Doc Code: **TR.PROV**
Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (11-08)
Approved for use through 01/31/2014 OMB 0651-0032
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Provisional Application for Patent Cover Sheet

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)

Inventor(s)

Inventor 1 Remove

Given Name	Middle Name	Family Name	City	State	Country i
Hyun		Lee	Ladera Ranch	CA	US

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

Title of Invention	MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION
Attorney Docket Number (if applicable)	NT003

Correspondence Address

Direct all correspondence to (select one):

● The address corresponding to Customer Number	○ Firm or Individual Name
Customer Number	79141

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

● No.

○ Yes, the name of the U.S. Government agency and the Government contract number are:

Doc Code: **TR.PROV**
Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (11-08)
Approved for use through 01/31/2014 OMB 0651-0032
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Entity Status

Applicant claims small entity status under 37 CFR 1.27

● Yes, applicant qualifies for small entity status under 37 CFR 1.27
○ No

Warning

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

Signature

Please see 37 CFR 1.4(d) for the form of the signature.

Signature	/Jamie J. Zheng/			Date (YYYY-MM-DD)	2012-07-27
First Name	Jamie	Last Name	Zheng	Registration Number (If appropriate)	51167

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **This form can only be used when in conjunction with EFS-Web. If this form is mailed to the USPTO, it may cause delays in handling the provisional application.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or paten. Accordingly, pursuant to the requirements of the Act, please be advised that : (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to an other federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application is related to commonly-owned U.S. patent applications No. 12/504,131, No. 12/761,179, No. 13/287,042, and No. 13/287,081, each of which is incorporated herein by reference in its entirety.

FIELD

[0002] The disclosure herein is related generally to memory modules, and more particularly to multi-rank memory modules and methods of operation.

BACKGROUND

[0003] With recent advancement of information technology and widespread use of the Internet to store and process information, more and more demands are placed on the acquisition, processing, storage and dissemination of vocal, pictorial, textual and numerical information by microelectronics-based combination of computing and communication means. In a typical computer or server system, memory modules are used to store data or information. A memory module usually includes multiple memory devices, such as dynamic random access memory devices (DRAM) or synchronous dynamic random access memory devices (SDRAM), packaged individually or in groups, and/or mounted on a printed circuit board (PCB). A processor or a memory controller accesses the memory module via a memory bus which, for a single-in-line memory module (SIMM), can have a 32-bit wide data path, or for a dual-in-line memory module (DIMM), can have a 64-bit wide data path.

[0004] The memory devices of a memory module are generally organized in ranks, with each rank of memory devices generally having a bit width. For example, a memory module in which each rank of the memory module is 64 bits wide is described as having an "x64" or "by 64" organization. Similarly, a memory module having 72-bit-wide ranks is described as having an "x72" or "by 72" organization.

[0005] The memory capacity or memory density of a memory module increases with the number of memory devices on the memory module. The number of memory devices of a memory module can be increased by increasing the number of memory devices per rank or by increasing the number of ranks.

[0006] In certain conventional memory modules, the ranks are selected or activated by control signals from a processor or memory controller during operation. Examples of such control signals include, but are not limited to, rank-select signals, also called chip-select signals. Most computer and server systems support a limited number of ranks per memory module, which limits the memory density of the memory modules that can be used in these computer and server systems.

[0007] For memory devices in such as a memory module to be properly accessed, distribution of control signals and a control clock signal in the memory module is subject to strict constraints. In some conventional memory modules, control wires are routed so there is an equal length to each memory component, in order to eliminate variation of the timing of the control signals and the control clock signal between different memory devices in the memory modules. The balancing of the length of the wires to each memory devices compromises system performance, limits the number of memory devices, and complicates their connections.

[0008] In some conventional memory systems, the memory controllers include leveling mechanisms for write and/or read operations to compensate for unbalanced wire lengths and memory device loading on the memory module. As memory operating speed and memory density continue to increase, however, such leveling mechanisms are also insufficient to insure proper timing of the control and/or data signals received and/or transmitted by the memory modules.

DESCRIPTION OF EMBODIMENTS

[0009] A memory module according to one embodiment includes memory devices organized in groups, a module control device, and buffer circuits. Command signals, including one or more read/write commands, control/address (C/A) signals and a system clock signal from a memory controller are received by the module control device, which generates module command signals and module control signals in response to the command signals. The module command signals are transmitted by the module control device to the memory devices via module command signal lines, and the module control signals are transmitted by the module control device to the buffer circuits via module control signal lines.

[0010] In one embodiment, each group of memory devices include at least two subgroups, each subgroup including at least one memory device. The buffer circuits are associated with respective groups of memory devices and are distributed across the memory module at positions corresponding to the respective groups of memory devices. During certain high speed operations, each module control signal arrives at different buffer circuits at different points of time across more than one clock cycle of the system clock. A buffer circuit associated with a respective group of memory devices is inserted into the data paths between the respective group of memory devices and the memory controller. Thus, the memory controller does not have direct control of the memory devices. Instead, each buffer circuit is configured to select a subgroup in the respective group of memory devices to communicate data with the memory controller in response to the module control signals. Thus, the memory module can have more ranks of memory devices than what is supported by the memory controller.

[0011] In one embodiment, each buffer circuits includes metastability detection circuits to detect metastability condition in the module control signals and signal adjustment circuits to adjust the module control signals and/or a module clock signal to mitigate any metastability condition in the module control signals.

[0012] Further, in one embodiment, each buffer circuit includes signal alignment circuits that determine, during a write operation, a time interval between a time when one or more module control signals are received from the module control circuit and a time when a strobe or data signal is received from the memory controller. This time interval is used during a subsequent read operation to time transmission of read data to the memory controller, such that the read data follows an associated read command by a read latency value associated with the memory system.

[0013] FIG. 1 shows a system 100 including a memory controller (MCH) 101 and one or more memory modules 110 coupled to the MCH by a memory bus 105, according to one embodiment. As shown, the memory bus includes command signal lines 120 and groups of system data/strobe signal lines 130. Also as shown, each memory module 110 has a plurality of memory devices 112 organized in a plurality of ranks 114. Each memory module 110 further includes a module control circuit (module controller or module control device) 116 coupled to the MCH 101 via the command signal lines 120, and a plurality of buffer circuits or isolation devices 118 coupled to the MCH 101 via respective groups of system data/strobe signal lines 130. In one embodiment, the memory devices 112, the module control circuit 116 and the isolation devices 118 can be mounted on a same side or different sides of a printed circuit board (module board) 119.

[0014] In the context of the present description, a rank refers to one or more memory devices that are controlled by a common set of address/control (C/A) signals. The number of ranks of memory devices in a memory module 110 may vary. For example, as shown, each memory module 110 may include four ranks of memory devices 112. In another embodiment, the memory module 110 may include 2 ranks of memory devices. In yet another embodiment, the memory module may include six or more ranks of memory devices 112.

[0015] In the context of the present description, a memory controller refers to any device capable of sending instructions or commands, or otherwise controlling the memory devices 112. Additionally, in the context of the present description, a memory bus refers to any component, connection, or groups of components and/or connections, used to provide electrical communication between a memory module and a memory controller. For example, in various embodiments, the memory bus 105 may include printed circuit board (PCB) transmission lines, module connectors, component packages, sockets, and/or any other components or connections that provide connections for signal transmission.

[0016] Furthermore, the memory devices 112 may include any type of memory devices. For example, in one embodiment, the memory devices 112 may include dynamic random access memory (DRAM) devices. Additionally, in one embodiment, each memory module 110 may include a dual in-line memory module (DIMM).

[0017] Referring to FIG. 2, which illustrates one memory module 110 according to an embodiment, the module control device 116 receives system command signals including control/address (C/A) signals and a system clock MCK from the MCH 101 via signal lines 120 and generates module command signals and module control signals based on system command signals. The module control device 116 also generates a module clock signal CK in response to the system clock signal MCK. The MCK signal may include a pair of complementary clock signals, MCK and $\overline{\text{MCK}}$, and the module clock signal may include a pair of complementary clock signals CK and $\overline{\text{CK}}$.

[0018] Examples of the system C/A signals include, but are not limited to, Chip Select (or /CS) signal, which is used to select a rank of memory devices to be accessed during a memory (read or write) operation; Row Address Select (or /RAS) signal, which is used mostly to latch a row address and to initiate a memory cycle; Column Address Select (or /CAS) signal, which is used mostly to latch a column address and to initiate a read or write operation; address signals, including bank address signals and row/column address signals, which are used to select a memory location on a memory device or chip; Write Enable (or /WE) signal, which is used to specify a read operation or a write operation, Output Enable (or /OE) signal, which is used to prevent data from appearing at the output until needed during a read operation, and the system clock signal MCK.

[0019] Examples of module command signals include, but are not limited to module /CS signals, which can be derived from the system /CS signals and one or more other system command signals, such as one or more bank address signals and/or one or more row/column address signals; a module /RAS signal, which can be, for example, a buffered version of the system /RAS signal; a module /CAS signal, which can be, for example, a buffered version of the system /CAS signal; module address signals, which can be, for example, buffered versions of some or all of the address signals; a module /WE signal, which can be, for example, a buffered version of the system /WE signal; a module /OE signal, which can be, for example a buffered version of the system /OE signal, and the module clock signal CK.

[0020] Examples of module control signals include, but are not limited to a mode signal (MODE), which specifies a mode of operation (e.g., test mode or operating mode) for the isolation devices 118; one or more enable signals, which are used by an isolation device to select one or more subgroups of memory devices to communicate data with the memory controller; and one or more ODT signals, which are used by the isolation devices to set up on-die termination for the data/strobe signals. In one embodiment, the module control signals are transmitted to the isolation devices 118 via respective module control signal lines 230. Alternatively, the module signals can be packetized before being transmitted to the isolation devices 118 via the module signal lines and decoded/processed at the isolation devices.

[0021] Module control device 116 transmits the module command signals to the memory devices 112 via module command signal lines 220. The memory devices 112 operate in response to the module command signals to receive write data or output read data as if the module command signals were from a memory controller. The module control device transmits the module control signals together with the module clock signal CK to the isolation devices 118 via module control signal lines 230. As shown in FIG. 2, at least some of the memory devices in each rank share a same set of module command signal lines 220, and at least some of the isolation devices 118 share a same set of module control signal lines 230.

[0022] As shown n FIGS. 2A and 2B, each rank 114 includes N memory devices, where N is an integer larger than one. For example, a first rank includes memory devices M_{11}, …, M_{i1}, $M_{i+1,1}$, …, M_N, a second rank includes memory devices M_{12}, …, M_{i2}, $M_{i+1,2}$, …, $M_{N,2}$, and so on. In one embodiment, the memory devices 112 are also organized in groups or sets, with each group corresponding to a respective group of system data/strobe signal lines 130 and including at least one memory device from each rank. For example, memory devices M_{11}, M_{12}, M_{13}, and M_{14} form a first group of memory devices, memory devices M_{i1}, M_{i2}, M_{i3}, and M_{i4} form an i^{th} group of memory devices, and so on.

[0023] As shown, the isolation devices 118 are associated with respective groups of memory devices and are coupled between respective groups of system data/strobe signal lines 130 and the respective groups of memory devices. For example, isolation device ID-1 among the isolation devices 118 is associated with the first group of memory devices M_{11}, M_{12}, M_{13}, and M_{14} and is coupled between the group of system data/strobe signal lines 130-1 and the first group of memory devices, isolation devices ID-*i* among the isolation devices 118 is associated with the i^{th} group of memory devices M_{i1}, M_{i2}, M_{i3}, and M_{i4} and is coupled between the group of system data/strobe signal lines 130-i and the i^{th} *group* of memory devices, and so on.

[0024] In one embodiment, each group or sets of memory devices are coupled to the associated isolation device 118 via a set of module data/strobe lines 210. Each group or set of memory devices is organized in subgroups or subsets, with each subgroup or subset including at least one memory device. The subgroups in a group of memory devices may be coupled to the associated isolation device 118 via a same set of module data/strobe lines 210 (as shown in FIG. 2A) or via respective subsets of module data/strobe lines 210 (as shown in FIG. 2B). For example, as shown in FIG. 2B, in the first group of memory devices, memory devices M_{11} and/or M_{13} form a first subgroup, and memory devices M_{12}

and/or M_{14} form a second subgroup; in the i^{th} group of memory devices, memory devices M_{i1} and/or M_{i3} form a first subgroup, and memory devices M_{i2} and/or M_{i4} form a second subgroup; and so on. The first subgroup of at least one memory device in each group of memory devices is coupled to the associated isolation device 118 via an associated first subset of module data/strobe lines YA, and the second subgroup of at least one memory device in each group of memory devices is coupled to the associated isolation device via an associated second subset of module data/strobe lines YB, as shown. For example, memory devices M_{11} and/or M_{13} form the first subgroup are/is coupled to the isolation device ID-1 via the corresponding first subset of module data/strobe lines YA-1, and memory devices M_{12} and/or M_{14} form the second subgroup are/is coupled to the isolation device ID-1 via the corresponding second subset of module data/strobe lines YA-2.

[0025] In one embodiment, the isolation devices 118 are added to the data paths between the MCH 101 and the memory module 110 and include data buffers between the MCH 101 and the respective groups of memory devices. In one embodiment, each isolation device 118 is configured to select a subgroup in the respective group of memory devices to communicate data with the MCH 101 in response to the module control signals, such that the memory module can include more ranks than what is supported by the MCH 101. Further, each isolation devices 118 is configured to isolate unselected subgroup(s) of memory devices from the MCH 101 during write operations, so that the MCH sees a load on each data line that is less than a load associated with the respective group of memory devices. In one embodiment, the MCH sees only a load associated with one memory device on each data/strobe signal line during write operations.

[0026] In one embodiment, the isolation devices 118 are distributed across the memory module 110 or the module board 119 in positions corresponding to the respective groups of memory devices. For example, isolation device ID-1 is disposed in a first position corresponding to the first group of memory devices M_{11}, M_{12}, M_{13}, and M_{14}, and isolation device ID-i is disposed in an i^{th} position separate from the first position and corresponding to the i^{th} group of memory devices M_{i1}, M_{i2}, M_{i3}, and M_{i4}. In one embodiment, the first position is between the first group of memory devices and an edge 201 of the module board

119 where connections (not shown) to the data/strobe signal lines 130 are placed, and i^{th} position is between the i^{th} group of memory devices and the edge 201 of the module board 119. In one embodiment, the isolation devices 118 are distributed along the edge 201 of the memory module 110. In one embodiment, each isolation device 118 is a separate integrated circuit device packaged either by itself or together with at least some of the respective group of memory devices. In one embodiment, the module data/strobe signal lines 210, the module command signal lines 220, and the module control signal lines 230 include signal traces formed on and/or in the module board 119.

[0027] As an option, memory module 110 may further include a serial-presence detect (SPD) device 240, which may include electrically erasable programmable read-only memory (EEPROM) for storing data that characterize various attributes of the memory module 110. Examples of such data include a number of row addresses, a number of column addresses, a data width of the memory devices, a number of ranks on the memory module 110, a memory density per rank, a number of memory device on the memory module 110, and a memory density per memory device, etc. A basic input/output system (BIOS) of system 100 can be informed of these attributes of the memory module 110 by reading from the SPD 240 and can use such data to configure the MCH 101 properly for maximum reliability and performance.

[0028] Thus, in the memory module 110, command signals are received and buffered by the module control circuit 116, so that the MCH sees only the module control circuit 116 as far as the command signals are concerned. Write data and strobe signals from the controller are received and buffered by the isolation devices 118 before being transmitted to the memory devices 112 by the isolation devices 118. On the other hand, read data and strobe signals from the memory devices are received and buffered by the isolation devices before being transmitted to the MCH via the system data/strobe signal lines 130. Thus, MCH 101 does not directly operate or control the memory devices 112. As far as data/strobe signals are concerned, the MCH 101 mainly sees the isolation devices 118, and the system 100 depends on the isolation devices 118 to properly time the transmission of the read data and strobe signals to the MCH 101.

[0029] In one embodiment, operation of the isolation devices are controlled by the module control signals from the module control circuit 116, which generates the module control signals according to the command signals received from the MCH. Thus, the module control signals need to be properly received by the isolation devices 118 to insure their proper operation. In one embodiment, the module control signals are transmitted together with the module clock signal CK, which is also generated by the module control circuit 116 based on the system clock signal MCK. The isolation circuits 118 buffers the module clock signal, which is used to time the sampling of the module control signals. Since the isolation devices 118 are distributed across the memory module, the module control signal lines 230 can stretch across the memory module 110, over a distance of several centimeters. As the module control signals travel over such a distance, they can become misaligned with the module clock signal, resulting in metastability in the received module control signals. Therefore, in one embodiment, the isolation circuits 118 includes metastability detection circuits to detect metastability condition in the module control signals and signal adjustment circuits to adjust the module control signals and/or the module clock signal to mitigate any metastability condition in the module control signals, as explained in further detail below.

[0030] Because the isolation devices 118 are distributed across the memory module 110, during high speed operations, it may take more than one clock cycle time of the system clock MCK for the module control signals to travel along the module control signals lines 230 from the module control device 116 to the farthest positioned isolation devices 118, such as isolation device ID-1 and isolation device ID-(*n*-1) in the exemplary configuration shown in FIG. 2. In other words, a same set of module control signals may reach different isolation devices 118 at different times across more than one clock cycle of the system clock. For example, when the clock frequency of the system clock is higher than 800 MHz, the clock cycle time is less than about 1.2 ns. With a signal travel speed of about 70ps per centimeter of signal line, a module control signal would travel about 15 cm during one clock cycle. When the clock frequency increases to 1600 MHz, a module control signal would travel less than 8 cm during one clock cycle. Thus, a module control signal line can have multiple module control signals on the line at the same time, i.e., before one module control signal reaches an end of the signal line, another module control signal appear on the signal line.

[0031] With the isolation devices 118 receiving module control signals at different times across more than one clock cycle, the module control signals alone are not sufficient to time the transmission of read data signals to the MCH 101 from the isolation devices 118. In one embodiment, each isolation devices includes signal alignment circuits that determine, during a write operation, a time interval between a time when one or more module control signals are received from the module control circuit 116 and a time when a write strobe or write data signal is received from the MCH 101. This time interval is used during a subsequent read operation to time the transmission of read data to the MCH 101, such that the read data follows a read command by a read latency value associated with the system 100, as explained in more detail below.

[0032] More illustrative information will now be set forth regarding various optional configurations, architectures, and features with which the foregoing framework may or may not be implemented, per the desires of the user. It should be strongly noted that the following information is set forth for illustrative purposes and should not be construed as limiting in any manner. Any of the following features may be optionally incorporated with or without the exclusion of other features described.

[0033] In one embodiment, as shown in FIG. 3, each group of signal lines 130 include a set of *n* data (DQ) signal lines 322 each for transmitting one of a set of data signals DQ_0, DQ_1, …, DQ_{n-1}, and at least one strobe (DQS) signal line 324 for transmitting at least one strobe signal DQS. Each set of module data/strobe lines Y include a set of *n* module data signal lines Y_0, Y_1, …, Y_{n-1} and at least one module strobe signal line Y_{DQS}. When the subsets of memory devices are coupled to the associated isolation device 118 via respective subsets of memory devices, each set of module data/strobe lines Y may include multiple subsets of module data/strobe lines, such as the subsets of module data/strobe lines YA and YB shown in FIG. 2B. Each subset of module data/strobe lines YA include a set of *n* first module data lines YA_0, YA_1, …, YA_n and at least one first module strobe signal line YA_{DQS}; and each subset of module data/strobe lines YB include a set of *n* second module data lines YB_0, YB_1, …, YB_n and at least one second module strobe signal line YB_{DQS}.

[0034] Each isolation device 118 includes a set of DQ routing circuits 320 coupled on one side to respective ones of the set of *n* DQ signal lines 322, and on another side to respective ones of the respective set of *n* module data lines, or respective ones of the respective subsets of module data lines, such as the first module data lines YA_0, YA_1, …, YA_n and the second module data lines YB_0, YB_1, …, YB_n. Each isolation device 118 further includes an ID control circuit 310 coupled on one side to the at least one DQS signal line 324, on another side to the one or more module strobe signal lines Y_{DQS}, or the first module strobe signal line YA_{DQS} and second module strobe signal line YB_{DQS}. The ID control circuit 310 also receives the module clock signal CK and the module control signals via the module control signal lines 230, and outputs ID control signals 330 to the DQ routing circuits 320, including, for example, one or more enable signals ENA and/or ENB, and some or all of the other received, decoded, and/or otherwise processed module control signals, a delay signal DS, a read DQS signal RDQS, a write DQS signal WDQS, and a buffer clock signal CK0. Each DQ routing circuit 320 is configured to enable data communication between the respective DQ signal line 322 with a selected subgroup of one or more memory devices in response to the module control signals, as explained in more detail below.

[0035] The memory devices 112 are coupled to the isolation devices 118 via a same set of module data/strobe signal lines or different subsets of module data/strobe signal lines. For example, as shown in FIG. 4A, memory devices M_{11}, M_{12}, M_{13}, and M_{14} in the first group of memory devices can be coupled to the isolation device ID-1 via a same set of module data lines $Y\text{-}1_0$, $Y\text{-}1_1$, …, $Y\text{-}1_{n-1}$ and module strobe line $Y\text{-}1_{DQS}$. In such embodiment, a subgroup in the group of memory devices can be selected by the isolation devices to communicated data with the MCH based on the phases of the data/strobe signals, which can be different with respect to different subgroups of memory devices.

[0036] Alternatively, as shown in FIG. 4B, memory devices M_{11} and M_{13}, which form a subgroup in the first group of memory devices, are coupled to the isolation device ID-1 via the module data lines YA-1_0, YA-1_1, …, YA-1_n and module strobe line YA-1_{DQS} and memory devices M_{12} and M_{14}, which form another subgroup in the first group of memory devices, are coupled to the isolation device ID-1 via the module data lines YB-1_0, YB-1_1, …, YB-1_n and module strobe line YB-1_{DQS}. Memory devices coupled to the same isolation devices can be disposed on a same side or different sides of the memory board 119. Memory devices coupled to the same isolation devices may be placed side by side, on opposite sides of the module boards 119, or stacked over each other, and/or over the associated isolation device.

[0037] Multiple memory devices having a data width that is less than a data width of the isolation devices 118 may be used in place of one of the memory devices 112, which has the same data width as that of the isolation devices. For example, as shown in FIG. 5A, two memory devices M_{11-1} and M_{11-2} may be used in place of the memory device M_{11}. Each of the two memory devices M_{11-1} and M_{11-2} has a data width of 4, and together they act like a memory device M_{11} of a data width of 8. Thus, memory device M_{11-1} is coupled to the isolation device ID-1 via module data lines YA-1_0, …, YA-1_3 and module strobe line YA-1_{DQS-1} while memory circuit M_{11-2} is coupled to the isolation device ID-1 via module data lines YA-1_4, …, YA-1_7 and module strobe line YA-1_{DQS-2}.

[0038] In another embodiment, as shown in FIG. 5B, four memory devices M_{11-1} to M_{11-4} may be used as the memory device M_{11}. Each of the four memory devices M_{11-1} to M_{11-4} has a data width of 4, and together they act like a memory device M_{11} of a data width of 16. Thus, memory device M_{11-1} is coupled to the isolation device ID-1 via module data lines YA-1_0, …, YA-1_3 and module strobe line YA-1_{DQS-1} while memory device M_{11-2} is coupled to the isolation device ID-1 via module data lines YA-1_4, …, YA-1_7 and module strobe line YA-1_{DQS-2}, and so on.

[0039] FIG. 6 illustrates the ID control circuit 310 in an isolation device 118. As shown, the ID control circuit 310 includes a clock buffer 610 to receive the module clock signal CK from the module control device 116, and to output a buffered clock signal CK0. The ID control circuit 310 further includes a strobe routing circuit 620 that are coupled on one side to the corresponding system DQS signal line 324 and on another side to the corresponding module DQS signal lines YA_{DQS} and YB_{DQS}. The ID control circuit 310 further includes a receiver circuit 630 with respect to each of at least some of the module control signals (MCS) to receive a respective one of the module control signals. The ID control circuit 310 further includes a command processing circuit 640 that provides the received, decoded, and/or otherwise processed module control signals 330 to the DQ routing circuits 320 and the strobe routing circuit 620 either directly or after further processing, if needed. The received/decoded/processed module control signals may include, for example, one or more enable signals ENA and/or ENB that are used by the DQ routing circuits 320 and the strobe routing circuit 620 to selectively enabling data communication between the MCH 101 and one of the subgroups in the respective group of memory devices, with which the isolation device is associated.

[0040] The strobe routing circuit 620 also buffers strobe signals received from either the MCH 101 or the memory devices 112, and output either a write strobe WDQS or read strobe RDQS to the DQ routing circuits 320. In one embodiment, the ID control circuit 310 further includes a delay control circuit 650 that receives one of the module control signals and either a data signal or a strobe signal and determines a delay amount to be used by the DQ routing circuit 320 and the strobe routing circuit 620. The delay amount is provided to the DQ routing circuit 320 and the strobe routing circuit in a delay signal DS.

[0041] In a receiver circuit 630, the respective MCS is received in accordance with the buffered clock signal CK0. In one embodiment, receiver circuit 630 samples the respective MCS using rising (or falling) edges of the buffered clock CK0. Since the isolation devices 118 are distributed across the memory module 110 at positions corresponding to the respective groups of memory devices, the module control signal lines 230 that carry the MCS to the isolation devices can stretch over a distance of more than 10 centimeters, as shown in FIG. 7. As the MCS and CK0 travel along their respective module control signal lines 710 and 720, they can become misaligned with each other when they reach the input pins 730 of an isolation device 118.

[0042] For example, a module control signal, like the MCS 810 shown in FIG. 8, can be perfectly aligned with the module clock signal CK, with a rising edge 801 of the module clock signal CK being at a center of a data eye 802, when the MCS signal and the clock signal leave the module control circuit 116. When the module control signal and the module clock signal reach an isolation device, however, their alignment can become shifted like the MCS 820 with respect to the CK signal, i.e., the rising edge 801 of the clock signal is near a left edge of a data eye of the MCS 820, barely providing enough set up time for proper sampling of the module control signal. Or, the module control signal, like the MCS 830, can be shifted with respect to the module clock signal such that a rising edge 801 of the clock signal is near a right edge of a data eye of the MCS, barely providing enough hold time for proper sampling of the module control signal. Or, ever worse, the module control signal, like the MCS 840, can be so shifted with respect to the module clock signal such that a rising edge 801 of the clock signal falls in the glitches 803 at the edge of a data eye of the MCS, meaning that the sampled results could be metastable.

[0043] In one embodiment, as shown in FIG. 9, a receiver circuit 630 includes a metastability detection circuit (MDC) 910 to determine a metastability condition in a corresponding module control signal MCS0. In one embodiment, the MDC 910 generates at least one delayed version of the buffered clock signal CK0 and at least one delayed version of the corresponding MCS0. The MDC 910 also generates one or more metastability indicators and outputs the one or more metastability indicators via lines 912 and/or 914.

[0044] The receiver circuit 630 further includes a signal selection circuit 920 that receives the buffered clock CK0 and the at least one delayed version of the buffered clock via signal lines 916. The signal selection circuit 920 also receives the corresponding MCS and the at least one delayed version of the corresponding MCS via signal lines 918. The signal selection circuit 920 selects a clock signal CK_i from among the buffered clock CK0 and the at least one delayed version of the buffered clock based on one or more of the metastability indicators. The signal selection circuit 920 may also select an MCS signal MCS_i from among the corresponding MCS and the at least one delayed version of the corresponding

MCS based on at least one other metastability indicator.

[0045] The receiver circuit 630 further includes a sampler or register circuit 930 that samples the selected module control signal MCS_i according to the selected clock signal CK_i and outputs the sampled signal as the received module control signal, which is provided to the command processing circuit 640 for further processing (if needed) before being provided to the DQ routing circuits 320 and DQS routing circuit 620.

[0046] FIG. 10A illustrates an MDC 910 according to one embodiment. As shown, the MDC 910 includes a delay circuit 1012 that generates a delayed version MCS1 of the corresponding MCS0 by adding a predetermined amount of delay (e.g., 10ps) to MCS0. MDC 910 also includes a delay circuit 1016 that generates a delayed version CK1 of the clock signal CK0 by adding a predetermined amount of delay to CK0. In one embodiment, CK1 is delayed from CK0 by about 1/10th of a clock cycle, e.g., 50-70 ps for an operating frequency of about 1600 MHz. The MDC 910 further includes a sampler circuit 1042 that samples MCS1 according to CK0 and outputs a sampled result A, a sampler circuit 1044 that samples MCS0 according to CK0 and outputs a sampled result B, and a sampler circuit 1046 that samples MCS0 according to CK1 and outputs a sampled result C. The MDC 910 further includes a logic circuit (e.g., a majority decision circuit) that generates metastability indicators Z1 and Z2 based on the sampled results A, B, and C.

[0047] In one embodiment, Z1 is the result of a logic operation (e.g., an XNOR operation) on the sampled result, e.g., $Z1 = \overline{A \oplus B}$, and Z2 is the result of another logic operation on the sampled results, e.g., $Z2 = \overline{B \oplus C}$. Thus, as shown in FIG. 10B and Table 1 below, when a metastability condition of insufficient hold time occurs, i.e., a rising clock edge 1061 of CK0 is close to the right side of a data eye where gliches at the edges of the data eyes can make C unpredicatable, A and B can be in agreement (i.e., Z1 is true) while B and C are likely not in agreement (i.e., Z2 is false). FIG. 10 C illustrates a metastability condition when there is insufficient set-up time. As shown in FIG. 10C and Table 1 below, a rising clock edge 1061 of CK0 is close to the left side of a data eye where gliches at the edges of the data eyes can make A unpredicatable. Thus, A and B can be in disagreement so Z1 is false while B and C can be in agreement so Z2 is true. Not shown in the figures is the situation that all A, B, and C are in agreement, meaning that both the rising clock edge 1061 of CK0 and the rising clock edge 1062 of CK1 are near the middle of an MCS0 data eye so there is no metastability issues and both Z1 and Z2 are true, as shown in Table 1.

[0048] FIG. 10D illustrates a signal selection circuit 920 according to an embodiment. As shown, in one embodiment, the signal selection circuit 920 includes a first multiplexor 1071 that selects between CK0 and CK1 based on the metastability indicator Z1, and a second multiplexor 1072 that selects between MCS0 and MCS1 based on the metastability indicator Z2. Thus, as shown in Table 1, where a metastability condition of insufficient hold time occurs, Z1=1 and Z2=0, and MCS1 is output from multiplexor 1071 while CK0 is output from multiplexor 1072. Sampler 930 thus samples MCS1 according to the rising edges of CK0. Thus, more hold time is provided to mitigate the metastability condition since MCS1 is shifted from MCS0 toward the right.

[0049] On the other hand, where a metastability condition of insufficient set-up time occurs, Z1=0 and Z2=1, and CK1 is output from multiplexor 1071 while MCS0 is output from multiplexor 1072. Sampler 930 thus samples MCS0 according to the rising edges of CK1. Since CK1 is shifted from CK0 toward the right, more set-up time is provided to mitigate the metastability condition.

[0050] In the case when no metastability is detected, Z1=1 and Z2=1, and CK0 is output from multiplexor 1071 while MCS0 is output from multiplexor 1072. So, the unshifted module control signal is sampled according to the unshifted module clock signal.

Table 1 Metastability Detection and Signal Selection

Sampler Output			MS Indicators		MS Condition	Signal Selection	
A	B	C	Z1	Z2		CK	MCS
D1	D1	D2	1	0	insufficient hold time	CK0	MCS1
D1	D2	D2	0	1	insufficient set-up time	CK1	MCS0
D1	D1	D1	1	1	no metastability	CK0	MCS0

[0051] FIGS. 10A-10D illustrate a relatively simple implementation of the metastability detection circuit (MDC) 910 where only three different sample points are provided to detect metastability condition in the module control signal. In general, the MDC 910 may generate more delayed versions of the module clock signal CK0 and/or the corresponding module control signal MCS0, and may include more sampler circuits to sample any additional delayed versions of the module control signal according to either the module clock signal or one of the delayed versions of the module clock signal. For example, as shown in FIG. 11A, the MDC 910 can include a plurality of delay circuits 1102 that generate *m* delayed versions of MCS0, e.g., MCS1, MCS2, … MCS*m*, and *m* delayed versions of CK0, e.g., CK1, CK2, …, CK*m*. The MDC 910 can include sampler circuits 1104 that sample MCS0 according to CK0, CK1, …, CK*m*, respectively, and sampler circuits 1104 that sample MCS0, MCS1, MCS2, … MCS*m* according to CK0, respectively. The outputs of the samplers 1104 are provided to a logic circuit 1120, which determines a metastability condition in MCK0 based on the sampler outputs using, for example, a majority decision logic. The logic circuit 1120 outputs a first metastability indicator on line(s) 912 and a second metastability indicator on line(s) 914.

[0052] FIG. 11B illustrates a signal selection circuit 920 according to an embodiment. As shown, in one embodiment, the signal selection circuit 920 includes a first multiplexor 1171 that selects between CK0, CK1, …, CK*m* based on the metastability indicator provided on line(s) 912, and a second multiplexor 1172 that selects between MCS0, MCS1, …, MCS*m* based on the metastability indicator provided on line(s) 914, such that the rising edges of the selected clock signal, e.g., Ck*i*, are close to the middle of the respective data eyes in the selected module control signal, e.g., MCS*i*. The selected signals MCS*i* and Ck*i* are provided to the sampler 930, which samples MCS*i* according to the rising edges of CK*i*.

[0053] As stated above, since the isolation devices 118 are inserted in the data paths between the MCH 101 and the respective groups of memory devices 112, the MCH 101 no longer has direct control of the memory devices 112. Thus, conventional read/write leveling techniques are not sufficient for managing read/write data timing. In one embodiment, the isolation devices 118 includes signal alignment mechanism to time the transmission of read data signals based on timing information derived from a prior write operation, as discussed further below.

[0054] FIG. 12A is a timing diagram for a write operation according to one embodiment. As shown, after a write command W/C associated with the write operation is received by the module control circuit 116 at time t1, the module control circuit 116 outputs one or more enable signals EN at time t2 in response to the write commands. The one or more enable signals are received by an isolation device 118 at time t3, which afterwards receives one or more strobe signal DQS from the MCH 101 at time t4. Note that the same enable signal may be received by another isolation device 118 at time t3′, which can be in a different cycle of the system clock MCK from the cycle which t3 is in. The time interval between t4 and t1 is consistent with a write latency W.L. associated with the system 100, and is controllable by the MCH 101 and knowable to the isolation device 118. The time interval between t4 and t3, referred to hereafter as an enable-to-write data delay EWD, can be determined by the isolation device 118 since both these signals are received by the isolation device. Based on such determination, the isolation device 118 can have knowledge of the time interval between t3 and t1, referred to hereafter as a command-to-enable delay CED, which can be used by the isolation device 118 to properly time transmission of read data to the MCH, as explained further below.

[0055] FIG. 12B is a timing diagram for a read operation according to one embodiment. As shown, after a read command R/C associated with the read operation is received by the module control circuit 116 at time t5, the module control circuit 116 outputs one or more enable signals EN at time t6 in response to the read commands. The one or more enable signals are received by an isolation device 118 at time t7, which afterwards receives at time t8 read data signals (not shown) and one or more strobe signal DQS from the respective group of memory devices. Note that the same enable signal may be received by another isolation device 118 at time t3′, which can be in a different cycle of the system clock MCK from the cycle which t3 is in. Thus, the enable signals alone cannot be used to time the transmission of the read signals by the isolation devices 118.

[0056] With knowledge of the time interval between t7 and t5, which should be about the same as the time interval between t3 and t1, i.e., the command-to-enable delay CED, however, the isolation device can add a proper amount of delay to the read data signals and the one or more DQS signal such that the read data signals and the one or more DQS signal are transmitted at time t9 by the isolation device to the MCH 101 via the respective group of data/strobe signal lines 130, with the time interval between t9 and t5 being consistent with a read latency R.L. associated with the system 100.

[0057] The time interval between t4 and t3, i.e., the enable to write data delay EWD, is determined by the delay control circuit 650 in the ID control circuit 310, as shown in FIG. 6. According to one embodiment, as shown in FIG. 13, the delay control circuit 650 includes a peramble detector 1310 to detect a write preamble in the DQS, a flip-flop circuit 1320 having an enable input EN receiving one of the module control signals and a clock input CK receiving the buffered module clock signal CK0, and a counter circuit 1330 having a Start input receiving the one of the module control signals, a Stop input receiving an output of the flip-flop circuit 1320. Thus, the output of the counter circuit, i.e., the delay signal DS, would indicate a time interval from when the write preamble is detected and when the one of the module control signal is received.

[0058] FIG. 14 illustrates a DQ or DQS routing circuit 320 or 620 according to an embodiment. As shown, the DQ/DQS routing circuit 320/620 includes a DQ/DQS pin 1401 that is coupled to the corresponding DQ/DQS signal line 322/324, a set of one or more DQS pins 1402 that is coupled to a corresponding module DQ/DQS line(s) Y/Y_{DQS}, or YA/YA_{DQS} and YB/YB_{DQS}. The DQ/DQS routing circuit 320/620 further includes a write strobe buffer 1410 that buffers write data/strobe, and a write data/strobe receiver 1420 that samples the write data/strobe. The DQ/DQS routing circuit 320/620 further includes a plurality of write paths 1430 that are selectable or can be selectively enabled by one or more of the module control signals, such as the enable signals ENA and ENB.

[0059] The DQS routing circuit further includes a plurality of read paths 1450 that are selectable by the one or more of the module control signals. Output from the seleted read path is delayed in a delay circuit 1460 by an amount controlled by the delay signal DS, and sampled by a sampler circuit 1470. The sampled read data/strobe is transmitted by transmitter 1480 onto the corresponding data/strobe signal line 322/324 via the DQ/DQS pin 1401.

[0060] FIG. 15 illustrates a DQS routing circuit 620 according to an embodiment. As shown, the DQS routing circuit 620 includes a first DQS pin 1501 that is coupled to a corresponding DQS signal line 324, a second DQS pin 1502A that is coupled to a corresponding module DQS line YA_{DQS}, a third DQS pin 1502B that is coupled to a corresponding module DQS line YB_{DQS}. The DQS routing circuit 620 further includes a first write strobe path coupled between the first DQS pin 1501 and the second DQS pin 1502A and a second write strobe path coupled between the first DQS pin 1501 and the third DQS pin 1502B. The first write strobe path includes a write strobe buffer 1510 that buffers a write strobe, a write strobe receiver 1520 that samples the write strobe according to the buffered module signal CK0. The sampled write strobe is provided to the DQ routing circuits 320 as the write strobe WDQS. The first write strobe path further includes a first write strobe transmitter 1530A that transmits the write strobe to one or more memory devices 112 coupled to the module strobe line YA_{DQS}. The second write strobe path includes the write strobe buffer 1510, the write strobe receiver 1520, and a second write strobe transmitter 1530B that transmits the write strobe to one or more memory devices 112 coupled to the module strobe line YB_{DQS}. The first and second write strobe transmitters, 1530A and 1530B, are controlled by two enable signals, ENA and ENB, respectively, such that the first write strobe path and the second write strobe path can be selectively enabled/disabled by the enable signals, ENA and ENB.

[0061] The DQS routing circuit further includes a read strobe path coupled between the first DQS pin 1501 and a selected one of the second and third DQS pins 1502A and 1502B. In the read strobe path, a select circuit 1550 (e.g., a multiplexor) selects either a read strobe signal received via DQS pin 1502A or a read strobe signal received via DQS pin 1502B based on one or both of the enable signals ENA or ENB. The selected read strobe signal is delayed in a delay circuit 1560 by an amount controlled by the delay signal DS, and sampled by a sampler circuit 1570 according to the buffered module clock signal CK0. The sampled read strobe is provided to the DQ routing circuits 320 as the read strobe RDQS and is transmitted by transmitter 1580 onto the corresponding strobe signal line 324 via the first DQS pin 1501.

[0062] FIG. 16 illustrates a DQ routing circuit 320 according to an embodiment. As shown, the DQ routing circuit 320 includes a first DQ pin 1601 that is coupled to a corresponding DQ signal line 130, a second DQ pin 1602A that is coupled to a corresponding module DQ line YA_{DQ}, a third DQ pin 1602B that is coupled to a corresponding module DQ line YB_{DQ}. The DQ routing circuit 320 further includes a first write data path coupled between the first DQ pin 1601 and the second DQ pin 1602A and a second write data path coupled between the first DQ pin 1601 and the third DQ pin 1602B. The first write data path includes a write data buffer 1610, a write data receiver 1620 that samples write data according to the write strobe WDQS from the DQS routing circuit 620, and a first write data transmitter 1630A that transmits the write data to one or more memory devices 112 coupled to the module data line YA_{DQ}. The second write data path includes the write data buffer 1610, the write data receiver 1620, and a second write data transmitter 1630B that transmits the write data to one or more memory devices 112 coupled to the module data line YB_{DQ}. The first and second write data transmitters, 1530A and 1530B, are controlled by two enable signals, ENA and ENB, respectively. Thus, the first write data path and the second write data path can be selectively enabled/disabled by the enable signals, ENA and ENB.

[0063] The DQ routing circuit further includes a read data path coupled between the first DQ pin 1601 and a selected one of the second and third DQ pins 1602A and 1602B. In the read data path, a select circuit 1650 (e.g., a multiplexor) selects either a read data signal received via DQ pin 1602A or a read data signal received via DQ pin 1602B based on one of the enable signals ENA or ENB. The selected read data signal is delayed in a delay circuit 1660 by an amount controlled by the delay signal DS. The delayed read data signal is then sampled by a receiver circuit 1670 according to the read strobe RDQS from the DQS routing circuit 620, and transmitted by transmitter 1680 onto the corresponding data signal line 130 via the first DQ pin 1601.

[0064] FIG. 17 illustrate a delay circuit 1560 or 1660 according to an embodiment. As shown, the delay circuit 1560 or 1660 includes a plurality of delay stages, such as delay stages 1710, 1720, and 1730, each delaying a read data or read strobe signal from the select circuit 1550/1650 by a predetermined amount. The delay circuit 1560 or 1660 further includes a select circuit 1740 (e.g., a multiplexor) that selects from among the read data or read strobe signal and the outputs from the delay stages according to the delay signal DS. The output of the select circuit 1740, is provided to the sampler circuit 1570 or 1670, either directly or after being buffered by a buffer circuit 1750.

[0065] Thus, as shown in FIG. 18, in one embodiment, a memory module 110 operates in the memory system 100 according to a method 1800. In the method, during a write operation, one or more first enable signals are received by an isolation device 118 from a module control circuit or module controller 116 (1810). The one or more first enable signals are generated by the module controller 116 in response to write command signals received from the MCH 101 and are used by the isolation device 118 to enable a write path to allow write data be communicated to a selected subgroup of memory devices among the group of memory devices coupled to the isolation device 118. After a time interval from receiving the one or more first enable signals, write data DQ and write strobe DQS are received by the isolation device 118 from the MCH 101 (1820). In one embodiment, upon receiving the one or more first enable signal, a counter is started, which is stopped when the write data DQ or write strobe DQS is received. Thus, a time interval EWD between receiving the one or more enable signals and receiving the write strobe signal DQS is recorded.

[0066] Since the time interval between the arrival of the command signals from the MCH 101 and the arrival of the write data/strobe signal DQ/DQS from the MCH 101 is a fixed time interval determined by a write latency parameter associated with the system 100, the time interval EWD can be used to ascertain a time interval CED between the time when a command signal is received by the memory module 110 and the time when the one or more enable signals are received by the isolation device 118. The time interval CED can be used by the isolation device 118 to properly time the transmission of read data to the MCH 101, as described above and explained further below.

[0067] As shown in FIG. 18, a delay signal DS is generated according to the time interval EWD (1830). Concurrent to receiving the write strobe signal DQS, the isolation device 118 also receives a set of write data signals DQ (1840). The received write data signals are transmitted to the subgroup of memory devices (1850), which are selected from the group of memory devices coupled to the isolation device 118 by the one or more first enable signals.

[0068] During a read operation, one or more second enable signals are received by the isolation device 118 from the module controller 116 (1860). The one or more second enable signals are generated by the module controller 116 in response to read command signals received from the MCH 101, and are used by the isolation device 118 to select a subgroup of memory devices from which to receive read data. Afterwards, a read strobe signal DQS and a set of read data signal DQ are received from the selected subgroup of memory devices (1870). To properly time the transmission of the DQS and DQ signals to the MCH 101, the DQS and DQ signals are adjusted (e.g., delayed) according to the delay signal DS, such that the DQS and DQ signals follow a read command by a time interval consistent with a read latency parameter associated with the system 100.

WE CLAIM:

1. A memory module to operate in a memory system with a memory controller, the memory system operating according to a system clock, comprising:

a module control device to receive command signals from the memory controller and to output module command signals and module control signals;

memory devices organized in groups, each group including at least two subgroups, each subgroup including at least one memory device, the memory devices receiving the module command signals from the module control device; and

a plurality of buffer circuits to receive the module control signals, a respective buffer circuit corresponding to a respective group of memory devices, the respective buffer circuit selecting a subgroup in the respective group of memory devices to communicate data with the memory controller in response to the module control signals; and

wherein the plurality of buffer circuits are distributed across a surface of the memory module such that each module control signal arrives at the plurality of buffer circuits at different points of time across more than one clock cycle of the system clock.

2. The memory module of claim 1, wherein each buffer circuit includes receiver circuits to receive corresponding ones of the module control signals, each receiver circuit including a metastability detection circuit to detect a metastability condition in the corresponding module control signal.

3. The memory module of claim 2, wherein the metastability detection circuit includes a plurality of sampler circuits each sampling one of the corresponding module control signal and one or more delayed versions of the corresponding module control signal according to one of a clock signal and one or more delayed versions of the clock signal, the clock signal being derived from the system clock, and wherein the metastability condition is detected based on outputs from the plurality of samplers.

4. The memory module of claim 2, wherein each receiver circuit further includes one or more clock signal delay circuits, one or more module control signal delay circuits, a clock signal selection circuit, and a module control signal selection circuit, and wherein:

each clock signal delay circuit adds a predetermined amount of delay to a clock signal, the clock signal being derived from the system clock;

each module control signal delay circuit adds a predetermined amount of delay to a corresponding module control signal;

the clock signal selection circuit receives the clock signal and outputs from the one or more clock delay circuits and outputs a selected clock signal based on the metastability condition; and

the module control signal selection circuit receives the corresponding module control signal and outputs from the one or more module control signal delay circuits and outputs a selected module control signal based on the metastability condition.

5. The memory module of claim 4, wherein each receiver circuit further includes a sampler circuit to sample the selected module control signal according to the selected clock signal.

6. The memory module of claim 1, wherein each buffer circuit comprises:

a time interval determination circuit to receive, during a write operation, a first signal from the module controller and a second signal from the memory controller and to generate a signal indicating a time interval between the first signal and the second signal;

a delay circuit to delay a read data signal according to the signal indicating the time interval.

7. A buffer circuit to be coupled between a memory controller and a group of memory devices, the group of memory devices including at least two subgroups, each subgroup including at least one memory device, comprising:

receiver circuits to receive control signals, each receiver circuit including:

a metastability detection circuit to detect a metastability condition in a corresponding control signal;

a selection circuit to receive the corresponding control signal and one or more delayed versions thereof and to output a selected control signal based on the metastability condition; and

a sampler circuit to generate a received control signal from the selected control signal; and

isolation circuits selecting a subgroup in the group of memory devices to communicate data with the memory controller in response to the received control signals.

8. The buffer circuit of claim 7, wherein the isolation circuits isolate one or more subgroup from the memory controller during a write operation such that the memory controller sees only a memory device load from the selected subgroup during the write operation.

9. The buffer circuit of claim 7, wherein each receiver circuit further includes one or more clock signal delay circuits, one or more control signal delay circuits, wherein the selection circuit includes a clock signal selection circuit and a control signal selection circuit, and wherein:

each clock signal delay circuit adds a predetermined amount of delay to a clock signal;

each control signal delay circuit adds a predetermined amount of delay to a corresponding control signal;

the clock signal selection circuit receives the clock signal and signals output from the one or more clock delay circuits, and outputs a selected clock signal based on the metastability condition; and

the control signal selection circuit receives the control signal and signals output from the one or more control signal delay circuits, and outputs a selected control signal based on the metastability condition.

10. The buffer circuit of claim 9, wherein the sampler circuit samples the selected control signal according to the selected clock signal to generate the received control signal.

11. A method of operating a memory module in a memory system having a memory controller, the memory module having memory devices, the memory system operating according to a system clock, comprising:

receiving command signals from the memory controller;

generating module signals from the command signals;

receiving the module signals at different points of time across more than one clock cycle of the system clock; and

selecting some of the memory devices to communicate data with the memory controller based on the module signals.

12. The method of claim 11, further comprising:

determining a metastability condition associated with a respective module control signal;

generating a received module control signal for each respective module control signal based on the associated metastability condition, the received module control signal selecting a first signal path for communicating data between one or more memory devices and the memory controller.

13. The method of claim 12, wherein the received module control signal disables a second signal path to isolate one or more memory devices from the memory controller.

14. The method of claim 11, further comprising:

receiving at least one first enable signal from within the memory module, the first enable signal for enabling a write signal path between the memory controller and a selected group of one or more memory devices;

receiving one of a write strobe signal and write data signal from the memory controller after a time interval from receiving the at least one first enable signal;

receiving at least one second enable signal for selecting a read signal path;

receiving a read data signal from the selected read signal path;

delaying the read data signal according to the time interval; and

transmitting the delayed read data signal to the memory controller.

15. A buffer circuit for use in a memory module coupled to a memory controller, the memory module including a module controller, comprising:

a time interval determination circuit to receive, during a write operation, a first signal from the module controller and a second signal from the memory controller and to generate a signal indicating a time interval between the first signal and the second signal;

a delay circuit to delay a read data signal according to the signal indicating the time interval.

16. The buffer circuit of claim 15, wherein the first signal is an enable signal to enable a write path between a selected group of one or more memory devices and the memory controller during the write operation.

17. The buffer circuit of claim 15, wherein the second signal is a write strobe signal.

18. The buffer circuit of claim 15, wherein the second signal has a preamble, and wherein the time interval determination circuit includes a preamble detector to detect the preamble of the second signal.

19. The buffer circuit of claim 18, wherein the time interval determination circuit further includes a counter to count a number of clock cycles between receipt of the first signal and detection of the preamble of the second signal.

20. A method of operating a memory module coupled to a memory controller, the memory module having memory devices, comprising:

receiving at least one first enable signal from within the memory module, the first enable signal for enabling a write signal path between the memory controller and a selected group of one or more memory devices;

receiving one of a write data signal and write strobe signal from the memory controller after a time interval from receiving the at least one first enable signal;

receiving at least one second enable signal for selecting a read signal path;

receiving a read data signal from the selected read signal path;

delaying the read data signal according to the time interval; and transmitting the read data signal to the memory controller.

21. The method of claim 20, wherein delaying the read data signal includes delaying the read data signal such that it is lagging the second enable signal by the time interval.

22. The method of claim 20, wherein the write signal is a write strobe signal, the method further comprising:

detecting a preamble of the write strobe signal to determine the time interval.

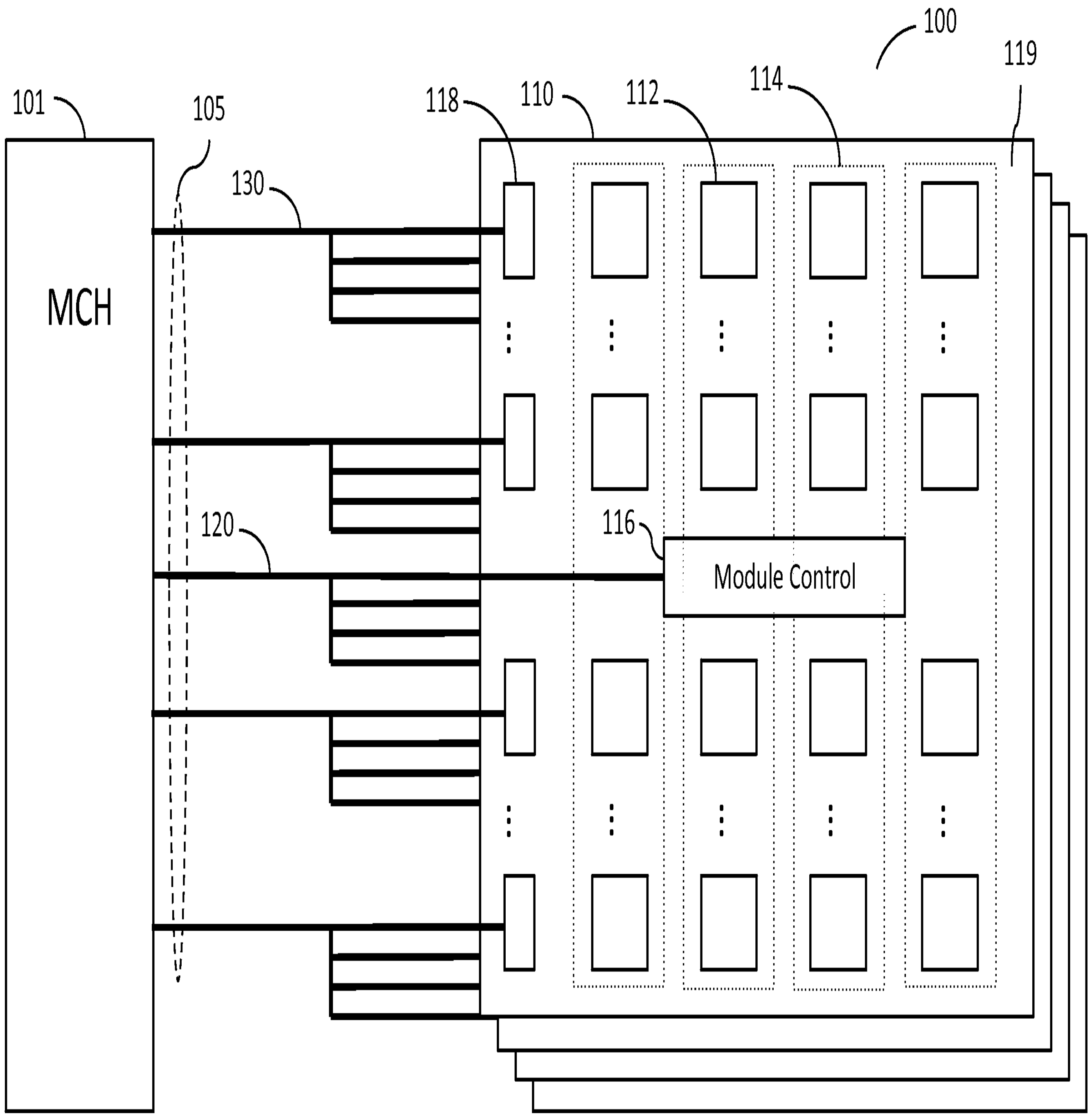


FIG. 1

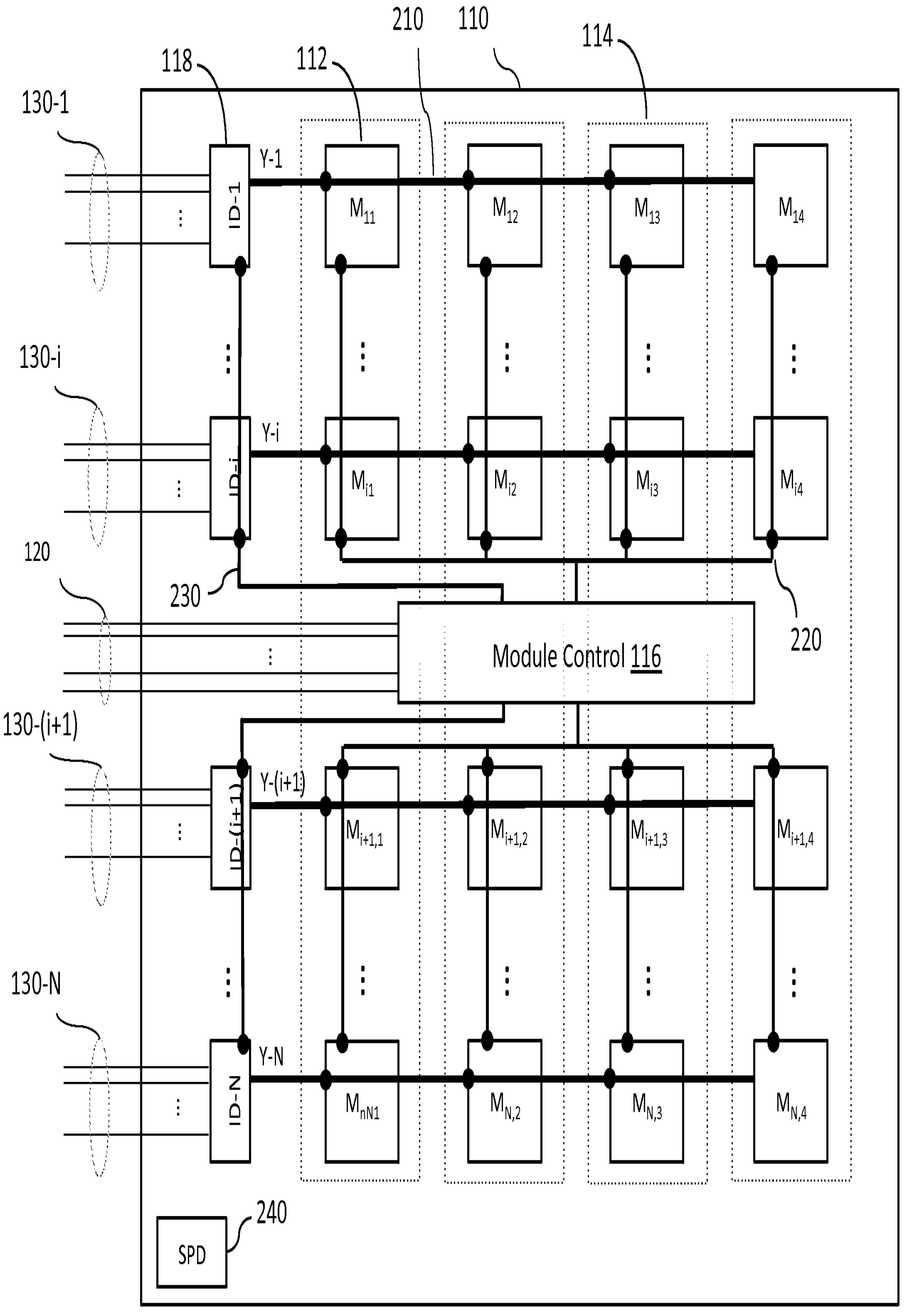
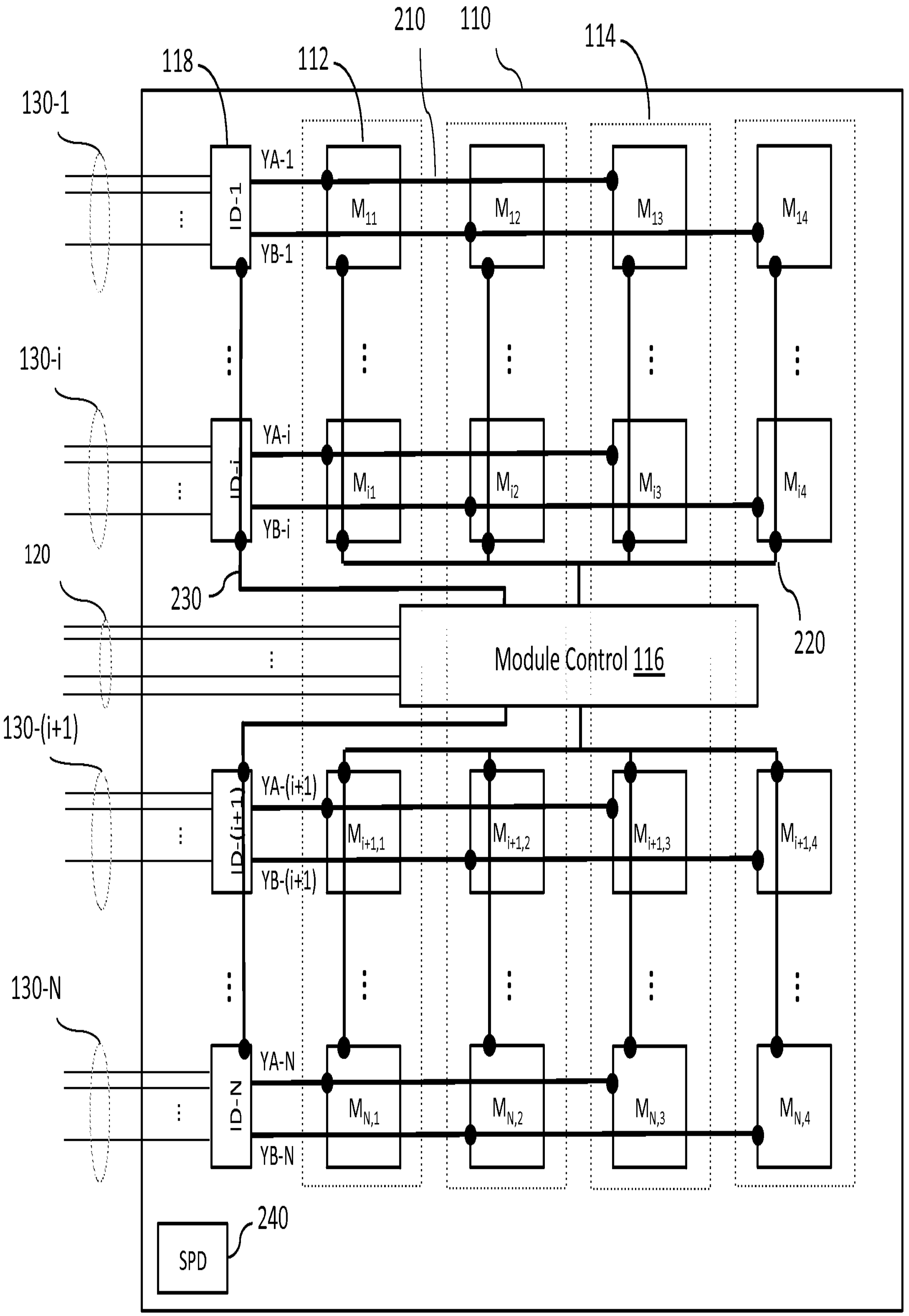
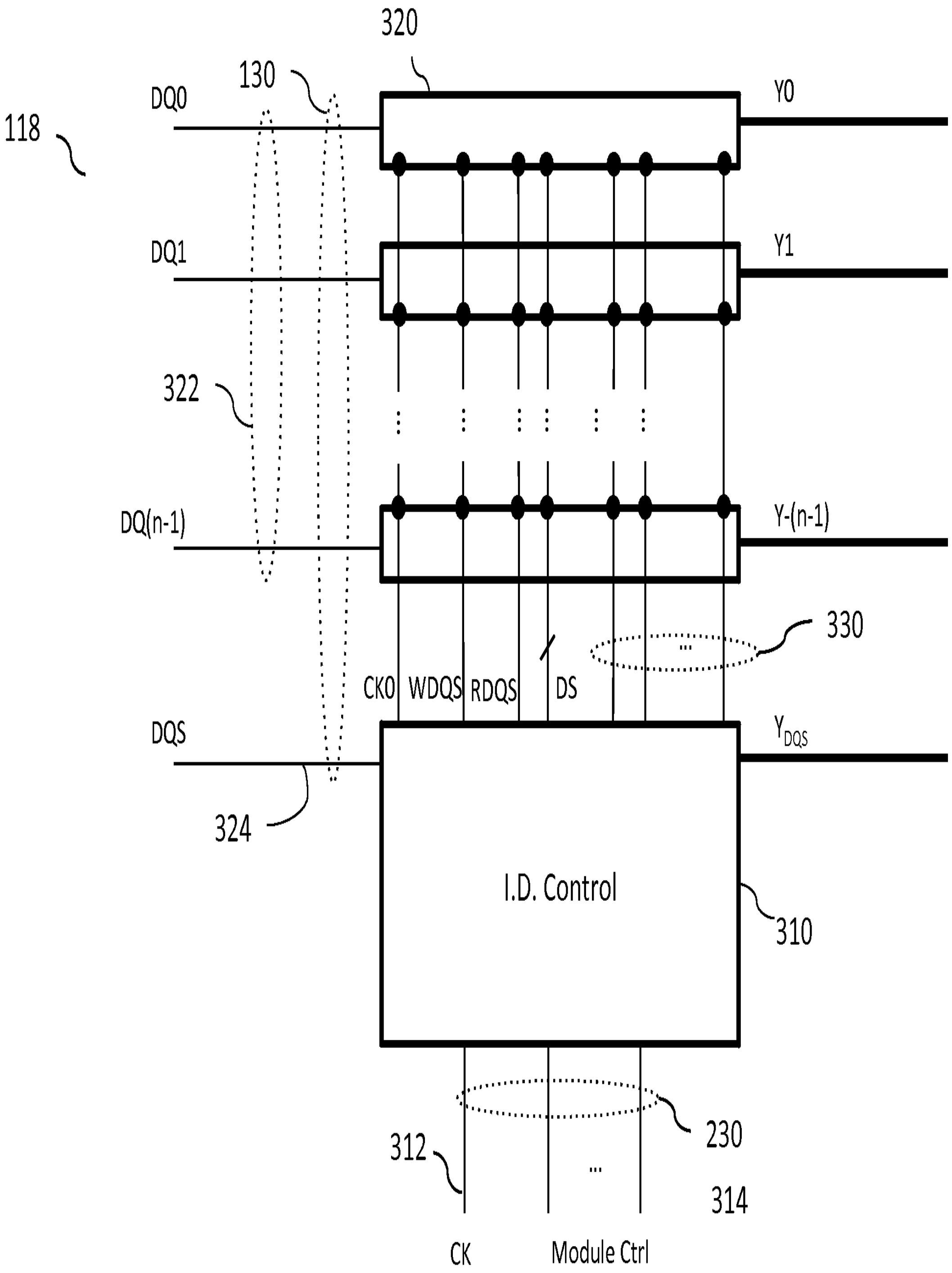

118
112
210
110
114
130-1
Y-1
ID-1
M11
M12
M13
M14
130-i
Y-i
ID-i
Mi1
Mi2
Mi3
Mi4
120
230
220
Module Control 116
130-(i+1)
Y-(i+1)
ID-(i+1)
Mi+1,1
Mi+1,2
Mi+1,3
Mi+1,4
130-N
Y-N
ID-N
MnN1
MN,2
MN,3
MN,4
SPD
240

FIG. 2A



210
110
114
118
112
130-1
YA-1
ID-1
M_{11}
M_{12}
M_{13}
M_{14}
YB-1
130-i
YA-i
ID-i
M_{i1}
M_{i2}
M_{i3}
M_{i4}
YB-i
120
230
220
Module Control 116
130-(i+1)
YA-(i+1)
ID-(i+1)
$M_{i+1,1}$
$M_{i+1,2}$
$M_{i+1,3}$
$M_{i+1,4}$
YB-(i+1)
130-N
YA-N
ID-N
$M_{N,1}$
$M_{N,2}$
$M_{N,3}$
$M_{N,4}$
YB-N
SPD
240

FIG. 2B



320
130
DQ0
118
Y0
DQ1
Y1
322
DQ(n-1)
Y-(n-1)
330
...
CK0
WDQS
RDQS
DS
DQS
Y_{DQS}
324
I.D. Control
310
230
312
...
314
CK
Module Ctrl

FIG. 3

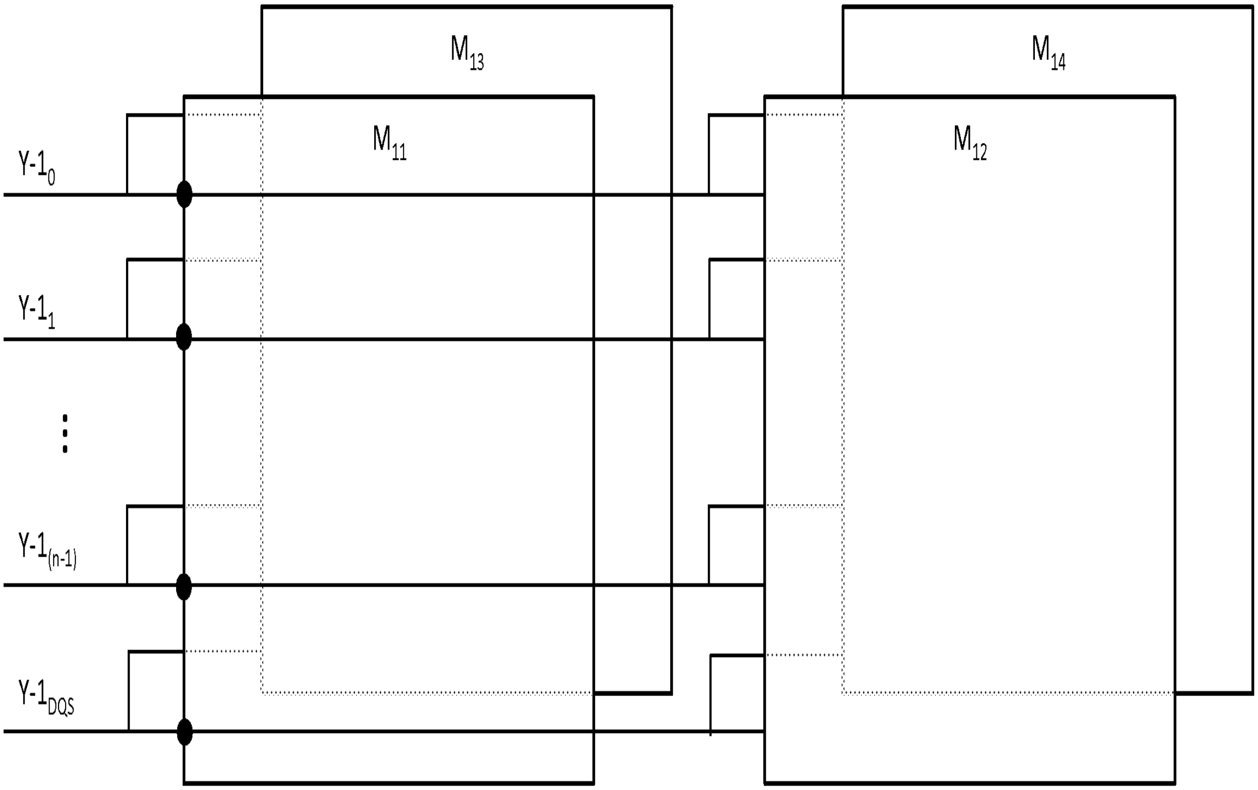


FIG. 4A

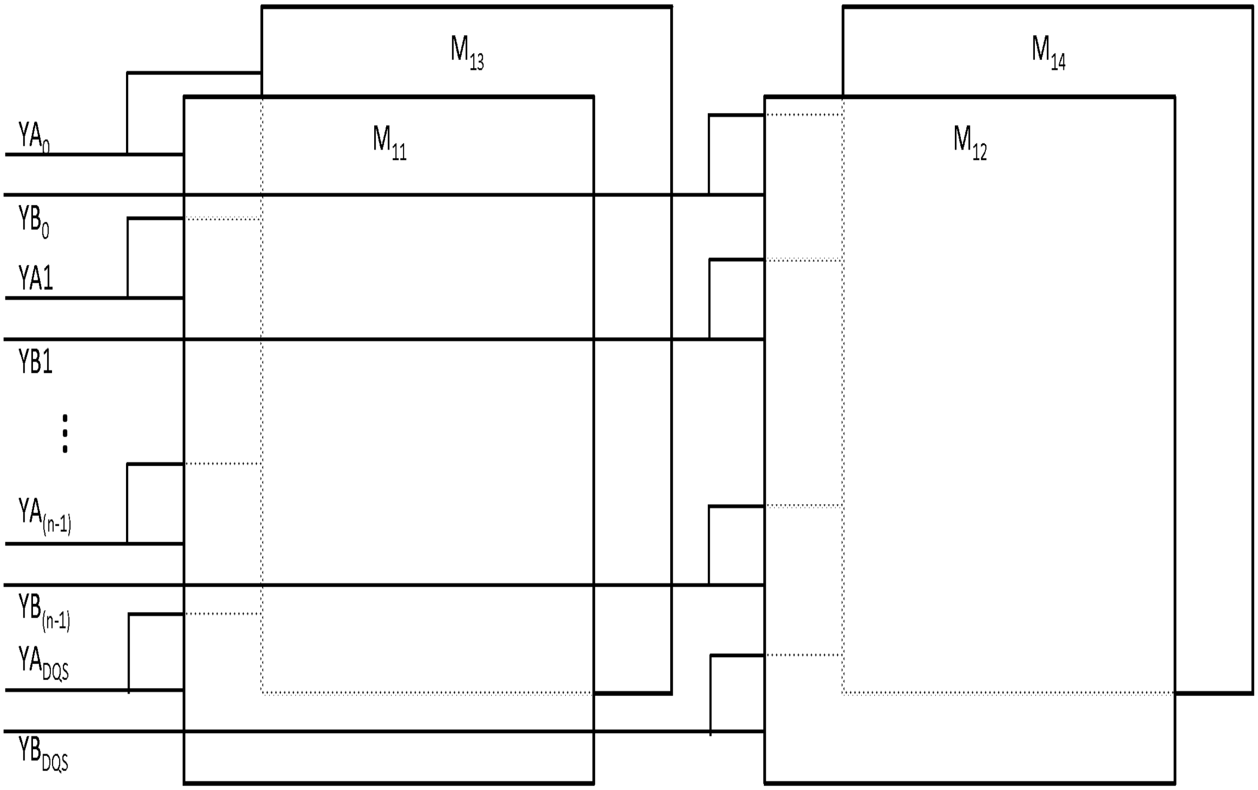


FIG. 4B

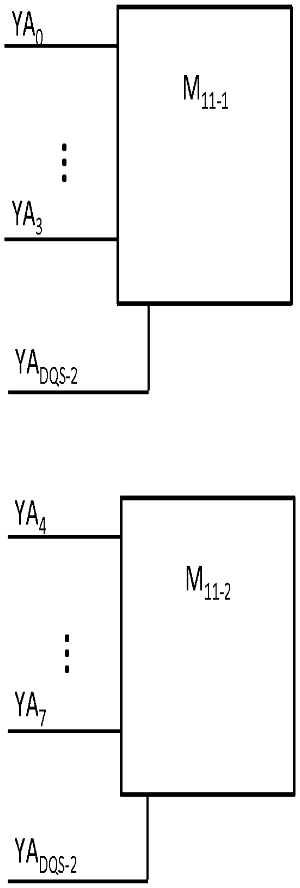

YA_0
$M_{11\text{-}1}$
YA_3
$YA_{DQS\text{-}2}$
YA_4
$M_{11\text{-}2}$
YA_7
$YA_{DQS\text{-}2}$

FIG. 5A

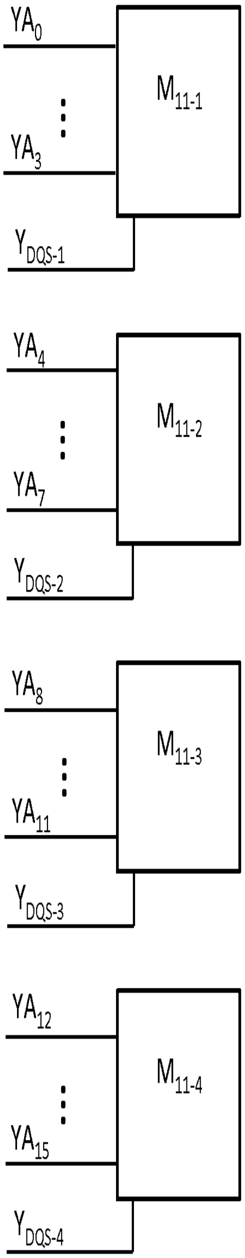

YA_0
$M_{11\text{-}1}$
YA_3
$Y_{DQS\text{-}1}$
YA_4
$M_{11\text{-}2}$
YA_7
$Y_{DQS\text{-}2}$
YA_8
$M_{11\text{-}3}$
YA_{11}
$Y_{DQS\text{-}3}$
YA_{12}
$M_{11\text{-}4}$
YA_{15}
$Y_{DQS\text{-}4}$

FIG. 5B

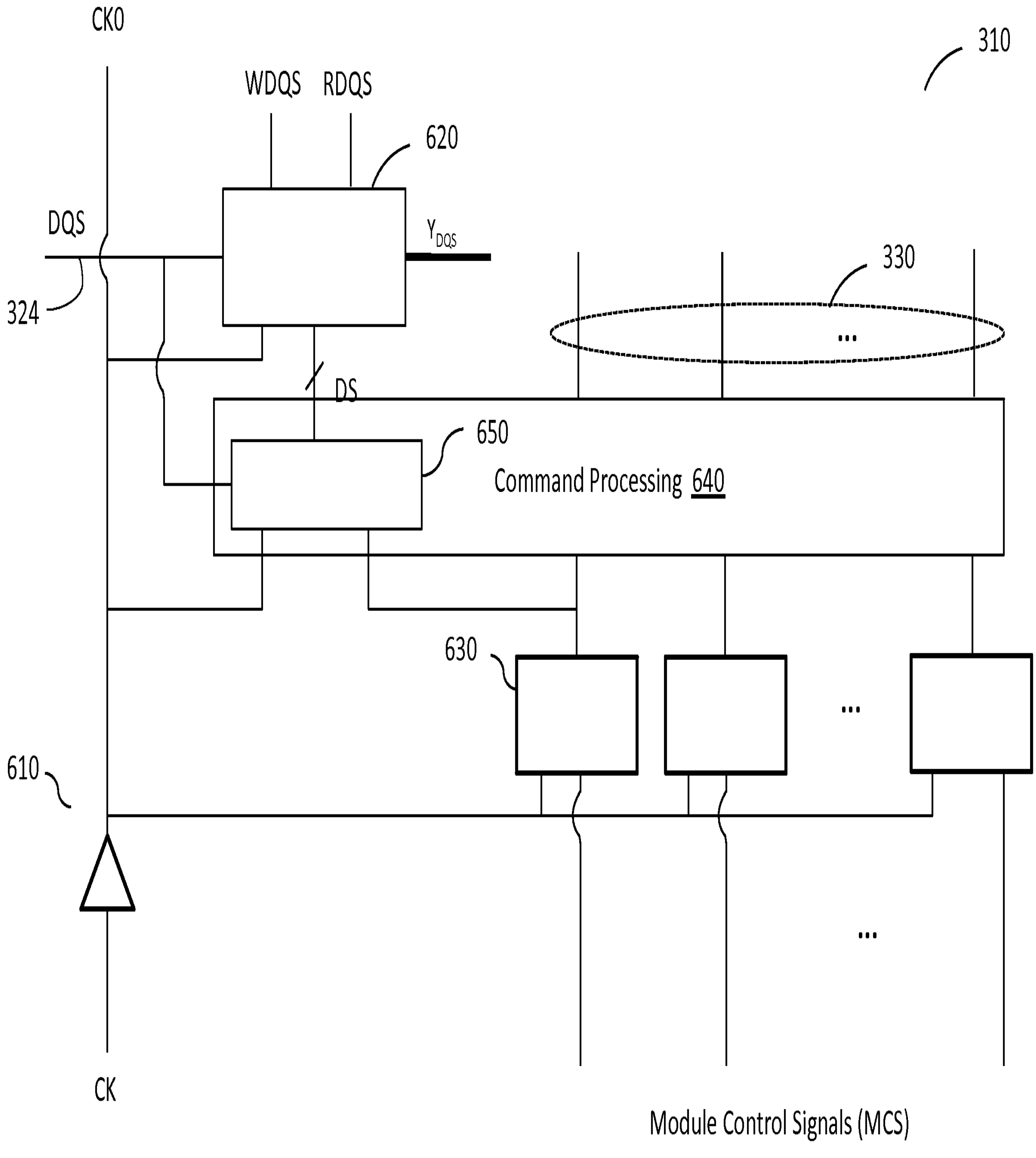

CK0
310
WDQS
RDQS
620
DQS
Y_{DQS}
330
324
...
DS
650
Command Processing 640
630
...
610
...
CK
Module Control Signals (MCS)

FIG. 6

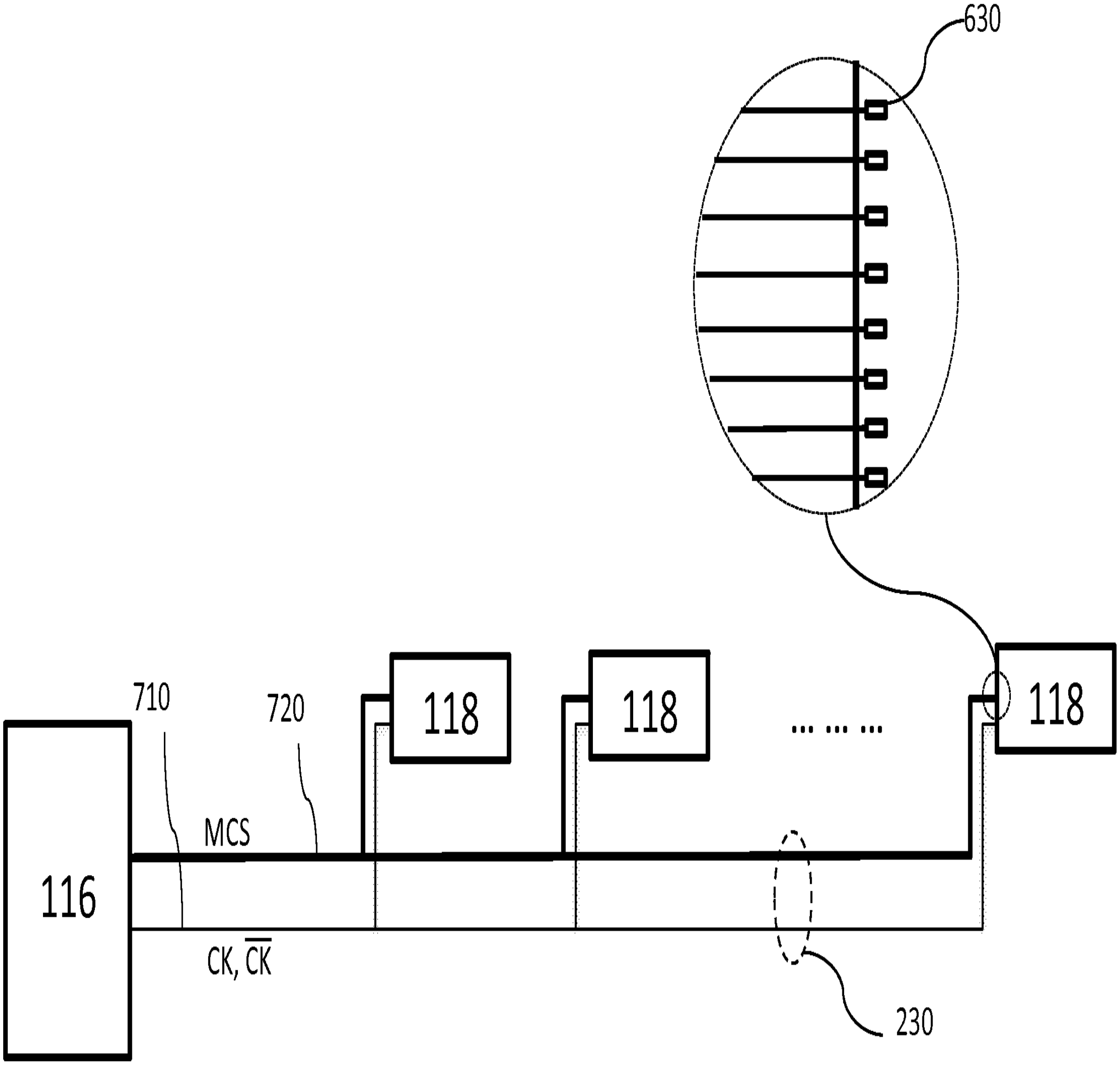


FIG. 7

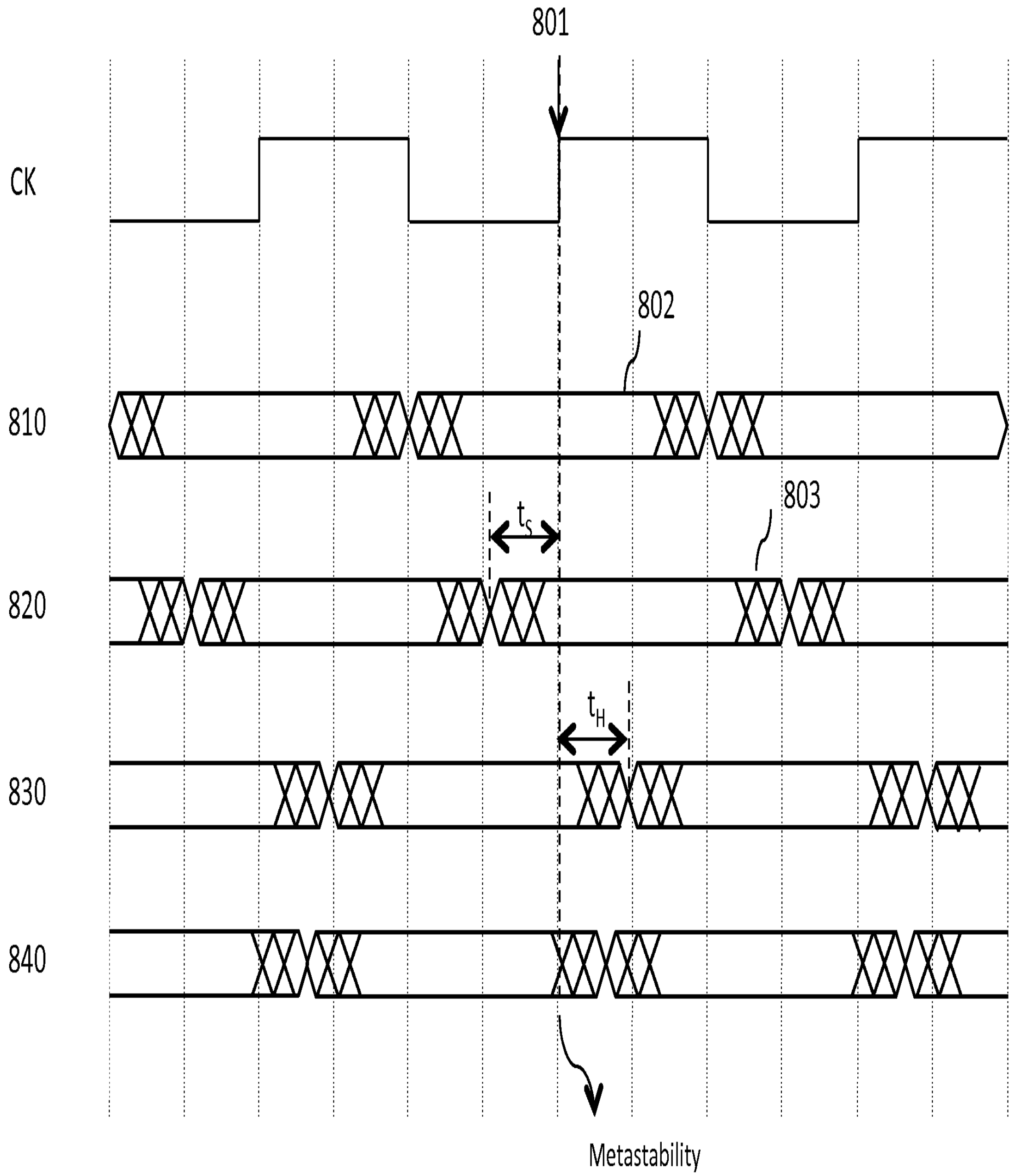


FIG. 8

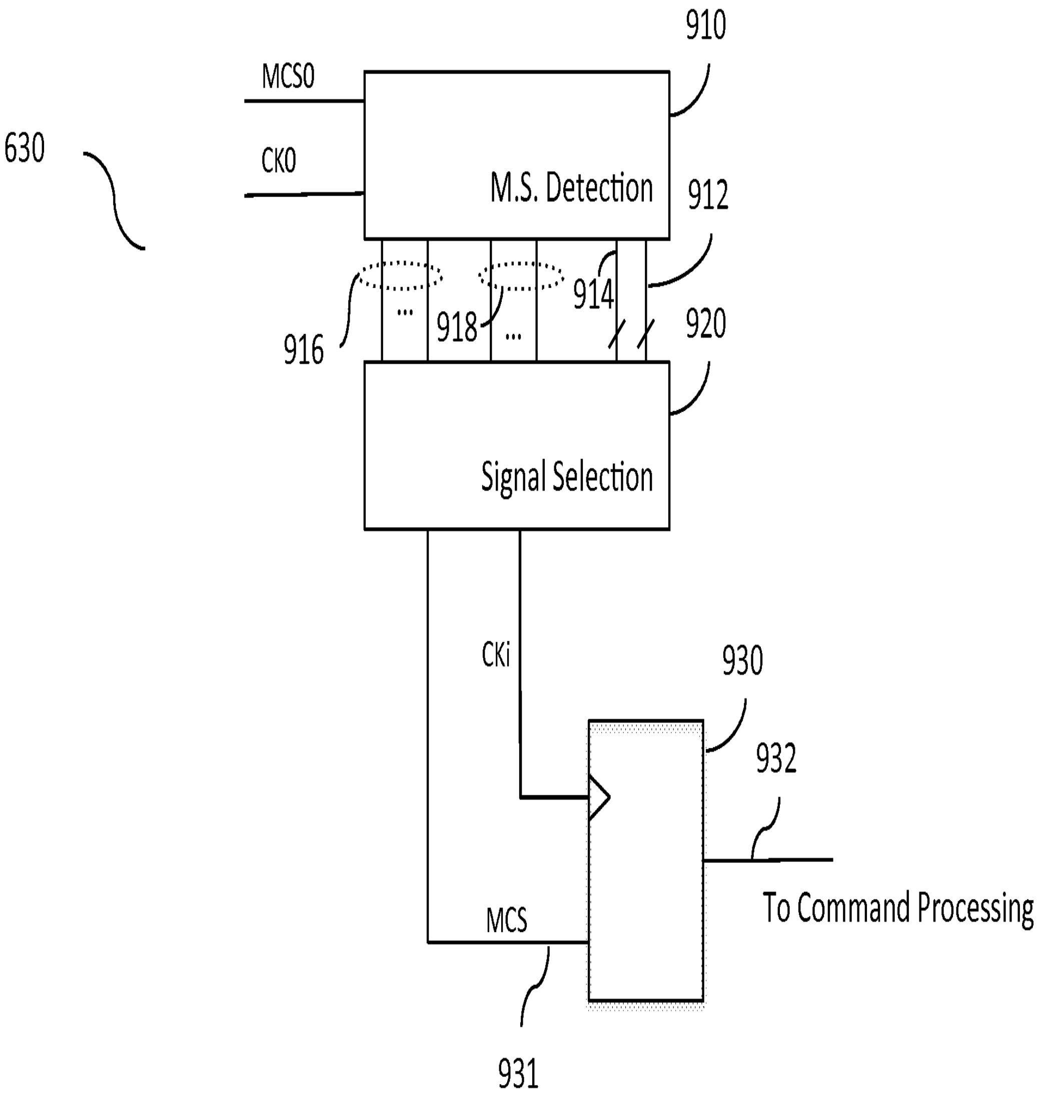

910
MCS0
CK0
630
M.S. Detection
912
914
916
918
920
Signal Selection
CKi
930
932
To Command Processing
MCS
931

FIG. 9

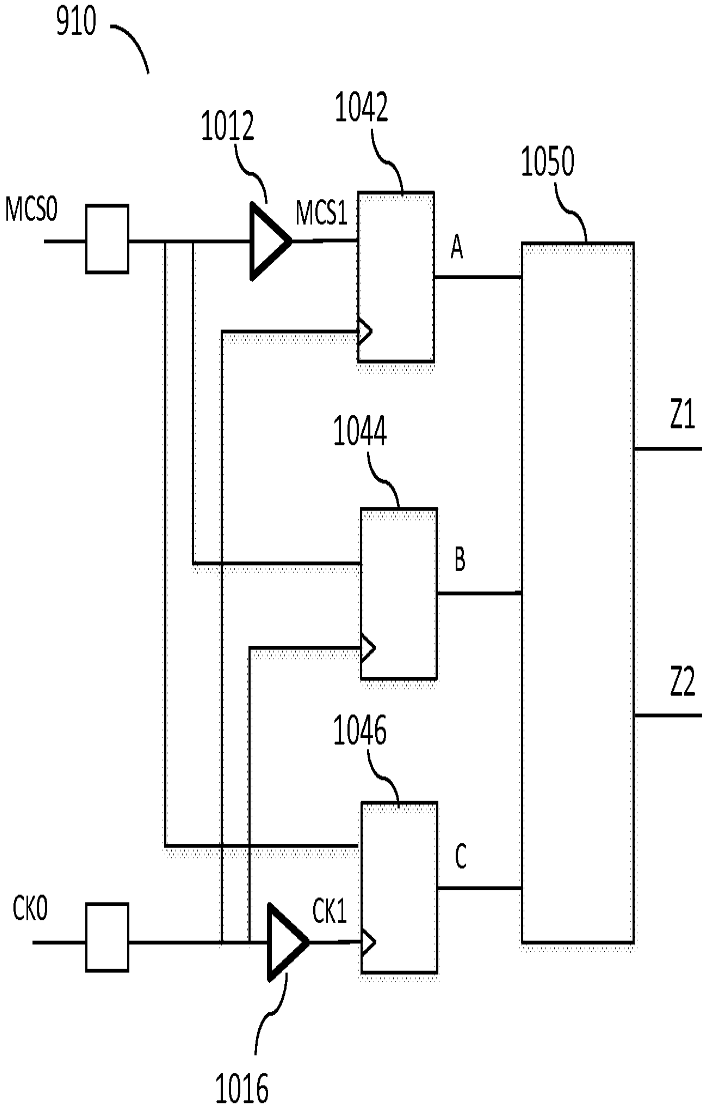


FIG. 10A

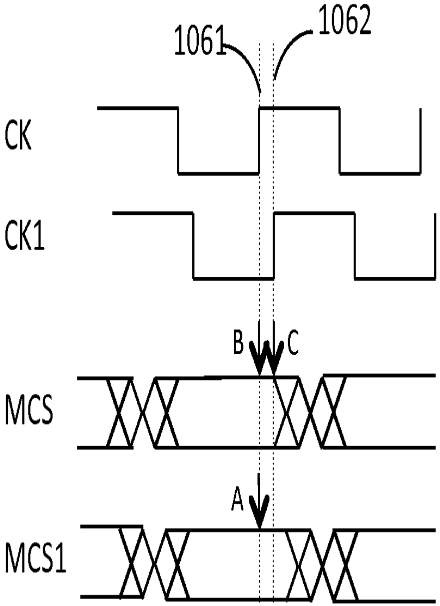


FIG. 10B

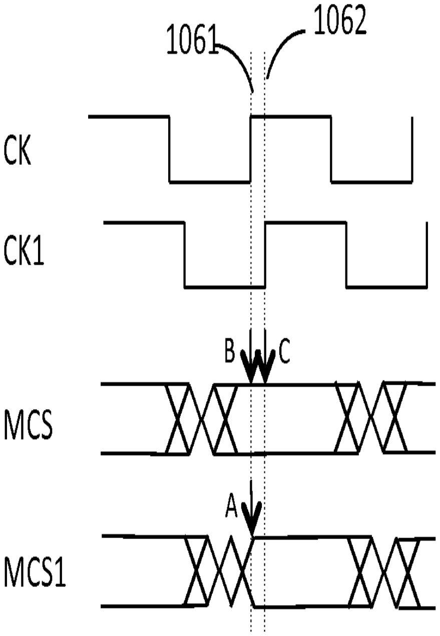


FIG. 10C

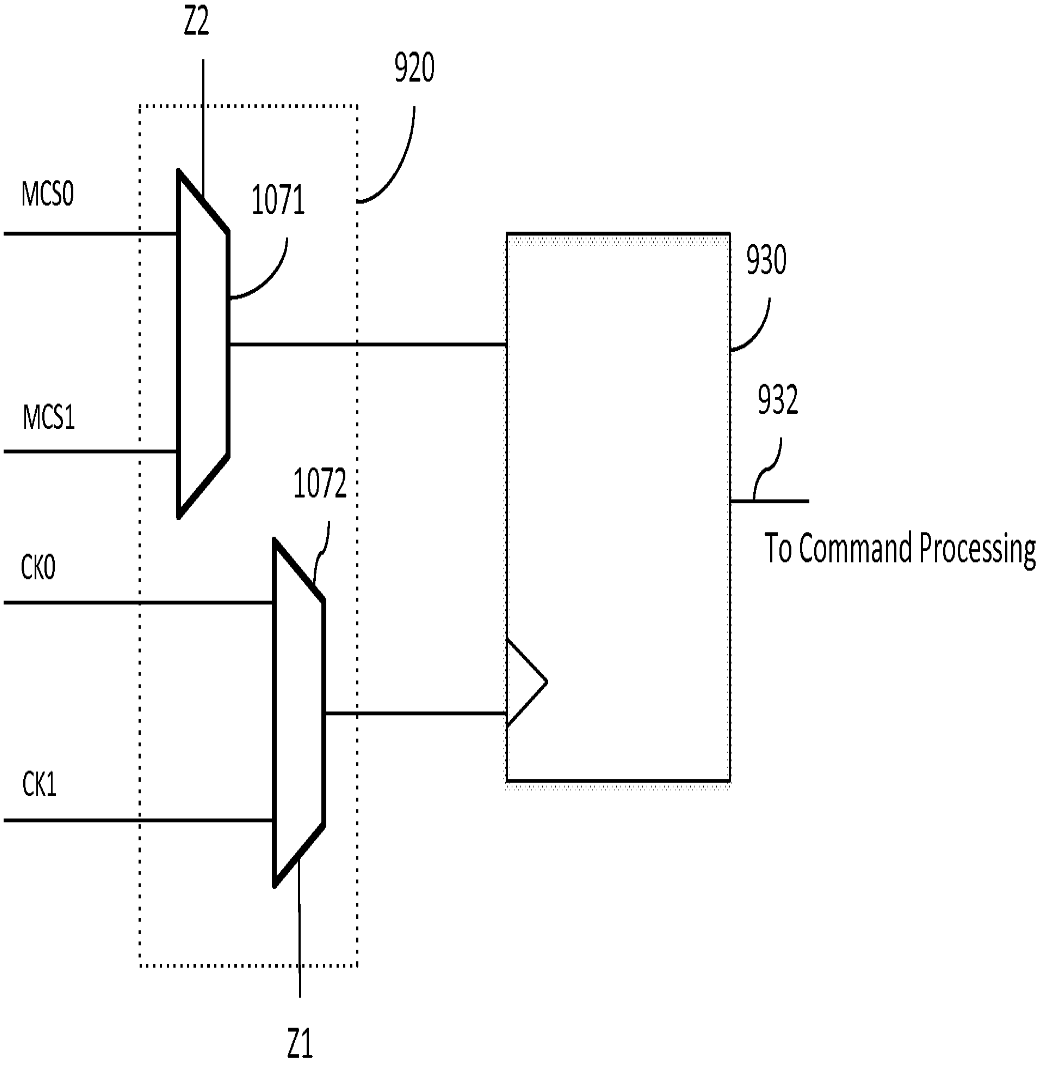

Z2
920
MCS0
1071
930
MCS1
932
1072
To Command Processing
CK0
CK1
Z1

FIG. 10D

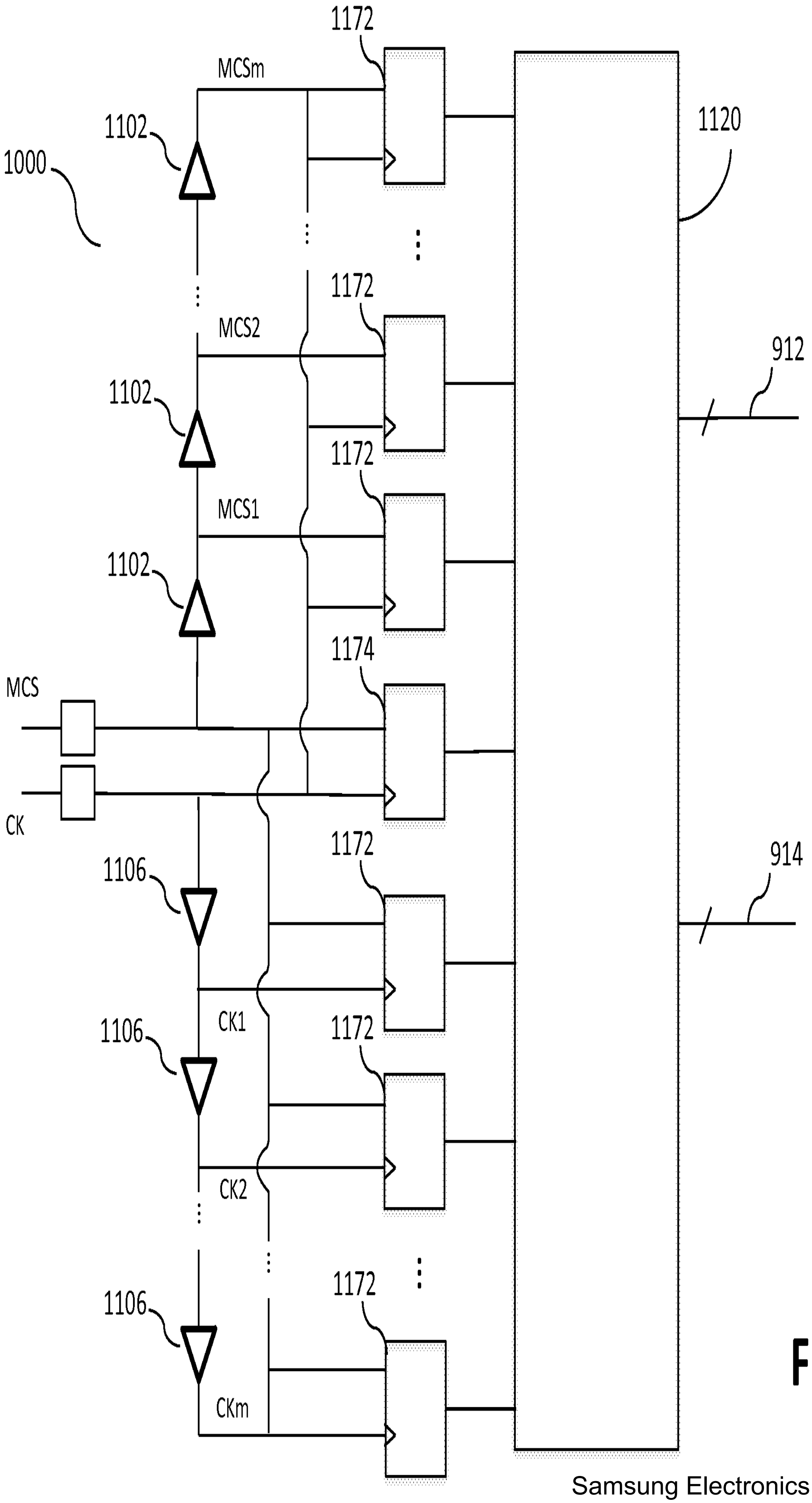

1172
MCSm
1102
1000
1120
1172
MCS2
1102
1172
912
MCS1
1102
1174
MCS
CK
1106
1172
914
CK1
1106
1172
CK2
1106
1172
CKm

FIG. 11A

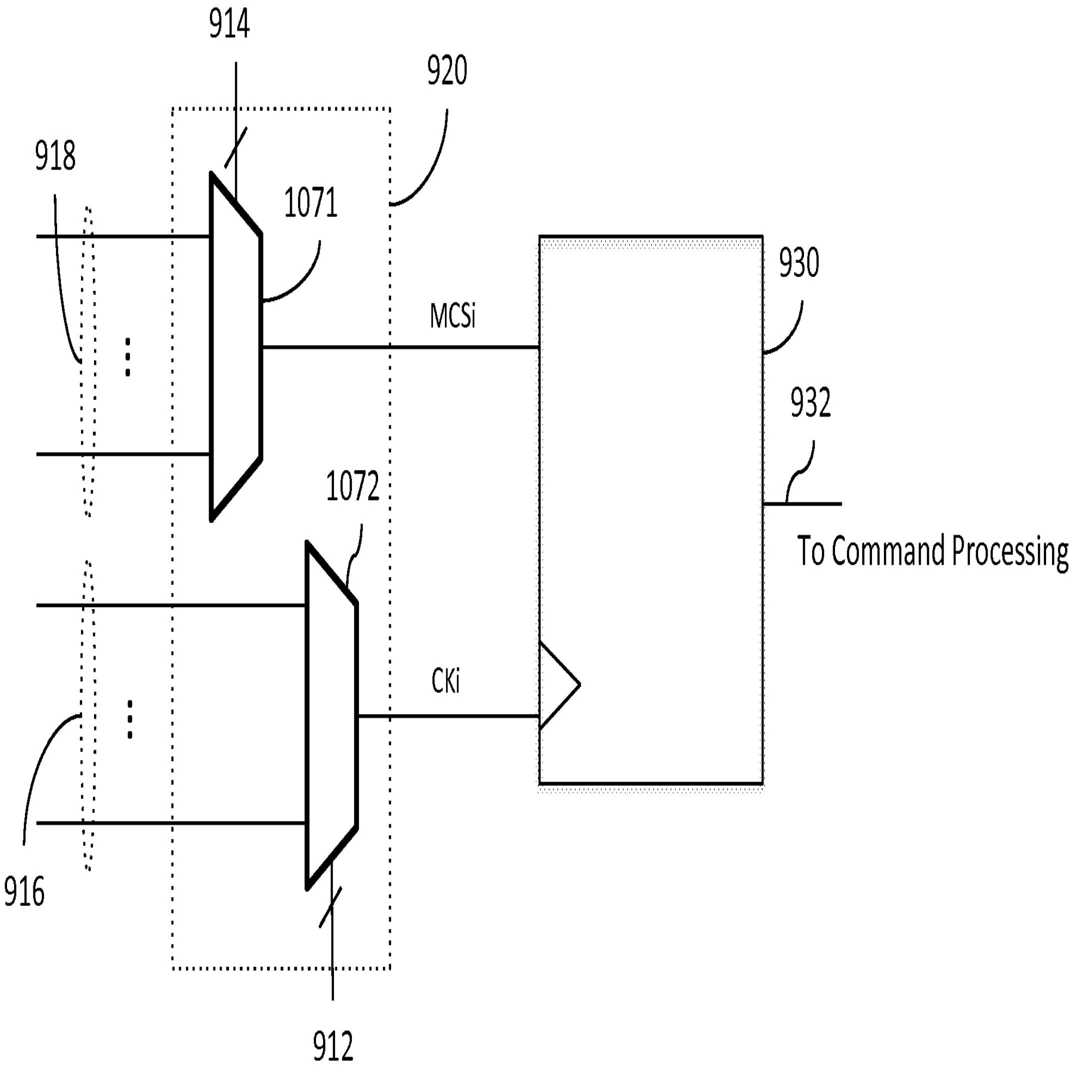

914
920
918
1071
930
MCSi
932
1072
To Command Processing
CKi
916
912

FIG. 11B

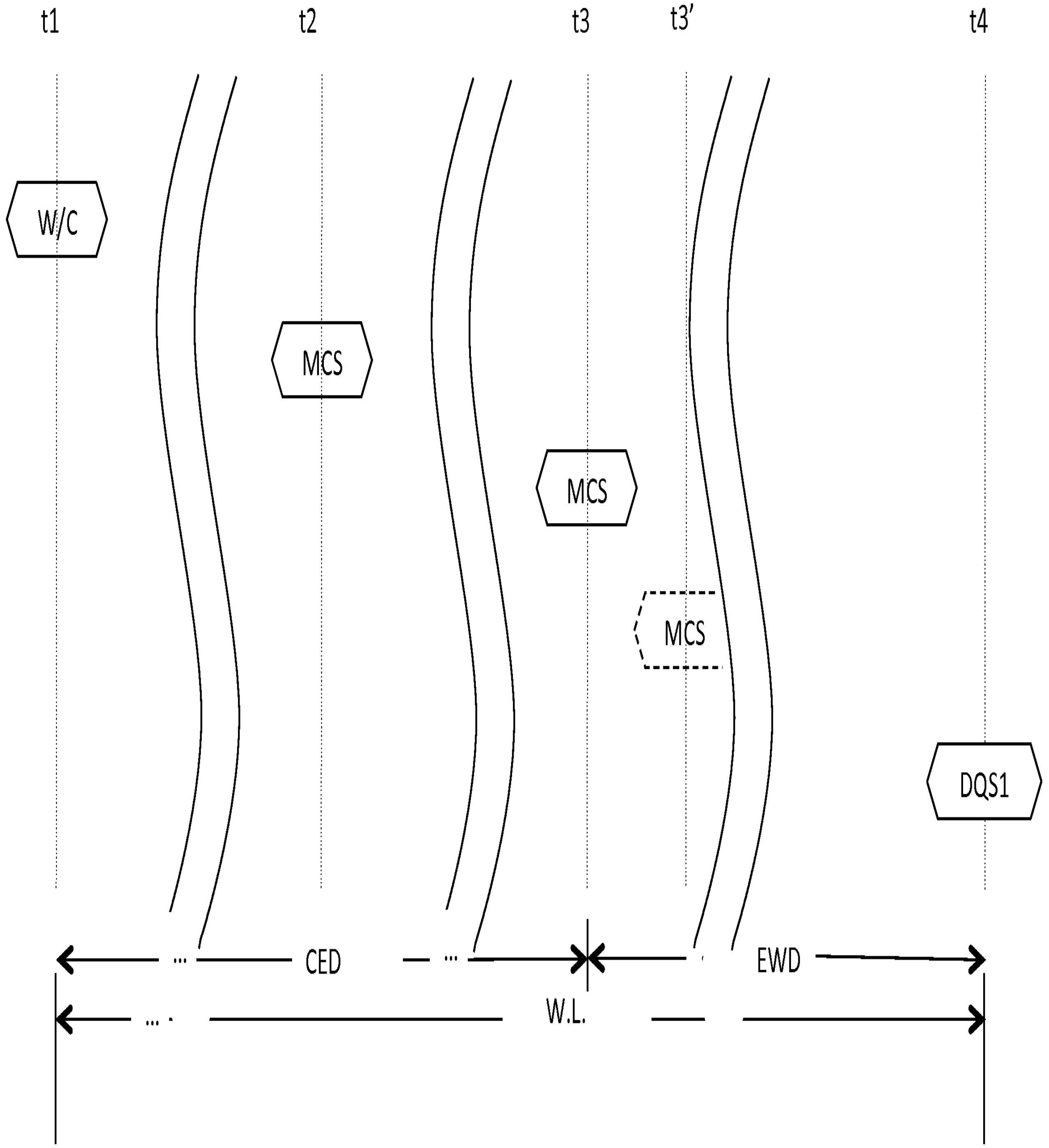


FIG. 12A

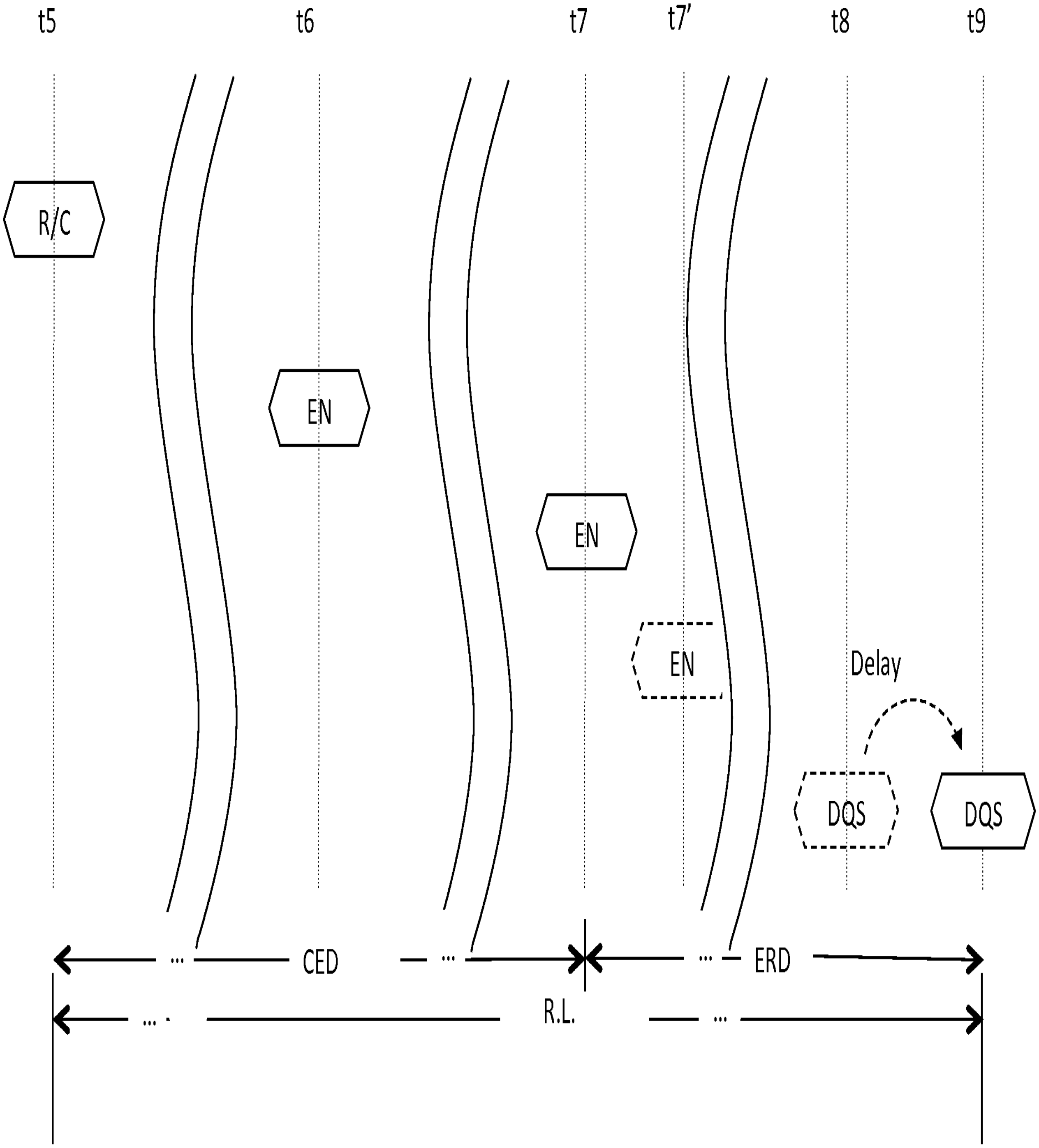


FIG. 12B

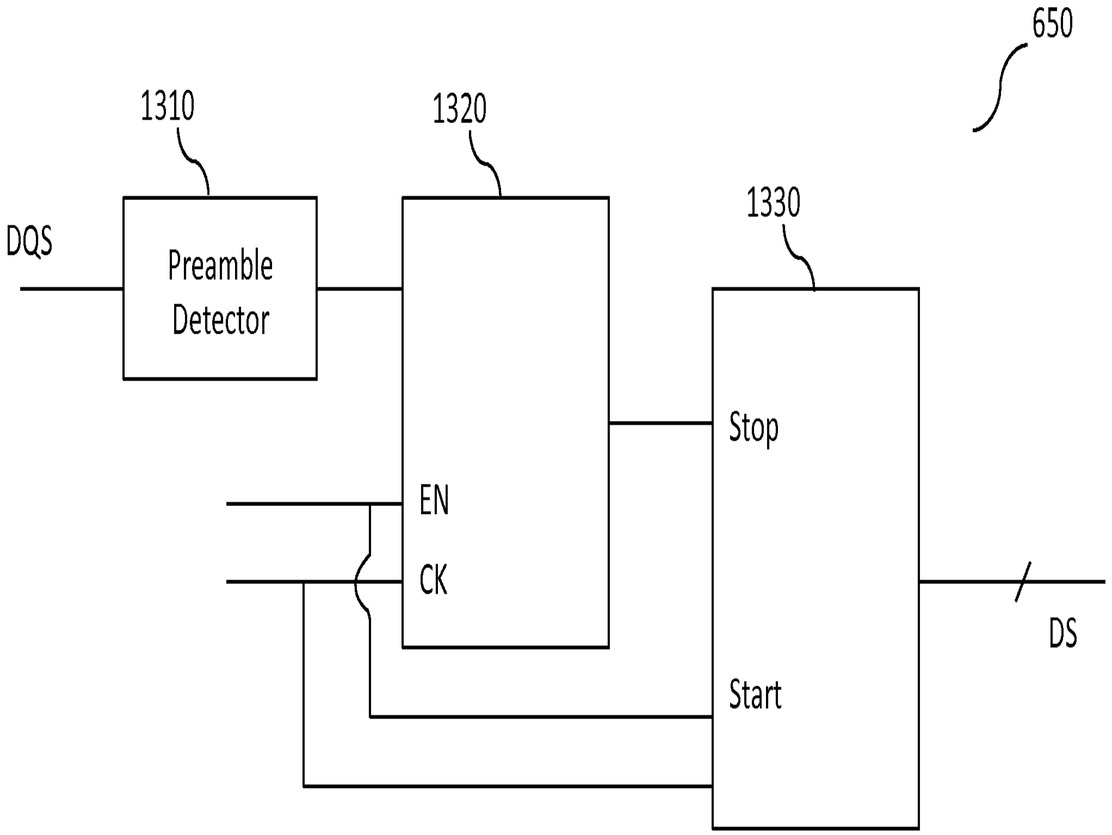


FIG. 13

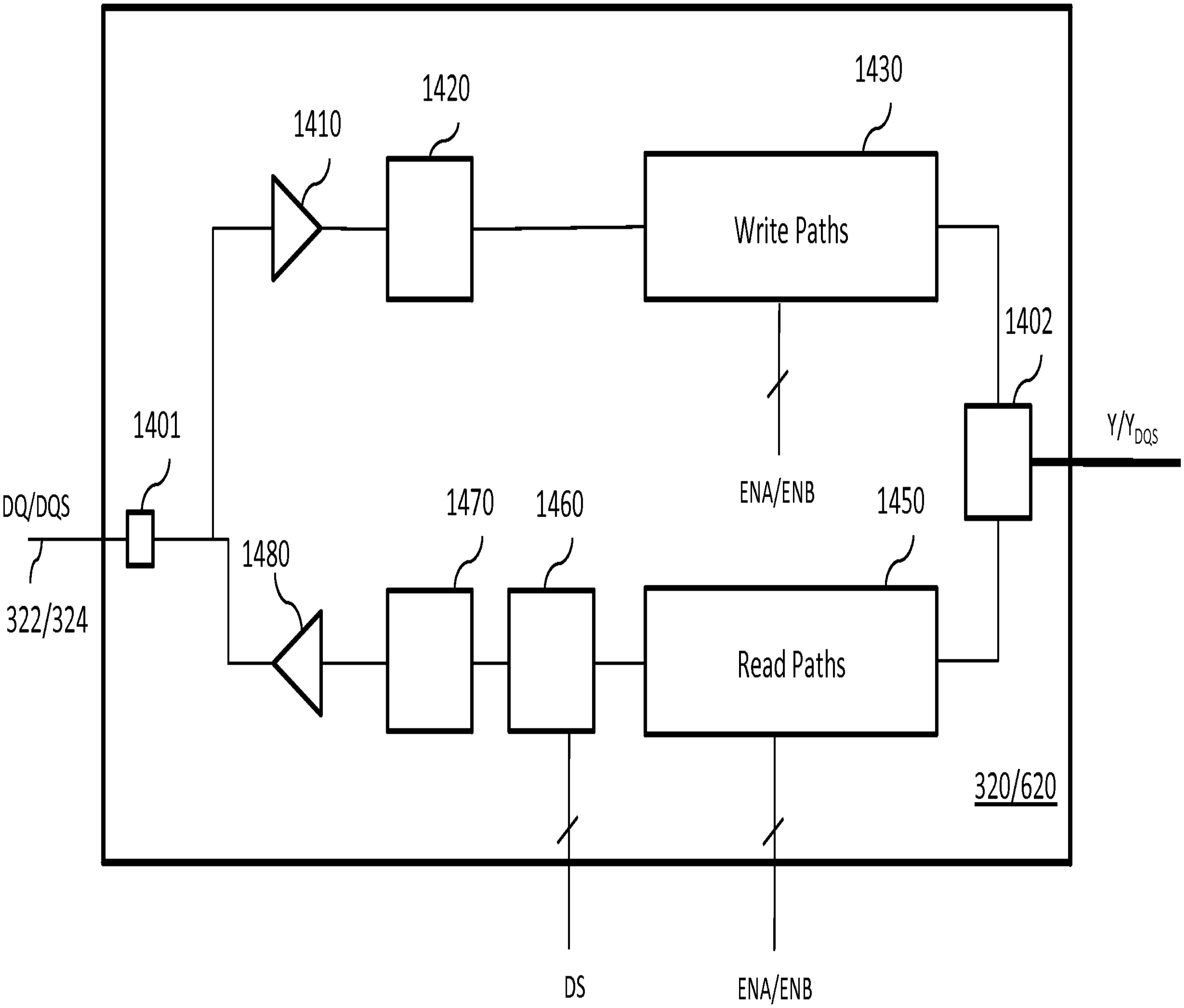

1430
1420
1410
Write Paths
1402
1401
Y/Y_{DQS}
DQ/DQS
ENA/ENB
1470
1460
1450
1480
322/324
Read Paths
320/620
DS
ENA/ENB

FIG. 14

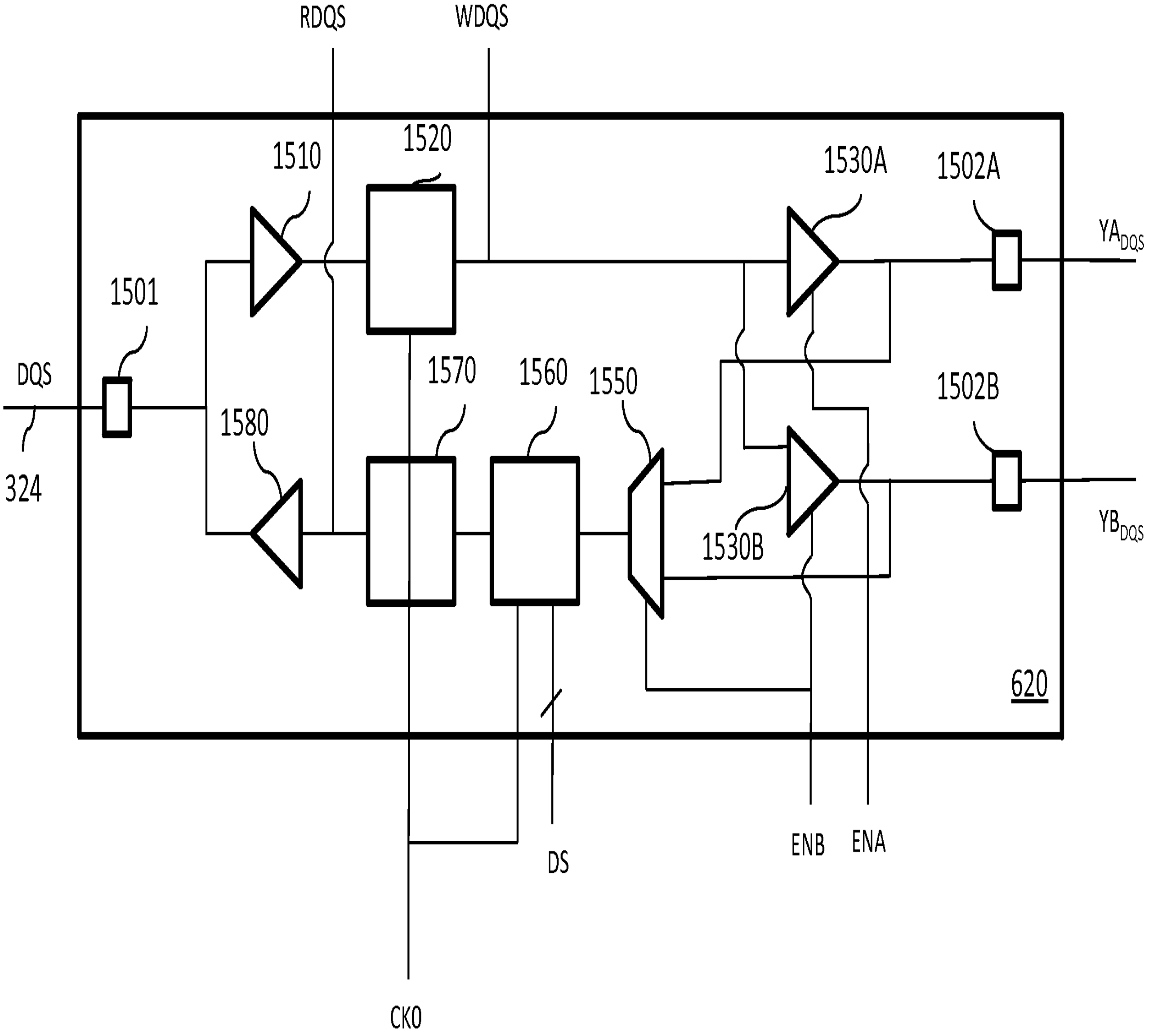

RDQS
WDQS
1520
1510
1530A
1502A
YA_{DQS}
1501
DQS
324
1570
1560
1550
1502B
1580
1530B
YB_{DQS}
620
DS
ENB
ENA
CK0

FIG. 15

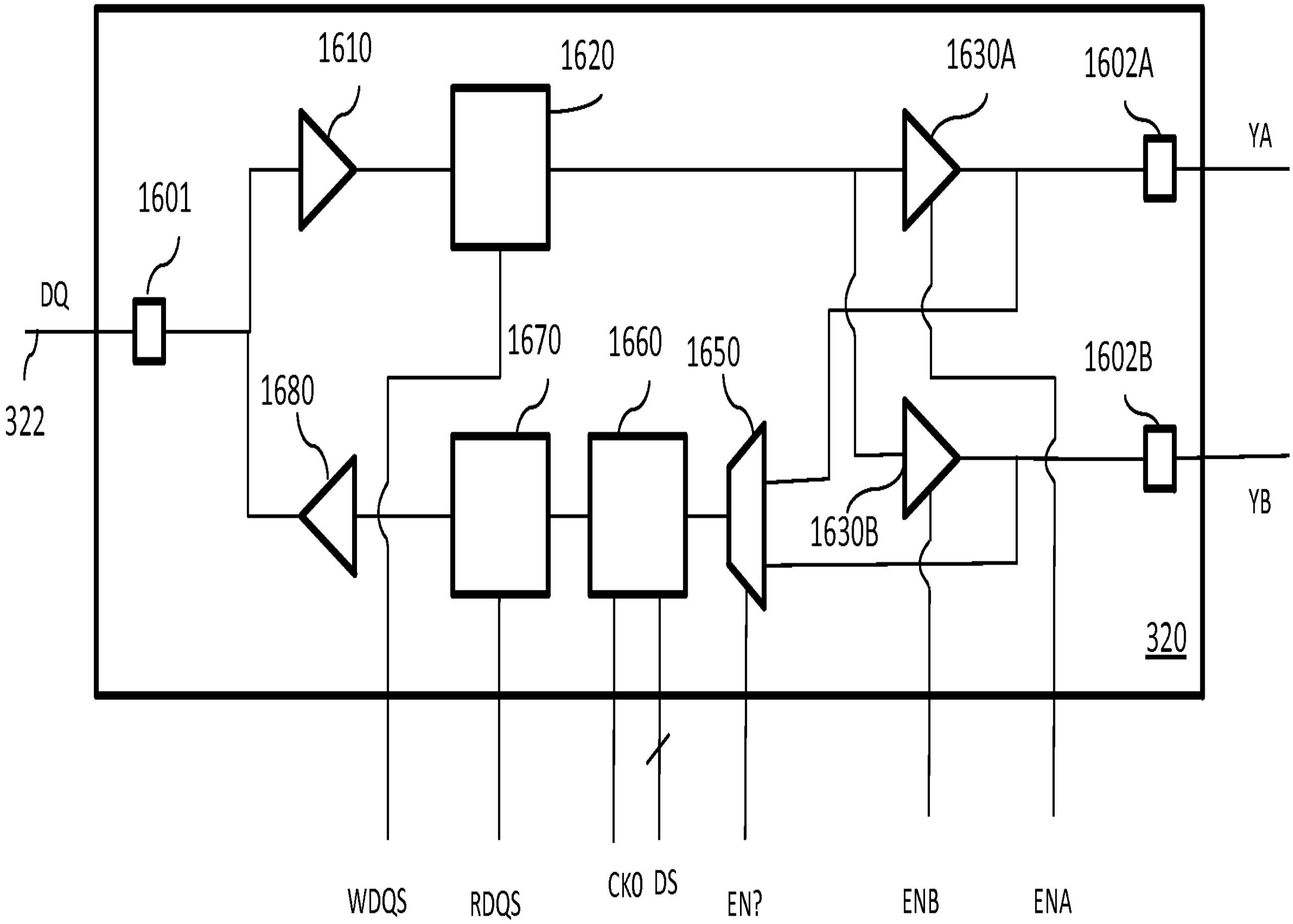

1610
1620
1630A
1602A
YA
1601
DQ
322
1670
1660
1650
1680
1602B
1630B
YB
320
WDQS
RDQS
CK0
DS
EN?
ENB
ENA

FIG. 16

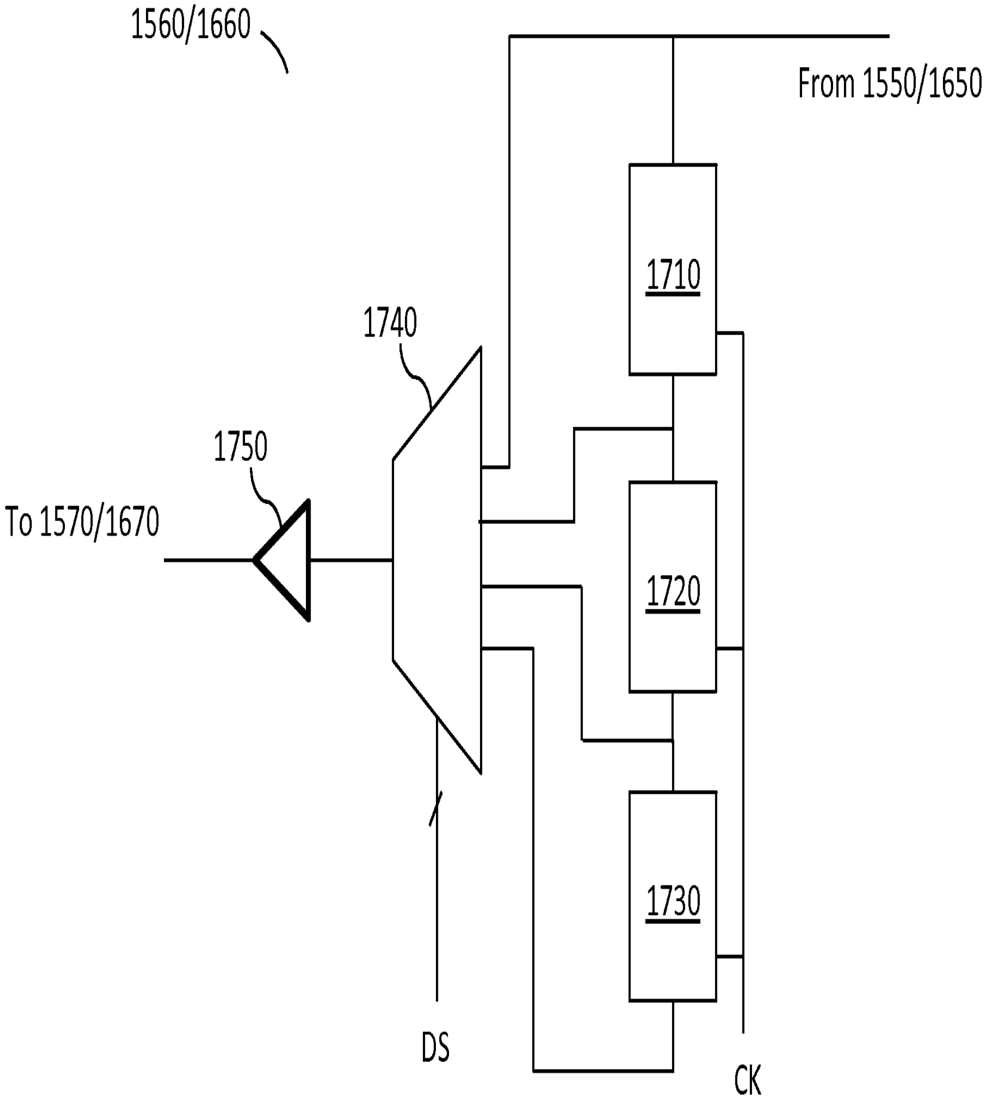

1560/1660
From 1550/1650
1710
1740
1750
To 1570/1670
1720
1730
DS
CK

FIG. 17

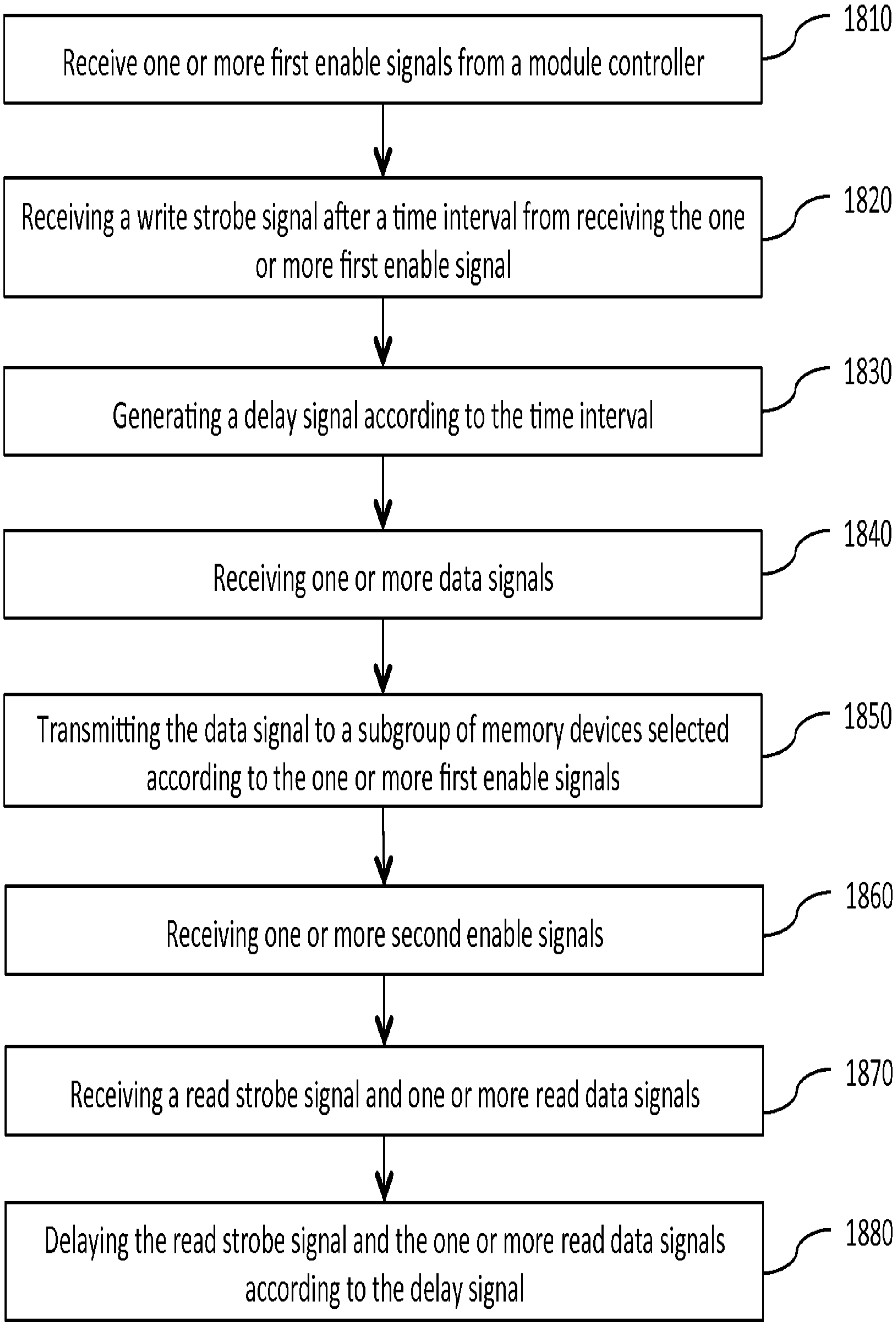

1810
Receive one or more first enable signals from a module controller
1820
Receiving a write strobe signal after a time interval from receiving the one or more first enable signal
1830
Generating a delay signal according to the time interval
1840
Receiving one or more data signals
1850
Transmitting the data signal to a subgroup of memory devices selected according to the one or more first enable signals
1860
Receiving one or more second enable signals
1870
Receiving a read strobe signal and one or more read data signals
1880
Delaying the read strobe signal and the one or more read data signals according to the delay signal

FIG. 18

Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION
First Named Inventor/Applicant Name:	Hyun Lee
Filer:	Jamie Jie Zheng
Attorney Docket Number:	NT003
Filed as Small Entity	

Provisional Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Provisional Application filing fee	2005	1	125	125
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
		Total in USD ($)		**125**

Electronic Acknowledgement Receipt

EFS ID:	13365080
Application Number:	61676883
International Application Number:	
Confirmation Number:	3331
Title of Invention:	MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION
First Named Inventor/Applicant Name:	Hyun Lee
Customer Number:	79141
Filer:	Jamie Jie Zheng
Filer Authorized By:	
Attorney Docket Number:	NT003
Receipt Date:	27-JUL-2012
Filing Date:	
Time Stamp:	23:07:27
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$125
RAM confirmation Number	9017
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1	Provisional Cover Sheet (SB16)	ProvisionalCoverSheet.pdf	2071285 ed2f7d02197750c207a62a8bfe37baf10f7a1fe8	no	3
Warnings:					
Information:					
2	Specification	NT003ProvisionalApplication.pdf	330488 081209270778ad5a3983f158c230315534ef46cc	no	30
Warnings:					
Information:					
3	Drawings-only black and white line drawings	NT003ProvisionalDrawings.pdf	214623 dad05d91f7b6ee22b26e86b685796b022780e7c8	no	23
Warnings:					
The page size in the PDF is too large. The pages should be 8.5 x 11 or A4. If this PDF is submitted, the pages will be resized upon entry into the Image File Wrapper and may affect subsequent processing					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	29497 133c71ed48d3fb90fdb478b15179157f9987bcf4	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	2645893		

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

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

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

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