

(51) International Patent Classification:  
**G06F 11/10** (2006.01) **G06F 11/267** (2006.01)(21) International Application Number:  
PCT/US2011/043648(22) International Filing Date:  
12 July 2011 (12.07.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
12/835,292 13 July 2010 (13.07.2010) US  
13/087,640 15 April 2011 (15.04.2011) US(71) Applicant (for all designated States except US): **SAN-DISK TECHNOLOGIES INC.** [US/US]; Two Legacy Town Center, 6900 North Dallas Parkway, Plano, Texas 75024 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **CHUNG, Chun Sing Jackson** [CA/US]; 181 W Weddell Dr. Apt. 32, Sunnyvale, California 94089-1525 (US). **CHENG, Steven Shisan** [US/US]; 1121 S. Bernardo Avenue, Sunnyvale, California 94087 (US). **EREZ, Eran** [IL/US]; 1956 Serge Avenue, San Jose, California 95130 (US).(74) Agents: **CLEVELAND, Michael G.** et al.; 505 Montgomery Street, Suite 800, San Francisco, California 94111 (US).

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

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

[Continued on next page]

(54) Title: DYNAMIC OPTIMIZATION OF BACK-END MEMORY SYSTEM INTERFACE

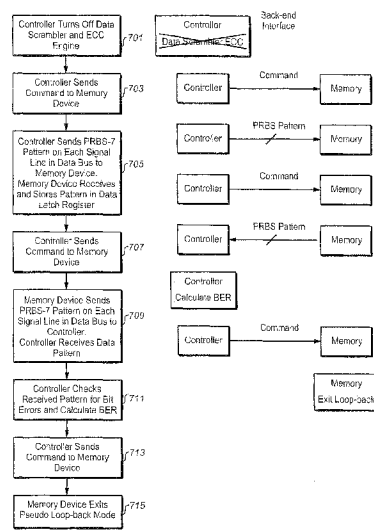


FIG. 15

(57) Abstract: A structure, and corresponding operating techniques, are presented for the internal controller to memory circuit interface for memory systems such as a flash memory card or other similarly structured devices. The interface between the controller circuit and memory circuit (or circuits) includes a feedback process where the amount of error that arises due to controller-memory transfers is monitored and the transfer characteristics (such as clock rate, drive strength, etc.) can be modified accordingly. Techniques are also presented for dynamically optimizing the performance of the controller-memory (or "back-end") interface of a non-volatile memory system. Memory systems are usually designed to have a certain amount of error tolerance for error that can then be corrected by ECC. In many circumstances, such as when a device is new, the ECC capabilities of the system exceed what is needed to correct data storage errors. In these circumstances the memory system internally allots a non-zero portion of this error correction capacity to the back-end interface. This allows for the interface to operate at, for example, higher speed or lower power, even though this will likely lead to transmission path error. The system can also calibrate the back-end interface to determine that amount of error that result from various operating conditions, allowing the operating parameters of the back-end interface to be set according to amount of error that is allotted to the transfer process.





---

**Published:**

— with international search report (Art. 21(3))

— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

## **DYNAMIC OPTIMIZATION OF BACK-END MEMORY SYSTEM INTERFACE**

### **BACKGROUND**

[0001] This application relates to the operation of re-programmable non-volatile memory systems, such as semiconductor flash memory, and, more specifically, to the internal interface between the memory system's controller and memory circuits.

[0002] Solid-state memory capable of nonvolatile storage of charge, particularly in the form of EEPROM and flash EEPROM packaged as a small form factor card, has recently become the storage of choice in a variety of mobile and handheld devices, notably information appliances and consumer electronics products. Unlike RAM (random access memory) that is also solid-state memory, flash memory is non-volatile, and retaining its stored data even after power is turned off. Also, unlike ROM (read only memory), flash memory is rewritable similar to a disk storage device. In spite of the higher cost, flash memory is increasingly being used in mass storage applications. Conventional mass storage, based on rotating magnetic medium such as hard drives and floppy disks, is unsuitable for the mobile and handheld environment. This is because disk drives tend to be bulky, are prone to mechanical failure and have high latency and high power requirements. These undesirable attributes make disk-based storage impractical in most mobile and portable applications. On the other hand, flash memory, both embedded and in the form of a removable card is ideally suited in the mobile and handheld environment because of its small size, low power consumption, high speed and high reliability features.

[0003] Flash EEPROM is similar to EEPROM (electrically erasable and programmable read-only memory) in that it is a non-volatile memory that can be erased and have new data written or "programmed" into their memory cells. Both utilize a floating (unconnected) conductive gate, in a field effect transistor structure, positioned over a channel region in a semiconductor substrate, between source and drain regions. A control gate is then provided over the floating gate. The threshold voltage characteristic of the transistor is controlled by the amount of charge that is retained on the floating gate. That is, for a given level of charge on the floating gate,

there is a corresponding voltage (threshold) that must be applied to the control gate before the transistor is turned “on” to permit conduction between its source and drain regions. In particular, flash memory such as Flash EEPROM allows entire blocks of memory cells to be erased at the same time.

[0004] The floating gate can hold a range of charges and therefore can be programmed to any threshold voltage level within a threshold voltage window. The size of the threshold voltage window is delimited by the minimum and maximum threshold levels of the device, which in turn correspond to the range of the charges that can be programmed onto the floating gate. The threshold window generally depends on the memory device’s characteristics, operating conditions and history. Each distinct, resolvable threshold voltage level range within the window may, in principle, be used to designate a definite memory state of the cell.

[0005] The transistor serving as a memory cell is typically programmed to a "programmed" state by one of two mechanisms. In "hot electron injection," a high voltage applied to the drain accelerates electrons across the substrate channel region. At the same time a high voltage applied to the control gate pulls the hot electrons through a thin gate dielectric onto the floating gate. In "tunneling injection," a high voltage is applied to the control gate relative to the substrate. In this way, electrons are pulled from the substrate to the intervening floating gate. While the term “program” has been used historically to describe writing to a memory by injecting electrons to an initially erased charge storage unit of the memory cell so as to alter the memory state, it has now been used interchangeable with more common terms such as “write” or “record.”

[0006] The memory device may be erased by a number of mechanisms. For EEPROM, a memory cell is electrically erasable, by applying a high voltage to the substrate relative to the control gate so as to induce electrons in the floating gate to tunnel through a thin oxide to the substrate channel region (i.e., Fowler-Nordheim tunneling.) Typically, the EEPROM is erasable byte by byte. For flash EEPROM, the memory is electrically erasable either all at once or one or more minimum erasable blocks at a time, where a minimum erasable block may consist of one or more sectors and each sector may store 512 bytes or more of data.

[0007] The memory device typically comprises one or more memory chips that may be mounted on a card. Each memory chip comprises an array of memory cells supported by peripheral circuits such as decoders and erase, write and read circuits. The more sophisticated memory devices also come with a controller that performs intelligent and higher level memory operations and interfacing.

[0008] There are many commercially successful non-volatile solid-state memory devices being used today. These memory devices may be flash EEPROM or may employ other types of nonvolatile memory cells. Examples of flash memory and systems and methods of manufacturing them are given in United States patents nos. 5,070,032, 5,095,344, 5,315,541, 5,343,063, and 5,661,053, 5,313,421 and 6,222,762. In particular, flash memory devices with NAND string structures are described in United States patent nos. 5,570,315, 5,903,495, 6,046,935. Also nonvolatile memory devices are also manufactured from memory cells with a dielectric layer for storing charge. Instead of the conductive floating gate elements described earlier, a dielectric layer is used. Such memory devices utilizing dielectric storage element have been described by Eitan et al., "NROM: A Novel Localized Trapping, 2-Bit Nonvolatile Memory Cell," IEEE Electron Device Letters, vol. 21, no. 11, November 2000, pp. 543-545. An ONO dielectric layer extends across the channel between source and drain diffusions. The charge for one data bit is localized in the dielectric layer adjacent to the drain, and the charge for the other data bit is localized in the dielectric layer adjacent to the source. For example, United States patents nos. 5,768,192 and 6,011,725 disclose a nonvolatile memory cell having a trapping dielectric sandwiched between two silicon dioxide layers. Multi-state data storage is implemented by separately reading the binary states of the spatially separated charge storage regions within the dielectric.

[0009] In order to improve read and program performance, multiple charge storage elements or memory transistors in an array are read or programmed in parallel. Thus, a "page" of memory elements are read or programmed together. In existing memory architectures, a row typically contains several interleaved pages or it may constitute one page. All memory elements of a page will be read or programmed together.

[0010] In flash memory systems, erase operation may take as much as an order of magnitude longer than read and program operations. Thus, it is desirable to have the

erase block of substantial size. In this way, the erase time is amortized over a large aggregate of memory cells.

[0011] The nature of flash memory predicates that data must be written to an erased memory location. If data of a certain logical address from a host is to be updated, one way is rewrite the update data in the same physical memory location. That is, the logical to physical address mapping is unchanged. However, this will mean the entire erase block contain that physical location will have to be first erased and then rewritten with the updated data. This method of update is inefficient, as it requires an entire erase block to be erased and rewritten, especially if the data to be updated only occupies a small portion of the erase block. It will also result in a higher frequency of erase recycling of the memory block, which is undesirable in view of the limited endurance of this type of memory device.

[0012] Data communicated through external interfaces of host systems, memory systems and other electronic systems are addressed and mapped into the physical locations of a flash memory system. Typically, addresses of data files generated or received by the system are mapped into distinct ranges of a continuous logical address space established for the system in terms of logical blocks of data (hereinafter the "LBA interface"). The extent of the address space is typically sufficient to cover the full range of addresses that the system is capable of handling. In one example, magnetic disk storage drives communicate with computers or other host systems through such a logical address space. This address space has an extent sufficient to address the entire data storage capacity of the disk drive.

[0013] There is an ongoing effort to improve the performance of memory devices, by reducing power consumption as well as increasing device speed. As noted above, non-volatile memory devices are commonly formed of a controller circuit and one or more memory chips connected to each other over a bus structure. The settings of the controller/memory device interface, such as the voltage values and frequencies used, are typically set according to the expected worst case scenario in order to have sufficient safety margin in order to avoid device failure. Consequently, in most circumstances the interface is being operated in less than optimal conditions. As this interface can be a limiting factor in device performance, there is consequently room for improvement in the design of this interface.

**[0014]** SUMMARY OF THE INVENTION

**[0015]** According to a general aspect of the invention, a method of operating a non-volatile memory system is presented. The non-volatile memory system includes: a controller circuit having a memory interface; a memory circuit having an array of non-volatile memory cells and a controller interface; and a bus structure connected to the memory interface of the controller circuit and to the controller interface of the memory circuit for the transfer of data and commands between them. The memory system can tolerate a first, non-zero amount of accumulated error from when data is transferred from the controller to be written to the memory array until the data is received at the controller after subsequently being read back from the memory array. The method includes the controller allotting a first, non-zero portion of the first amount of error to the transfer of the data between the controller circuit and the memory circuit via the bus structure, where the remainder of the first amount of error being allotted to write, storage, and read of the data on the memory circuit. The controller circuit sets the transfer characteristics between the controller circuit and the memory circuit to operate to allow up to the first portion of error.

**[0016]** In other aspects, methods of operating a non-volatile memory system including a memory circuit and a controller circuit are presented. The controller circuit performs a transmission error calibration by performing, for each of a plurality of values for each of one or more operating parameters for a bus structure connecting the controller with the memory circuit. This process includes transferring a data set of a known data pattern from the controller through transmission circuitry on the controller to the bus structure and receiving the data set from the bus structure through receiving circuitry on the memory circuit. The data set as received is stored in buffer memory on the memory circuit and then the data set, as stored in buffer memory on the memory circuit and without being written into the array, is transferred through transmission circuitry on the memory circuit to the bus structure. Based The data set is received from the bus structure through receiving circuitry on the controller and a comparison is performed of the data set as received with the know pattern. Based on the comparison, determining the amount of error associated with the transmission process is determined for the used one or more operating parameter. The memory system is subsequently operated to allow a first, non-zero amount of error in the

transmission of data between the controller circuit and the memory circuit, where the controller circuit selects values of the operating parameters according to the transmission error calibration process based on the determined associated amounts of error.

[0017] According to another general aspect of the invention, a non-volatile memory system has a controller circuit, including a memory interface and logic circuitry, and a memory circuit, including an array of non-volatile memory cells, a controller interface, and logic circuitry. The memory system also includes a bus structure connected to the memory interface of the controller circuit and to the controller interface of the memory circuit for the transfer of data and commands between the controller and memory circuit. A feedback processing circuit is connected to the logic circuitry of the receiving one the controller and the memory circuit during a transfer of data between them to receive information on the amount of error occurring as a result of the transfer, and connected to one or both of the memory interface and the controller interface to adjust the characteristics of the transfer between them in response to the amount of error.

[0018] In still other aspects, methods of operating a non-volatile memory system including a non-volatile memory circuit and a controller circuit are presented. A first hash value is generated from a data set in logic circuitry on a first of the controller circuit and the memory circuit. The data set and the first hash value are transmitted to a bus structure through an interface on the first of the controller circuit and the memory circuit and received from the bus structure through an interface on the second of the controller circuit and the memory circuit. A second hash value is then generated from the data set as received in logic circuitry on the second of the controller circuit and the memory circuit and the first hash value, as received, is then compared to the second hash value on the second of the controller circuit and the memory circuit. Based on the comparison of the first hash value as received and the second hash value by the logic circuitry on the second of the controller circuit and the memory circuit, the system determines whether to alter characteristics of the transfer of data between the controller circuit and the memory circuit.

[0019] Various aspects, advantages, features and embodiments of the present invention are included in the following description of exemplary examples thereof,



which description should be taken in conjunction with the accompanying drawings. All patents, patent applications, articles, other publications, documents and things referenced herein are hereby incorporated herein by this reference in their entirety for all purposes. To the extent of any inconsistency or conflict in the definition or use of terms between any of the incorporated publications, documents or things and the present application, those of the present application shall prevail.

### BRIEF DESCRIPTION OF THE DRAWINGS

[0020] **FIG. 1** illustrates schematically the main hardware components of a memory system suitable for implementing the present invention.

[0021] **FIG. 2** illustrates schematically a non-volatile memory cell.

[0022] **FIG. 3** illustrates the relation between the source-drain current  $I_D$  and the control gate voltage  $V_{CG}$  for four different charges Q1-Q4 that the floating gate may be selectively storing at any one time.

[0023] **FIG. 4A** illustrates schematically a string of memory cells organized into an NAND string.

[0024] **FIG. 4B** illustrates an example of an NAND array 210 of memory cells, constituted from NAND strings 50 such as that shown in **FIG. 4A**.

[0025] **FIG. 5** illustrates a page of memory cells, organized for example in the NAND configuration, being sensed or programmed in parallel.

[0026] **FIG. 6(0) - 6(2)** illustrate an example of programming a population of 4-state memory cells.

[0027] **FIGs. 7A-7E** illustrate the programming and reading of the 4-state memory encoded with a given 2-bit code.

[0028] **FIG. 8** illustrates the memory being managed by a memory manager with is a software component that resides in the controller.

[0029] **FIG. 9** illustrates the software modules of the back-end system.

[0030] **FIGs. 10A(i) – 10A(iii)** illustrate schematically the mapping between a logical

group and a metablock. **FIG. 10B** illustrates schematically the mapping between logical groups and metablocks.

[0031] **FIG. 11** is a block diagram showing a feedback mechanism for determining interface integrity based on an existing infrastructure.

[0032] **FIG. 12** is a block diagram illustrating embodiments where the feedback mechanism uses a hash engine to determine interface integrity.

[0033] **FIG. 13** is a diagram showing an example for transmitting the data and the generated hash value over the bus interface.

[0034] **FIG. 14** schematically illustrates the contributions to the bit errors in a memory system.

[0035] **FIG. 15** can be used to illustrate the operation of a pseudo loop-back method in the back-end interface.

[0036] **FIGs. 16 and 17** respectively correspond to blocks 705 and 709 of **FIG. 15**.

[0037] **FIG. 18** is an example of a shmoo plot showing transmission BER vs. data bus voltage and data transfer rate.

[0038] **FIG. 19** is a block diagram that illustrates this crosstalk in a memory system whose bus structure uses multiple memory data buses.

## DETAILED DESCRIPTION

### MEMORY SYSTEM

[0039] **FIG. 1 to FIG. 7** provide example memory systems in which the various aspects of the present invention may be implemented or illustrated.

[0040] **FIG. 8 to FIG. 10** illustrate preferred memory and block architectures for implementing the various aspects of the present invention.

[0041] **FIGs. 11-13** illustrate the use of an adaptive internal interface between the controller and the memory circuit or circuits.

[0042] FIG. 1 illustrates schematically the main hardware components of a memory system suitable for implementing the present invention. The memory system 90 typically operates with a host 80 through a host interface. The memory system is typically in the form of a memory card or an embedded memory system. The memory system 90 includes a memory 200 whose operations are controlled by a controller 100. The memory 200 comprises of one or more array of non-volatile memory cells distributed over one or more integrated circuit chip. The controller 100 includes an interface 110, a processor 120, an optional coprocessor 121, ROM 122 (read-only-memory), RAM 130 (random access memory) and optionally programmable nonvolatile memory 124. The interface 110 has one component interfacing the controller to a host and another component interfacing to the memory 200. Firmware stored in nonvolatile ROM 122 and/or the optional nonvolatile memory 124 provides codes for the processor 120 to implement the functions of the controller 100. Error correction codes may be processed by the processor 120 or the optional coprocessor 121. In an alternative embodiment, the controller 100 is implemented by a state machine (not shown.) In yet another embodiment, the controller 100 is implemented within the host.

#### Physical Memory Structure

[0043] FIG. 2 illustrates schematically a non-volatile memory cell. The memory cell 10 can be implemented by a field-effect transistor having a charge storage unit 20, such as a floating gate or a dielectric layer. The memory cell 10 also includes a source 14, a drain 16, and a control gate 30.

[0044] There are many commercially successful non-volatile solid-state memory devices being used today. These memory devices may employ different types of memory cells, each type having one or more charge storage element.

[0045] Typical non-volatile memory cells include EEPROM and flash EEPROM. Examples of EEPROM cells and methods of manufacturing them are given in United States patent no. 5,595,924. Examples of flash EEPROM cells, their uses in memory systems and methods of manufacturing them are given in United States patents nos. 5,070,032, 5,095,344, 5,315,541, 5,343,063, 5,661,053, 5,313,421 and 6,222,762. In particular, examples of memory devices with NAND cell structures are described in

United States patent nos. 5,570,315, 5,903,495, 6,046,935. Also, examples of memory devices utilizing dielectric storage element have been described by Eitan et al., "NROM: A Novel Localized Trapping, 2-Bit Nonvolatile Memory Cell," IEEE Electron Device Letters, vol. 21, no. 11, November 2000, pp. 543-545, and in United States patents nos. 5,768,192 and 6,011,725.

[0046] In practice, the memory state of a cell is usually read by sensing the conduction current across the source and drain electrodes of the cell when a reference voltage is applied to the control gate. Thus, for each given charge on the floating gate of a cell, a corresponding conduction current with respect to a fixed reference control gate voltage may be detected. Similarly, the range of charge programmable onto the floating gate defines a corresponding threshold voltage window or a corresponding conduction current window.

[0047] Alternatively, instead of detecting the conduction current among a partitioned current window, it is possible to set the threshold voltage for a given memory state under test at the control gate and detect if the conduction current is lower or higher than a threshold current. In one implementation the detection of the conduction current relative to a threshold current is accomplished by examining the rate the conduction current is discharging through the capacitance of the bit line.

[0048] FIG. 3 illustrates the relation between the source-drain current  $I_D$  and the control gate voltage  $V_{CG}$  for four different charges Q1-Q4 that the floating gate may be selectively storing at any one time. The four solid  $I_D$  versus  $V_{CG}$  curves represent four possible charge levels that can be programmed on a floating gate of a memory cell, respectively corresponding to four possible memory states. As an example, the threshold voltage window of a population of cells may range from 0.5V to 3.5V. Seven possible memory states "0", "1", "2", "3", "4", "5", "6", respectively representing one erased and six programmed states may be demarcated by partitioning the threshold window into five regions in interval of 0.5V each. For example, if a reference current,  $I_{REF}$  of 2  $\mu A$  is used as shown, then the cell programmed with Q1 may be considered to be in a memory state "1" since its curve intersects with  $I_{REF}$  in the region of the threshold window demarcated by  $V_{CG} = 0.5V$  and 1.0V. Similarly, Q4 is in a memory state "5".

[0049] As can be seen from the description above, the more states a memory cell is made to store, the more finely divided is its threshold window. For example, a memory device may have memory cells having a threshold window that ranges from  $-1.5\text{V}$  to  $5\text{V}$ . This provides a maximum width of  $6.5\text{V}$ . If the memory cell is to store 16 states, each state may occupy from  $200\text{mV}$  to  $300\text{mV}$  in the threshold window. This will require higher precision in programming and reading operations in order to be able to achieve the required resolution.

[0050] **FIG. 4A** illustrates schematically a string of memory cells organized into an NAND string. An NAND string 50 comprises of a series of memory transistors M1, M2, ... Mn (e.g.,  $n = 4, 8, 16$  or higher) daisy-chained by their sources and drains. A pair of select transistors S1, S2 controls the memory transistors chain's connection to the external via the NAND string's source terminal 54 and drain terminal 56 respectively. In a memory array, when the source select transistor S1 is turned on, the source terminal is coupled to a source line (see **FIG. 4B**). Similarly, when the drain select transistor S2 is turned on, the drain terminal of the NAND string is coupled to a bit line of the memory array. Each memory transistor 10 in the chain acts as a memory cell. It has a charge storage element 20 to store a given amount of charge so as to represent an intended memory state. A control gate 30 of each memory transistor allows control over read and write operations. As will be seen in **FIG. 4B**, the control gates 30 of corresponding memory transistors of a row of NAND string are all connected to the same word line. Similarly, a control gate 32 of each of the select transistors S1, S2 provides control access to the NAND string via its source terminal 54 and drain terminal 56 respectively. Likewise, the control gates 32 of corresponding select transistors of a row of NAND string are all connected to the same select line.

[0051] When an addressed memory transistor 10 within an NAND string is read or is verified during programming, its control gate 30 is supplied with an appropriate voltage. At the same time, the rest of the non-addressed memory transistors in the NAND string 50 are fully turned on by application of sufficient voltage on their control gates. In this way, a conductive path is effectively created from the source of the individual memory transistor to the source terminal 54 of the NAND string and likewise for the drain of the individual memory transistor to the drain terminal 56 of

the cell. Memory devices with such NAND string structures are described in United States patent nos. 5,570,315, 5,903,495, 6,046,935.

[0052] **FIG. 4B** illustrates an example of an NAND array 210 of memory cells, constituted from NAND strings 50 such as that shown in **FIG. 4A**. Along each column of NAND strings, a bit line such as bit line 36 is coupled to the drain terminal 56 of each NAND string. Along each bank of NAND strings, a source line such as source line 34 is couple to the source terminals 54 of each NAND string. Also the control gates along a row of memory cells in a bank of NAND strings are connected to a word line such as word line 42. The control gates along a row of select transistors in a bank of NAND strings are connected to a select line such as select line 44. An entire row of memory cells in a bank of NAND strings can be addressed by appropriate voltages on the word lines and select lines of the bank of NAND strings. When a memory transistor within a NAND string is being read, the remaining memory transistors in the string are turned on hard via their associated word lines so that the current flowing through the string is essentially dependent upon the level of charge stored in the cell being read.

[0053] **FIG. 5** illustrates a page of memory cells, organized for example in the NAND configuration, being sensed or programmed in parallel. **FIG. 5** essentially shows a bank of NAND strings 50 in the memory array 210 of **FIG. 4B**, where the detail of each NAND string is shown explicitly as in **FIG. 4A**. A “page” such as the page 60, is a group of memory cells enabled to be sensed or programmed in parallel. This is accomplished by a corresponding page of sense amplifiers 212. The sensed results are latches in a corresponding set of latches 214. Each sense amplifier can be coupled to a NAND string via a bit line. The page is enabled by the control gates of the cells of the page connected in common to a word line 42 and each cell accessible by a sense amplifier accessible via a bit line 36. As an example, when respectively sensing or programming the page of cells 60, a sensing voltage or a programming voltage is respectively applied to the common word line WL3 together with appropriate voltages on the bit lines.

#### Physical Organization of the Memory

[0054] One important difference between flash memory and of type of memory is that

a cell must be programmed from the erased state. That is the floating gate must first be emptied of charge. Programming then adds a desired amount of charge back to the floating gate. It does not support removing a portion of the charge from the floating to go from a more programmed state to a lesser one. This means that update data cannot overwrite existing one and must be written to a previous unwritten location.

[0055] Furthermore erasing is to empty all the charges from the floating gate and generally takes appreciably time. For that reason, it will be cumbersome and very slow to erase cell by cell or even page by page. In practice, the array of memory cells is divided into a large number of blocks of memory cells. As is common for flash EEPROM systems, the block is the unit of erase. That is, each block contains the minimum number of memory cells that are erased together. While aggregating a large number of cells in a block to be erased in parallel will improve erase performance, a large size block also entails dealing with a larger number of update and obsolete data. Just before the block is erased, a garbage collection is required to salvage the non-obsolete data in the block.

[0056] Each block is typically divided into a number of pages. A page is a unit of programming or reading. In one embodiment, the individual pages may be divided into segments and the segments may contain the fewest number of cells that are written at one time as a basic programming operation. One or more pages of data are typically stored in one row of memory cells. A page can store one or more sectors. A sector includes user data and overhead data. Multiple blocks and pages distributed across multiple arrays can also be operated together as metablocks and metapages. If they are distributed over multiple chips, they can be operated together as megablocks and megapage.

#### Examples of Multi-level Cell ("MLC") Memory Partitioning

[0057] A nonvolatile memory in which the memory cells each stores multiple bits of data has already been described in connection with **FIG. 3**. A particular example is a memory formed from an array of field-effect transistors, each having a charge storage layer between its channel region and its control gate. The charge storage layer or unit can store a range of charges, giving rise to a range of threshold voltages for each field-effect transistor. The range of possible threshold voltages spans a threshold

window. When the threshold window is partitioned into multiple sub-ranges or zones of threshold voltages, each resolvable zone is used to represent a different memory states for a memory cell. The multiple memory states can be coded by one or more binary bits. For example, a memory cell partitioned into four zones can support four states which can be coded as 2-bit data. Similarly, a memory cell partitioned into eight zones can support eight memory states which can be coded as 3-bit data, etc.

#### All-bit, Full-Sequence MLC Programming

[0058] **FIG. 6(0) - 6(2)** illustrate an example of programming a population of 4-state memory cells. **FIG. 6(0)** illustrates the population of memory cells programmable into four distinct distributions of threshold voltages respectively representing memory states “0”, “1”, “2” and “3”. **FIG. 6(1)** illustrates the initial distribution of “erased” threshold voltages for an erased memory. **FIG. 6(2)** illustrates an example of the memory after many of the memory cells have been programmed. Essentially, a cell initially has an “erased” threshold voltage and programming will move it to a higher value into one of the three zones demarcated by verify levels  $vV_1$ ,  $vV_2$  and  $vV_3$ . In this way, each memory cell can be programmed to one of the three programmed state “1”, “2” and “3” or remain un-programmed in the “erased” state. As the memory gets more programming, the initial distribution of the “erased” state as shown in **FIG. 6(1)** will become narrower and the erased state is represented by the “0” state.

[0059] A 2-bit code having a lower bit and an upper bit can be used to represent each of the four memory states. For example, the “0”, “1”, “2” and “3” states are respectively represented by “11”, “01”, “00” and “10”. The 2-bit data may be read from the memory by sensing in “full-sequence” mode where the two bits are sensed together by sensing relative to the read demarcation threshold values  $rV_1$ ,  $rV_2$  and  $rV_3$  in three sub-passes respectively.

#### Bit-by-Bit MLC Programming and Reading

[0060] **FIGs. 7A-7E** illustrate the programming and reading of the 4-state memory encoded with a given 2-bit code. **FIG. 7A** illustrates threshold voltage distributions of the 4-state memory array when each memory cell stores two bits of data using the 2-bit code. Such a 2-bit code has been disclosed in US Patent No. 7,057,939.



[0061] **FIG. 7B** illustrates the lower page programming (lower bit) in a 2-pass programming scheme using the 2-bit code. The fault-tolerant LM New code essentially avoids any upper page programming to transit through any intermediate states. Thus, the first pass lower page programming has the logical state (upper bit, lower bit) = (1, 1) transits to some intermediate state (x, 0) as represented by programming the “unprogrammed” memory state “0” to the “intermediate” state designated by (x, 0) with a programmed threshold voltage greater than  $D_A$  but less than  $D_C$ .

[0062] **FIG. 7C** illustrates the upper page programming (upper bit) in the 2-pass programming scheme using the 2-bit code. In the second pass of programming the upper page bit to “0”, if the lower page bit is at “1”, the logical state (1, 1) transits to (0, 1) as represented by programming the “unprogrammed” memory state “0” to “1”. If the lower page bit is at “0”, the logical state (0, 0) is obtained by programming from the “intermediate” state to “3”. Similarly, if the upper page is to remain at “1”, while the lower page has been programmed to “0”, it will require a transition from the “intermediate” state to (1, 0) as represented by programming the “intermediate” state to “2”.

[0063] **FIG. 7D** illustrates the read operation that is required to discern the lower bit of the 4-state memory encoded with the 2-bit code. A readB operation is first performed to determine if the LM flag can be read. If so, the upper page has been programmed and the readB operation will yield the lower page data correctly. On the other hand, if the upper page has not yet been programmed, the lower page data will be read by a readA operation.

[0064] **FIG. 7E** illustrates the read operation that is required to discern the upper bit of the 4-state memory encoded with the 2-bit code. As is clear from the figure, the upper page read will require a 3-pass read of readA, readB and readC, respectively relative to the demarcation threshold voltages  $D_A$ ,  $D_B$  and  $D_C$ .

[0065] In the bit-by-bit scheme for a 2-bit memory, a physical page of memory cells will store two logical data pages, a lower data page corresponding to the lower bit and an upper data page corresponding to the upper bit.

#### Binary and MLC Memory Partitioning

[0066] **FIG. 6** and **FIG. 7** illustrate examples of a 2-bit (also referred to as “D2”) memory. As can be seen, a D2 memory has its threshold range or window partitioned into 4 regions, designating 4 states. Similarly, in D3, each cell stores 3 bits (low, middle and upper bits) and there are 8 regions. In D4, there are 4 bits and 16 regions, etc. As the memory’s finite threshold window is partitioned into more regions, the resolution and for programming and reading will necessarily become finer. Two issues arise as the memory cell is configured to store more bits.

[0067] First, programming or reading will be slower when the threshold of a cell must be more accurately programmed or read. In fact in practice the sensing time (needed in programming and reading) tends to increase as the square of the number of partitioning levels.

[0068] Secondly, flash memory has an endurance problem as it ages with use. When a cell is repeatedly programmed and erased, charges is shuttled in and out of the floating gate 20 (see **FIG. 2**) by tunneling across a dielectric. Each time some charges may become trapped in the dielectric and will modify the threshold of the cell. In fact over use, the threshold window will progressively narrow. Thus, MLC memory generally is designed with tradeoffs between capacity, performance and reliability.

[0069] Conversely, it will be seen for a binary memory, the memory’s threshold window is only partitioned into two regions. This will allow a maximum margin of errors. Thus, binary partitioning while diminished in storage capacity will provide maximum performance and reliability.

[0070] The multi-pass, bit-by-bit programming and reading technique described in connection with **FIG. 7** provides a smooth transition between MLC and binary partitioning. In this case, if the memory is programmed with only the lower bit, it is effectively a binary partitioned memory. While this approach does not fully optimize the range of the threshold window as in the case of a single-level cell (“SLC”) memory, it has the advantage of using the same demarcation or sensing level as in the operations of the lower bit of the MLC memory. As will be described later, this approach allows a MLC memory to be “expropriated” for use as a binary memory, or vice versa. How it should be understood that MLC memory tends to have more

stringent specification for usage.

#### Binary Memory and Partial Page Programming

[0071] The charge programmed into the charge storage element of one memory cell produces an electric field that perturbs the electric field of a neighboring memory cell. This will affect the characteristics of the neighboring memory cell which essentially is a field-effect transistor with a charge storage element. In particular, when sensed the memory cell will appear to have a higher threshold level (or more programmed) than when it is less perturbed.

[0072] In general, if a memory cell is program-verified under a first field environment and later is read again under a different field environment due to neighboring cells subsequently being programmed with different charges, the read accuracy may be affected due to coupling between neighboring floating gates in what is referred to as the “Yupin Effect”. With ever higher integration in semiconductor memories, the perturbation of the electric field due to the stored charges between memory cells (Yupin effect) becomes increasingly appreciable as the inter-cellular spacing shrinks.

[0073] The Bit-by-Bit MLC Programming technique described in connection with **FIG. 7** above is designed to minimize program disturb from cells along the same word line. As can be seen from **FIG. 7B**, in a first of the two programming passes, the thresholds of the cells are moved at most half way up the threshold window. The effect of the first pass is overtaken by the final pass. In the final pass, the thresholds are only moved a quarter of the way. In other words, for D2, the charge difference among neighboring cells is limited to a quarter of its maximum. For D3, with three passes, the final pass will limit the charge difference to one-eighth of its maximum.

[0074] However, the bit-by-bit multi-pass programming technique will be compromised by partial-page programming. A page is a group of memory cells, typically along a row or word line, that is programmed together as a unit. It is possible to program non overlapping portions of a page individually over multiple programming passes. However, owing to not all the cells of the page are programmed in a final pass together, it could create large difference in charges programmed among the cells after the page is done. Thus partial-page programming would result in more program disturb and would require a larger margin for sensing

accuracy.

[0075] In the case the memory is configured as binary memory, the margin of operation is wider than that of MLC. In the preferred embodiment, the binary memory is configured to support partial-page programming in which non-overlapping portions of a page may be programmed individually in one of the multiple programming passes on the page. The programming and reading performance can be improved by operating with a page of large size. However, when the page size is much larger than the host's unit of write (typically a 512-byte sector), its usage will be inefficient. Operating with finer granularity than a page allows more efficient usage of such a page.

[0076] The example given has been between binary versus MLC. It should be understood that in general the same principles apply between a first memory with a first number of levels and a second memory with a second number of levels more than the first memory.

#### Logical and Physical Block Structures

[0077] **FIG. 8** illustrates the memory being managed by a memory manager with is a software component that resides in the controller. The memory 200 is organized into blocks, each block of cells being a minimum unit of erase. Depending on implementation, the memory system may operate with even large units of erase formed by an aggregate of blocks into "metablocks" and also "megablocks". For convenience the description will refer to a unit of erase as a metablock although it will be understood that some systems operate with even larger unit of erase such as a "megablock" formed by an aggregate of metablocks.

[0078] The host 80 accesses the memory 200 when running an application under a file system or operating system. Typically, the host system addresses data in units of logical sectors where, for example, each sector may contain 512 bytes of data. Also, it is usual for the host to read or write to the memory system in unit of logical clusters, each consisting of one or more logical sectors. In some host systems, an optional host-side memory manager may exist to perform lower level memory management at the host. In most cases during read or write operations, the host 80 essentially issues a command to the memory system 90 to read or write a segment containing a string of

logical sectors of data with contiguous addresses.

[0079] A memory-side memory manager 300 is implemented in the controller 100 of the memory system 90 to manage the storage and retrieval of the data of host logical sectors among metablocks of the flash memory 200. The memory manager comprises a front-end system 310 and a back-end system 320. The front-end system 310 includes a host interface 312. The back-end system 320 includes a number of software modules for managing erase, read and write operations of the metablocks. The memory manager also maintains system control data and directory data associated with its operations among the flash memory 200 and the controller RAM 130.

[0080] **FIG. 9** illustrates the software modules of the back-end system. The Back-End System mainly comprises two functional modules: a Media Management Layer 330 and a Dataflow and Sequencing Layer 340.

[0081] The media management layer 330 is responsible for the organization of logical data storage within a flash memory meta-block structure. More details will be provided later in the section on “Media management Layer”.

[0082] The dataflow and sequencing layer 340 is responsible for the sequencing and transfer of sectors of data between a front-end system and a flash memory. This layer includes a command sequencer 342, a low-level sequencer 344 and a flash Control layer 346. More details will be provided later in the section on “Low Level System Spec”.

[0083] The memory manager 300 is preferably implemented in the controller 100. It translates logical addresses received from the host into physical addresses within the memory array, where the data are actually stored, and then keeps track of these address translations.

[0084] **FIGS. 10A(i) – 10A(iii)** illustrate schematically the mapping between a logical group and a metablock. The metablock of the physical memory has  $N$  physical sectors for storing  $N$  logical sectors of data of a logical group. **FIG. 10A(i)** shows the data from a logical group  $LG_i$ , where the logical sectors are in contiguous logical order  $0, 1, \dots, N-1$ . **FIG. 10A(ii)** shows the same data being stored in the

metablock in the same logical order. The metablock when stored in this manner is said to be “sequential.” In general, the metablock may have data stored in a different order, in which case the metablock is said to be “non-sequential” or “chaotic.”

[0085] There may be an offset between the lowest address of a logical group and the lowest address of the metablock to which it is mapped. In this case, logical sector address wraps round as a loop from bottom back to top of the logical group within the metablock. For example, in **FIG. 10A(iii)**, the metablock stores in its first location beginning with the data of logical sector  $k$ . When the last logical sector  $N-1$  is reached, it wraps around to sector  $0$  and finally storing data associated with logical sector  $k-1$  in its last physical sector. In the preferred embodiment, a page tag is used to identify any offset, such as identifying the starting logical sector address of the data stored in the first physical sector of the metablock. Two blocks will be considered to have their logical sectors stored in similar order when they only differ by a page tag.

[0086] **FIG. 10B** illustrates schematically the mapping between logical groups and metablocks. Each logical group 380 is mapped to a unique metablock 370, except for a small number of logical groups in which data is currently being updated. After a logical group has been updated, it may be mapped to a different metablock. The mapping information is maintained in a set of logical to physical directories, which will be described in more detail later.

#### ADAPTIVE CONTROLLER-MEMORY INTERFACE

[0087] This section presents the use of a feedback mechanism and processing unit that monitors transfer integrity of the internal controller-memory interface of the memory system and can adjust the interface settings accordingly. This allows for system to optimize interface performance. For example, the system's power may be able be reduced or the bus clock for the interface sped up, which, as this can often be an internal performance bottleneck, allows for an increase in performance as seen from outside of the memory system (i.e., from the host). In case of transmission errors, assisted by the interface integrity feedback and, depending on the embodiment, other sensors or parameters, the feedback processing unit can decide whether to adjust the interface settings, perform a transmission retry or to ignore the error. The discussion below will also be given in the context of a memory card using a NAND-type architecture for memory arrays as shown in **FIGs 4A, 4B and 5**, but readily

extends to similar internal interfaces for other architectures, other forms of memory and non-card uses, such as embedded systems, SSD, and so on.

[0088] Although the following discussion may be based on various exemplary embodiments to provide concrete examples, the techniques and structures here can be applied fairly generally to memory systems having a controller and multiple banks that can independently operated, where the banks include some amount of non-volatile memory, whether flash or other variety, that can be used to store system data that the controller can use to manage the memory system. In addition to the other referenced cited above, theses can include the various memory systems presented in the following US patent, patent publication and application numbers: 7,480,766; US-2005-0154819-A1; US-2007-0061581-A1; US-2007-0061597-A1; US-2007-0113030-A1; US-2008-0155178-A1; US-2008-0155228-A1; US-2008-0155176-A1; US-2008-0155177-A1; US-2008-0155227-A1; US-2008-0155175-A1; 12/348,819; 12/348,825; 12/348,891; 12/348,895; 12/348,899; 12/642,584; 12/642,611; US 12/642,649; 12/642,728; 12/642,740; and 61/142,620.

[0089] Before discussion of the exemplary embodiment, this section will begin by further considering the problem being addressed. The controller-memory device interface is used to transfer data between a controller (100, **FIG. 1**) and one or more NAND (in the exemplary embodiment) devices (200, **FIG. 1**). (Note that this discussion relates to the internal interface on the memory system 90 between the controller 100 and the Flash memory 200, whereas the interface 110 is the host interface that the controller uses for communication with outside of the memory system.) Different NAND interface modes have been developed to increase the interface performance trading off speed, power consumption, and so on. As this interface is often a performance bottleneck, these interfaces are pushed to the limits to maximize system performances. To avoid data error, interface settings (such as voltage, frequency, drive strength and slew rate control) are being set for worst case scenario (extreme temperatures, extreme load capacitances, extreme voltages, and so on). Consequently, devices are typically designed to have worst-case safety margins, which translates to big margins in typical conditions. In such typical conditions the interface settings can be optimized to much higher interface performances without compromising the product reliability. Without a mechanism such as is presented in

the following, memory devices will continue to operate in the worst-case performances settings.

[0090] For example, a simple comparison between the burst data transfer times for a 16-bit Normal Mode at the nominal bus frequency of 33MHz to the accelerated 40MHz, to the hyper-accelerated 50MHz and the super hyper-accelerated 60MHz yields significant latency reductions of about 17%, 33% and 45%, respectively This is shown in Table 1, where the columns are the frequency, corresponding cycle ( $t_{cyc}$ ), time transfer 2142 Bytes of data, and speed ratio relative to that at 33MHz.

<b>Freq</b>	<b><math>t_{cyc}</math></b>	<b>2142B xfr time</b>	<b>Ratio</b>
33MHz	30.3ns	32454.5ns	1.00
40MHz	25.0ns	26775.0ns	0.83
50MHz	20.0ns	21420.0ns	0.66
60MHz	16.7ns	17850.0ns	0.55

TABLE 1

[0091] In the prior art, Flash interface performance is typically set to a fixed performance for a given product. The design then takes into account the worst-case design. In some products, the Flash interface is designed for a “close to worst-case”, allowing some interface performance optimization, but at the risk for some lower device yield or increased data error.

[0092] This section presents a feedback mechanism and processing unit that monitors the interface transfer integrity and adjusts the interface settings accordingly in order to optimize interface performance. In case of transmission errors, the feedback processing unit (assisted by the interface integrity feedback and possibly by other sensors or parameters) can decide whether to adjust the interface settings, perform a transmission retry or to ignore the error. In case of no transmission errors, the feedback processing unit may decide to leave interface settings as they are or modify interface settings in order increase the interface performance. Additionally, the interface integrity feedback mechanism can be designed in such a way that the feedback processing unit can get different grades of information, such as a binary pass/fail indication, a pass/fail plus number of errors, or a pass/fail plus number of errors plus error locations.

[0093] According to the embodiment, the feedback mechanism can utilize existing



device infrastructure or be further optimized by dedicated mechanism such as hash engine. Such dedicated mechanism may be implemented in hardware, software, or a combination of these. The hash engine may also be complemented by an error correction engine capable of correcting transmission errors. Such method would allow the interface to cope with a level of bit error rate, while still reaching optimal performance. Transmission correction ability is valuable because the design of ECC in the prior art for NAND bit failures only considers errors on the memory itself, and does not take into account the interface errors that may occur as the data is transferred between the controller and the memory device. As interface performance goes up the likelihood of transmission errors goes up. Having the legacy ECC to deal with interface errors degrade the legacy ECC capabilities in terms of performance and probability for unrecoverable error. Designing a dedicated interface error correction engine can allow for a “divide and rule”, letting the legacy ECC to focus only on the NAND generated errors. (Additional background detail on ECC can be found in the following US patents, patent publications, and patent application numbers: 2009/0094482; 7,502,254; 2007/0268745; 2007/0283081; 7,310,347; 7,493,457; 7,426,623; 2007/0220197; 2007/0065119; 2007/0061502; 2007/0091677; 2007/0180346; 2008/0181000; 2007/0260808; 2005/0213393; 6,510,488; 7,058,818; 2008/0244338; 2008/0244367; 2008/0250300; and 2008/0104312.)

[0094] **FIG. 11** is a block diagram showing such a feedback mechanism, but based on a typical prior art existing NAND/controller infrastructure. This will help to further illustrate some of the concepts involved as well as providing an alternate embodiment of an adaptive interface. In **FIG. 11** only the elements relevant to the present discussion are explicitly shown, the other being suppressed to simply the present discussion. On the controller 100 are the ASIC core 411, ECC circuitry 413, an output buffer 415, an input buffer 425, transmitting circuitry 417, and receiving circuitry 427. Although shown as separate here, this may not be so in an actual implementation: the input and output buffers may be overlap or be the same; the transmitting and receiving be share elements of even be the same; the ECC circuitry may be implemented as software in the ASIC core; and so on. On the memory side 200, the shown elements are read circuitry 431 and transfer circuitry 441 (which again may partially or completely overlap), an input data buffer 433 and an output data buffer 443 (which may similarly be a single buffer) and NAND core 435. The

controller 100 and the memory circuit are then connected by the bus structure 401.

[0095] A typical flow for a set of host data once it is received at the controller 100 is from the ASIC core 411 to the output data buffer 415, through the transmission circuitry 417 and onto the bus structure 401. On the memory 200 the data is transferred from the bus by the receiving circuitry 431 into the input data buffer 433 and then written into the NAND core 435. Subsequently, when the host wants to access the data it is read out of the NAND core 435 to the output data buffer 443, transferred onto the bus structure 401 by the transmission circuitry 441, and then read off the bus into the controller's input data buffer 425 by the receiving circuitry 427. Memory systems typically use error correction code (ECC) to detect and correct for error that may enter the data, where the controller generates the corresponding ECC that is transmitted and written into the NAND core along with the data and then read back with the data. The ECC engine 413 then has access to the data and its corresponding ECC, allowing the data to be checked and corrected as needed before it is passed on to the host.

[0096] Although the ECC can be used to correct data error, it can only correct a limited amount of error, where the amount is a design choice. Within these capabilities, the ECC engine 413 can correct for any error accumulated during the round trip, including transmission errors as well as error associated with the NAND core 435 itself, such as write error, read error and disturbs and other degradation of the data while stored; however the choice of ECC is usually based just on considerations of error related to the NAND core 435. In some arrangements, such as with "strong ECC" as disclosed in some of the references cited above for ECC, the code is based on the properties of the memory and how the data states are mapping into the memory. The transmission between the controller and memory is largely overlooked and taken to add no error. Accordingly, the interface needs to be set accordingly, leading the parameters to be set according to the worst, or near worst, case, as described above.

[0097] A first set of embodiments is based on the elements of **FIG. 11** to supply the feedback used to optimize the interface characteristics. A set of data, along with corresponding ECC, is sent on a round trip from the control to the memory and back to the controller, much as with the standard write followed by a read described above,

except that the data (and corresponding ECC) are not actually written into the memory core. After a write transfer is being issued from the controller 100 to the memory circuit 200, the controller can use the buffer latches 433 and 443 to read back the data. This is represent by the path 437, although if the input and output buffers are the same, there would not be an actual transfer. As this round trip removes any error associated with the array of 435 itself, this isolates the effects of the transmission and allows the ECC engine 413 to determine the memory interface's integrity. The interface parameters may then be modified and the process can be re-issued. This way both the write and read interface parameters may be optimized.

[0098] **FIG. 12** is a block diagram illustrating another set of embodiments, but where the feedback mechanism uses a hash engine and optional data correction engine specific to the interface. Rather than refer to the controller and the memory chip, **FIG. 12** is presented in terms of the circuitry on the transmitter side 520 and the receiver side 530 since, as described further below, either one of these may be the controller and the other the memory depending on whether it is a read process or write process and the two sides need not be symmetric.

[0099] The transmitter side 520 will again include a write data buffer 521 and transmission interface circuitry 529. It will now also include a hash values generator 525 and a multiplexer 527. In a transfer process, the data to be written (523) is transferred out of the write buffer 520 to both the hash value generator 525 and the MUX 527. The hash value generator 525 correspondingly generates a hash value from the data, which is then also passed on to MUX 527. The multiplexer then supplies the data followed by its hash value to the transmission interface circuitry 529 and then onto the bus structure 550.

[00100] The receiver side again includes receiver interface circuitry and a read data buffer 535, plus some additional elements. After the read interface circuitry 532 takes the data and corresponding hash value off the bus 550, the de-multiplexing circuitry 533 separates the hash value from the data, the read data is sent to the buffer 535 and also to a receiver side hash value generator 539 that again generates a hash value from the data set. The receiver side generated hash value is then compared to the received hash value in the comparison circuitry 541. Depending on the embodiment, the result of the comparison can just determine whether these values

match or further determine the amount of error due to the transfer process. A data correction engine 537 can also be included in some embodiments to correct interface errors without having to perform a data retransmission. In the exemplary embodiment, the hash generators (and the optional data correction engine on the receiver side) separate from ECC for used NAND core error, although there may be some overlap in circuitry; and, in fact, both may be implemented on the same logic circuitry of the controller but by different firmware code. (Although considered separate for this discussion, the two error detection/correction parts can also be interactive in a more general embodiment as described below.) Typically the hash value will be generated based on the entirety of the information being sent (user data, corresponding ECC, header information, etc.), but in alternate embodiments, it could be generated from just a portion by, say, stripping off the various overhead and just using the user data itself for generating the hash value.

**[0100] FIG. 12** also includes feedback processing unit 560 connected to receive the output of the hash comparison circuitry 541. This feed back is then analyzed at 561, which, depending on the embodiment may consider one or more of temperature, supply voltage level, and the processing related quality of the NAND core. At 563 the results of this feedback can then be used to adjust the transmission process and is correspondingly connected to one or both of the transmission interface circuitry 529 and read interface circuitry 531. For a write operation (where the controller is the transmitter side), after a write transfer is issued from the controller to the memory device, the feedback processing unit may just read back the comparison of the generated hash values and by that determine the write-direction memory interface integrity. Based on this, interface write parameters may be modified and the process can be re-issued if desired. Symmetrically, the same operation may be employed to the read-direction where the memory is the transmitter side.

**[0101] FIG. 13** is a diagram showing an example for transmitting the data and the generated hash value over the bus interface. As shown at top, the corresponding hash value is automatically appended to the data, so that these will be transmitted together when a device operates in this mode. In the second option, shown at bottom, the data payload is transmitted, the receiving side requests the corresponding hash value, which is then generated and transmitted. The data payload can be of predefined

length or of random length. If data payload length is predefined then the hash value can be appended to the data, as in the first option, or sent on request. If data payload length is random, then the hash value can be send after issuing a specific command.

[0102] A number of variations are possible for the techniques and corresponding circuitry described with respect to **FIG. 12**. With respect the hash value engines and hash values, the hash engine may be parity code (cyclic redundancy check, or CRC), ECC, and so on. For example, a “binary” embodiment can be used that would return Pass/Fail, can be built based on Error Bit Count (CRC) and has the benefit of a low gate count to implement. Alternately, a “soft” embodiment can an return Error Bit Count (EBC), and optionally the locations of the failed bits, and can be built based on ECC codes such BCH or Reed–Solomon codes, providing more information to assist the system with accurate decisions. The hash engine may optionally also have a complementary feature of correcting the interface failures, for example similarly to the correcting flipped bits from the memory core, as represented by data correction engine 537 of **FIG. 12**. Based on the feedback from the transfer, the system may repeat the transfer. Transfer retry may be decided based on a binary transfer status or on a soft transfer status. Further, transfer retry may be decided based on a combination of transfer status and number of NAND bit flips; for example, if the interface introduced N errors and the NAND introduced M errors and the controller error correcting capability is P, and  $P > N + M$ , then the system may decide to not re-transmit.

[0103] The system can also be configured in various different ways. The configuration can be symmetric, where the hash engines at controller and memory sides are the same, or asymmetric. In an asymmetric configuration, different configurations are used for different transmission directions; for example, a faster mechanism can be designed for read transmissions while a more reliable mechanism is designed for the write transmissions. Also, it should be noted that even if the interface is configured symmetrically, as the settings may be changed during the interval between the initial write and subsequent read of the data, it may function asymmetrically with respect to a given set of data.

[0104] The feedback processing unit 530 may be variously located on the controller 100, on the memory 200, on both, or distributed between the two. In can also be

formed on a separate circuit. In many applications, it will be most practical to implement the feedback processing unit on the controller, since the controller circuit often includes more advanced processing capabilities and also as memory system frequently of formed of multiple memory chips, but the techniques presented here are not so limited. In any of these variations, the checking for data transfer status phase is the responsibility of the feedback processing unit.

[0105] Considering further an example where the feedback processing unit is located on the controller side: in the read direction, after the controller reads the data and hash value, these will pass through the feedback mechanism and the controller will determine the pass/fail status and can adjust (or not) the interface setting accordingly. As the controller has already read the data and hash value, there is no further need of information from the flash side to determine the status, as this can be done in the controller's logic. In the write direction, the data payload and corresponding hash value are sent to the memory side and the controller can then operate in several different ways: Read the Pass/Fail status from the memory side; read back the hash value and determine Pass/Fail; read back the Error Bit Count (EBC) from the memory; read back the EBC and error locations from the NAND; or read Pass/Fail status and number of corrected bits from the memory side.

[0106] The feedback processing unit may decide to modify interface settings. For example the following interface settings may be modified: drive strength; bus frequency or other timing parameters; interface voltage; interface mode (e.g. switch from a Normal/Conventional-mode to a Toggle-mode); and so on. These interface settings may then be modified in an adaptive feedback fashion. As factors such as process variations, supply voltage levels, and temperature affect the likelihood of interface error, these factors can also included as inputs to the feedback analysis 561 on **FIG. 12**.

[0107] The bus frequency and other parameter settings can be based on earlier remissions, with nominal parameter settings also being settable in various ways. For example, a Look Up Table (LUT) having different values for different bus capacitance/NAND configurations can be used. Such a Look Up Table (LUT) can also have different values for different operating process parameters, voltage supply levels, temperature, and so on. Process parameters, voltage supply levels, and

temperature can also be a variable in a function (formula), instead of predefined in a LUT.

[0108] The interface settings optimization task may be operated in the background. Special events, such as voltage supply or temperature change, can also be used to trigger an interface settings training task. The interface training task may use also a known pattern transmitted across and not written to the NAND core, such as was described above with respect to **FIG. 11** and path 437. Interface settings can also be different, and based on, read direction and write direction, or on different data retention requirements.

[0109] The preceding discussion has mainly considered the memory system as having a controller and a single memory device circuit. More generally, the system may include several memory chips that can be connected to the controller (and the feedback processing unit if a separate circuit) using various bus topologies. For example, all the memory chips may share a single system bus; or each memory circuit may have its own controller-memory bus; or various hybrid arrangements could be used. Different interface settings can then be applied to this plurality of NAND devices (e.g., if interfacing several devices, this could be done in parallel). Different interface setting can also be used based on the particular NAND device being accessed, as interface quality may be a function of the particular NAND device's load and/or cell/block quality). Further, within a given memory device, different interface setting can also be applied to the blocks within the NAND core, as interface quality may be a function of the particular block's quality.

[0110] More detail on the techniques of the forgoing section can be found in US patent application number 12/835,292 filed on July 13, 2010.

#### DYNAMIC OPTIMIZATION OF BACK-END MEMORY SYSTEM INTERFACE

[0111] This section will consider the controller-memory (or "back-end") interface of the memory system further and present some methods for dynamically optimizing the back-end read and write performance suitable for high-speed memory systems, including those with multiple memory data buses. As discussed above, a memory system is usually designed to have a certain amount of error tolerance; and although this error can arise both in the controller-memory transmission process and in actual

on-memory storage process, traditionally only the last of these is considered for the ECC process and the back-end interface is typically optimized to eliminate, or at least minimize as far as possible, transmission channel error. In many cases, though, the data error resulting from the storage process (including read and write error) may be well below the ECC capabilities of the system. For example, although a heavily cycled device may need the full available data correction, a fresh device may have relatively little error, leaving the system with excess error correction capability. This section presents methods whereby the memory system internally allots a non-zero portion of this error correction capacity to the transmission channel. This allows for the interface to operate at , for example, higher speed or lower power, even though this will likely lead to transmission path error. When the memory portion requires a higher amount of error correction, the allotment can dynamically be adjusted. In complementary aspects, the system can calibrate transmission path to determine the amount of resultant transmission error for different operating parameters and then select the parameters based upon how much is to be allowed.

**[0112]** Considering the back-end interface between the controller and the memory section further, a typical memory system consists of a memory controller and a memory device, such as a NAND flash memory module. The back-end interface is the data bus between the memory and its controller. The interface is typically established in one of two ways. In the first, if the controller and memory devices are discrete components, the back-end interface is established by conducting traces on printed circuit boards (PCB), onto which these components are mounted. In the second, the controller and memory can be encapsulated in a single package, such as system-in-package (SIP) or multi-chip package (MCP). In this second case, the back-end interface is established by the package substrate.

**[0113]** As discussed in the previous section, the overall bit error rate (BER) in memory systems can be attributed to two main factors: the reliability of data retention in memory devices, such as NAND flash memory; and imperfection of back-end interface, which can causes transmission error. Error correction coding (ECC) can then be employed in memory systems to address this overall BER. **FIG. 14** schematically illustrates the contributions to the bit errors in a memory system.



[0114] As shown in FIG. 14., one of the main sources of the overall bit error rate BER 605. The effect of this error, due to data degradation for the stored data (from charge leakage, disturbs and so on), as well as any error introduced in the read and write process, is shown as NAND Retention 601. Traditionally, the data correction is used to account just for this factor, which is observed when the data is read. Error due to channel imperfection is shown at 603 and affects both data read and write, but the impact will again be observed at read. The sources of the channel impaction error can include inter-symbol interference (ISI), same data bus (intra-bus) crosstalk, inter-bus cross talk (on multi-data buses designs), printed circuit board (PCB) noise, silicon die noise, package noise, and so on. On the other side, ECC 607 can correct error up to a certain level of error.

[0115] As the data transfer rate between the controller and the memory increases, the back-end interface becomes more susceptible to signal integrity related issues contributing to 603, such as crosstalk among signals within the same data bus (intra-memory data bus crosstalk) and inter-symbol interference (ISI). In addition, the introduction of memory topologies where controllers can access multiple memory devices simultaneously (multi-memory data buses design) subjects the back-end interface to simultaneously switching noise and crosstalk among data buses (inter-memory data bus crosstalk). In addition to bus speed, factors such as the voltage amplitude of the data bus and temperature (ambient temperature for PCB traces and junction temperature for system-in-package (SIP) or multi-chip package (MCP)) can also affect signal integrity of the back-end interface. Therefore, the intrinsic imperfections of the back-end interface becomes a bottleneck in determining overall system performance for high-speed memory systems. The pin capacitance of memory devices increases with the number of memory dies. High capacity memory devices constructed with multiple memory dies exhibit high capacitance on their data input/outputs (I/Os), which further degrades the edge rate and signal integrity of the data bus structure.

[0116] Signal integrity related issues on signal traces can be minimized by increasing the spacing the signal traces apart from one another to minimize crosstalk; but this approach is limited by the available area on the PCB or substrate. It can also be reduced by selecting PCB materials with low dielectric constant and low dissipation

factor (loss tangent); but these PCB materials are more costly than the typical materials. So although there are ways to reduce this error with output lowering bus speed or otherwise degrading operating bus parameters, these ways suffer from drawbacks.

[0117] This section presents a dynamic optimization technique to address these signal integrity issues in the back-end interface and also account for process variations among controller and memory devices. In addition to process variations, the voltage settings and temperature under which the memory system operates may vary. A static solution does not account for variations in process, voltage, and temperature, and thus may not be the optimal approach.

[0118] This section uses a pseudo loop-back manner to dynamically optimize the back-end performance of memory systems, including those with multiple memory data buses. This can be done with the sort mechanisms similar to those described in the preceding sections or by using predetermined data pattern. The exemplary embodiment will use a pseudo-random bit pattern (PRBS). Dynamically optimizing the data bus settings can help to maximize the reliability of data transmission between the controller and the memory devices. This can allow the memory system to differentiate transmission error from error caused on the memory devices. These aspects can be particularly advantageous for products equipped with high-speed back-end memory interfaces and multiple memory data buses.

[0119] Conducting traces on PCB or package substrate have limited bandwidth, which causes inter-symbol interference (ISI). The effect of ISI depends on the edge rate (rise time and fall time), data rate, and data pattern. In digital communications, the pseudo-random bit pattern (PRBS) pattern is sometimes used to exploit the worst-case ISI impact of a data link because such pattern is rich in frequency components. A PRBS pattern is a repeating pattern that has properties similar to a random sequence and is used to measure jitter and eye mask of transmitted data in electrical data links. The PRBS is typically denoted as a  $2^X - 1$  PRBS or PRBS-X, where the power (X) indicates the shift register length used to create the pattern. Each  $2^X - 1$  PRBS contains every possible combination of X number of bits (except one). It is desirable to use the longest PRBS pattern practical as it exerts the greatest stress on the signal link and provides a better representation of random data.

[0120] Although the exemplary embodiment uses a pseudo-random bit pattern, other patterns can be used, as long as the system knows the pattern of the data set that was used so that this can be compared to the data that comes back at the end of the loop-back process. The exemplary embodiment uses the PRBS pattern because its random-like character can maximize the ISI impact of the signal link. In addition to the PRBS pattern, other types of data pattern can be utilized in this invention, and each individual pattern may generate a different result.

[0121] A PRBS pattern can be applied to each of the signal link in the parallel back-end interface. Ideally, although the pattern would be repeated infinitely, which is not practically feasible in the memory system, this should not be a major drawback provided that the pattern can be repeated a sufficient number of time by using a short pattern. For example, if the page size of a NAND flash memory is 16kB, a PRBS-7 pattern with pattern length of 127b can be repeated completely 129 times on each signal link of an 8-bit data bus. The remaining bit ( $16384b - 127b \times 129 = 1b$ ) constitutes an incomplete copy of the PRBS-7. This incomplete PRBS pattern at the end should not cause a significant as most of the transmission link effects are accounted for by the 129 cycles of complete PRBS pattern.

[0122] FIG. 15 can be used to illustrate the operation of a pseudo loop-back method in the back-end interface. In FIG. 15, the left hand side is a flow, with the right hand side schematically illustrating the corresponding controller-memory interactions. At 701, the controller turns off its data scrambling and error correction coding (ECC) capabilities, where on the right hand side, this is represented by these elements being X-ed out. Consequently, all data transferred out of and into the controller are in their raw format without any scrambling or correction. At 703, the controller sends a command to the memory device, telling it to store and hold the data to be received in its data latch register without transferring them to the memory cells. That is, there is no programming of this data set into the memory cells. At 705, the controller sends a known data pattern (here an independent PRBS-7 pattern) on each of the signal links in the data bus structure to the memory device. The memory device holds the data in the data latch register until it becomes full.

[0123] At 707, the controller sends a command to the memory device, telling it to send back the data stored in its data latch register continuously until the controller

instructs it to stop. That is, after the data latch register has dumped all 16kb of data on each link, it will start sending back the same data again. Thus, the PRBS-7 pattern repeats on each signal link of the data bus. The reason for the exemplary embodiment using the continuous operation of the PRBS pattern transfer is that bit error introduced by a signal link is a probabilistic event. The greater the amount of data transferred across a link, the more accurate and representative of the transmission bit error rate (BER) as a statistical measurement of the intrinsic link performance. At 709, the controller receives the repeating PRBS-7 pattern (or other used data pattern) on each signal link of the data bus.

[0124] At 711, the controller compares the received data to the sent data pattern (here the standard PRBS-7 pattern) and reports any error as transmission BER. The controller then sends a command to the memory device to stop sending the PRBS-7 pattern (713) and exit the pseudo loop-back mode (715).

[0125] **FIG. 16** corresponds to block 705 of **FIG. 15** where the controller sends the data pattern memory device, with **FIG. 17** corresponding to block 709. On these two figures, one particular example of a bus structure 811 is shown, where several lines are shown at top for command and control signals and a number of data lines shown below. Here, CLE = command latch enable, ALE = address latch enable, RE = read enable, WE = write enable, DQS = data strobe and there are eight input/output lines (IO0-IO7). These box diagrams are simplified for purposes of this discussion, with only an ECC block 805 and a PRBS generator 803 shown on the memory controller 801 and only the data register REG 833 is represented on the memory device 831, with other elements (including the non-volatile memory array on 831) not being explicitly shown. When the controller 801 is sending the data pattern across to the memory device 831 in **FIG. 16**, the write enable signal and the data strobe will be asserted and each signal line will carry the data pattern. Here, each IO line carries an independent pattern. Here, these are all individual copies of PRBS, but can have differing timing, as indicated by their relative skew in the figure. On the memory 831 the data pattern is then stored as received in the register 833. In **FIG. 17** the data is sent back from the register 833 across the bus structure 811 to the controller 801, so the read enable signal and the data strobe are now asserted. Once on the data pattern has completed the round trip (without being written into the non-volatile memory) and

is back on the controller 801, it can be checked against its original form and see how much corruption has occurred. Although the bus structure 811 is a parallel bus interface with multiple signal lines, this is just a particular example other bus structures can be used for the transmission channel, such as serial data arrangements.

[0126] The performance of a memory system is characterized by the data transfer rate (bus operating frequency) with respect to power consumption, which is directly related to the data bus voltage. By varying the voltage amplitude of the data bus (determined by the I/O of the device driving the data bus) and the data transfer rate, a 3-dimensional representation as a shmoo plot can be created with the transfer rate plotted along the x-axis, data bus voltage along the y-axis, and transmission BER along the z-axis. Here, the data transfer rate can refer to either the one applied during write operations, as in 705 of FIG. 15, when data is transferred from the controller circuit to the memory circuit, or to the read operation, as in 709, when data is transferred from the memory to the controller. The data represented in such a shmoo plot can be measured at various output drive impedance (drive strength) and temperatures to cover the worst-case, typical, and best-case scenarios. As a result, for a given amount of allowable transmission error, the optimal point of operation can be determined at a given combination of parameters such data bus voltage, output drive impedance, slew rate, line capacitance, transfer rate, temperature, and power consumption. An example of the shmoo plot is shown in FIG. 18.

[0127] FIG. 18 is an example of a shmoo plot showing transmission BER vs. data bus voltage and data transfer rate at a fixed output drive impedance, slew rate, line capacitance, and temperature for a particular example of a memory system. The data bus voltage is  $V_{DD}$  on the vertical axis and the transfer rate on the horizontal. The amount of transmission BER is represented by color on the graph, with the key on the right of the figure. In this black and white representation the representation of very low and very high error amount appear the same, but the lower error region in the main figure is to the left of the light-colored dividing regions, with the higher error regions to the right. Based on this sort of data, for amount of allowed transmission data, a combination of operating parameters can be selected, where, as usual, this will often involve a trade-off. For example, if the desired amount of allowable BER is  $10^{-5}$ , if maximum speed is the primary concern,  $V_{DD}$  would be taken about 3.1-3.2 V,

allowing a transfer rate of about 170-180 Mb/s. If power consumption is a more important concern, a lower  $V_{DD}$  value could be used, say 2.8 V, would then allow a transfer rate of about 150 Mb/s for about the same transfer BER. If the BER allocated to the transmission channel is re-allotted to a different value, based, for example, on the how much the memory has cycled or the ECC indicating that the combined contributions to the BER are approaching the maximum capabilities of the system, the operating parameters for the bus system can then be adjusted by the controller based on this data.

[0128] Consequently, after calibrating the system by capturing the data represented in the shmoo plots at various output drive impedance, slew rate, line capacitance, temperature and so on, the memory system can operate according to various cases. For example, given a desired transmission BER, the memory system looks up and selects the optimal data bus voltage, data transfer rate, output drive impedance, and slew rate. (This data from the calibration process can be kept in either the non-volatile memory or in the memory space (RAM) in the controller circuit.) For example, it may select the lowest data bus voltage, highest data transfer rate, and lowest output drive impedance that result in the desired transmission BER. In another example, given a particular combination of data bus voltage, data transfer rate, output drive impedance, slew rate, line capacitance, and temperature, the memory system knows what transmission BER it can expect. Alternately, the memory system may select an operating condition that balances all factors – data bus voltage, data transfer rate, output drive strength, and transmission BER.

[0129] Since the design of the I/O buffers in the controller and the memory device may be different, the optimal read and write performance of the memory system can be determined separately. In addition to the differences resulting from differing memory system design, there will also be differences for individual examples of the same device, due to process variations as well as by differences in operating conditions. To account for device aging, changes in operating conditions, and so on, the calibration process can also be repeated. For example, an initial calibration could be performed before the device is shipped at test time, and then the controller could recalibrate system periodically or in response to an event, such as device cycling, error results, notable changes in operating conditions and so on. Thus, in addition to

change the proportion of total error allotted to the transmission channel, the corresponding operating parameter for a given allocation may change dynamically.

[0130] As noted above, performance can be optimized during both the read and write processes. Going back to **FIG. 15**, for performance optimization during memory read, at 705 the system slows down the transfer rate of the data pattern being written into the memory device's data latch register to maximize the integrity of the transfer for the data pattern. For example, at 10 MHz transfer rate, it will take 1.6 ms to fill up a 16kb data latch register. At 709 and 711, the system measures the transmission BER incurred during the read operation where the memory device's I/O is the driver and the controller is the receiver. The shmoo plot data then show the relationship between the memory device's I/O voltage and the read frequency.

[0131] For performance optimization during memory write, at 705, the system varies the voltage and transfer rate of the data pattern being written into the memory device's data latch register. At 709, the system will then slow down the transfer rate of the data from memory device to controller to prevent injecting additional bit errors by the signal link. The transmission BER measured is then the one incurred during the write operation in 705. The shmoo plot data will consequently represent the relationship between the controller's I/O voltage and the write frequency.

[0132] So far, the various aspects presented here have been presented in a context where this is only a single bus between a controller circuit and a single memory circuit. However, a memory system may include multiple devices with various bus topologies; and when there are multiple busses, interactions between these busses can lead to additional sources of error. The techniques here can provide the ability to skew each of the signal links within a data bus in the back-end interface to a specified resolution, for example, 100 ps. Such skewing capability can be introduced by either the controller or the memory device in either the driver or the receiver. Introducing skew into the data bus allows the system to compensate for length mismatch of the signal traces in the PCB or package substrate. Introducing skew can reduce the impact of near-end and far-end crosstalk in the back-end interface and, consequently, lower the transmission BER. Two types of crosstalk are involved: intra-memory data bus crosstalk; and inter-memory data bus crosstalk. Such crosstalk causes jitter in the data bus. A typical memory system uses the clock sent by the driver to sample each

individual signal in the parallel data bus at the same moment. Thus, an increase in jitter on each signal in the data bus will cause an increase in transmission BER. By skewing the data across multiple memory data buses so that they do not align relative to one another, inter-memory data bus cross talk can be reduced.

[0133] FIG. 19 is a block diagram that illustrates this crosstalk in a memory system whose bus structure uses multiple memory data buses. The memory system includes controller 901 and the multiple, here four, memory devices 931-1, 931-2, 932-3, 934-4 that are connected to the controller through respective busses 911-1, 911-2, 911-3, 911-4. For each of the busses 911, they will have one or more IO lines as shown in the detail as IO1 to IOX. As before, these individual buses can various numbers of IO lines operated in a parallel, serial, or a combination of these for the transfer of data. This sort of multi-bus arrangement is often implemented in SSD type devices (see, for example US patent 7,376,034, US patent 7,765,339, or the paper “A High Performance Controller for NAND Flash-based Solid State Disk (NSSD)” Park et al, Samsung, Non-volatile Semiconductor Workshop, 2006, IEEE, NVSMW 2006, 21<sup>st</sup>, vol. no., pp. 17-20, 12-16 Feb, 2006) in order to improvement performance, but is also found in the design of some memory cards and other memory systems. In addition to the sort of intra-memory data bus cross talk between the IO lines of a given bus, there will also now be inter-memory data bus crosstalk signal on the different busses. When combined with the use of the PRBS pattern and pseudo loop-back mode described above, the optimal skew that generates the least crosstalk, and consequently the lowest transmission BER, can be determined given a certain combination of data bus voltage, transfer rate, temperature, output drive impedance, slew rate, line capacitance, and power consumption.

[0134] The various aspects presented here provide a low-cost solution that optimizes the back-end interface performance under the presence of various signal integrity issues. By dynamically allotting “unused” ECC capability to the transfer process, the performance can be improved as described. As noted above, some memory systems use a type of “strong” ECC that exploits the properties of multi-state memory devices, in which case the error correction capabilities transferred for use of the transmission channel may not transfer in a 1 to 1 fashion. It should also be noted that although the memory system incorporates ECC to compensate for data error, there is typically not



the same provision for commands and the memory device will often not accept corrupted commands, so that although error may intentionally allowed in the transmission of data, this will not be the case for commands. Consequently, although a higher transfer rate may be allowed for the data by these mechanisms, a slower, safer setting for the transfer rate (or other parameters) can be incorporated so that no error is incurred for the control signals.

### Conclusion

[0135] The foregoing detailed description of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. The described embodiments were chosen in order to best explain the principles of the invention and its practical application, to thereby enable others skilled in the art to best utilize the invention in various embodiments and with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the claims appended hereto.

IT IS CLAIMED:

1. A method of operating a non-volatile memory system, the non-volatile memory system including a controller circuit having a memory interface, a memory circuit having an array of non-volatile memory cells and a controller interface, and a bus structure connected to the memory interface of the controller circuit and to the controller interface of the memory circuit for the transfer of data and commands therebetween, wherein the memory system can tolerate a first, non-zero amount of accumulated error from when data is transferred from the controller to be written to the memory array until the data is received at the controller after subsequently being read back from the memory array, the method comprising:

allotting by the controller circuit of a first, non-zero portion of the first amount of error to the transfer of the data between the controller circuit and the memory circuit via the bus structure, the remainder of the first amount of error being allotted to write, storage, and read of the data on the memory circuit; and

setting by the controller circuit of the transfer characteristics between the controller circuit and the memory circuit to operate to allow up to the first portion of error.

2. The method of claim 1, wherein said transfer characteristics include the voltage amplitude of the bus structure.

3. The method of claim 1, wherein said transfer characteristics include the data transfer rate on the bus structure.

4. The method of claim 1, wherein said transfer characteristics include signal drive strength.

5. The method of claim 1, wherein said transfer characteristics include signal slew rate.

6. The method of claim 1, wherein the memory system includes error code and correction (ECC) circuitry and the first amount of error is based on the capability of the ECC circuitry.

7. The method of claim 6, where the ECC circuitry is on the controller circuit.
8. The method of claim 6, the method further including:
  - receiving at the controller circuit said data from a host;
  - generating corresponding ECC code for the data;
  - transferring the data and corresponding ECC code from the controller circuit to memory circuitry on the bus structure according to said transfer characteristics; and
  - subsequently writing the data and corresponding ECC code as received at the memory circuit into the array of memory cells.
9. The method of claim 1, further comprising:
  - subsequently re-allotting by the controller circuit of a second portion of the first amount of error to the transfer of the data between the controller circuit and the memory circuit via the bus structure; and
  - setting by the controller circuit of the transfer characteristics between the controller circuit and the memory circuit to operate to allow up to the second portion of error.
10. The method of claim 9, wherein the controller circuit re-allots the portion of error from the first amount to the second amount in response to the number of program-erase cycles experienced by the memory circuit.
11. The method of claim 9, wherein the controller circuit re-allots the portion of error from the first amount to the second amount in response to the amount of error detected in data read back from the memory array
12. The method of claim 1, wherein the memory system maintains a correspondence between values of one or more operational parameters for the bus structure and the amount of resultant error to the transfer of the data between the controller circuit and the memory circuit via the bus structure, wherein setting the transfer characteristics includes:
  - selecting values for the one or more operational parameters by the controller circuit based upon the correspondence.

13. The method of claim 12, wherein the correspondences is for a plurality of operational parameters and selecting values for the plurality of operational parameters includes choosing between a plurality of combinations of the plurality of parameters that allow up to the first portion of error according to one or more predetermined performance criteria.

14. The method of claim 13, wherein the operational parameters include the voltage amplitude of the bus structure. and the data transfer rate on the bus structure.

15. The method of claim 13, wherein the operational parameters include signal drive strength.

16. The method of claim 13, wherein the operational parameters include signal slew rate.

17. The method of claim 12, further comprising:  
prior to selecting said values for the one or more operational parameters, establishing by the controller circuit of the correspondence.

18. The method of claim 17, wherein establishing the correspondence includes:

for each of said values of the operational parameters, transferring a known pattern of data over the bus structure from the controller circuit to the memory circuit and back without writing the data into the array of non-volatile memory cells, and comparing the data as received back at the controller circuit with the data pattern as sent.

19. The method of claim 17, wherein the memory system include multiple memory circuits and the bus structure includes a corresponding multiplicity of buses whereby each of the memory circuits is individually connected to the controller circuit, and wherein establishing by the controller circuit of the correspondence includes determining inter-bus cross-talk error.

20. A method of operating a non-volatile memory system having a controller circuit and a memory circuit including an array of non-volatile memory cells, the method comprising:

performing by the controller circuit of a transmission error calibration by performing, for each of a plurality of values for each of one or more operating parameters for a bus structure connecting the controller with the memory circuit, a process including:

transferring a data set of a known data pattern from the controller through transmission circuitry on the controller to the bus structure;

receiving the data set from the bus structure through receiving circuitry on the memory circuit;

storing the data set as received in buffer memory on the memory circuit;

transferring the data set as stored in buffer memory on the memory circuit, and without being written into the array, through transmission circuitry on the memory circuit to the bus structure;

receiving the data set from the bus structure through receiving circuitry on the controller;

performing a comparison of the data set as received with the known pattern; and

based on the comparison, determining the amount of error associated with the transmission process for the used one or more operating parameter; and

subsequently operating the memory system to allow a first, non-zero amount of error in the transmission of data between the controller circuit and the memory circuit, wherein the controller circuit selects values of the operating parameters according to the transmission error calibration process based on the determined associated amounts of error.

21. The method of claim 20, wherein the operating parameters include the voltage amplitude of the bus structure.

22. The method of claim 20, wherein the operating parameters include the data transfer rate on the bus structure.

23. The method of claim 20, wherein the operating parameters include signal drive strength.

24. The method of claim 20, wherein the operating parameters include signal slew rate.

25. The method of claim 20, further comprising:

subsequently operating the memory the memory system to allow a second amount of error in the transmission of data between the controller circuit and the memory circuit, wherein the controller circuit selects values of the operating parameters according to the transmission error calibration process based on the determined associated amounts of error.

26. The method of claim 20, further comprising:

subsequently re-performing the transmission error calibration process.

27. The method of claim 20, wherein subsequent to performing the transmission error calibration process, the memory system stores the result thereof in non-volatile memory.

28. The method of claim 20, wherein the transmission error calibration is performed for a plurality of operating parameters and the controller circuit chooses between a plurality of combinations of the plurality of operating parameters that allow up to the first amount error according to one or more predetermined performance criteria.

29. The method of claim 20, wherein the memory system include multiple memory circuits and the bus structure includes a corresponding multiplicity of buses whereby each of the memory circuits is individually connected to the controller circuit, and wherein transferring the data set of a known data pattern includes data for determining inter-bus cross-talk error.

30. A non-volatile memory system, comprising:

a controller circuit, including a memory interface and logic circuitry;

a memory circuit, including an array of non-volatile memory cells, a controller interface, and logic circuitry;

a bus structure connected to the memory interface of the controller circuit and to the controller interface of the memory circuit for the transfer of data and commands therebetween; and

a feedback processing circuit connected to the logic circuitry of the receiving one the controller and the memory circuit during a transfer of data therebetween to receive information on the amount of error occurring as a result of the transfer, and connected to one or both of the memory interface and the controller interface to adjust the characteristics of the transfer therebetween in response to the amount of error.

31. The non-volatile memory system of claim 30, wherein said transfer of data is from the controller circuit to the memory circuit.

32. The non-volatile memory system of claim 30, wherein said transfer of data is from the memory circuit to the controller circuit.

33. The non-volatile memory system of claim 30, wherein the logic circuitry of each of the controller and the memory circuit include a hash value generator,

wherein in the transfer process from the first one of the memory circuit and the controller to the second one thereof, the first transmits over the bus structure a data set and a first hash value generated from the data set by first one's logic circuit, and the second one receives from the bus structure the data set and the first hash value and generates from the received data set a second hash value from the data set by second one's logic circuit, and

wherein the logic circuitry of the second one of the memory circuit and the controller further includes comparison circuit connected to receive and perform a comparison of the first hash value as received and the second hash value, said amount of error being based on the result of the comparison.

34. The non-volatile memory system of claim 33, wherein the circuitry used for generating the respective hash values on the memory circuit and the controller are equivalent.

35. The non-volatile memory system of claim 33, wherein the circuitry used for generating the respective hash values on the memory circuit and the controller are not equivalent.

36. The non-volatile memory system of claim 30, wherein the comparison determines whether or not the first and second hash values are equal, and, in response to determining that the first and second values are not equal, the logic of circuitry of the second one of the memory circuit and the controller further quantifies the amount of error.

37. The non-volatile memory system of claim 30, wherein the feedback processing circuit is formed on the same integrated circuit as the controller.

38. The non-volatile memory system of claim 30, wherein the feedback processing circuit is formed on the same integrated circuit as the memory circuit.

39. The non-volatile memory system of claim 30, wherein the feedback processing circuit is formed on a separate integrated circuit from both of the controller and the memory circuit.

40. The non-volatile memory system of claim 30, wherein the memory circuit is formed of a plurality of integrated circuits, each including an array of non-volatile memory cells, a controller interface, and logic circuitry.

41. The non-volatile memory system of claim 40, wherein each integrated circuit of the memory circuit is connected to the controller by a distinct bus.

42. The non-volatile memory system of claim 40, wherein one or more of integrated circuit of the memory circuit are connected to the controller by a shared bus.

43. The non-volatile memory system of claim 40, wherein characteristics of the transfer between the controller and a plurality of the integrated circuits of the memory circuit can be independently adjusted.



44. A method of operating a non-volatile memory system including a controller circuit and a non-volatile memory circuit, the method comprising:

generating a first hash value from a data set in logic circuitry on a first of the controller circuit and the memory circuit;

transmitting to a bus structure the data set and the first hash value through an interface on the first of the controller circuit and the memory circuit;

receiving the data set and the first hash value from the bus structure through an interface on the second of the controller circuit and the memory circuit

generating a second hash value from the data set as received in logic circuitry on the second of the controller circuit and the memory circuit;

comparing the first hash value as received and the second hash value on the second of the controller circuit and the memory circuit; and

based on the comparison of the first hash value as received and the second hash value by the logic circuitry on the second of the controller circuit and the memory circuit, determining whether to alter characteristics of the transfer of data between the controller circuit and the memory circuit.

45. The method of claim 44, wherein the first of the controller circuit and the memory circuit is the controller circuit.

46. The method of claim 44, wherein the first of the controller circuit and the memory circuit is the memory circuit.

47. The method of claim 44, wherein the logic circuitry used for generating the respective hash values on the memory circuit and the controller are equivalent.

48. The method of claim 44, wherein the logic circuitry used for generating the respective hash values on the memory circuit and the controller are not equivalent.

49. The method of claim 44, wherein the characteristics of the transfer of data between the controller circuit and the memory circuit include a frequency at which data is transferred.

50. The method of claim 44, wherein the characteristics of the transfer of data between the controller circuit and the memory circuit include a clock frequency for the bus structure..

51. The method of claim 44, wherein the characteristics of the transfer of data between the controller circuit and the memory circuit include a slew rate used in the transfer.

52. The method of claim 44, wherein the characteristics of the transfer of data between the controller circuit and the memory circuit include an interface voltage at which data is transferred.

53. The method of claim 44, wherein the characteristics of the transfer of data between the controller circuit and the memory circuit include a drive strength at which data is transferred.

54. The method of claim 44, wherein the hash values are built based on a cyclic redundancy check..

55. The method of claim 44, wherein the hash values are built based on an error correction code.

56. The method of claim 44, wherein prior transmitting the data set and the first hash value through the interface on the first of the controller circuit and the memory circuit, the characteristics of the transfer of data are set to an initial set of values determined from a look up table based upon the characteristics of the bus structure and the memory circuit of the memory system.

57. The method of claim 44, wherein prior transmitting the data set and the first hash value through the interface on the first of the controller circuit and the memory circuit, the characteristics of the transfer of data are set to an initial set of values determined from a look up table based upon the one or more parameters.

58. The method of claim 44, wherein prior transmitting the data set and the first hash value through the interface on the first of the controller circuit and the memory circuit, the characteristics of the transfer of data are set to an initial set of values determined from the quality of a previous transmission of data between the controller circuit and the memory circuit.

59. The method of claim 58, wherein the one or more parameters include a supply voltage level.

60. The method of claim 58, wherein the one or more parameters include a temperature.

61. The method of claim 58, wherein the one or more parameters include processing values for one or both of the controller circuit and the memory circuit.

62. The method of claim 44, the generating and transmitting of the first hash value is in response to a request from one of the controller circuit and the memory circuit.

63. The method of claim 44, further comprising:

in response to said determining, altering the characteristics of the transfer of data between the controller circuit and the memory circuit. wherein the altering is done symmetrically with respect to transfers from the controller circuit to the memory circuit and transfers from the memory circuit to the controller circuit.

64. The method of claim 44, further comprising:

in response to said determining, altering the characteristics of the transfer of data between the controller circuit and the memory circuit. wherein the altering is done asymmetrically with respect to transfers from the controller circuit to the memory circuit and transfers from the memory circuit to the controller circuit.

65. A method of operating a non-volatile memory system having a controller circuit and a memory circuit including an array of non-volatile memory cells, the method comprising:

transferring a set of data from buffer memory on the controller through transmission circuitry on the controller to a bus structure connecting the controller to the memory circuit;

receiving the set of data from the bus structure through receiving circuitry on the memory circuit;

storing the set of data as received in buffer memory on the memory circuit;

transferring the set of data as stored in buffer memory on the memory circuit, and without being written into the array, through transmission circuitry on the memory circuit to the bus structure;

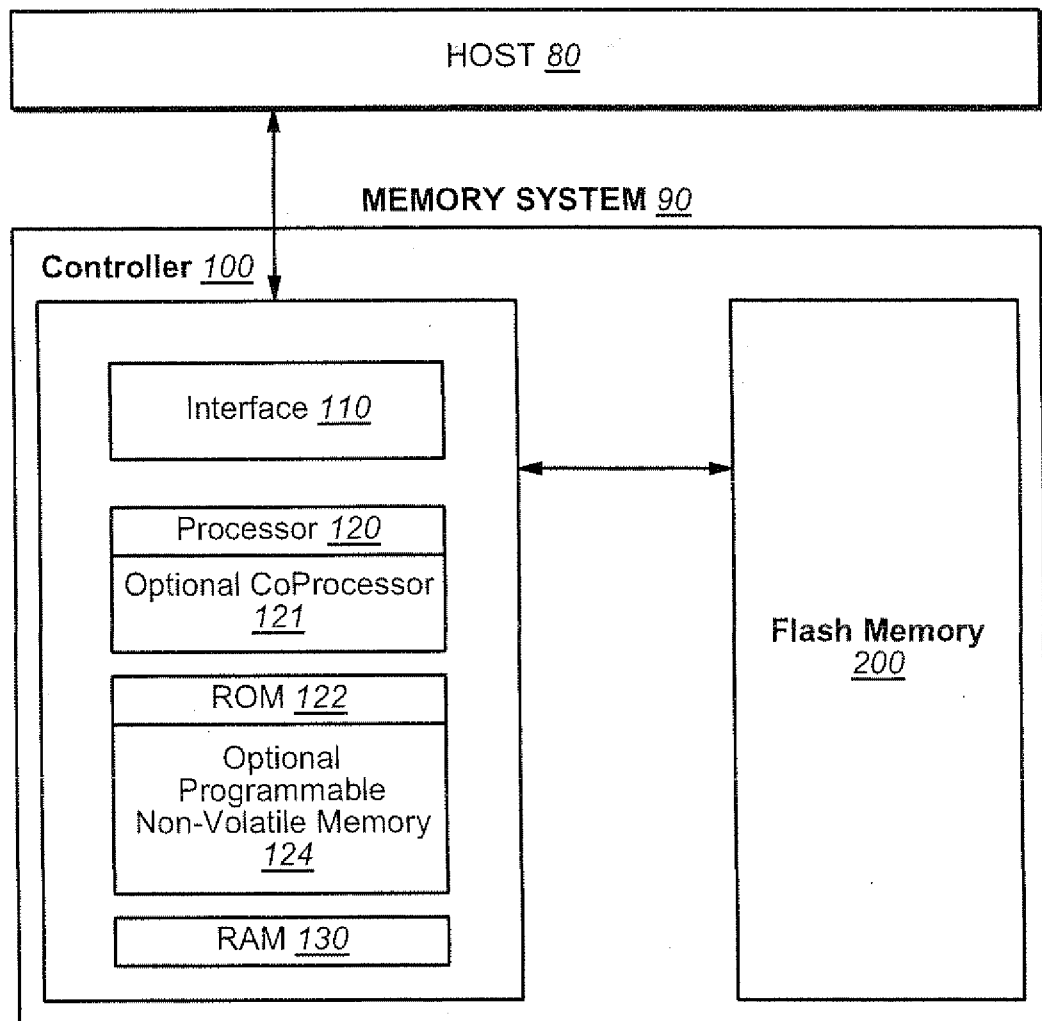
receiving the set of data from the bus structure through receiving circuitry on the controller;

subsequently storing the set of data as received in buffer memory on the controller;

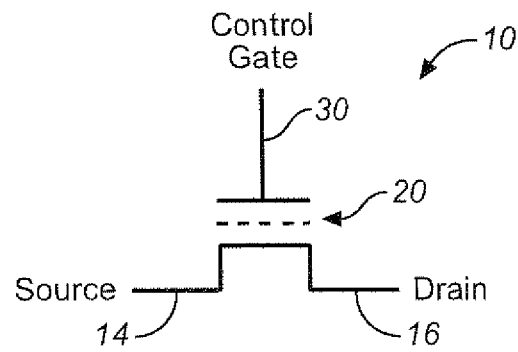
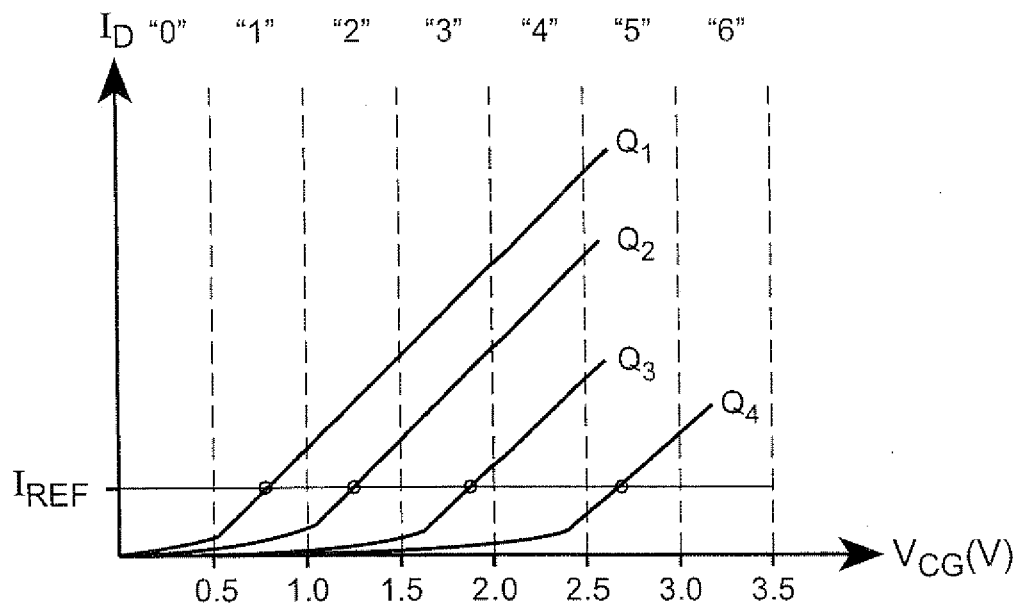
subsequently adjusting characteristics of the transfer of data between the controller circuit and the memory circuit based upon amount of error the set of data as received and stored in the buffer memory on the controller.

66. The method of claim 65, wherein the amount of error is determined by ECC circuitry on the controller.

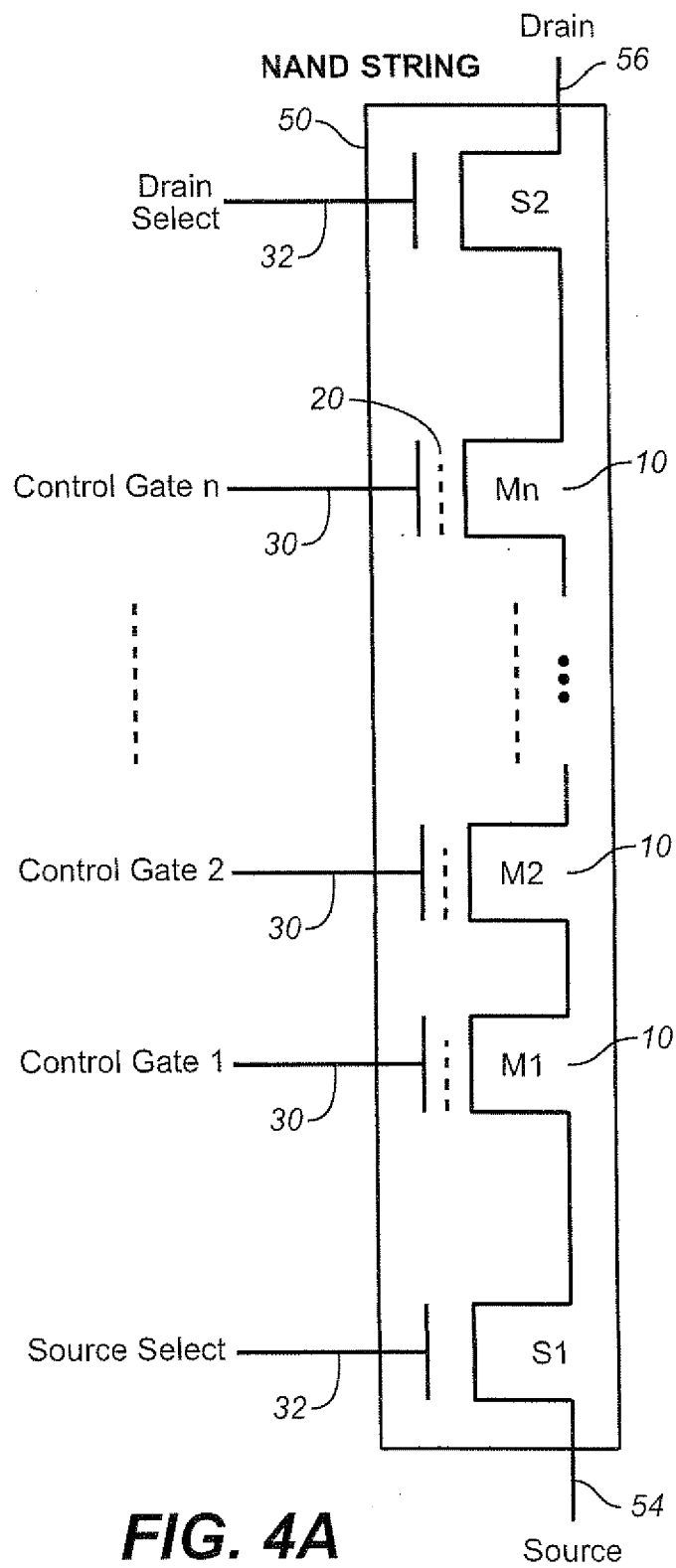
1 / 19

**FIG. 1**

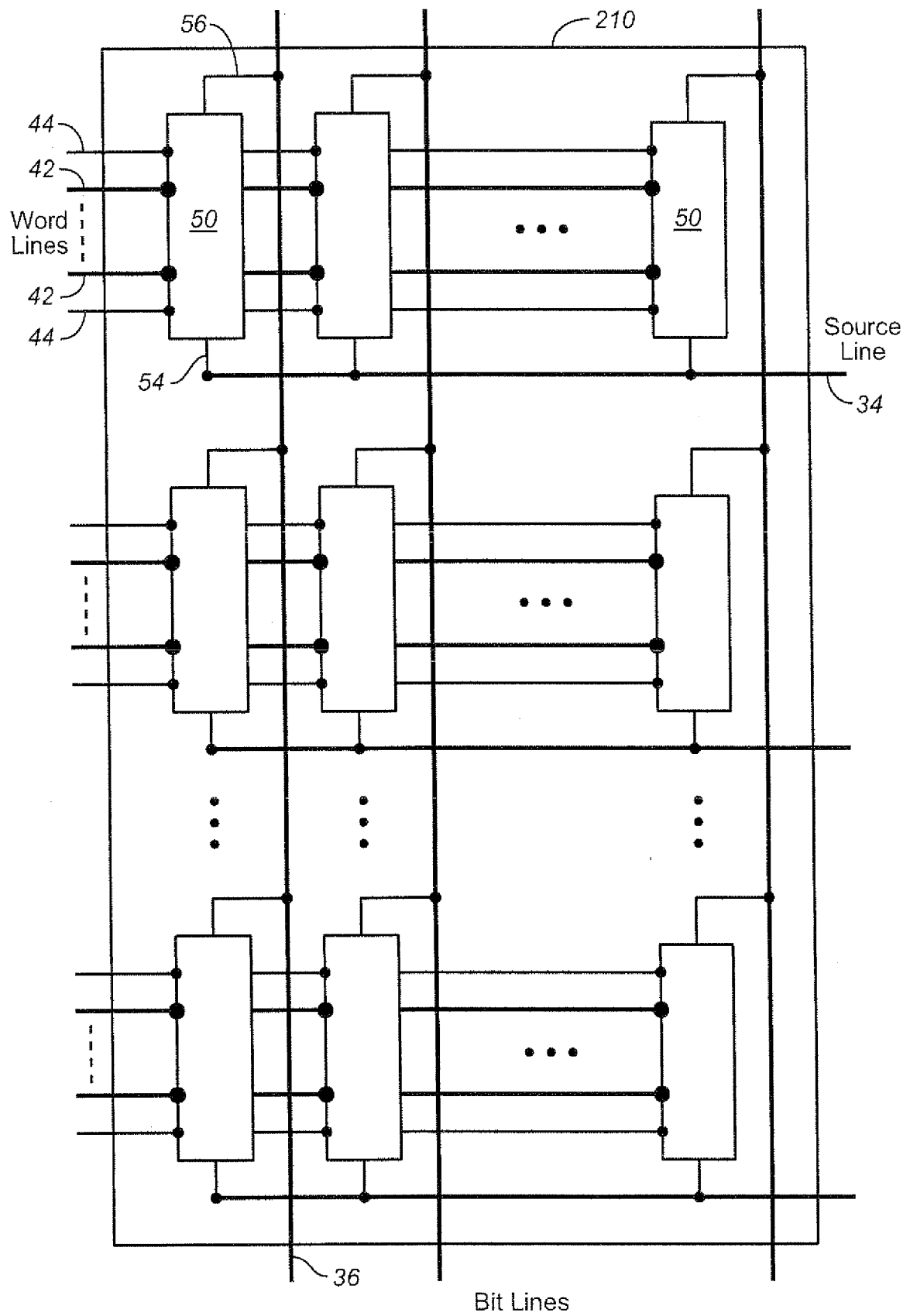
2 / 19

**FIG. 2****FIG. 3**

3 / 19



4 / 19

**FIG. 4B**



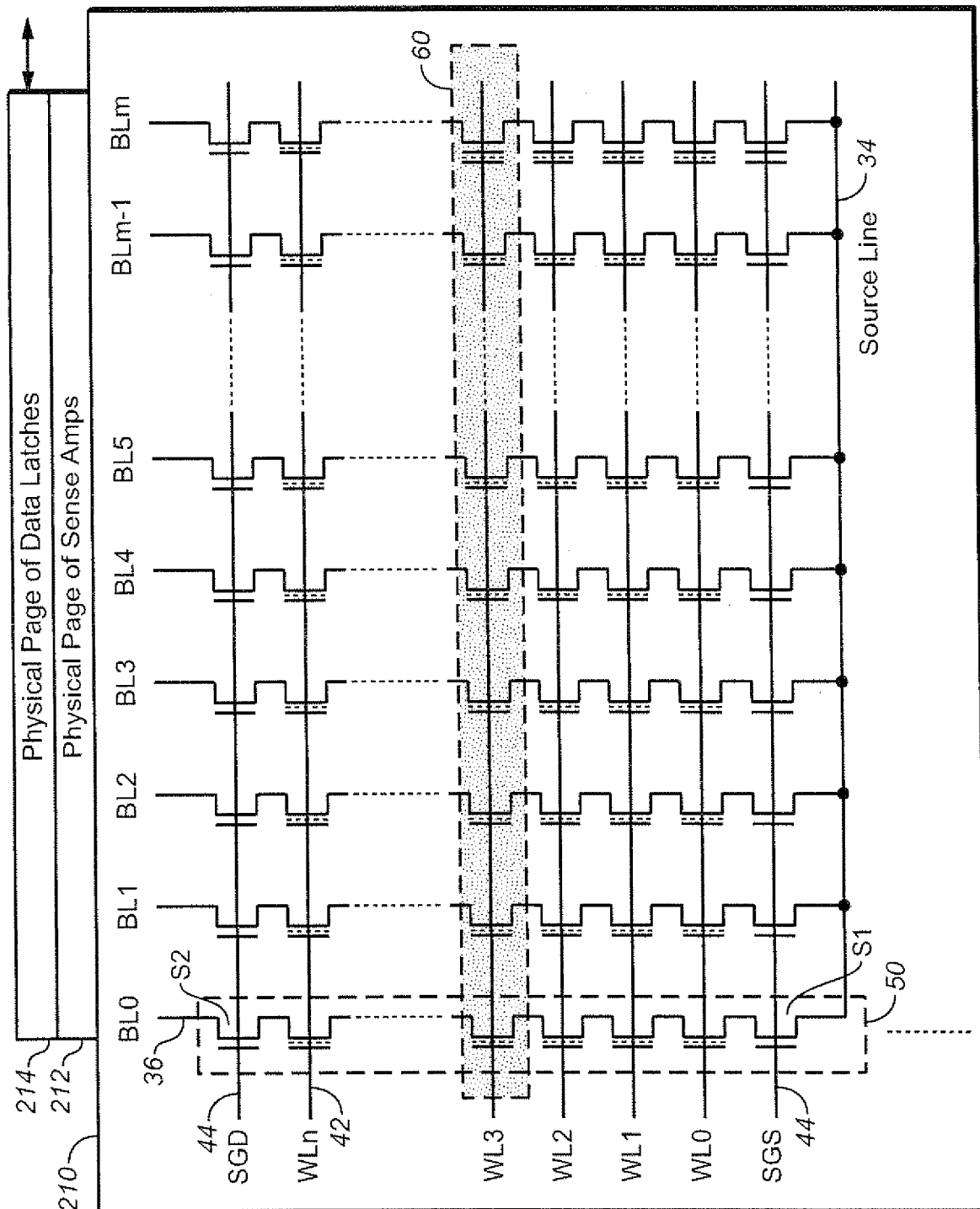
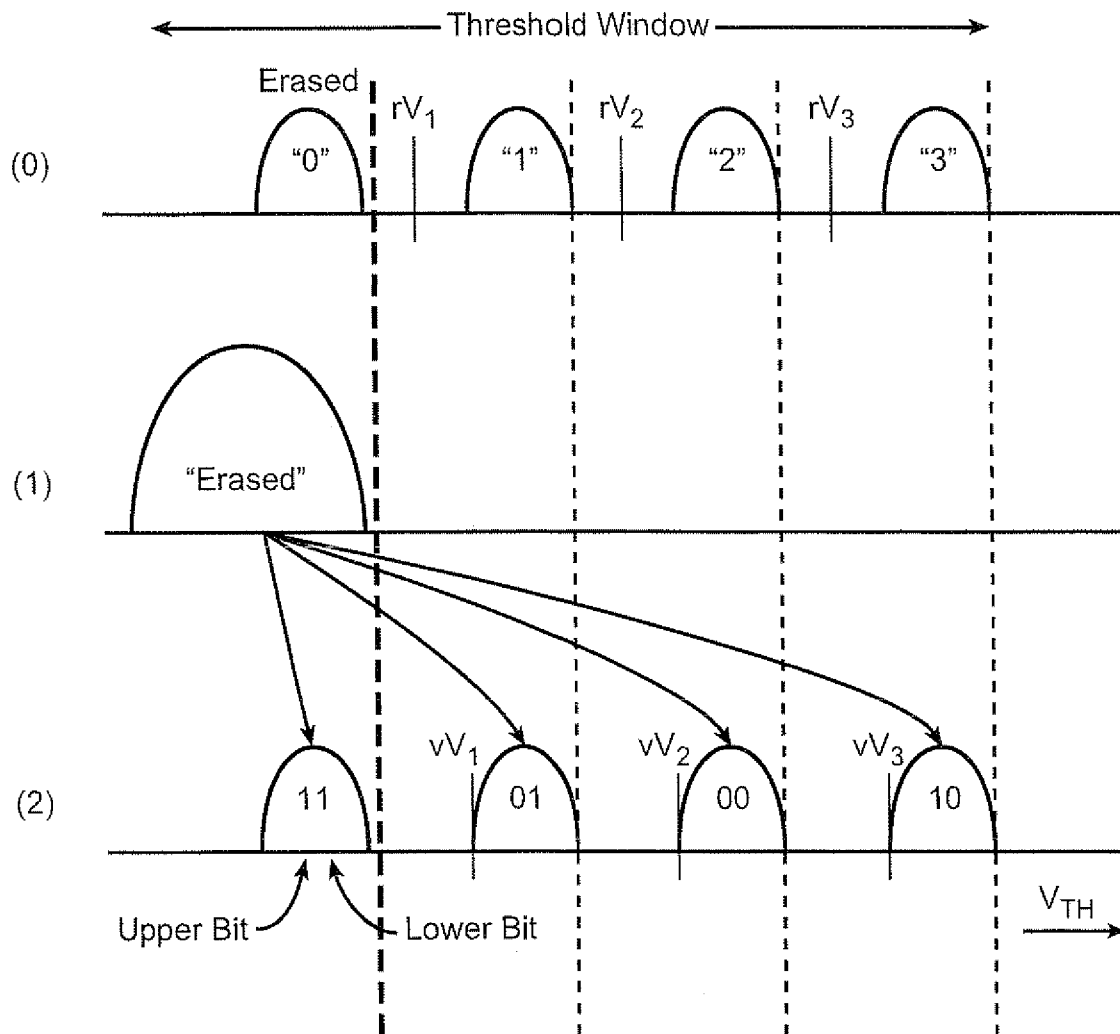


FIG. 5

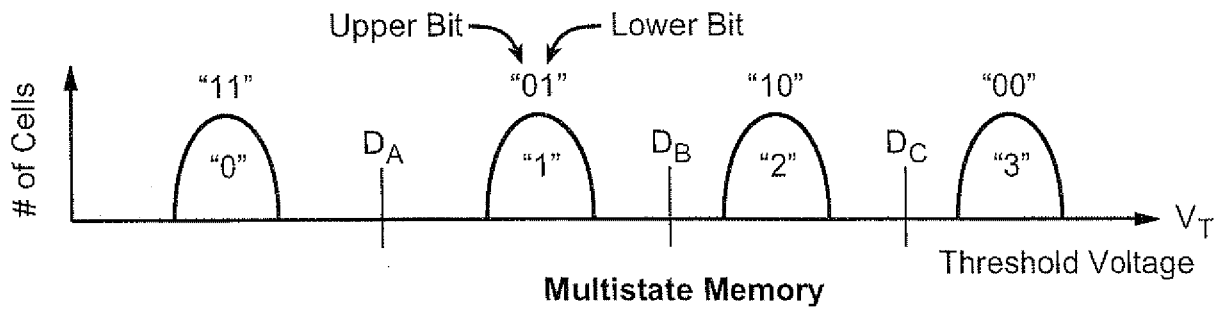
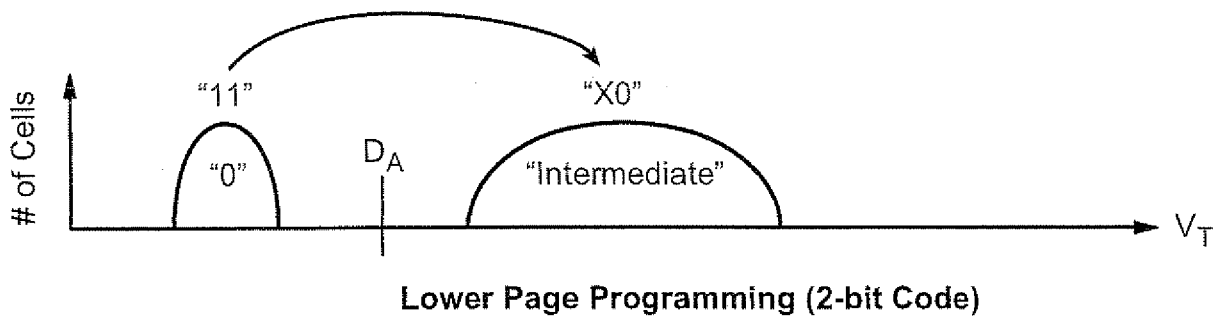
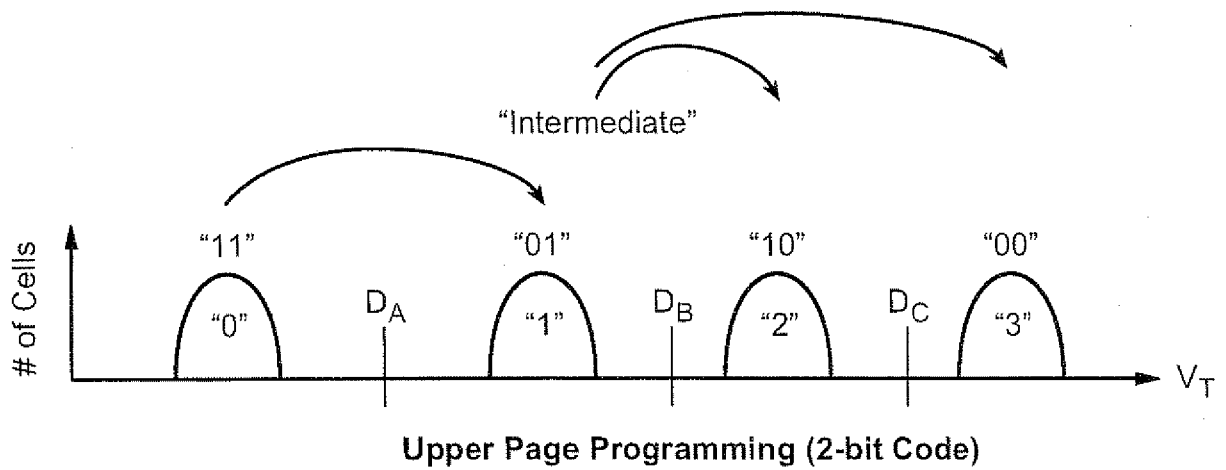
6 / 19



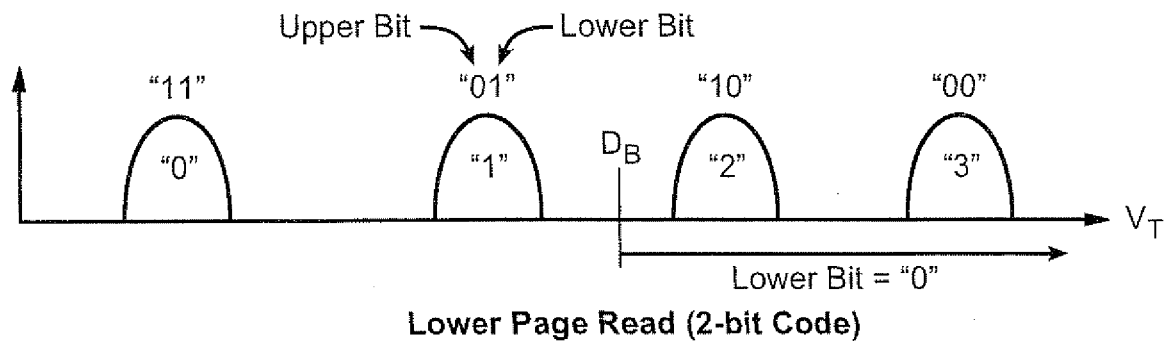
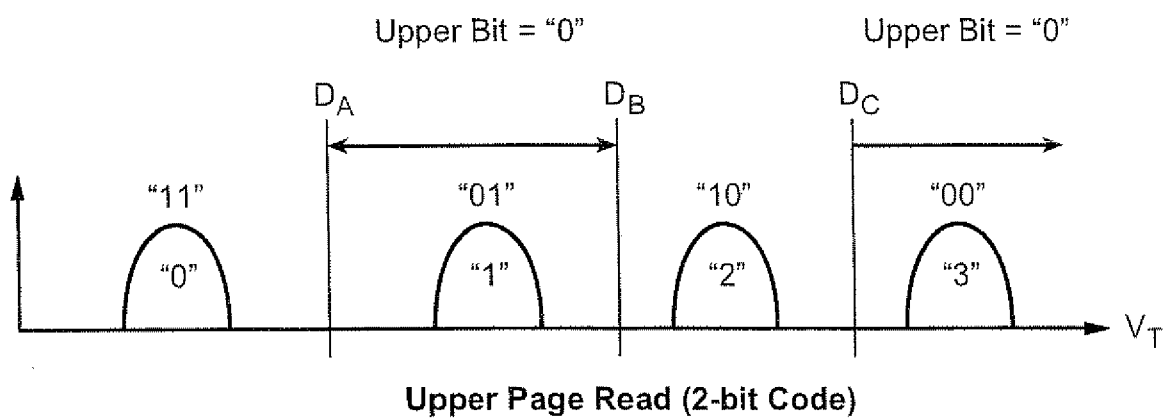
Programming into Four States Represented by a 2-bit Code

**FIG. 6**

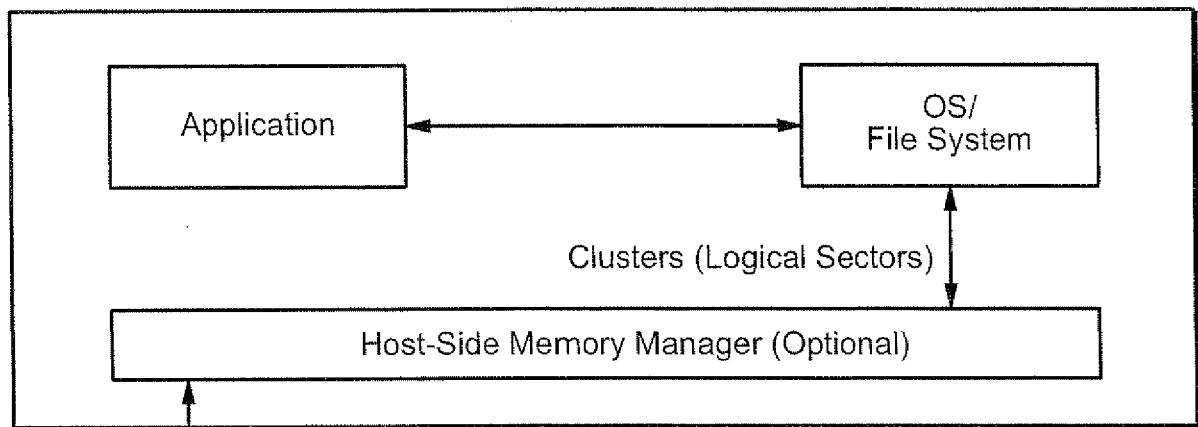
7 / 19

**FIG. 7A****FIG. 7B****FIG. 7C**

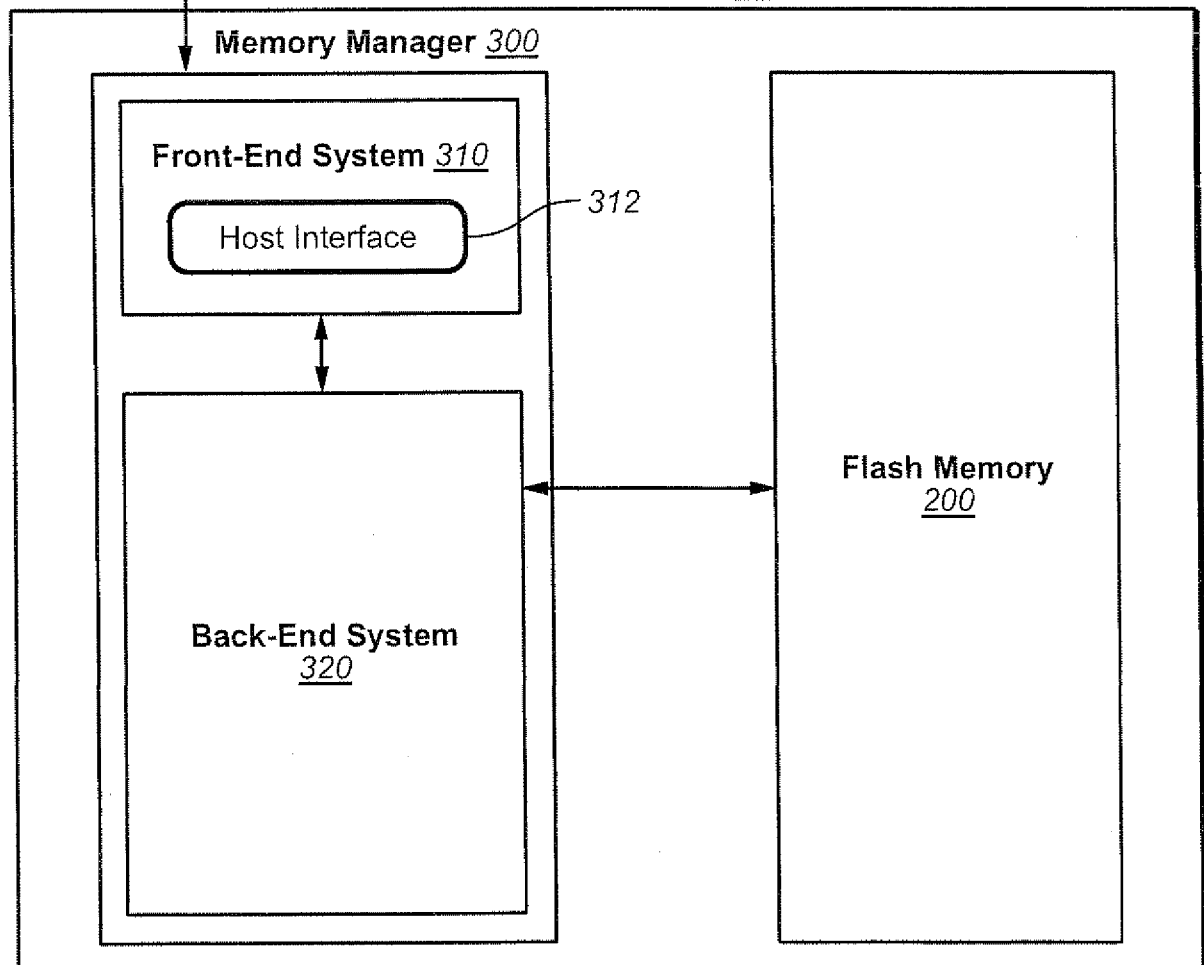
8 / 19

**FIG. 7D****FIG. 7E**

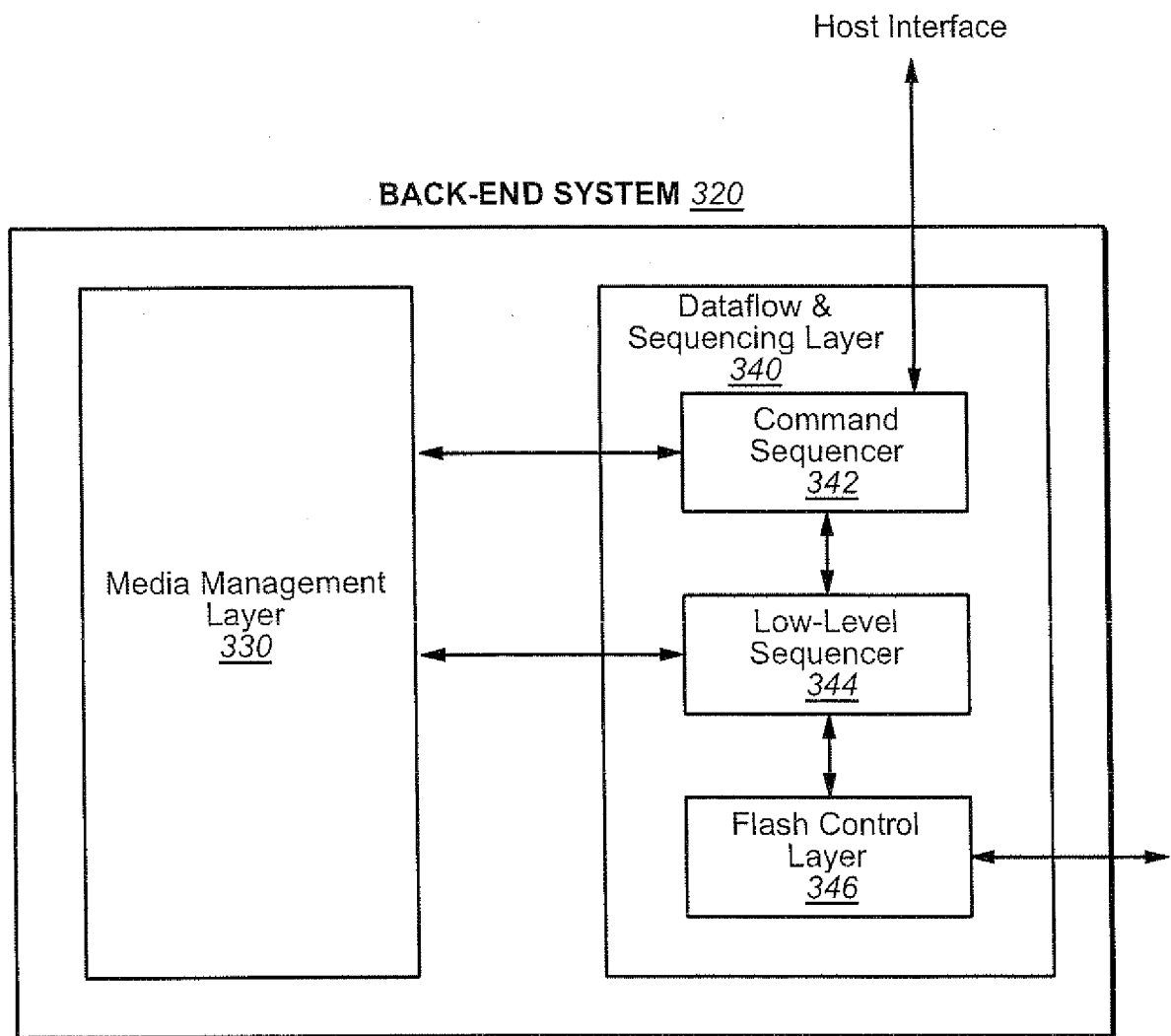
9 / 19

**HOST 80**

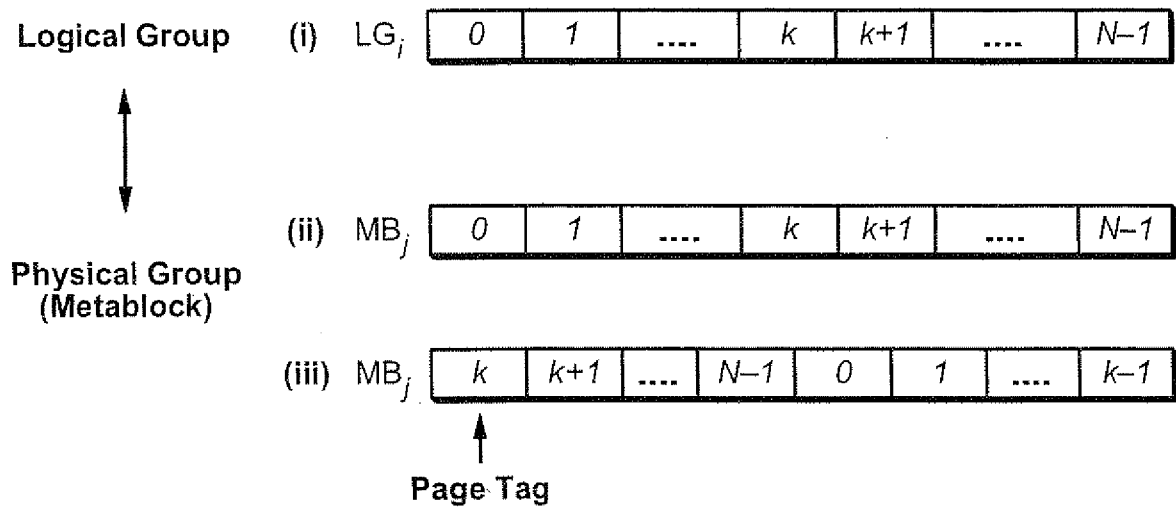
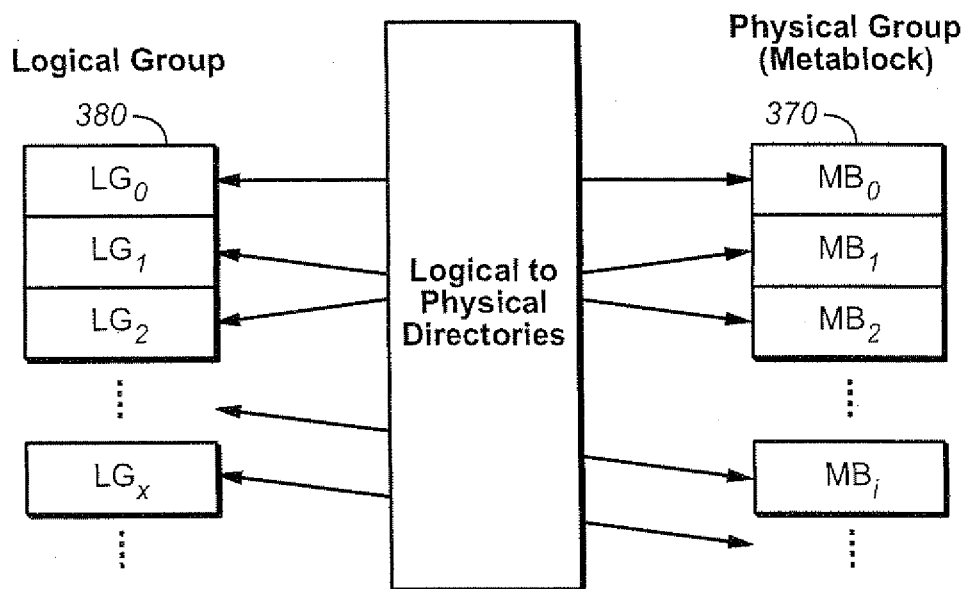
Logical Sectors

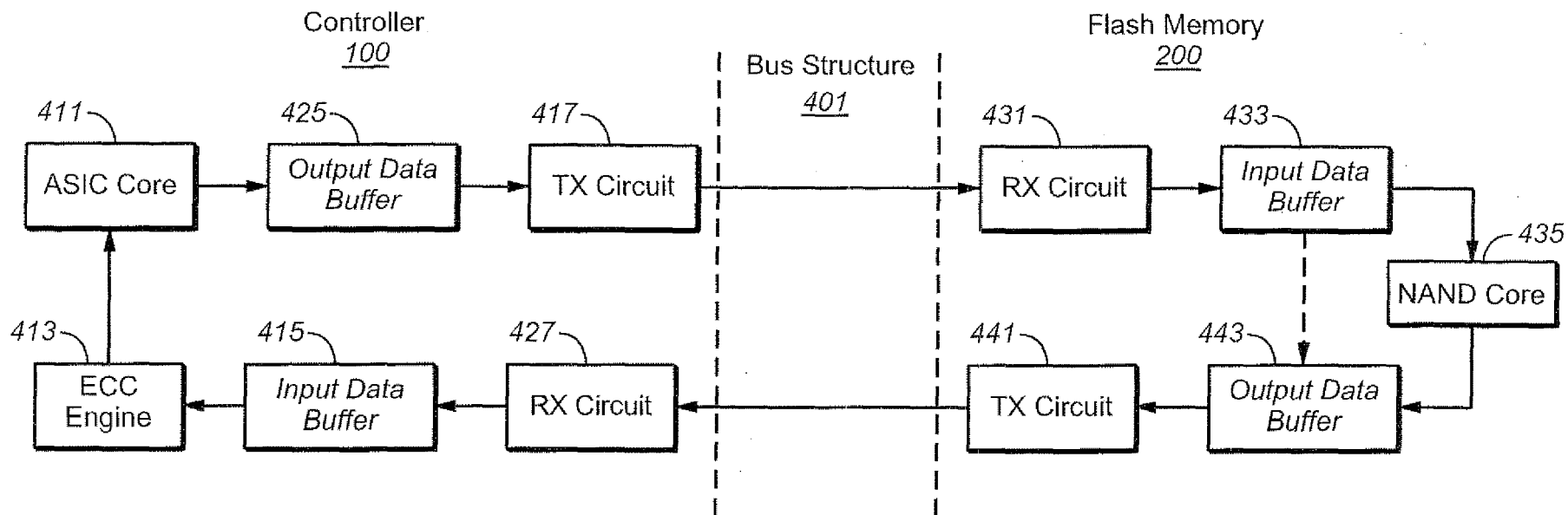
**MEMORY SYSTEM 90****FIG. 8**

10 / 19

**FIG. 9**

11 / 19

**FIG. 10A****FIG. 10B**



**FIG. 11**





Option 1: Hash Value Automatically Appended

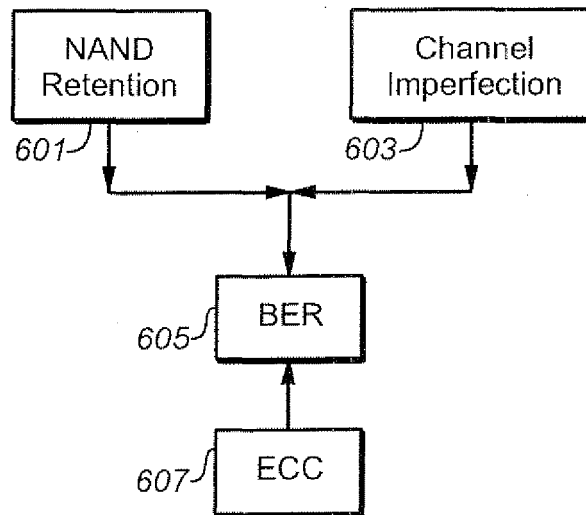


Option 2: Hash Value Provided on Request

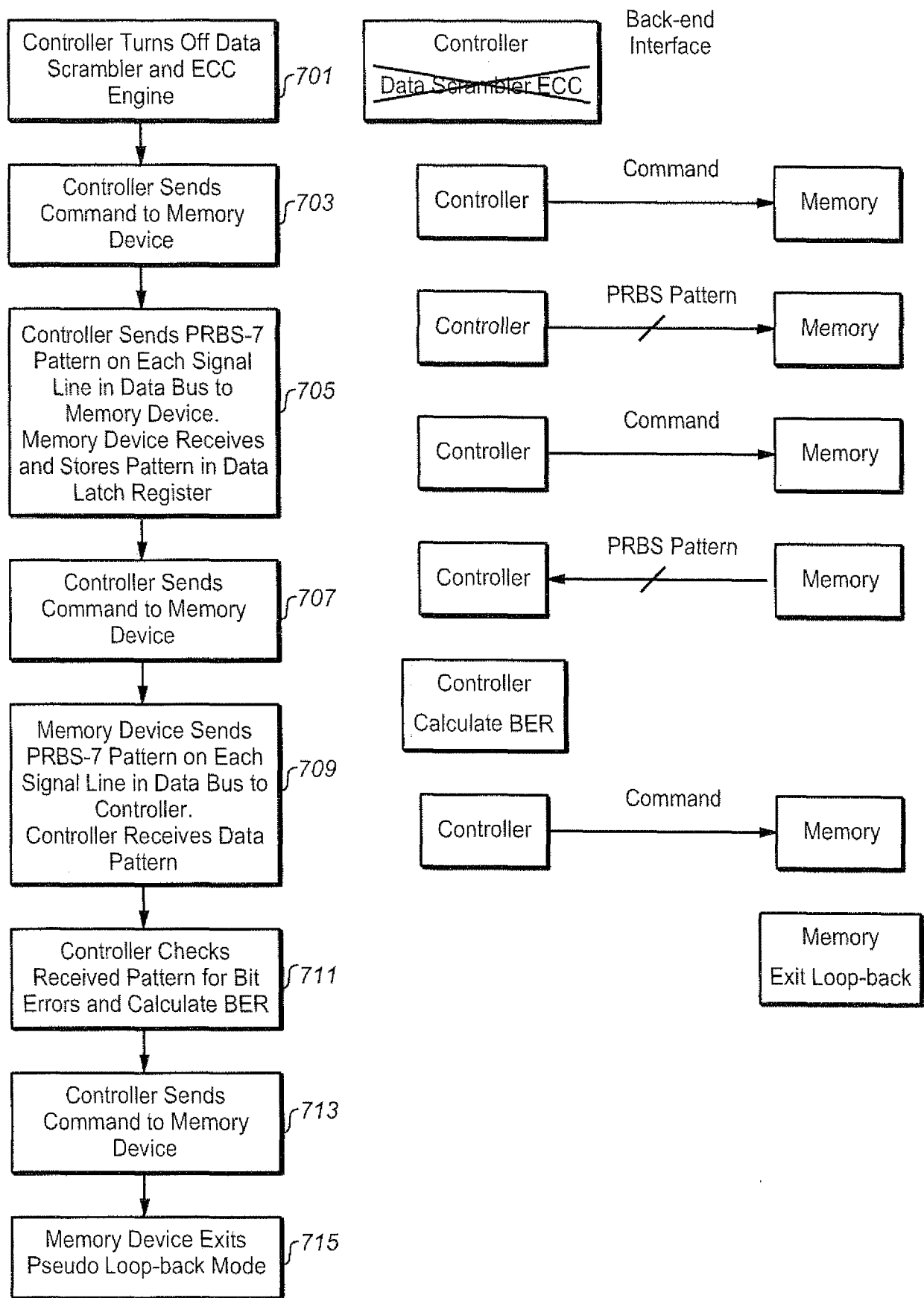


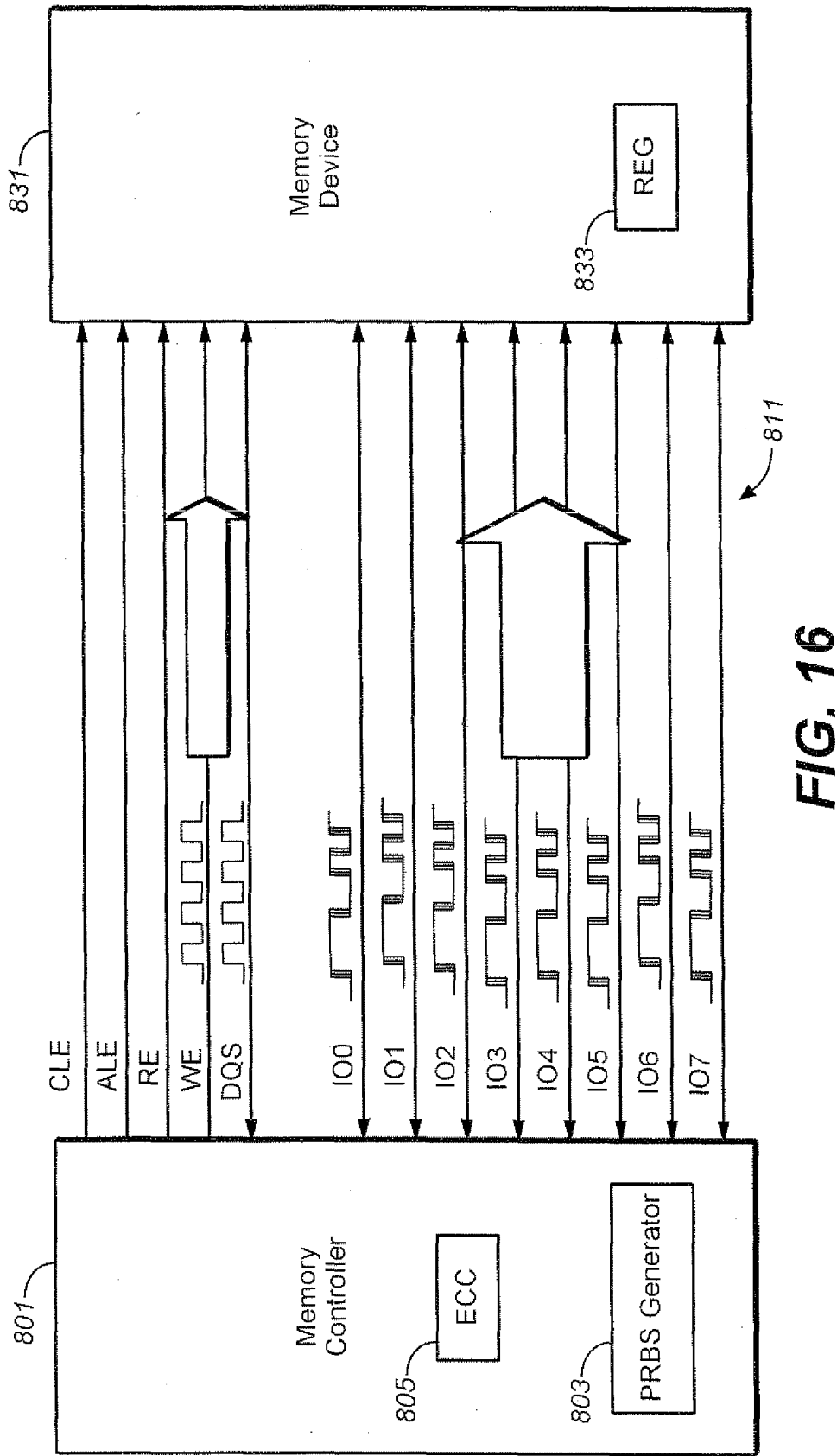
**FIG. 13**

**FIG. 14**



15 / 19

**FIG. 15**



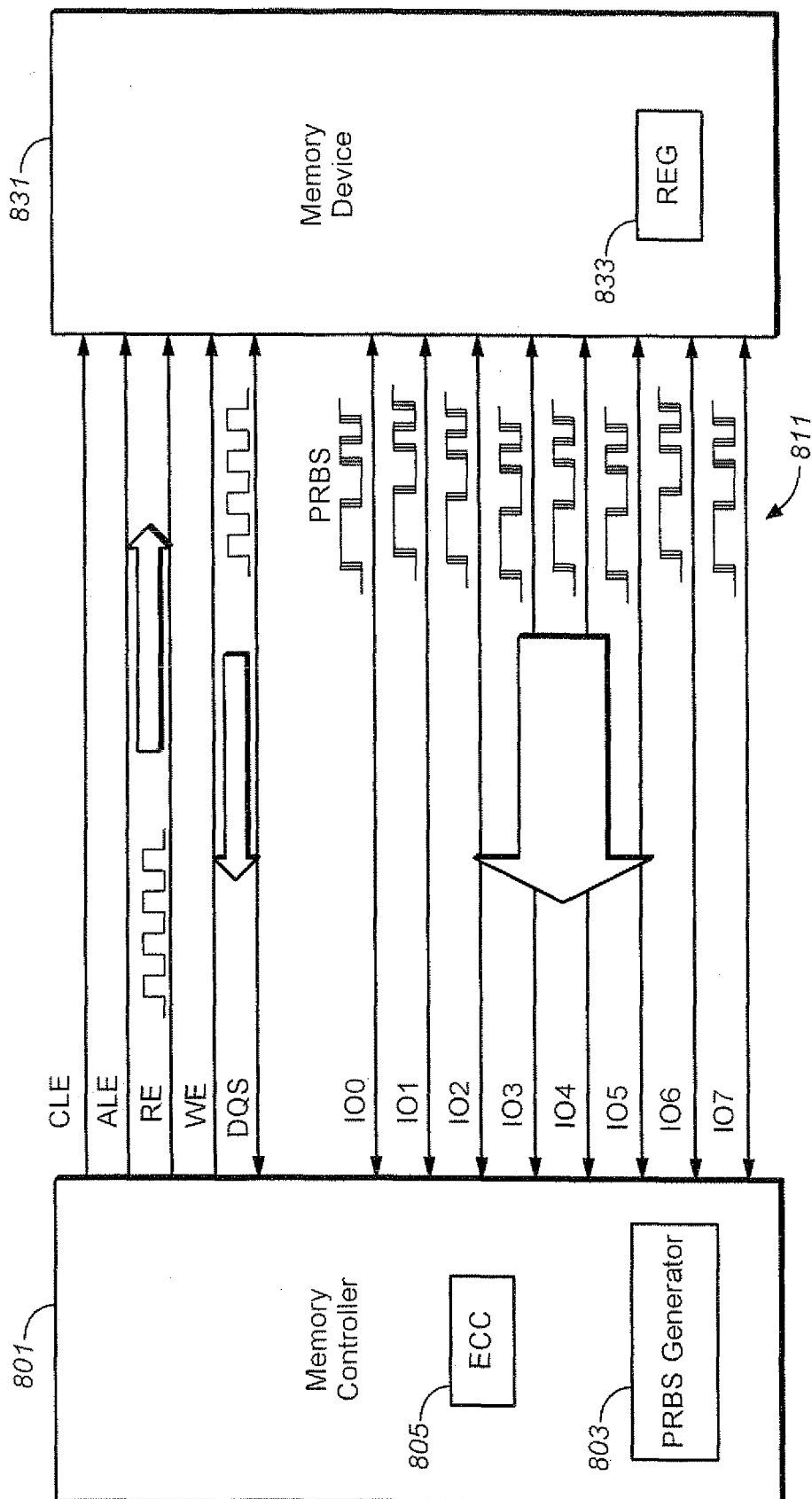
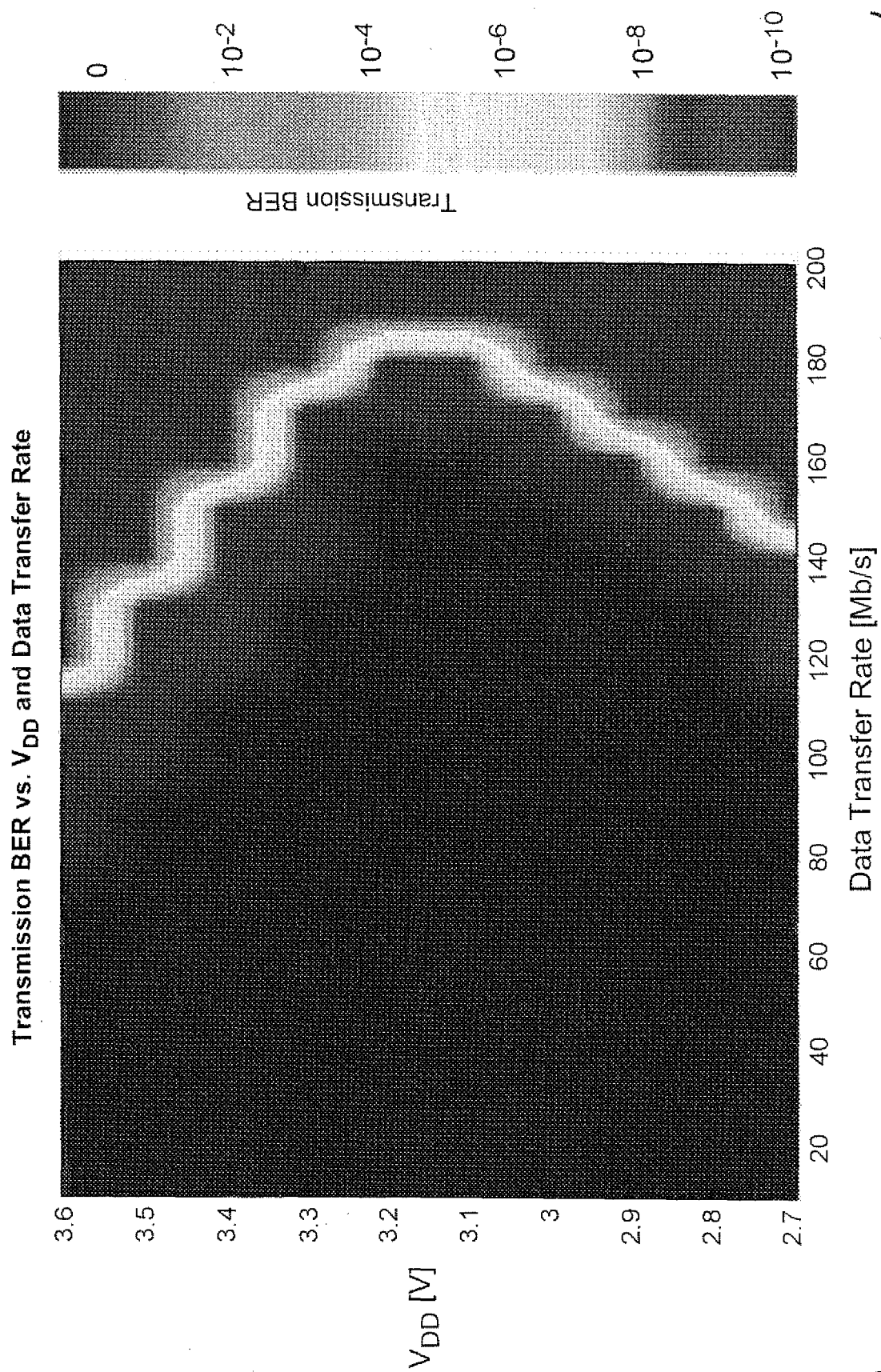
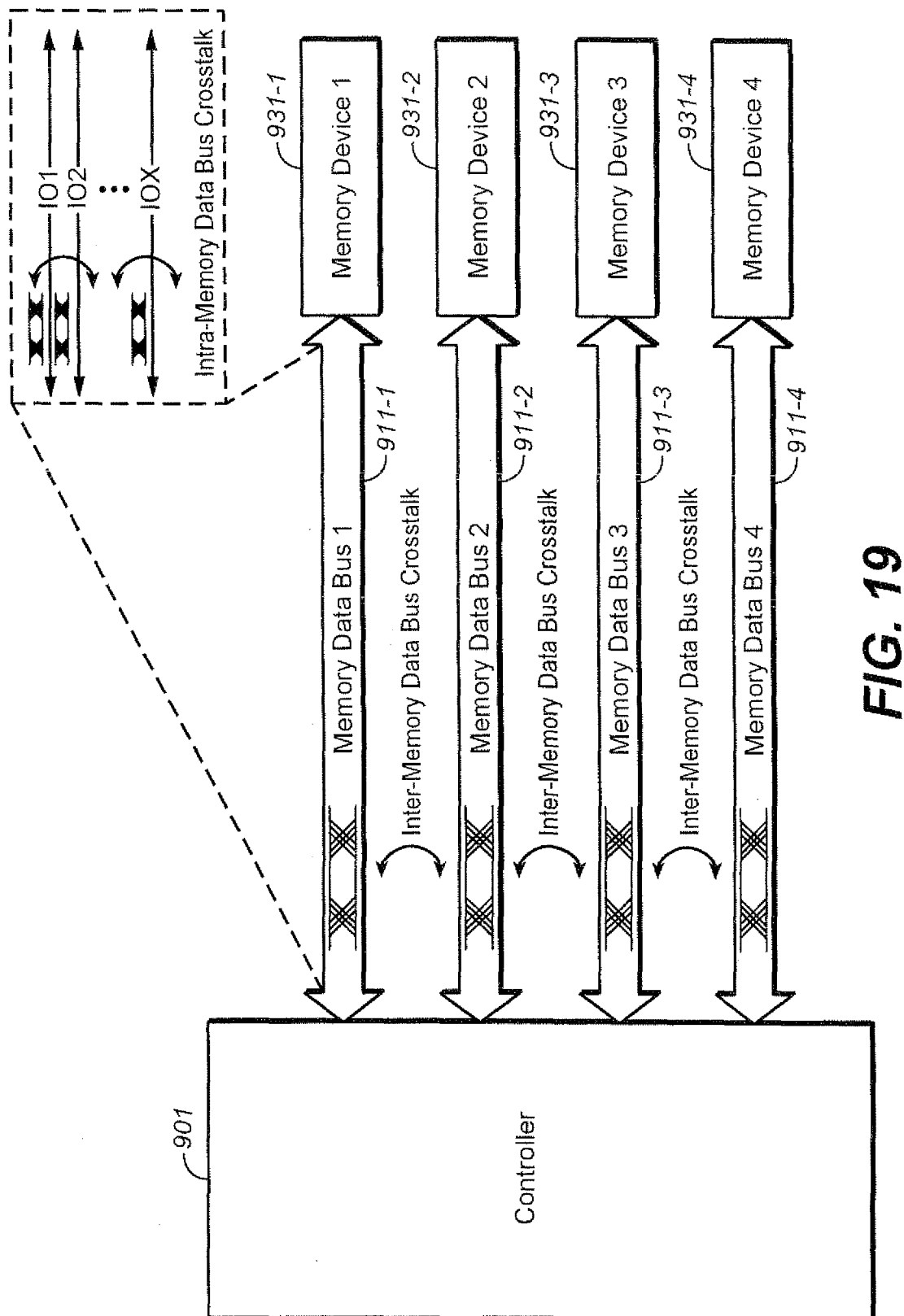


FIG. 17

**FIG. 18**



# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2011/043648

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. G06F11/10 G06F11/267  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
G06F G11C

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EP0-Internal

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 2009/271678 A1 (SCHNEIDER ANDREAS [DE] ET AL) 29 October 2009 (2009-10-29) the whole document	1-19, 30-64 20-29, 65,66
X A	----- US 2008/112255 A1 (NYGREN AARON JOHN [DE] ET AL) 15 May 2008 (2008-05-15) page 1, paragraph 1 - paragraph 5 page 2, paragraph 24 - page 4, paragraph 40; figures 1-3	1-43,65, 66 44-64
A	----- US 2007/271495 A1 (SHAEFFER IAN [US] ET AL) 22 November 2007 (2007-11-22) the whole document	1-66
A	----- US 2007/089035 A1 (ALEXANDER JAMES W [US] ET AL) 19 April 2007 (2007-04-19) the whole document	1-66

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance  
"E" earlier document but published on or after the international filing date  
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
"O" document referring to an oral disclosure, use, exhibition or other means  
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  
"&" document member of the same patent family

Date of the actual completion of the international search

18 November 2011

Date of mailing of the international search report

30/11/2011

Name and mailing address of the ISA/  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Bauer, Regine



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2011/043648

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009271678 A1	29-10-2009	NONE	
-----			
US 2008112255 A1	15-05-2008	DE 102007053978 A1	21-05-2008
		US 2008112255 A1	15-05-2008
-----			
US 2007271495 A1	22-11-2007	EP 2024834 A2	18-02-2009
		JP 2009537899 A	29-10-2009
		KR 20090028538 A	18-03-2009
		US 2007271495 A1	22-11-2007
		US 2008163007 A1	03-07-2008
		WO 2007136655 A2	29-11-2007
-----			
US 2007089035 A1	19-04-2007	NONE	
-----			