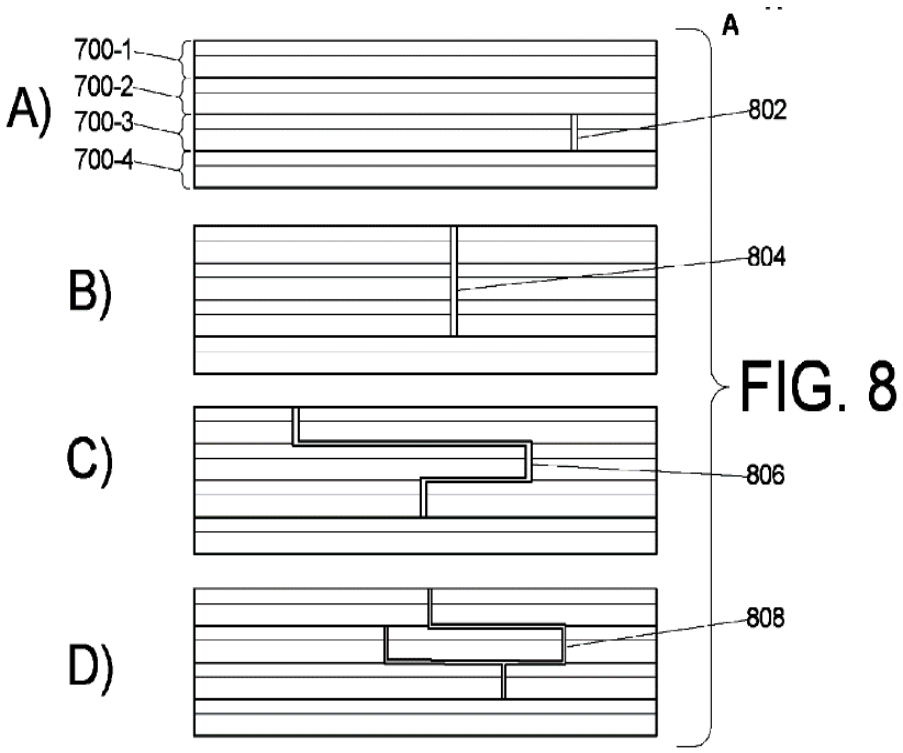


FIG. 8

Wyman explains that "FIG. 8A illustrates, in simplified form, a through-chip via connection (802) extending between the mother chip (700-4) and daughter 1 (700-3) … requir[ing] minimal drive … FIG. 8B illustrates, in simplified form, a longer through-chip via connection (804) extending directly between the mother chip (700-4) and daughter 3 (700-1) in a straight line. The increased resistance, capacitance and impedance of such a connection (804) might require additional drive than that referred to in connection with FIG. 8A. …" *Id*. at 6:15-30. Thus, "the output drive current [can] be modified to meet the requirements of a particular output load." *Id.* at 7:13-18. However, Wyman recognizes that even with the "significantly longer connection (808) between the mother chip (700-4) with a fan

out to both daughter 2 (700-2) and daughter 3 (7001)," as shown in Fig. 8.D, using "the total drive … would be overkill." *Id.* at 6:44-50.

180. A Skilled Artisan would have understood from this disclosure that the longer the connection in a stack of chips, the more drive it requires, but using the "total drive" would be wasteful. Ex. 1017 (Wyman) at 1:22-24 ("In instances where chips are stacked, creating shorter runs, utilizing the full capacity of a driver would be both wasteful and inefficient since far less power is required."). Therefore, a Skilled Artisan would have been motivated to use drivers of different strengths (i.e., a "*first driver*" having a "*first driver size*" and a "*second driver*" having a "*second driver size*") for Kim's different sized through-silicon vias TSV1 and TSV2. *See also supra* at ¶¶135-137 (explaining that different strengths are achieved by transistors of different sizes). In particular, a Skilled Artisan would have been motivated to drive the shorter TSV1 with a smaller driver and the longer TSV2 with a larger driver in order to use power in an efficient manner as taught by Wyman. Making such a combination was well within the level of skill, since Wyman itself teaches drivers of different sizes, and the combination would have provided nothing more than expected, a power efficient drive of signals through Kim's TSVs.

181. Therefore, it would have been obvious to a Skilled Artisan to implement Wyman's teaching of using drivers optimized for the length of the corresponding die interconnects in the memory package of Kim to save power.

B. Claims 1-34 Are Unpatentable Over Grounds 1-3

182. As explained below, claims 1-6, 8-14, 16-19, and 29-34 are unpatentable over Ground 1 (Kim in view of Rajan); claims 1-14, 16-19, and 29-34 are unpatentable over Ground 2 (Ground 1 in view of Riho); and claims 1-6 and 8-34 are unpatentable over Ground 3 (Ground 1 in view of Wyman).

1. Claim 1 is Unpatentable

183. Claim 1 is unpatentable as the prior art references in Grounds 1-3 render the claim obvious.

a) [1.a] Preamble

184. The preamble of claim 1 requires "*[a] memory package, comprising*."

185. Kim discloses a "*memory package*" including main chip C0 and stacked memory chips C1 and C2 as illustrated in Kim's Fig. 5. *See* Ex. 1014 at Fig. 5 (reproduced below, annotated).

2 (reproduced bellow, disclosing that the basic organization of DRAMs includes a memory array); Ex. 1022 (Jacob et al.) at 316-320 (same).

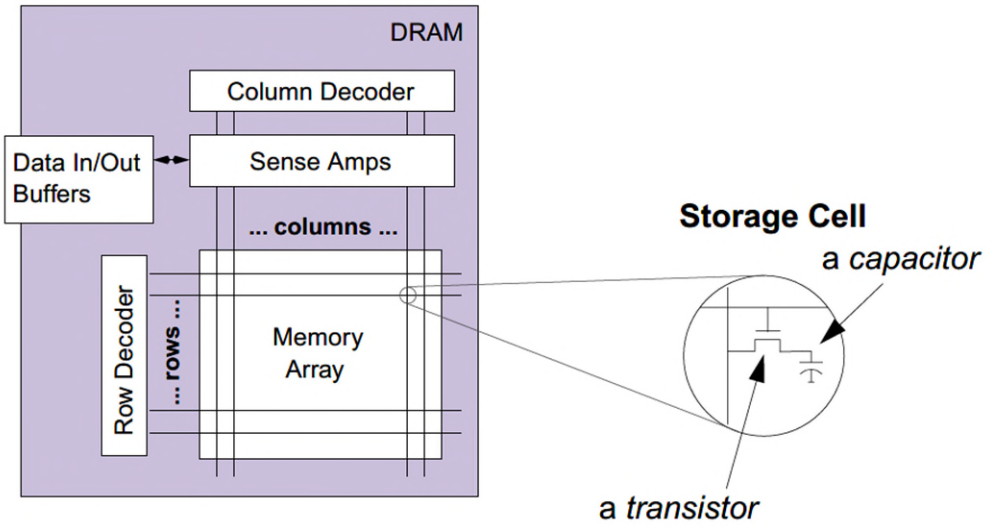


Figure 2: Basic Organization of DRAM Internals
The DRAM memory array is a grid of storage cells, where one bit of data is stored at each intersection of a *row* and a *column*.

Therefore, Kim also discloses and renders obvious to a Skilled Artisan that the slave chips C1 and C2 are "*array dies*" as required by the claim.

218. Kim also discloses that slave chip C1 is in a "*first group of array dies*" coupled to through-silicon via TSV1, and slave chip C2 is in a "*second group of at least one array die*" coupled to through-silicon via TSV2. Ex. 1014 at Fig. 5 (reproduced above, annotated).

219. To the extent one might argue that the "*first group of array dies*" requires multiple slave chips being coupled to the through-silicon via TSV1, Kim discloses and renders obvious such an arrangement to a Skilled Artisan. For example, Kim explains that "although only one main chip and two slave chips are

shown [in Fig. 5], it is to be understood that any number of main and slave chips may be used." Ex. 1014 at [0048]. Furthermore, Kim explains that "the shared data input/output section 1000 according to the embodiment of the present invention can be applied to a semiconductor memory apparatus in which a plurality of chips are stacked. While the semiconductor memory apparatuses having two ranks are explained with reference to FIGS. 2 and 5, a person having ordinary skill in the art will appreciate that the technical concept of the present invention can be applied to a semiconductor memory apparatus which are divided into three or more ranks." *Id.* at [0050].

220. A Skilled Artisan would have understood from this disclosure that Kim's invention is not limited to the exemplary two slave chips C1 and C2 and two ranks and that the stack can include more slave chips and/or more ranks. A Skilled Artisan would have also understood that some of the additional slave chips proposed by Kim can be in the same group as chips C1 or C2 and connected to TSV1 or TSV2, respectively. Indeed, such an arrangement of multiple groups of the stacked dies, where each die in a group is connected to the same data interconnect, was commonly known and used to implement multiple ranks in memory stacks at the time. *See, e.g.*, Ex. 1024 (Ruckerbauer (US7,796,446)) at Fig. 5 (reproduced below, annotated), 9:14-18 ("[W]ithin the memory stack 100, a first sub set of memory dies is coupled to the first transfer data path 12a and a

second sub set is coupled to the second transfer data path 12b. These different sub sets could, for example, be associated to different ranks within the stack 100.").

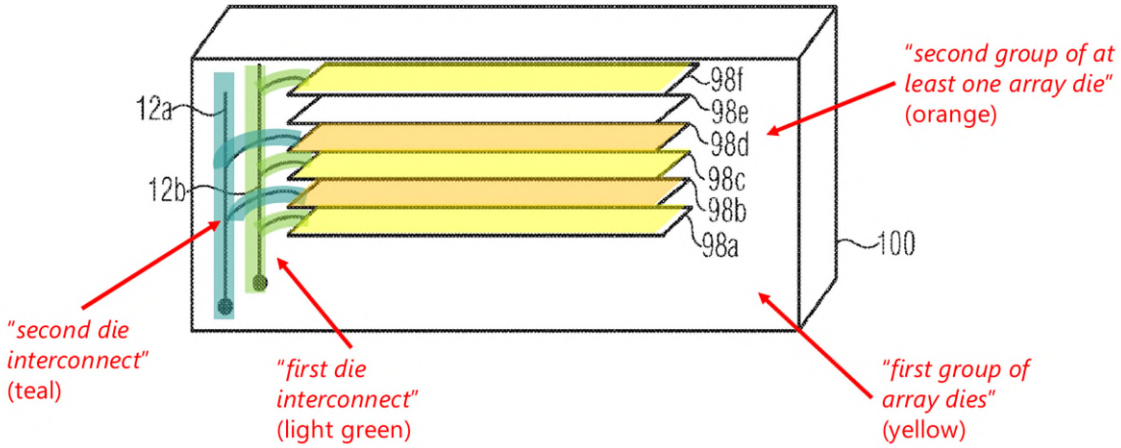


FIG 5

221. To the extent one might argue that Kim alone does not render obvious the additional chips sharing the data connections TSV1 and TSV2, it would have been obvious in view of Rajan. As discussed above, a Skilled Artisan would have been motivated to look to Rajan for details about adding more chips in the stack as proposed by Kim. *Supra* ¶¶150-168. For example, Rajan explains that "FIG. 4 illustrates another embodiment of a system 411 in which the buffer chip 413 is interfaced to a host system 404 and is coupled to the DRAM chips 417A-417D somewhat differently than in the system of FIG. 2 or 3. There is a shared address, control, and clock bus 419, and a plurality of shared data busses 415A and 415B.

chip." Ex. 1016 at [0050]. Riho describes that "in this embodiment, each I/O group is formed by ×32 DQ signal bits and the controller chip can perform communication control of the plurality of I/O groups (×256 DQ signal bits) with a single synchronization signal and with high accuracy." Ex. 1016 at [0050]; *see also id.* at [0053-54].

341. A Skilled Artisan would have understood from this disclosure that, in the combination of Kim and Riho, "*a first number of* [two] *array dies in the first group of array dies* [sharing a data signal TSV] *and a second number of at least one array die* [i.e., two dies] *in the second group of at least one array die* [sharing another data signal TSV] *are selected in consideration of a load of the first die interconnect and a load of the second die interconnect* [including the load of each of those data signal TSVs] *so as to reduce a difference between a first load on the first data conduit and a second load on the second data conduit* [which a Skilled Artisan understood reduces skew between the data signals]." *Supra* at ¶¶135-137. For example, a Skilled Artisan would have understood that the larger the capacitive load, the larger the delay in the transition of the driven signal. Ex. 1030 (Brown) at p.134. Accordingly, a difference in load may result in different signal timing, resulting in a skew between different signals. For example, DQ and DQS signals should be synchronized, but a skew between them may result in an erroneous read of the data value. Ex. 1019 (JESD79-3C) at 13 ("The data strobes DQS, DQSL,

and DQSU are paired with differential signals DQS#, DQSL#, and DQSU#, respectively, to provide differential pair signaling to the system during reads and writes."). Thus, a Skilled Artisan would have also understood that, when each data (and strobe) signal TSV is made to have substantially the same length, and each data (and strobe) signal TSV is connected to the same number of two SDRAMs, the difference in load for driving the data (and strobe) signals on different TSVs is reduced. As a result, for example, "the skew between data signals DQ and data strobe signals DQS/B can be minimized." Ex. 1016 at [0050]. Riho further explains that "[t]his structure (equi-length interconnections) is very important in the stacked structure in which the plurality of DRAM sets are stacked in sequence with respect to the controller chip. This is because, in this embodiment, each I/O group is formed by ×32 DQ signal bits and the controller chip can perform communication control of the plurality of I/O groups (×256 DQ signal bits) with a single synchronization signal and with high accuracy." *Id*.

342. Therefore, Kim renders obvious in view of Riho this limitation to a Skilled Artisan.

343. Accordingly, the prior art references in Ground 2 satisfy the "*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 between a first load on the first data conduit and a second load on the second data conduit" limitation.

b) [7.b] The First and Second Load

344. Claim 7 requires "*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*."

345. The prior art references in Ground 2 teach this limitation to a Skilled Artisan.

346. The combination of Kim and Riho discloses "*the first load including a load of the first die interconnect* [on one data signal TSV]*, and a load of the first group of array dies* [including a pair of SDRAM chips]*, and the second load including a load of the second die interconnect* [on another data signal TSV] *and a load of the second group of at least one array die* [including another pair of SDRAM chips]."

347. For example, Riho discloses that, "when the controlled chips are divided into a plurality of groups and the penetrating through substrate vias are commonly used by the controlled chips of the respective groups, it is possible to reduce the load of the interconnections as compared with the case where the penetrating through substrate vias are not commonly used." Ex. 1016 at [0013].

454. As discussed above with reference to claim 5, the prior art references in each of Grounds 1, 2, and 3 satisfy "*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.*" *Supra* ¶¶308-313.

15. Claim 15 is Unpatentable

455. Claim 15 is unpatentable as the prior art references in Ground 3 render the claim obvious.

a) [15.a] First and Second Data Conduit and First and Second Driver

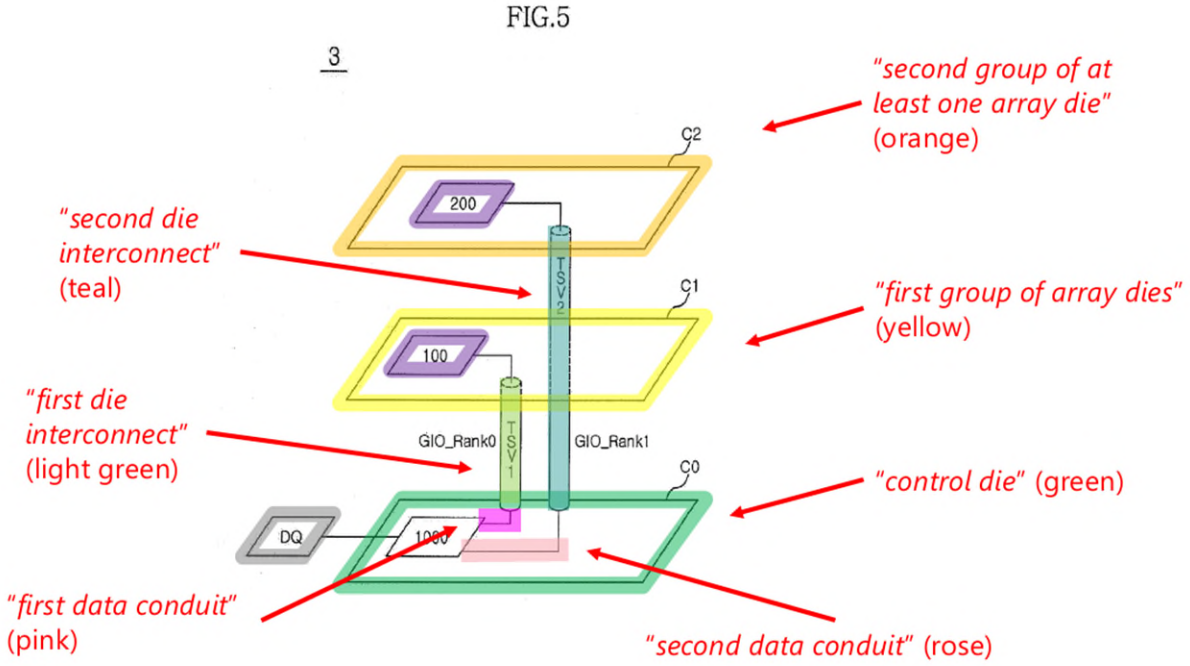
456. Claim 15 requires "*[t]he memory package of claim 11, wherein: the first data conduit comprises at least a first driver having a first driver size, and the second data conduit comprises at least a second driver having a second driver size*."

457. The prior art references in Ground 3 teach this limitation to a Skilled Artisan.

458. Kim in combination with Wyman renders obvious that "*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*."

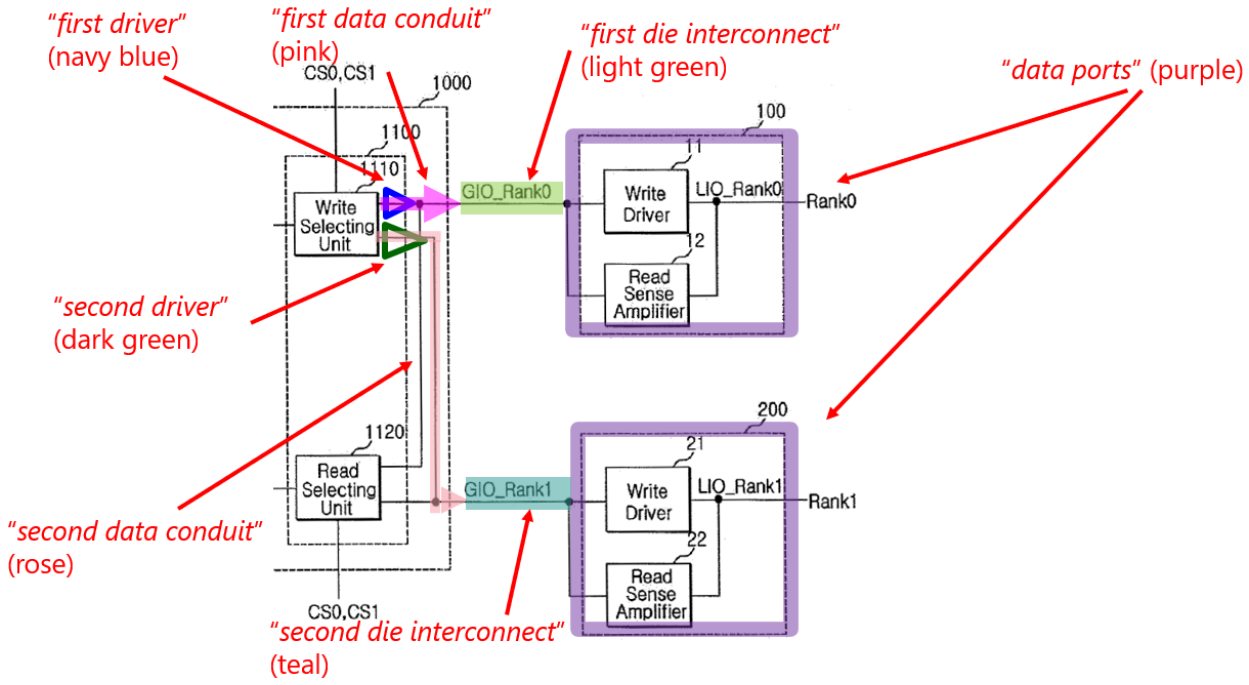
459. As discussed above with reference to claim limitations [1.e.2] and [1.e.3], Kim discloses the claimed "*first data conduit*" (annotated pink below) to

TSV1 and the "*second data conduit*" (annotated rose below) to TSV2. *Supra* ¶¶248-264; *see, e.g.*, Ex. 1014 at Fig. 5 (reproduced below, annotated).



460. As expressly shown in this figure of Kim, TSV1 coupled to the "*first data conduit*" (pink) is shorter than TSV2 coupled to the "*second data conduit*" (rose). A Skilled Artisan would have understood that, due to this difference in length of TSV1 and TSV2, each of the "*first data conduit*" and the "*second data conduit*" can include a respective driver that is optimized for driving the data signal through the corresponding TSV in an efficient way. For example, Wyman explains that "[i]n instances where chips are stacked, creating shorter runs, utilizing the full capacity of a driver would be both wasteful and inefficient since far less power is required." Ex. 1017 at 1:22-24. Therefore, a Skilled Artisan would have been motivated to follow the teaching of Wyman in the system of Kim

and avoid using full capacity drivers, thus saving power and not being "wasteful and inefficient." *Id*. Such an arrangement of having separate drivers (e.g., a "*first driver*," navy blue below, and a "*second driver*, dark green below) for each data conduit is illustrated in annotated (partial) Fig. 3 of Kim, reproduced below. *See* Ex. 1014 at Fig. 3 (reproduced below, in part, annotated).



461. As discussed above, a Skilled Artisan would have understood and been motivated to turn to analogous art like Wyman to implement the stacked devices of Kim with efficient driving signals between the stacks. *Supra* ¶¶174-181. For example, Wyman discloses that longer connections in a stacked arrangement require larger driving force. Ex. 1017 at Figs. 7, 8 (reproduced below).

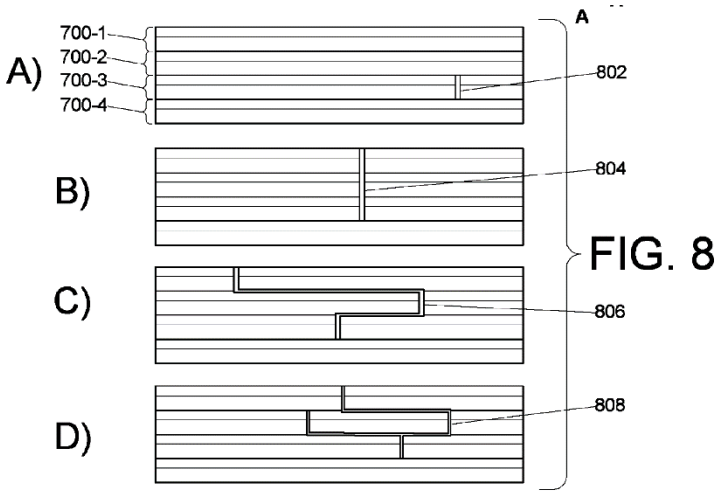


FIG. 8

Wyman explains that "FIG. 8A illustrates, in simplified form, a through-chip via connection (802) extending between the mother chip (700-4) and daughter 1 (700-3) … requir[ing] minimal drive … FIG. 8B illustrates, in simplified form, a longer through-chip via connection (804) extending directly between the mother chip (700-4) and daughter 3 (700-1) in a straight line. The increased resistance, capacitance and impedance of such a connection (804) might require additional drive than that referred to in connection with FIG. 8A. …" *Id*. at 6:15-30. However, Wyman recognizes that even with the "significantly longer connection (808) between the mother chip (700-4) with a fan out to both daughter 2 (700-2) and daughter 3 (7001)," as shown in Fig. 8.D, using "the total drive … would be overkill." *Id*. at 6:44-50.

462. A Skilled Artisan would have understood from this disclosure that the longer the connection in a stack of chips, the more drive it requires, but using the "total drive" would be wasteful. Ex. 1017 at 1:22-24 ("In instances where chips are

stacked, creating shorter runs, utilizing the full capacity of a driver would be both wasteful and inefficient since far less power is required."). Therefore, a Skilled Artisan would have been motivated to use drivers of different strengths (i.e., a "*first driver*" having a "*first driver size*" and a "*second driver*" having a "*second driver size*") for Kim's different sized through-silicon vias TSV1 and TSV2. As discussed above, a Skilled Artisan would have been motivated to drive the shorter TSV1 with a smaller driver and the longer TSV2 with a larger driver to use power in an efficient manner as taught by Wyman. *Supra* ¶¶174-181; *see also* annotated partial Fig. 3 of Kim above. Making such a combination was well within the level of skill, since Wyman itself teaches drivers of different strengths, and the combination would have provided nothing more than expected: a power efficient drive of signals through Kim's TSVs.

463. Therefore, the combination of Kim and Wyman renders obvious this limitation to a Skilled Artisan.

464. Accordingly, the prior art references in Ground 3 satisfy the "*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*" limitation.

b) [15.b] First and Second Driver Size Less Than Driver Size Sufficient to Drive Signal

465. Claim 15 requires "*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*."

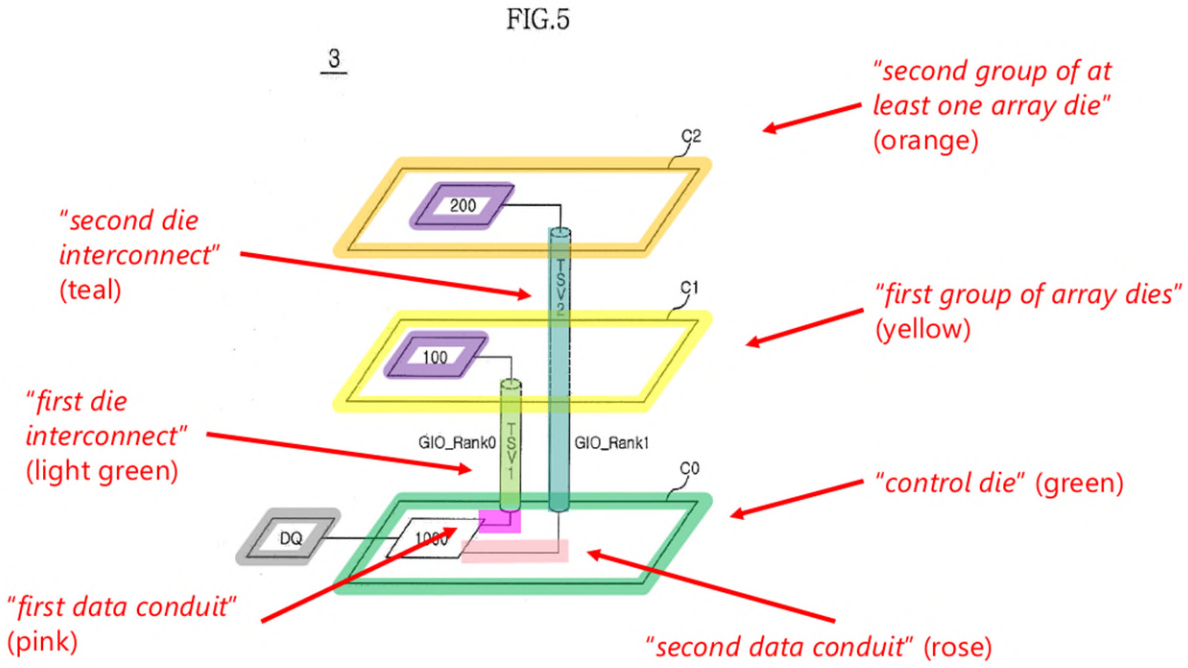
466. The prior art references in Ground 3 teach this limitation to a Skilled Artisan.

467. Kim in combination with Wyman renders obvious to a Skilled Artisan that "*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.*"

468. For example, Wyman discloses to a Skilled Artisan that using a driver strength that is "overkill" is "wasteful and inefficient since far less power is required." Ex. 1017 at 6:47-50, 1:22-24. Therefore, for the reasons discussed above with reference to claim limitation [15.a], a Skilled Artisan would have been motivated to use a "*driver size*" that is required for efficient signal communication, but that would not produce a drive that is "overkill." *Supra* ¶¶456-464.

469. A Skilled Artisan would have also understood that the "*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

Kim's stacked system is larger than the "*driver size*" that is required for efficient signal communication through TSV1 and TSV2. *Supra,* ¶¶135-137. Indeed, each of TSV1 and TSV2 are coupled to circuitry in some, but not all, of the slave chips in the stack, thus requiring less strength than driving circuitry in all the slave chips. *See* Ex. 1014 at Fig. 5 (reproduced below, annotated).



Furthermore, Wyman explains that the "reduced path distances of stacked hybridized chips therefore require less drive to adequately drive a signal to an off-chip board and/or chip." Ex. 1017 (Wyman) at 1:45-47; *see also id*. at 6:57-60 ("By using the above approach, a designer could assess the amount of drive needed for to adequately supply the signal to the devices (902, 904, 906) as needed while using less than the total drive a drive circuit of the prior art might supply.").

470. Therefore, the combination of Kim and Wyman renders obvious this limitation to a Skilled Artisan.

471. Accordingly, the prior art references in Ground 3 satisfy the "*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*" limitation.

16. Claim 16 is Unpatentable

472. Claim 16 is unpatentable as the prior art references in Grounds 1-3 render the claim obvious.

473. Claim 16 requires "*[t]he 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*."

474. The prior art references in Grounds 1-3 teach this limitation to a Skilled Artisan.

475. As shown in the table below, claim 16 is very similar to claim limitations [1.e.4] and [8.b.3]: