



(51) International Patent Classification:

G06F 11/10 (2006.01) G06F 12/02 (2006.01)

G06F 12/00 (2006.01) G06F 3/00 (2006.01)

(21) International Application Number:

PCT/US2021/016201

(22) International Filing Date:

02 February 2021 (02.02.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

16/893,011 04 June 2020 (04.06.2020) US

(71) Applicant: WESTERN DIGITAL TECHNOLOGIES, INC. [US/US]; 5601 Great Oaks Parkway, San Jose, CA 95119 (US).

(72) Inventors: MUTHIAH, Ramanathan; c/o Western Digital Technologies, Inc., 5601 Great Oaks Parkway, San Jose, CA 95119 (US). SHETTY, Abhishek; c/o Western Digital Technologies, Inc., 5601 Great Oaks Parkway, San Jose, CA 95119 (US).

(74) Agent: HETZ, Joseph, F.; Brinks Gilson & Lione, P.O. Box 10087, Chicago, IL 60610 (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: STORAGE SYSTEM AND METHOD FOR CRASH ANALYSIS

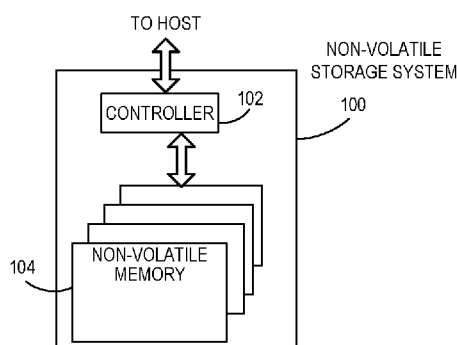


FIG. 1A

(57) Abstract: A storage system and method for crash analysis are provided. In one embodiment, a storage system comprises a memory and a controller. The controller is configured to store a crash log in a buffer in the controller; determine whether the storage system has crashed; and in response to determining that the storage system has crashed, send the crash log from the buffer in the controller to a host in communication with the storage system. Other embodiments are provided.

Storage System and Method for Crash Analysis

Background

[0001] A storage system can contain a controller and a non-volatile memory for storing user data, such as documents, images, videos, and music. The controller can store diagnostic logs in the memory for later analysis if the storage system crashes or fails.

Brief Description of the Drawings

[0002] Figure 1A is a block diagram of a non-volatile storage system of an embodiment.

[0003] Figure 1B is a block diagram illustrating a storage module of an embodiment.

[0004] Figure 1C is a block diagram illustrating a hierarchical storage system of an embodiment.

[0005] Figure 2A is a block diagram illustrating components of the controller of the non-volatile storage system illustrated in Figure 1A according to an embodiment.

[0006] Figure 2B is a block diagram illustrating components of the non-volatile memory storage system illustrated in Figure 1A according to an embodiment.

[0007] Figure 3 is a block diagram of a host and storage system of an embodiment.

[0008] Figure 4 is a block diagram of a host and storage system of an embodiment that can be used for crash analysis.

[0009] Figure 5 is a block diagram of two storage systems of an embodiment that can be used for crash analysis.

[0010] Figure 6 is a flow chart of a method performed by a storage system of an embodiment for crash analysis.

[0011] Figure 7 is a flow chart of a method performed by a host of an embodiment for crash analysis.

Detailed Description

[0012] Overview

[0013] By way of introduction, the below embodiments relate to a storage system and method for crash analysis. In one embodiment, a storage system is provided comprising a memory and a controller. The controller is configured to store a crash log in a buffer in

the controller; determine whether the storage system has crashed; and in response to determining that the storage system has crashed, send the crash log from the buffer in the controller to a host in communication with the storage system.

[0014] In some embodiments, the crash log is sent to the host for storage in a host memory buffer in the host.

[0015] In some embodiments, the storage system further comprises a host data path configured to route data between the memory and the host, and wherein the crash log is sent from the buffer in the controller to the host outside of the host data path.

[0016] In some embodiments, the host data path comprises one or more of the following components: an error correction code module, a memory interface module, and a host interface module.

[0017] In some embodiments, the crash log comprises a pre-determined data structure known to the host and the storage system.

[0018] In some embodiments, the pre-determined data structure is communicated between the host and the storage system during initialization of the storage system.

[0019] In some embodiments, a watchdog module in the controller is configured to send the crash log from the buffer in the controller to the host in response to determining that the storage system has crashed.

[0020] In some embodiments, the memory comprises a three-dimensional memory.

[0021] In another embodiment, a method is provided that is performed in a first storage system in communication with a second storage system. The method comprises: storing diagnostic data in a controller memory buffer in the first storage system; and in response to determining that the first storage system has failed, sending the diagnostic data from the controller memory buffer in the first storage system to a second storage system in communication with the first storage system.

[0022] In some embodiments, the diagnostic data is sent to the second storage device for storage in a controller memory buffer in the second storage system.

[0023] In some embodiments, the first storage system comprises a host data path configured to route data to and from the memory, and wherein the diagnostic data is sent from the controller memory buffer in the first storage system to the second storage system outside of the host data path.

[0024] In some embodiments, the host data path comprises one or more of the following components: an error correction code module, a memory interface module, and a host interface module.

[0025] In some embodiments, the diagnostic data comprises a pre-determined data structure known to first and second storage systems.

[0026] In some embodiments, a watchdog module in the controller in the first storage system is configured to send the diagnostic data from the controller memory buffer to the second storage system in response to determining that the first storage system has failed.

[0027] In some embodiments, the diagnostic data is sent from the controller memory buffer in the first storage system to the second storage system via a direct memory access operation.

[0028] In another embodiment, a storage system is provided comprising: a memory; means for storing a crash log in a controller memory buffer in the storage system; and means for, in response to determining that the storage system has crashed, sending the crash log from the controller memory buffer to an external device for analysis.

[0029] In some embodiments, the external device comprises a host.

[0030] In some embodiments, the external device comprises a second storage system.

[0031] In some embodiments, the crash log is sent from the controller memory buffer to the external device outside of a host data path to the memory.

[0032] In some embodiments, the crash log comprises a pre-determined data structure known to the host and the external device.

[0033] Other embodiments are possible, and each of the embodiments can be used alone or together in combination. Accordingly, various embodiments will now be described with reference to the attached drawings.

[0034] Embodiments

[0035] Storage systems suitable for use in implementing aspects of these embodiments are shown in Figures 1A-1C. Figure 1A is a block diagram illustrating a non-volatile storage system 100 according to an embodiment of the subject matter described herein. Referring to Figure 1A, non-volatile storage system 100 includes a controller 102 and non-volatile memory that may be made up of one or more non-volatile memory die 104. As used herein, the term die refers to the collection of non-volatile memory cells, and

associated circuitry for managing the physical operation of those non-volatile memory cells, that are formed on a single semiconductor substrate. Controller 102 interfaces with a host system and transmits command sequences for read, program, and erase operations to non-volatile memory die 104.

[0036] The controller 102 (which may be a non-volatile memory controller (e.g., a flash, resistive random-access memory (ReRAM), phase-change memory (PCM), or magnetoresistive random-access memory (MRAM) controller)) can take the form of processing circuitry, a microprocessor or processor, and a computer-readable medium that stores computer-readable program code (e.g., firmware) executable by the (micro)processor, logic gates, switches, an application specific integrated circuit (ASIC), a programmable logic controller, and an embedded microcontroller, for example. The controller 102 can be configured with hardware and/or firmware to perform the various functions described below and shown in the flow diagrams. Also, some of the components shown as being internal to the controller can also be stored external to the controller, and other components can be used. Additionally, the phrase “operatively in communication with” could mean directly in communication with or indirectly (wired or wireless) in communication with through one or more components, which may or may not be shown or described herein.

[0037] As used herein, a non-volatile memory controller is a device that manages data stored on non-volatile memory and communicates with a host, such as a computer or electronic device. A non-volatile memory controller can have various functionality in addition to the specific functionality described herein. For example, the non-volatile memory controller can format the non-volatile memory to ensure the memory is operating properly, map out bad non-volatile memory cells, and allocate spare cells to be substituted for future failed cells. Some part of the spare cells can be used to hold firmware to operate the non-volatile memory controller and implement other features. In operation, when a host needs to read data from or write data to the non-volatile memory, it can communicate with the non-volatile memory controller. If the host provides a logical address to which data is to be read/written, the non-volatile memory controller can convert the logical address received from the host to a physical address in the non-volatile memory. (Alternatively, the host can provide the physical address.) The non-

volatile memory controller can also perform various memory management functions, such as, but not limited to, wear leveling (distributing writes to avoid wearing out specific blocks of memory that would otherwise be repeatedly written to) and garbage collection (after a block is full, moving only the valid pages of data to a new block, so the full block can be erased and reused).

[0038] Non-volatile memory die 104 may include any suitable non-volatile storage medium, including resistive random-access memory (ReRAM), magnetoresistive random-access memory (MRAM), phase-change memory (PCM), NAND flash memory cells and/or NOR flash memory cells. The memory cells can take the form of solid-state (e.g., flash) memory cells and can be one-time programmable, few-time programmable, or many-time programmable. The memory cells can also be single-level cells (SLC), multiple-level cells (MLC), triple-level cells (TLC), or use other memory cell level technologies, now known or later developed. Also, the memory cells can be fabricated in a two-dimensional or three-dimensional fashion.

[0039] The interface between controller 102 and non-volatile memory die 104 may be any suitable flash interface, such as Toggle Mode 200, 400, or 800. In one embodiment, storage system 100 may be a card based system, such as a secure digital (SD) or a micro secure digital (micro-SD) card. In an alternate embodiment, storage system 100 may be part of an embedded storage system.

[0040] Although, in the example illustrated in Figure 1A, non-volatile storage system 100 (sometimes referred to herein as a storage module) includes a single channel between controller 102 and non-volatile memory die 104, the subject matter described herein is not limited to having a single memory channel. For example, in some storage system architectures (such as the ones shown in Figures 1B and 1C), 2, 4, 8 or more memory channels may exist between the controller and the memory device, depending on controller capabilities. In any of the embodiments described herein, more than a single channel may exist between the controller and the memory die, even if a single channel is shown in the drawings.

[0041] Figure 1B illustrates a storage module 200 that includes plural non-volatile storage systems 100. As such, storage module 200 may include a storage controller 202 that interfaces with a host and with storage system 204, which includes a plurality of non-

volatile storage systems 100. The interface between storage controller 202 and non-volatile storage systems 100 may be a bus interface, such as a serial advanced technology attachment (SATA), peripheral component interconnect express (PCIe) interface, or double-data-rate (DDR) interface. Storage module 200, in one embodiment, may be a solid state drive (SSD), or non-volatile dual in-line memory module (NVDIMM), such as found in server PC or portable computing devices, such as laptop computers, and tablet computers.

[0042] Figure 1C is a block diagram illustrating a hierarchical storage system. A hierarchical storage system 250 includes a plurality of storage controllers 202, each of which controls a respective storage system 204. Host systems 252 may access memories within the storage system via a bus interface. In one embodiment, the bus interface may be a Non-Volatile Memory Express (NVMe) or fiber channel over Ethernet (FCoE) interface. In one embodiment, the system illustrated in Figure 1C may be a rack mountable mass storage system that is accessible by multiple host computers, such as would be found in a data center or other location where mass storage is needed.

[0043] Figure 2A is a block diagram illustrating components of controller 102 in more detail. Controller 102 includes a front end module 108 that interfaces with a host, a back end module 110 that interfaces with the one or more non-volatile memory die 104, and various other modules that perform functions which will now be described in detail. For example, as will be discussed below, in one embodiment, the controller 102 comprises a watchdog module 111, which can be implemented in hardware and/or software/firmware, that can be used in crash/debug/diagnostic analysis in conjunction with a controller memory buffer (CMB) 112 or other form of memory in the controller 102. A module may take the form of a packaged functional hardware unit designed for use with other components, a portion of a program code (e.g., software or firmware) executable by a (micro)processor or processing circuitry that usually performs a particular function of related functions, or a self-contained hardware or software component that interfaces with a larger system, for example.

[0044] Referring again to modules of the controller 102, a buffer manager/bus controller 114 manages buffers in random access memory (RAM) 116 and controls the internal bus arbitration of controller 102. A read only memory (ROM) 118 stores system

boot code. Although illustrated in Figure 2A as located separately from the controller 102, in other embodiments one or both of the RAM 116 and ROM 118 may be located within the controller. In yet other embodiments, portions of RAM and ROM may be located both within the controller 102 and outside the controller.

[0045] Front end module 108 includes a host interface 120 and a physical layer interface (PHY) 122 that provide the electrical interface with the host or next level storage controller. The choice of the type of host interface 120 can depend on the type of memory being used. Examples of host interfaces 120 include, but are not limited to, SATA, SATA Express, serially attached small computer system interface (SAS), Fibre Channel, universal serial bus (USB), PCIe, and NVMe. The host interface 120 typically facilitates transfer for data, control signals, and timing signals.

[0046] Back end module 110 includes an error correction code (ECC) engine 124 that encodes the data bytes received from the host, and decodes and error corrects the data bytes read from the non-volatile memory. A command sequencer 126 generates command sequences, such as program and erase command sequences, to be transmitted to non-volatile memory die 104. A RAID (Redundant Array of Independent Drives) module 128 manages generation of RAID parity and recovery of failed data. The RAID parity may be used as an additional level of integrity protection for the data being written into the memory device 104. In some cases, the RAID module 128 may be a part of the ECC engine 124. A memory interface 130 provides the command sequences to non-volatile memory die 104 and receives status information from non-volatile memory die 104. In one embodiment, memory interface 130 may be a double data rate (DDR) interface, such as a Toggle Mode 200, 400, or 800 interface. A flash control layer 132 controls the overall operation of back end module 110.

[0047] The storage system 100 also includes other discrete components 140, such as external electrical interfaces, external RAM, resistors, capacitors, or other components that may interface with controller 102. In alternative embodiments, one or more of the physical layer interface 122, RAID module 128, media management layer 138 and buffer management/bus controller 114 are optional components that are not necessary in the controller 102.

[0048] Figure 2B is a block diagram illustrating components of non-volatile memory die 104 in more detail. Non-volatile memory die 104 includes peripheral circuitry 141 and non-volatile memory array 142. Non-volatile memory array 142 includes the non-volatile memory cells used to store data. The non-volatile memory cells may be any suitable non-volatile memory cells, including ReRAM, MRAM, PCM, NAND flash memory cells and/or NOR flash memory cells in a two dimensional and/or three dimensional configuration. Non-volatile memory die 104 further includes a data cache 156 that caches data. Peripheral circuitry 141 includes a state machine 152 that provides status information to the controller 102.

[0049] Returning again to Figure 2A, the flash control layer 132 (which will be referred to herein as the flash translation layer (FTL) or, more generally, the “media management layer,” as the memory may not be flash) handles flash errors and interfaces with the host. In particular, the FTL, which may be an algorithm in firmware, is responsible for the internals of memory management and translates writes from the host into writes to the memory 104. The FTL may be needed because the memory 104 may have limited endurance, may only be written in multiples of pages, and/or may not be written unless it is erased as a block. The FTL understands these potential limitations of the memory 104, which may not be visible to the host. Accordingly, the FTL attempts to translate the writes from host into writes into the memory 104.

[0050] The FTL may include a logical-to-physical address (L2P) map and allotted cache memory. In this way, the FTL translates logical block addresses (“LBAs”) from the host to physical addresses in the memory 104. The FTL can include other features, such as, but not limited to, power-off recovery (so that the data structures of the FTL can be recovered in the event of a sudden power loss) and wear leveling (so that the wear across memory blocks is even to prevent certain blocks from excessive wear, which would result in a greater chance of failure).

[0051] Turning again to the drawings, Figure 3 is a block diagram of a host 300 and storage system 100 of an embodiment. The host 300 can take any suitable form, including, but not limited to, a computer, a mobile phone, a tablet, a wearable device, a digital video recorder, a surveillance system, etc. The host 300 in this embodiment (here, a computing device) comprises a processor 330 and a memory 340. In one embodiment,

computer-readable program code stored in the host memory 340 configures the host processor 330 to read data from and write data to the storage system 100, as well as perform error analysis on crash logs from the storage system 100.

[0052] As mentioned above, a controller of a storage system can store a crash log in the same non-volatile memory that is used for storing user data, such as documents, images, videos, and music. A crash log can contain diagnostic information that can assist in the debugging process to analyze why the storage system crashed (e.g., to determine the root cause of what caused the bug). Examples of diagnostic information can include, but are not limited to, information about the state of the storage system, the call stack, log data, firmware state machine information, and memory status.

[0053] In theory, the crash log can be analyzed by a host or other device if the storage system crashes/fails to determine the cause of the problem and potentially fix it to avoid future crashes. Since the crash log is stored in the non-volatile memory, the host or other device would need to access the memory. Reading the crash logs from the memory can be a tedious or impossible task because a crash typically means that the firmware in the storage system will not respond to host commands. In such a situation, memory access is typically not available until the storage system is reinitialized. Unfortunately, reinitializing the storage system can result in the loss of state machine and register values, which would hinder the ability to fully analyze the error and diagnostic data in the logs. So, a crash can result in a fatal system error and make the storage system non-responsive during that run or power cycle.

[0054] The following embodiments can be used to address this issue. In one embodiment, the storage system (sometimes referred to herein as “the device”) 100 stores the crash log in a memory in the controller 102 (e.g., the controller memory buffer (CMB) 112), and the controller 102 has a watchdog module 111 that sends the crash log from the CMB 112 to an external device (e.g., the host 300 or another storage system) for analysis. Because the crash logs are not stored in the memory 104, the inability of the host 300 or other device to access the memory 104 does not prevent accessing the crash logs. Also, as other modules (e.g., an error correction code module, a memory interface module, and a host interface module) in the host data path to and from the memory 104

may not be functional after the crash, the watchdog module 111 can send the crash logs to the external device outside of the host data path.

[0055] The following paragraphs provide example implementations of embodiments. It should be understood that these are merely examples, and other implementations are possible.

[0056] Turning again to the drawings, Figure 4 is a block diagram that illustrates an embodiment in which the storage system 100 stores the crash log in a controller memory buffer (CMB) 420, and a watchdog module in the storage system 100 sends debug data from the crash log to the host 300 (in a channel other than the standard control/data path) after a crash is detected. In this embodiment, the host 300 stores the crash log in a host memory buffer (HMB) 410, which is memory that is designated for used by the storage system 100 in the NVMe specification. So, the controller 102 of the storage system 100 can issue read and write commands to the HMB 410 as if it were its own memory (with potentially some restrictions). Because the HMB 410 is outside of the failing device and is an existing interface in the NVMe specification, the HMB 410 can be used to remove the dependency on the failing device for diagnostic data. It should be noted HMBs and CMBs have been used in prior storage system as a “luxury” to improve performance. However, in one embodiment, the use of these components would be more than just a “luxury,” as they would store data that would otherwise be difficult or impossible to access from the memory 104.

[0057] Figure 6 is a flow chart 600 that illustrates the operation of this embodiment. As shown in Figure 6, upon initialization of the storage system 100 (or at some time prior to a crash), the storage system 100 and the host 300 share a predetermined data structure for crash analysis (act 610). During operation of the storage system 100, the controller 102 can log all data required to evaluate the reason for the crash into the CMB 420. Next, the controller 102 (e.g., the watchdog module 112) determines if the storage system 100 has crashed (act 620). Any suitable technique can be used to determine if the storage system 100 crashed/failed. For example, the controller 102 of the storage system 100 can receive an error signal from one or more components in the storage system 100 indicating a failure of that component, or the controller 102 can infer a failure by not receiving expected information in an expected amount of time. Other failure detection techniques

can be used. If the storage system 100 has not crashed, the controller 102 continues with its normal operations (act 630). However, if the storage system 100 has crashed, the watchdog module 112 moves the diagnostic data from the CMB 400 to the HMB 410 according to the predetermined data structure (act 640)

[0058] Figure 7 is a flow chart 700 that illustrates operations on the host 300 side. As shown in Figure 7, when the host 300 does not get a response from the storage system 100 to a command (e.g., a memory read/write command) (act 710), the host 300 determines if the storage system 100 has crashed (act 720). For example, the host 300 can receive an error signal from one or more components in the storage system 100 indicating a failure of that component, or the host 300 can infer a failure. If the host 300 does not detect a failure, it continues to wait for the storage system's response (act 730). However, if the host 300 detects a crash, it can assume that the watchdog module 111 in the storage system 100 sent the crash log to the HMB 410 and will analyze the crash log stored therein to analyze the error (act 740). For example, the host 300 can attempt to trigger certain debug options in the storage system 100 (e.g., using an offline parser). So, unlike with prior storage systems, the host 300 can parse the crash logs without having to depend on access to the memory 104 of the storage system 100. Hence, the storage system 100 does not need to be re-initialized to access the crash log, perform analysis of the diagnostic data, or perform specific debug operations (such as live monitoring some registers (e.g., by attaching a debugger)) that may not have been considered in a generic dump, if even possible, in prior storage systems. Because these embodiments keep the storage system's state machine intact, further probing is possible of controller and memory parameters. To establish memory states, tools can be used to probe failed device registers or even a portion of static RAM (SRAM) in the storage system. This makes identifying a failure simpler by not requiring the device be reinitialized.

[0059] The embodiments are usable with the storage systems 100 connected over fabrics as well. The debug data, along with the undisturbed device and memory, can assist faster determination of the root cause of the bug. The method may be desired by original equipment manufacturer (OEM) customers who may have to pass on the debug information generated in their site.

[0060] Further, as mentioned above, instead of sending the crash log to the host 300, the storage system 100 can send the crash log to another device for analysis. This alternative is shown in Figure 5, wherein the other device is a second storage system 500, and the crash log is stored in the CMB 520 in the second storage system 500. The crash log can be sent to the CMB 520 in the second storage system 500 via a direct memory access (DMA) operation. The functioning of this embodiment is similar to the one discussed above, with a tool/parser running in the second storage system 500 performing necessary analysis to assist in fixing the failure of the first storage system 100. Again, it is not necessary to initialize the failing device, thereby keeping its state machine intact for further probing controller and memory parameters.

[0061] Finally, as mentioned above, any suitable type of memory can be used. Semiconductor memory devices include volatile memory devices, such as dynamic random access memory (“DRAM”) or static random access memory (“SRAM”) devices, non-volatile memory devices, such as resistive random access memory (“ReRAM”), electrically erasable programmable read only memory (“EEPROM”), flash memory (which can also be considered a subset of EEPROM), ferroelectric random access memory (“FRAM”), and magnetoresistive random access memory (“MRAM”), and other semiconductor elements capable of storing information. Each type of memory device may have different configurations. For example, flash memory devices may be configured in a NAND or a NOR configuration.

[0062] The memory devices can be formed from passive and/or active elements, in any combinations. By way of non-limiting example, passive semiconductor memory elements include ReRAM device elements, which in some embodiments include a resistivity switching storage element, such as an anti-fuse, phase change material, etc., and optionally a steering element, such as a diode, etc. Further by way of non-limiting example, active semiconductor memory elements include EEPROM and flash memory device elements, which in some embodiments include elements containing a charge storage region, such as a floating gate, conductive nanoparticles, or a charge storage dielectric material.

[0063] Multiple memory elements may be configured so that they are connected in series or so that each element is individually accessible. By way of non-limiting

example, flash memory devices in a NAND configuration (NAND memory) typically contain memory elements connected in series. A NAND memory array may be configured so that the array is composed of multiple strings of memory in which a string is composed of multiple memory elements sharing a single bit line and accessed as a group. Alternatively, memory elements may be configured so that each element is individually accessible, e.g., a NOR memory array. NAND and NOR memory configurations are examples, and memory elements may be otherwise configured.

[0064] The semiconductor memory elements located within and/or over a substrate may be arranged in two or three dimensions, such as a two dimensional memory structure or a three dimensional memory structure.

[0065] In a two dimensional memory structure, the semiconductor memory elements are arranged in a single plane or a single memory device level. Typically, in a two dimensional memory structure, memory elements are arranged in a plane (e.g., in an x-z direction plane) which extends substantially parallel to a major surface of a substrate that supports the memory elements. The substrate may be a wafer over or in which the layer of the memory elements are formed or it may be a carrier substrate which is attached to the memory elements after they are formed. As a non-limiting example, the substrate may include a semiconductor such as silicon.

[0066] The memory elements may be arranged in the single memory device level in an ordered array, such as in a plurality of rows and/or columns. However, the memory elements may be arrayed in non-regular or non-orthogonal configurations. The memory elements may each have two or more electrodes or contact lines, such as bit lines and wordlines.

[0067] A three dimensional memory array is arranged so that memory elements occupy multiple planes or multiple memory device levels, thereby forming a structure in three dimensions (i.e., in the x, y and z directions, where the y direction is substantially perpendicular and the x and z directions are substantially parallel to the major surface of the substrate).

[0068] As a non-limiting example, a three dimensional memory structure may be vertically arranged as a stack of multiple two dimensional memory device levels. As another non-limiting example, a three dimensional memory array may be arranged as

multiple vertical columns (e.g., columns extending substantially perpendicular to the major surface of the substrate, i.e., in the y direction) with each column having multiple memory elements in each column. The columns may be arranged in a two dimensional configuration, e.g., in an x-z plane, resulting in a three dimensional arrangement of memory elements with elements on multiple vertically stacked memory planes. Other configurations of memory elements in three dimensions can also constitute a three dimensional memory array.

[0069] By way of non-limiting example, in a three dimensional NAND memory array, the memory elements may be coupled together to form a NAND string within a single horizontal (e.g., x-z) memory device levels. Alternatively, the memory elements may be coupled together to form a vertical NAND string that traverses across multiple horizontal memory device levels. Other three dimensional configurations can be envisioned wherein some NAND strings contain memory elements in a single memory level while other strings contain memory elements which span through multiple memory levels. Three dimensional memory arrays may also be designed in a NOR configuration and in a ReRAM configuration.

[0070] Typically, in a monolithic three dimensional memory array, one or more memory device levels are formed above a single substrate. Optionally, the monolithic three dimensional memory array may also have one or more memory layers at least partially within the single substrate. As a non-limiting example, the substrate may include a semiconductor such as silicon. In a monolithic three dimensional array, the layers constituting each memory device level of the array are typically formed on the layers of the underlying memory device levels of the array. However, layers of adjacent memory device levels of a monolithic three dimensional memory array may be shared or have intervening layers between memory device levels.

[0071] Then again, two dimensional arrays may be formed separately and then packaged together to form a non-monolithic memory device having multiple layers of memory. For example, non-monolithic stacked memories can be constructed by forming memory levels on separate substrates and then stacking the memory levels atop each other. The substrates may be thinned or removed from the memory device levels before stacking, but as the memory device levels are initially formed over separate substrates,

the resulting memory arrays are not monolithic three dimensional memory arrays. Further, multiple two dimensional memory arrays or three dimensional memory arrays (monolithic or non-monolithic) may be formed on separate chips and then packaged together to form a stacked-chip memory device.

[0072] Associated circuitry is typically required for operation of the memory elements and for communication with the memory elements. As non-limiting examples, memory devices may have circuitry used for controlling and driving memory elements to accomplish functions such as programming and reading. This associated circuitry may be on the same substrate as the memory elements and/or on a separate substrate. For example, a controller for memory read-write operations may be located on a separate controller chip and/or on the same substrate as the memory elements.

[0073] One of skill in the art will recognize that this invention is not limited to the two dimensional and three dimensional structures described but cover all relevant memory structures within the spirit and scope of the invention as described herein and as understood by one of skill in the art.

[0074] It is intended that the foregoing detailed description be understood as an illustration of selected forms that the invention can take and not as a definition of the invention. It is only the following claims, including all equivalents, that are intended to define the scope of the claimed invention. Finally, it should be noted that any aspect of any of the embodiments described herein can be used alone or in combination with one another.

What is claimed is:

1. A storage system comprising:
a memory; and
a controller configured to:
store a crash log in a buffer in the controller;
determine whether the storage system has crashed; and
in response to determining that the storage system has crashed, send the
crash log from the buffer in the controller to a host in communication with the
storage system.
2. The storage system of Claim 1, wherein the crash log is sent to the host for
storage in a host memory buffer in the host.
3. The storage system of Claim 1, further comprising a host data path configured to
route data between the memory and the host, and wherein the crash log is sent from the
buffer in the controller to the host outside of the host data path.
4. The storage system of Claim 3, wherein the host data path comprises one or more
of the following components: an error correction code module, a memory interface
module, and a host interface module.
5. The storage system of Claim 1, wherein the crash log comprises a pre-determined
data structure known to the host and the storage system.

6. The storage system of Claim 5, wherein the pre-determined data structure is communicated between the host and the storage system during initialization of the storage system.
7. The storage system of Claim 1, wherein a watchdog module in the controller is configured to send the crash log from the buffer in the controller to the host in response to determining that the storage system has crashed.
8. The storage system of Claim 1, wherein the memory comprises a three-dimensional memory.
9. A method comprising:
performing the following in a first storage system in communication with a second storage system:
storing diagnostic data in a controller memory buffer in the first storage system; and
in response to determining that the first storage system has failed, sending the diagnostic data from the controller memory buffer in the first storage system to a second storage system in communication with the first storage system.
10. The method of Claim 9, wherein the diagnostic data is sent to the second storage device for storage in a controller memory buffer in the second storage system.

11. The method of Claim 9, wherein the first storage system comprises a host data path configured to route data to and from the memory, and wherein the diagnostic data is sent from the controller memory buffer in the first storage system to the second storage system outside of the host data path.
12. The method of Claim 11, wherein the host data path comprises one or more of the following components: an error correction code module, a memory interface module, and a host interface module.
13. The method of Claim 9, wherein the diagnostic data comprises a pre-determined data structure known to first and second storage systems.
14. The method of Claim 9, wherein a watchdog module in the controller in the first storage system is configured to send the diagnostic data from the controller memory buffer to the second storage system in response to determining that the first storage system has failed.
15. The method of Claim 9, wherein the diagnostic data is sent from the controller memory buffer in the first storage system to the second storage system via a direct memory access operation.

16. A storage system comprising:
- a memory;
- means for storing a crash log in a controller memory buffer in the storage system;
- and
- means for, in response to determining that the storage system has crashed, sending the crash log from the controller memory buffer to an external device for analysis.
17. The storage system of Claim 16, wherein the external device comprises a host.
18. The storage system of Claim 16, wherein the external device comprises a second storage system.
19. The storage system of Claim 16, wherein the crash log is sent from the controller memory buffer to the external device outside of a host data path to the memory.
20. The storage system of Claim 16, wherein the crash log comprises a pre-determined data structure known to the host and the external device.

1/8

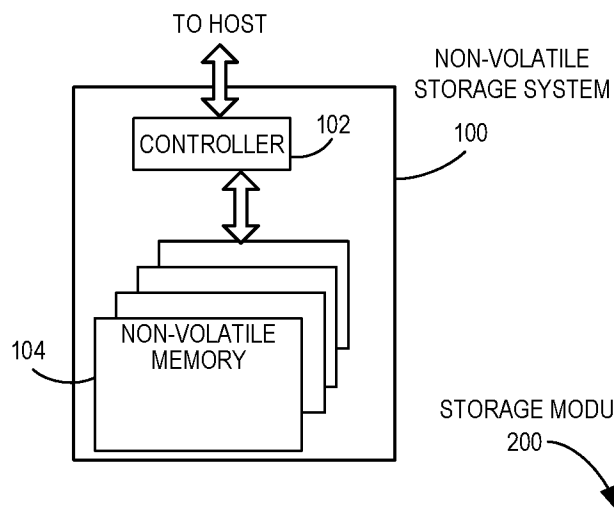


FIG. 1A

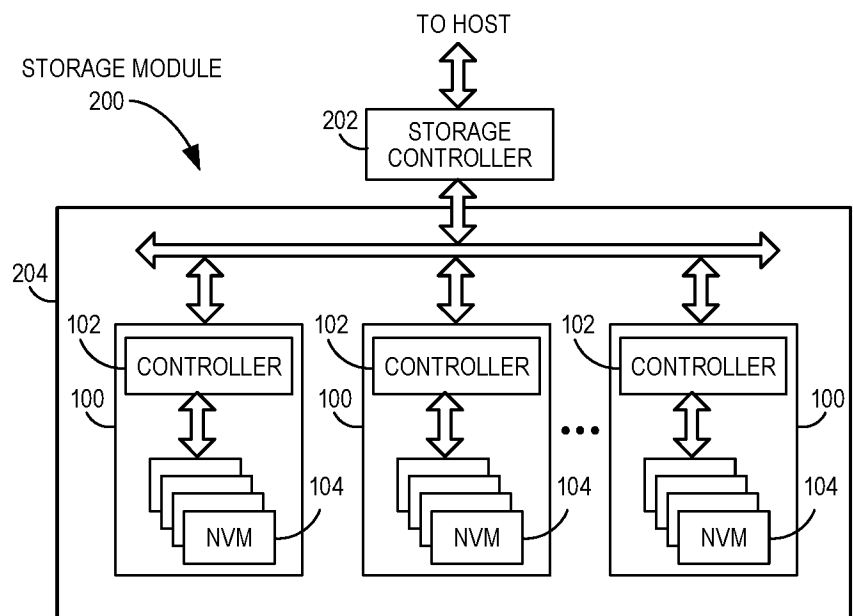


FIG. 1B

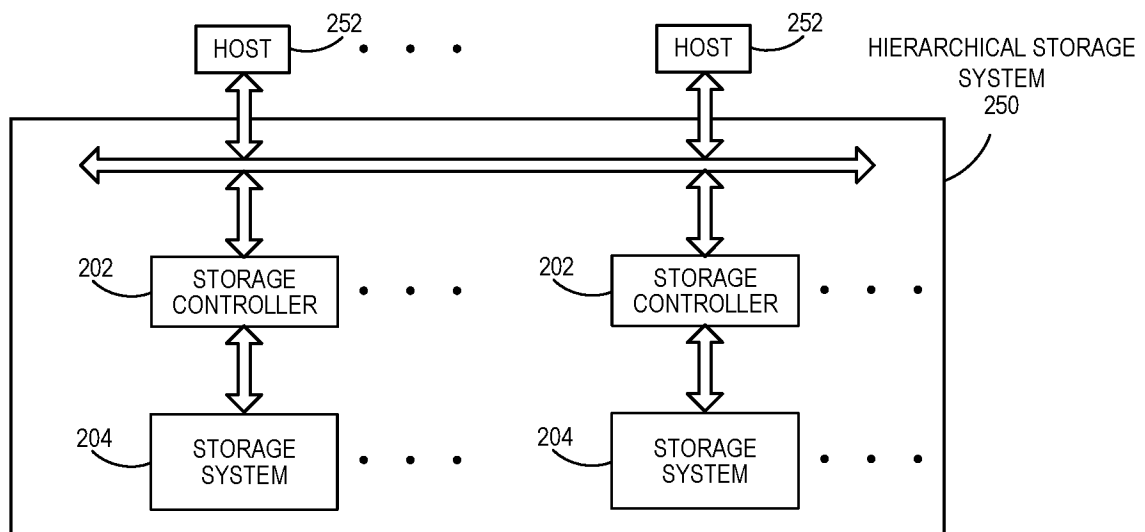


FIG. 1C

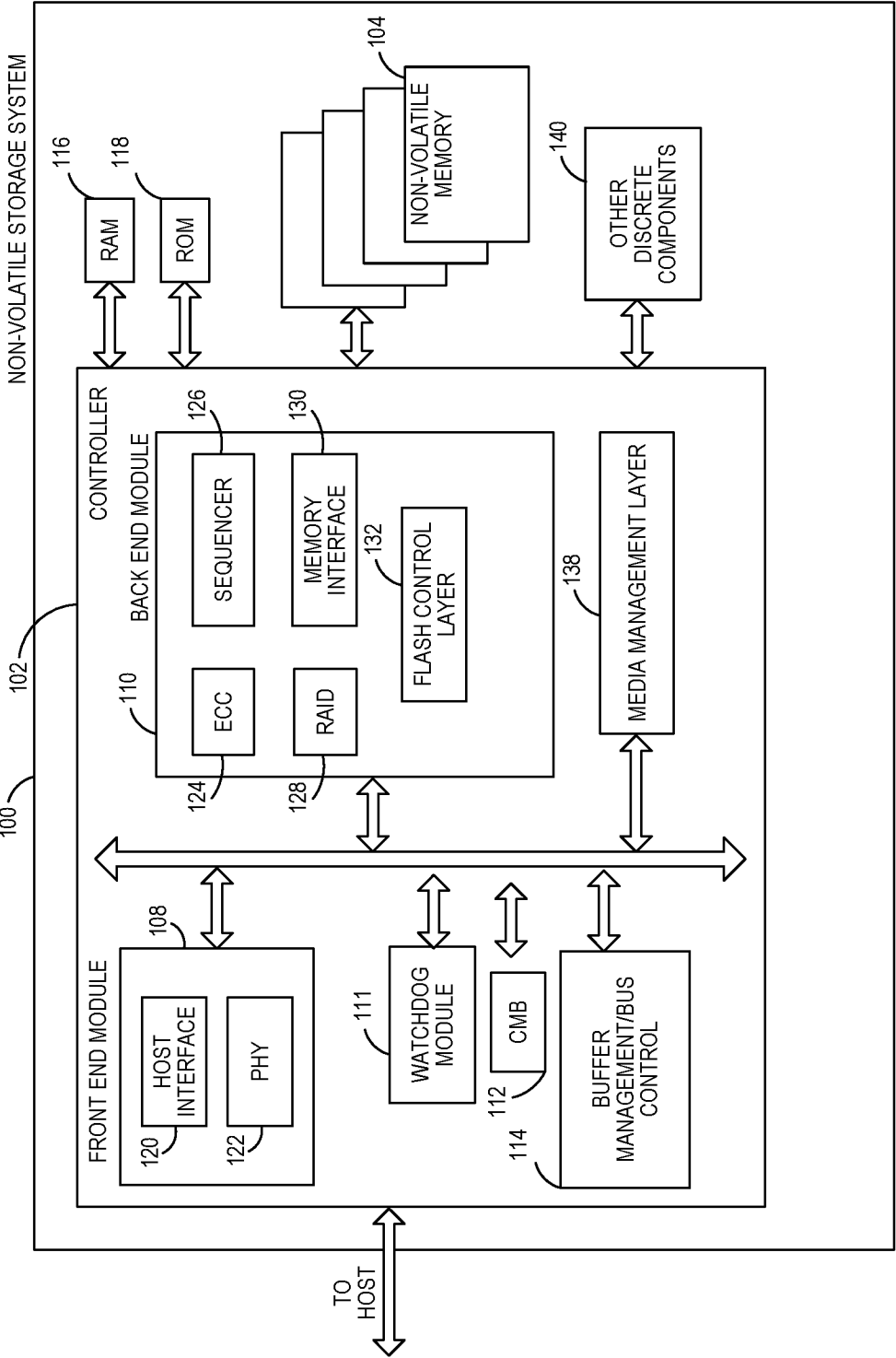


FIG. 2A

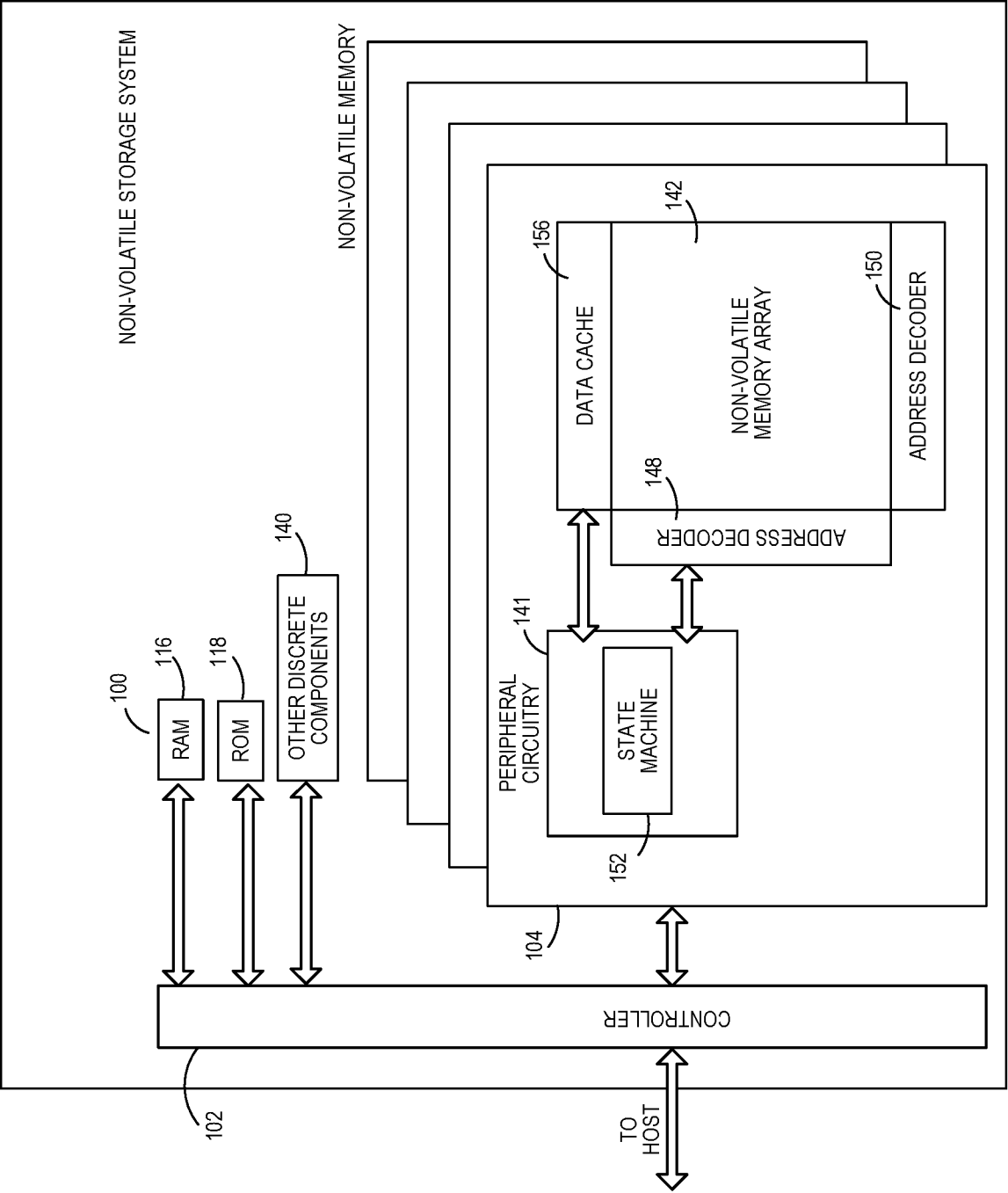


FIG. 2B

4/8

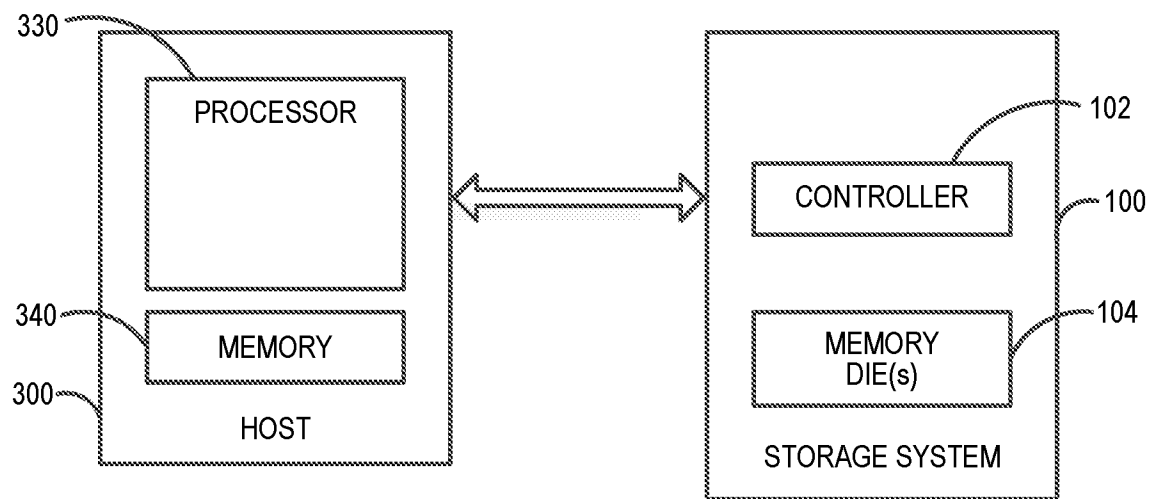


FIG. 3

5/8

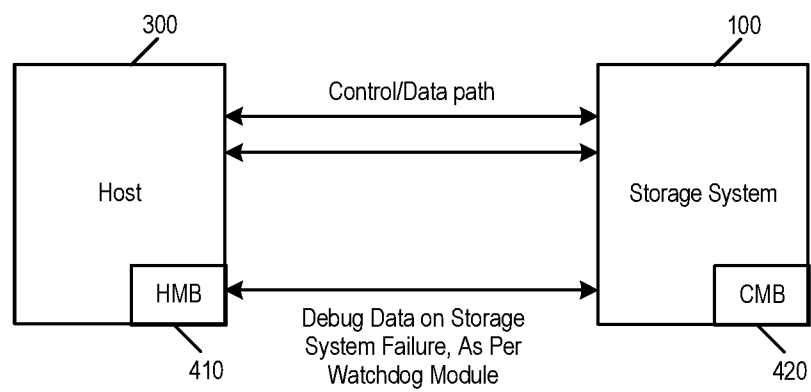


FIG. 4

6/8

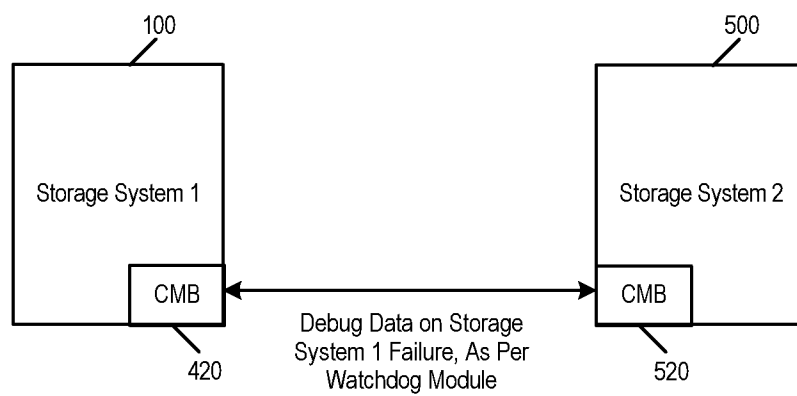


FIG. 5

7/8

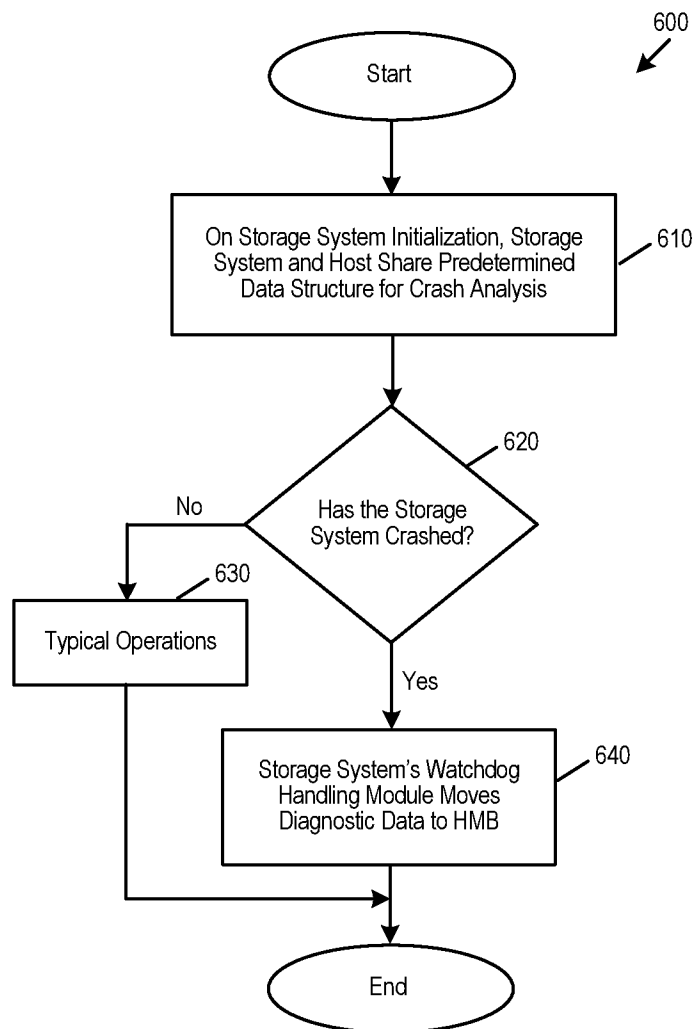


FIG. 6

8/8

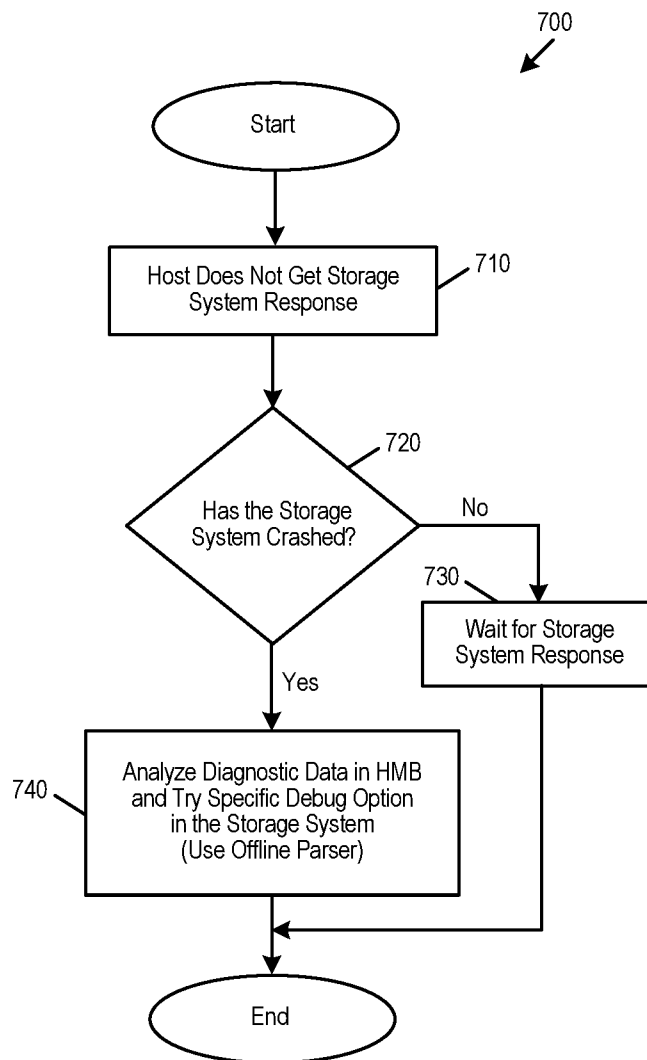


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/016201

A. CLASSIFICATION OF SUBJECT MATTER IPC (20210101) G06F 11/10, G06F 12/00, G06F 12/02, G06F 3/00 CPC (20130101) G06F 11/10, G06F 12/00, G06F 12/02, G06F 3/005 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) IPC (20210101) G06F 11/00, G06F 12/00 CPC (20180501) G06F 11/00, G06F 12/00 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Databases consulted: Esp@cenet, Google Patents		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 9760480 B1 Amazon Tech inc. [US] 12 Sep 2017 (2017/09/12) whole document	1-20
X	US 2020133911 A1 Dell products [US] 30 Apr 2020 (2020/04/30) whole document	1-20
☐ 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 "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 12 May 2021		Date of mailing of the international search report 12 May 2021
Name and mailing address of the ISA: Israel Patent Office Technology Park, Bldg.5, Malcha, Jerusalem, 9695101, Israel Email address: pctoffice@justice.gov.il		Authorized officer ZOZULYA Irina Telephone No. 972-73-3927253

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2021/016201

Patent document cited search report			Publication date	Patent family member(s)		Publication Date
US	9760480	B1	12 Sep 2017	US	9760480 B1	12 Sep 2017
-----			-----	-----		-----
US	2020133911	A1	30 Apr 2020	US	2020133911 A1	30 Apr 2020
-----			-----	-----		-----