



(51) International Patent Classification:
G06F 3/06 (2006.01)

(21) International Application Number:
PCT/US2018/065813

(22) International Filing Date:
14 December 2018 (14.12.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
15/849,014 20 December 2017 (20.12.2017) US

(71) Applicant: **MICRON TECHNOLOGY, INC.** [US/US];
8000 South Federal Way, Boise, Idaho 83716 (US).

(72) Inventors: **SUHLER, Paul A.**; 10050 Great Hills
Trail, Austin, Texas 78732 (US). **KAUL, Ram Kris-**
han; 1312 Cassiopeia Way, Austin, Texas 78732 (US).
DANIELSON, Michael B.; 500 Pipe Creek Lane, George-
town, Texas 78633 (US).

(74) Agent: **PORTNOVA, Marina** et al.; Lowenstein Sandler
LLP, One Lowenstein Drive, Roseland, New Jersey 07068
(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, 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, 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, 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).

(54) Title: UTILIZING WRITE STREAM ATTRIBUTES IN STORAGE WRITE COMMANDS

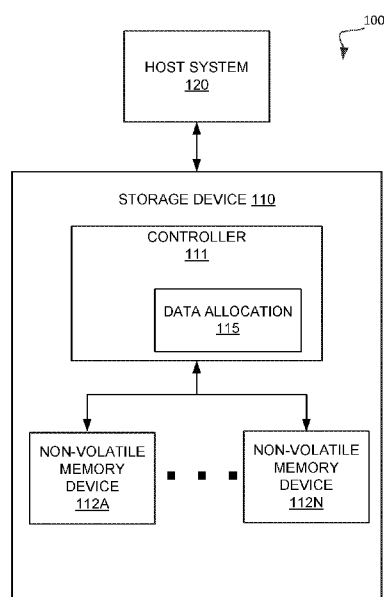


FIG. 1

(57) Abstract: An example method of determining storage operation parameters based on data stream attributes may include: receiving, by a controller, a write command specifying a data item and an identifier of a data stream comprising the data item, wherein a part of the identifier of the data stream encodes a data attribute shared by data items comprised by the data stream; determining, using the data attribute, a storage operation parameter; and transmitting, to a memory device, an instruction specifying the data item and the storage operation parameter.

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

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))*

UTILIZING WRITE STREAM ATTRIBUTES IN STORAGE WRITE COMMANDS

TECHNICAL FIELD

[001] The present disclosure generally relates to storage devices, and more specifically, relates to specifying and utilizing write stream attributes in storage write commands.

BACKGROUND

[002] A storage device, such as a solid-state drive (SSD), may include one or more non-volatile memory devices. The SSD may further include a controller that may manage allocation of data on the memory devices and provide an interface between the storage devices and the host computer system.

BRIEF DESCRIPTION OF THE DRAWINGS

[003] The present disclosure will be understood more fully from the detailed description given below and from the accompanying drawings of various implementations of the disclosure.

[004] Fig. 1 schematically illustrates an example computing environment operating in accordance with one or more aspects of the present disclosure;

[005] Fig. 2 schematically illustrates a programming model which may be implemented by the host system in communication with the storage device controller managing one or more storage devices, in accordance with one or more aspects of the present disclosure;

[006] Fig. 3 schematically illustrates an example structure of the write stream command, in accordance with one or more aspects of the present disclosure;

[007] Fig. 4 schematically illustrates an example data placement strategy implemented by the storage device controller operating in accordance with one or more aspects of the present disclosure;

[008] Fig. 5 is a flow diagram of an example method 500 of determining storage operation parameters based on data stream attributes, in accordance with one or more aspects of the present disclosure;

[009] Fig. 6 is a flow diagram of an example method 600 of providing data stream attributes within the data stream identifier field of stream write commands, in accordance with one or more aspects of the present disclosure;

[0010] Fig. 7 is a block diagram of an example storage device controller operating in accordance with one or more aspects of the present disclosure; and

[0011] Fig. 8 schematically illustrates a block diagram of an example computer system in which implementations of the present disclosure may operate.

DETAILED DESCRIPTION

[0012] Aspects of the present disclosure are directed to specifying and utilizing write stream attributes in storage write commands that are transmitted by a host system to a storage device controller. The host system may group, into several data streams, the data to be written to the storage device, such that each data stream would contain data items belonging to the same group of associated data (e.g., the data associated with a single data structure, such as a file or a database). Thus, the data items contained by a single data stream may share one or more attributes reflecting anticipated media usage patterns, e.g., the anticipated retention time (also referred to as the “stream temperature”) or the workload type. In certain implementations, the data stream may be identified by a dedicated field in each write command transmitted by the host system to the storage device controller. The storage device controller may utilize the stream identifying information in order to optimize the usage of the storage media (e.g., the negative-and (NAND) flash memory), e.g., by placing the data items of the same data stream in a contiguous section of the storage media.

[0013] In accordance with one or more aspects of the present disclosure, one or more bits of the stream identifier field of the write command may be utilized for specifying one or more data attributes shared by the data items of the data stream. Thus, each write command transmitted by the host system to the storage device controller may not only identify the stream, but also indicate the data attributes which are shared by the data items of the data stream. The storage device controller may utilize the stream identifying information enhanced by the data attributes in order to further optimize the usage of the storage media, e.g., by placing the data items of two or more data streams sharing one or more data attributes in the same or physically proximate sections of the storage media and/or avoiding the placement of two or more data streams having substantially different data attributes in the same or physically proximate sections of the storage media. Such placement strategies may be directed to distributing the programming and erasing cycles uniformly across the media in order to maximize the endurance of the storage media, as explained in more detail herein below.

[0014] Thus, aspects of the present disclosure represent significant improvements over various common implementations of storage devices and systems, by enhancing each write command with the data stream attributes in order to further optimize the usage of the storage

media. Various aspects of the above referenced methods and systems are described in details herein below by way of examples, rather than by way of limitation.

[0015] Fig. 1 schematically illustrates an example computing environment 100 operating in accordance with one or more aspects of the present disclosure. In general, the computing environment 100 may include a host system 120 that uses the storage device 110. For example, the host system 120 may write data to the storage device 110 and read data from the storage device 110. The host system 120 may be a computing device such as a desktop computer, laptop computer, network server, mobile device, or such computing device that includes a memory and a processing device. The host system 120 may include or be coupled to the storage device 110 so that the host system 120 may read data from or write data to the storage device 110. For example, the host system 120 may be coupled to the storage device 110 via a physical host interface. Examples of a physical host interface include, but are not limited to, a serial advanced technology attachment (SATA) interface, a peripheral component interconnect express (PCIe) interface, universal serial bus (USB) interface, an NVM Express (NVMe), Fibre Channel, Serial Attached SCSI (SAS), etc. The physical host interface may be used to transmit data between the host system 120 and the storage device 110. In an illustrative example, the host system 120 may be represented by the computer system 800 of Fig. 8.

[0016] As shown in Fig. 1, the storage device 110 may include a controller 111 and storage media, such as memory devices 112A to 112N. In certain implementations, the memory devices 112A to 112N may be provided by non-volatile memory devices, such as NAND flash memory. Each of the memory devices 112A to 112N may include one or more arrays of memory cells such as single level cells (SLCs), multi-level cells (MLCs), or quad-level cells (QLCs). Each of the memory cells may store bits of data (e.g., data blocks) used by the host system 120. Although non-volatile memory devices such as NAND flash memory are described, the memory devices 112A to 112N may be based on any other type of memory. For example, the memory devices 112A to 112N may be provided by random access memory (RAM), read-only memory (ROM), dynamic random access memory (DRAM), synchronous dynamic random access memory (SDRAM), phase change memory (PCM), magneto random access memory (MRAM), negative-or (NOR) flash memory, and electrically erasable programmable read-only memory (EEPROM). Furthermore, the memory cells of the memory devices 112A to 112N may be grouped as memory pages or data blocks that may refer to a unit of the memory device used to store data.

[0017] The controller 111 may communicate with the memory devices 112A to 112N to

perform operations including reading data from or writing data to the memory devices 112A-112N. The controller 111 may include hardware such as one or more integrated circuits, firmware, or a combination thereof. In operation, the controller 111 may receive commands or operations from the host system 120 and may convert the commands or operations into instructions or appropriate commands to achieve the desired access to the memory devices 112A-112N. In various illustrative examples, the controller 111 may be responsible for other operations such as wear leveling, garbage collection, error detection and error-correcting code (ECC), encryption, caching, and address translations between a logical block address and a physical block address that are associated with the memory devices 112A-112N.

[0018] In order to implement the systems and methods of the present disclosure, the controller 111 may include a data allocation functional component 115 that may be employed to allocate the incoming data to particular locations on memory devices 112A-112N. It should be noted that the component designation is of a purely functional nature, i.e., the functions of the data allocation component may be implemented by one or more hardware components and/or firmware modules of the controller 111, as described in more detail herein below. The storage device 110 may include additional circuitry or components that are omitted from Fig. 1 for clarity and conciseness.

[0019] Fig. 2 schematically illustrates a programming model which may be implemented by the host system 120 in communication with the storage device controller 111 managing one or more memory devices 112A-112N, in accordance with one or more aspects of the present disclosure. As schematically illustrated by Fig. 2, the host system may execute one or more applications 210A-210B. In an illustrative example, the application 210A may be in communication with the file system driver 220, which may be running in the kernel space of the host system 120 and may be employed for processing certain system calls, such as read and write calls initiated by one or more applications 210, including the application 210A, running in the user space of the host system 120. The file system driver 220 may be employed to translate the read, write, and other system calls issued by the application 210A into low-level application programming interface (API) calls to the storage driver 230, which, in turn may communicate to the device controller 111 controlling one or more memory devices 112A-112N. The storage driver 230 may be running in the kernel mode of the host system and may be employed to process API calls issued by the file system driver 220 and/or system calls issued by the application 210B into storage interface commands to be processed by the storage the device controller 111 managing one or more memory devices 112A-112N.

[0020] In an illustrative example, the storage driver 230 may implement a block storage

model, in which the data is grouped into blocks of one or more pre-defined sizes and is addressable by a block number. The block storage model may implement “read” and “write” command for storing and retrieving blocks of data. In an illustrative example, the storage driver 230 may implement a key-value storage model, in which the data is represented by the “value” component of a key-value pair is addressable by the “key” component of the key-value pair. The key value storage model may implement “put and get” commands, which are functionally similar to the “write” and “read” commands of the block storage model. Thus, the term “data item” as used herein may refer to a data block or to a key-value pair.

[0021] The application 210A-210B and/or the storage driver 230 executed by the host system 120 may group, into several data streams, the data to be written to the memory devices 112, such that the data items belonging to the same data stream would share one or more attributes. In an illustrative example, a data attribute may reflect the anticipated retention time of the data stream (also referred to as the “stream temperature”), such that a “hot” data stream would comprise short-living data items which are likely to be overwritten within a relatively short period of time (e.g., a period of time falling below a pre-defined low threshold), while a “cold” data stream comprise static data items which are not likely to be overwritten for a relatively long period of time (e.g., a period of time exceeding a pre-defined high threshold). In an illustrative example, the data stream temperature may be communicated to the storage driver 230 by the application 210 which produces the data stream and thus is presumably aware of its anticipated retention time. The data stream temperature may be communicated to the storage driver 230, e.g., via an Input/Output Control (IOCTL) system call. Alternatively, the data stream temperature may be determined by the storage driver 230, which may buffer the incoming data to be written to the memory devices 112A-112N, and may estimate the stream temperature based on the average frequency of overwrite operations requested by the application 210 with respect to one or more data items to be written to the memory devices 112A-112N. The storage driver 230 may then group the buffered data to be written to the storage device 110 into two or more data streams, and may issue stream write commands indicating the data stream temperature to the storage device controller 111, as described in more detail herein below.

[0022] In another illustrative example, a data attribute may reflect the workload type of the data stream, e.g., the “log data” attribute indicating that the data represents the logging data related to one or more databases and/or file systems or “user data” attribute indicating that the data represents other (not related to database or file system logs) types of data. The data stream workload type may be communicated to the storage driver 230 by the application

210 which produces the data stream and thus is presumably aware of its workload type. The data stream workload type may be communicated to the storage driver 230, e.g., via an Input/Output Control (IOCTL) system call. The storage driver may group the data labelled with the “log data” attribute into one or more data streams, and may issue stream write commands indicating the workload type to the storage device controller 111, as described in more detail herein below.

[0023] In certain implementations, the data stream may be identified by a dedicated field in each write command transmitted by the host system to the storage device controller. Fig. 3 schematically illustrates an example structure of the write stream command, in accordance with one or more aspects of the present disclosure. The write stream command 300 may include, among other fields, the operation code field 310 specifying the command type (e.g., the write stream command). The write stream command 300 may further include the flags field 320 specifying one or more parameters of the command. The write stream command 300 may further include the logical block address (LBA) field 330 specifying the LBA of the data being stored on the storage device. The write stream command 300 may further include the stream identifier field 340 represented by a bit string, which may be interpreted as an unsigned integer value. One or more bits (such as a group of one or more most significant bits or a group of or more least significant bits) of the stream identifier field 340 may be utilized for specifying one or more data stream attributes 350 shared by the data items of the data stream. In an illustrative example, one or more bits of the stream identifier field 340 may be utilized for specifying the data stream temperature (e.g., “0” indicating a cold stream and “1” indicating a hot stream, or “00” indicating unknown stream temperature, “01” indicating a cold stream, “10” indicating medium stream temperature, and “11” indicating a hot stream). In an illustrative example, one or more bits of the stream identifier field 340 may be utilized for specifying the workload type of the data stream (e.g., “1” indicating the “log data” workload type and “0” indicating “user data” workload type). The write stream command 300 may include various other fields which are omitted from Fig. 3 for clarity and conciseness.

[0024] Thus, each write command transmitted by the host system to the storage device controller may not only identify the stream, but also indicate the data attributes which are shared by the data items of the data stream. The storage device controller may utilize the stream identifying information enhanced by the data attributes in order to determine storage operation parameters (such as one or more parameters defining the data placement on the storage media) that would optimize the usage of the storage media. The storage device controller may implement one or more wear leveling methods directed to distributing the

programming and erasing cycles uniformly across the media. The wear leveling methods implemented by the storage device controller may involve avoiding placing the “hot” data to the physical blocks that have experienced relatively heavy wear. The storage device controller may place the “cold” data and/or move the data that has not been modified for at least a certain period of time (e.g., a period of time exceeding a certain threshold) out of blocks that have experienced a low number of programming/erasing cycles into more heavily worn blocks. This strategy frees up the low-worn blocks for the “hot” data, while reducing the expected wear on the heavily-worn blocks.

[0025] In an illustrative example, erasing one or more data items of one data stream may require erasing one or more data items which are stored within the same or physically proximate sections of the storage media. Therefore, placing the data streams having substantially different expected retention time within the same or physically proximate sections of the storage media may result in excessive number of programming and erasing cycles to be performed by the controller on the storage media. Conversely, placing the data streams having similar expected retention time within the same or physically proximate sections of the storage media may result in reducing the number of programming and erasing cycles to be performed by the controller on the storage media. Accordingly, a storage device controller operating in one or more aspects of the present disclosure may implement a data placement strategy which is directed to distributing the programming and erasing cycles uniformly across the media in order to maximize the endurance of the storage media.

[0026] Fig. 4 schematically illustrates an example data placement strategy implemented by the storage device controller operating in accordance with one or more aspects of the present disclosure. In an illustrative example, the storage device controller may place the data items of two or more data streams sharing one or more data attributes (such as the data stream temperature and/or data stream workload type) in the same or physically proximate sections of the storage media. In an illustrative example, “section of the storage media” may be represented by a group of one or more memory cells such as single level cells (SLCs), multi-level cells (MLCs), or quad-level cells (QLCs) of NAND type flash memory. In another illustrative example, “section of the storage media” may be represented by groups of memory units addressable by the same signal (such as a word line or a bit line).

[0027] As shown in Fig. 4, data streams 410 and 420, including data items 410A-410N and 420A-420K, respectively, may share the stream temperature 412 (e.g., “H” denoting “hot”). Accordingly, the storage device controller may issue one or more device-level instructions to the data storage devices in order to place the data items of the data streams 410

and 420, including, for example, data items 410A, 410B, and 420A, into the same section 450A of the storage media 400.

[0028] In another illustrative example, the storage device controller may avoid placing two or more data streams having substantially different data attributes (such as the data stream temperature and/or data stream workload type) in the same or physically proximate sections of the storage media. As shown in Fig. 4, the data stream 430 including data items 430A-430M may have a stream temperature 432 (e.g., “C” denoting “cold”) which is different from the stream temperature 412 shared by the data streams 410 and 420.

Accordingly, the storage device controller may issue one or more device-level instructions to the data storage devices in order to place the data items of the data stream 430, including, for example, data items 430A, 430B, and 430C, into the section 450B of the storage media 400.

[0029] Fig. 5 is a flow diagram of an example method 500 of determining storage operation parameters based on data stream attributes, in accordance with one or more aspects of the present disclosure. The method 500 may be performed by processing logic that may include hardware (e.g., processing device, circuitry, dedicated logic, programmable logic, microcode, hardware of a device, integrated circuit, etc.), software (e.g., instructions run or executed on a processing device), or a combination thereof. In some embodiments, the method 500 may be performed by the storage device controller 111 of Fig. 1.

[0030] As shown in Fig. 5, at block 510, the processing logic implementing the method may receive, from a host system, a write command specifying a data item to be written to a memory device managed by the storage device controller. The write command may further specify an identifier of a data stream to which the write command belongs. In an illustrative example, the identifier of the data stream is provided by an unsigned integer value. A portion of the identifier of the data stream may encode one or more data attributes shared by the data items of the data stream. In an illustrative example, the data attribute may include a value reflecting an anticipated retention time of the data items of the data stream. In another illustrative example, the data attribute may include a value reflecting a workload type of the data items of the data stream, as described in more detail herein above.

[0031] At block 520, the processing logic may parse the identifier of the data stream to determine a data attribute shared by data items comprised by the data stream. In an illustrative example, parsing the identifier of the data stream may involve identifying a bit string of a pre-defined size starting from a pre-defined position within the data stream identifier.

[0032] At block 530, the processing logic may determine, based on the data attribute, one

or more storage operation parameters (such as one or more parameters defining the data placement on the storage media) that would optimize the usage of the storage media, e.g., by uniformly distributing programming cycles across the storage media. In an illustrative example, a storage operation parameter may identify the section of the memory device to be utilized for storing the data item. In another illustrative example, the identified section may be located in a physical proximity of another section, which is used for storing another data stream having the same attribute as the data items being stored, as described in more detail herein above.

[0033] At block 540, the processing logic may transmit, to the storage device, an instruction specifying the data item and the storage operation parameters, as described in more detail herein above.

[0034] Fig. 6 is a flow diagram of an example method 600 of providing data stream attributes within the data stream identifier field of stream write commands, in accordance with one or more aspects of the present disclosure. The method 600 may be performed by processing logic that may include hardware (e.g., processing device, circuitry, dedicated logic, programmable logic, microcode, hardware of a device, integrated circuit, etc.), software (e.g., instructions run or executed on a processing device), or a combination thereof. In some embodiments, the method 600 may be performed by the host system 120 of Fig. 1 (e.g., by the storage driver 230 of Fig. 2).

[0035] As shown in Fig. 6, at block 610, the processing logic implementing the method may receive a plurality of data items to be written to a storage device. The plurality of data items may be produced by an application running on the host system, as described in more detail herein above with references to Fig. 2.

[0036] At block 620, the processing logic may group the received data items into one or more data streams, such that the data items contained by a single data stream may share one or more attributes reflecting anticipated media usage patterns, e.g., the anticipated retention time (also referred to as the “stream temperature”) or the workload type. In an illustrative example, the processing logic may identify, among the plurality of data items, two or more data items sharing one or more data attributes. Based on the data attribute values, the processing logic may append the identified data items to a newly created or an existing data stream. In an illustrative example, the data attribute may include a value reflecting an anticipated retention time of the data items of the data stream. In another illustrative example, the data attribute may include a value reflecting a workload type of the data items of the data stream, as described in more detail herein above.

[0037] At block 630, the processing logic may generate a data stream identifier which includes an encoded form of the data attribute. In an illustrative example, the data stream identifier may be provided by an unsigned integer value, one or more bits of which may be utilized for encoding the data attributes shared by the data items of the data stream. In an illustrative example, the bit string encoding the data attributes may have a pre-defined size and may start from a pre-defined position within the data stream identifier. In an illustrative example, the data attribute may include a value reflecting an anticipated retention time of the data items of the data stream. In another illustrative example, the data attribute may include a value reflecting a workload type of the data items of the data stream, as described in more detail herein above.

[0038] At block 640, the processing logic may transmit, to a controller of the storage device, one or more write commands specifying the data comprised by the first data item and the second data item. Each write command may further specify the data stream identifier, the reserved part of which encodes the data attribute.

[0039] Fig. 7 is a block diagram of an example storage device controller 700, which may implement the functionality of the controller 111 of Fig. 1. As shown in Fig. 7, the controller 700 may include a host interface circuitry 714 to interface with a host system via a physical host interface 706. The host interface circuitry 714 may be employed for converting commands received from the host system into device-level instructions. The host interface circuitry 714 may be in communication with the host-memory translation circuitry 716, which may be employed for translating host addresses to memory device addresses. For example, the host-memory translation circuitry 716 may convert logical block addresses (LBAs) specified by host system read or write operations to commands directed to non-volatile memory units identified by logical unit numbers (LUNs) 750. The host-memory translation circuitry 716 may include error detection/correction circuitry, such as exclusive or (XOR) circuitry that may calculate parity information based on information received from the host interface circuitry 714.

[0040] The memory management circuitry 718 may be coupled to the host-memory translation circuitry 716 and the switch 720. The memory management circuitry 718 may control various memory management operations including, but not limited to, initialization, wear leveling, garbage collection, reclamation, and/or error detection/correction. The memory management circuitry 718 may include block management circuitry 740 which may be employed for retrieving data from the volatile memory 717 and/or non-volatile memory identified by LUNs 750. For example, the block management circuitry 740 may retrieve

information such as identifications of valid data blocks, erase counts, and/or other status information of the LUNs 750. The memory management circuitry 718 may further include data allocation component 115 that may be employed to allocate the incoming data to particular locations on logical units identified by LUNs 750. It should be noted that the component designation is of a purely functional nature, i.e., the functions of the data allocation component may be implemented by one or more hardware components and/or firmware modules of the controller 700, such as the processor 728, which may be employed for implementing at least some of the above-referenced memory management operations.

[0041] The switch 720 may be coupled to the host-memory translation circuitry 716, the memory management circuitry 718, the non-volatile memory control circuitry 722, and/or the volatile memory control circuitry 724. The switch 720 may include and/or be coupled to a number of buffers. For example, the switch 720 may include internal static random access memory (SRAM) buffers (ISBs) 725. The switch may be coupled to DRAM buffers 727 that are included in the volatile memory 717. In some embodiments, the switch 720 may provide an interface between various components of the controller 700.

[0042] The non-volatile memory control circuitry 722 may store, in one of the buffers (e.g., the ISBs 725 or the buffer 727), information corresponding to a received read command. Furthermore, the non-volatile memory control circuitry 722 may retrieve the information from one of the buffers and write the information to a logical unit of the non-volatile memory identified by a LUN 750. The logical units identified by LUNs 750 may be coupled to the non-volatile memory control circuitry 722 by a number of channels. In some embodiments, the number of channels may be controlled collectively by the non-volatile memory control circuitry 722. In some embodiments, each memory channel may be coupled to a discrete channel control circuit 748. A particular channel control circuit 748 may control and be coupled to more than one memory unit 750 by a single channel.

[0043] The non-volatile memory control circuitry 722 may include a channel request queue (CRQ) 747 that is coupled to each of the channel control circuits 748. Furthermore, each channel control circuit 748 may include a memory unit request queue (RQ) 744 that is coupled to multiple memory unit command queues (CQs) 746. The CRQ 747 may be configured to store commands (e.g., write requests or read requests) shared between channels, the RQ 744 may be configured to store commands between the memory units 750 on a particular channel, and the CQ 746 may be configured to queue a current command and a next command to be executed subsequent to the current command.

[0044] The CRQ 747 may be configured to receive a command from the switch 720 and

relay the command to one of the RQs 744 (e.g., the RQ 744 associated with the channel that is associated with the particular logical unit identified by the LUN 750 for which the command is targeted). The RQ 744 may be configured to relay a first number of commands for a particular memory unit 750 to the CQ 746 that is associated with the particular logical unit identified by the LUN 750 in an order that the first number of commands were received by the RQ 744. A command pipeline may be structured such that commands to the logical unit move in a particular order (e.g., in the order that they were received by the RQ 744). The RQ 744 may be configured to queue a command for a particular logical unit in response to the CQ 746 associated with the particular logical unit being full and the CRQ 747 may be configured to queue a command for a particular RQ 744 in response to the particular RQ 744 being full.

[0045] The RQ 744 may relay a number of commands for different logical units identified by LUNs 750 to the CQs 746 that are associated with the logical units in an order according to the status of the logical units. For example, the logical unit status may be a ready/busy status. The command pipeline is structured such that the commands between different logical units may move out of order (e.g., in an order that is different from the order in which they were received by the RQ 744 according to what is efficient for overall memory operation at the time). For example, the RQ 744 may be configured to relay a first one of the second number of commands to a first CQ 746 before relaying a second command from the second number of commands to a second CQ 746 in response to the status of the different logical unit associated with the second CQ 746 being busy, where the first command is received later in time than the second command. The RQ 744 may be configured to relay the second command to the second CQ 746 in response to the status of the logical unit associated with the second CQ 746 being ready (e.g., subsequent to relaying the first command).

[0046] In some embodiments, the control circuits for each channel may include discrete error detection/correction circuitry 737 (e.g., error correction code (ECC) circuitry), coupled to each channel control circuit 748 and/or a number of error detection/correction circuits 737 that can be used with more than one channel. The error detection/correction circuitry 737 may be configured to apply error correction such as Bose–Chaudhuri–Hocquenghem (BCH) error correction to detect and/or correct errors associated with information stored in the logical unit identified by the LUN 750. The error detection/correction circuitry 737 may be configured to provide differing error correction schemes for SLC, MLC, or QLC operations.

[0047] Fig. 8 illustrates an example computer system 800 within which a set of instructions, for causing the computer system to perform any one or more of the

methodologies discussed herein, may be executed. In an illustrative example, the computer system 800 may implement the functions of the host system 120 of Fig. 1. In alternative implementations, the computer system may be connected (e.g., networked) to other computer systems in a LAN, an intranet, an extranet, and/or the Internet. The computer system may operate in the capacity of a server or a client computer system in client-server network environment, as a peer computer system in a peer-to-peer (or distributed) network environment, or as a server or a client computer system in a cloud computing infrastructure or environment.

[0048] The computer system may be a personal computer (PC), a tablet PC, a set-top box (STB), a Personal Digital Assistant (PDA), a cellular telephone, a web appliance, a server, a network router, a switch or bridge, or any computer system capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that computer system. Further, while a single computer system is illustrated, the term “computer system” shall also be taken to include any collection of computer systems that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein.

[0049] The example computer system 800 includes a processing device 802, a main memory 804 (e.g., read-only memory (ROM), flash memory, dynamic random access memory (DRAM)), a static memory 806 (e.g., flash memory, static random access memory (SRAM), etc.), and a data storage device 818, which communicate with each other via a bus 830. In an illustrative example, the data storage device 818 may implement the functions of the storage device 110 of Fig. 1.

[0050] Processing device 802 represents one or more general-purpose processing devices such as a microprocessor, a central processing unit, or the like. More particularly, the processing device may be complex instruction set computing (CISC) microprocessor, reduced instruction set computing (RISC) microprocessor, very long instruction word (VLIW) microprocessor, or processor implementing other instruction sets, or processors implementing a combination of instruction sets. Processing device 802 may also be one or more special-purpose processing devices such as an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), network processor, or the like. The processing device 802 is configured to execute instructions 826 for performing the operations and steps discussed herein.

[0051] The computer system 800 may further include a network interface device 808 to communicate over the network 820. The computer system 800 also may include a video

display unit 810 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)), an alphanumeric input device 812 (e.g., a keyboard), a cursor control device 814 (e.g., a mouse), a graphics processing unit 822, a signal generation device 816 (e.g., a speaker), graphics processing unit 822, video processing unit 828, and audio processing unit 832.

[0052] The data storage device 818 may include computer-readable storage medium 824 on which is stored one or more sets of instructions or software 826 embodying any one or more of the methodologies or functions described herein. The instructions 826 may also reside, completely or at least partially, within the main memory 804 and/or within the processing device 802 during execution thereof by the computer system 800, the main memory 804 and the processing device 802 also constituting computer-readable storage media. The computer-readable storage medium 824, data storage device 818, and/or main memory 804 may correspond to the storage device 110 of Fig. 1.

[0053] In one implementation, the instructions 826 include instructions to implement functionality corresponding to a data allocation component (e.g., data allocation component 115 of Fig. 1). While the computer-readable storage medium 824 is shown in an example implementation to be a single medium, the term “computer-readable storage medium” should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term “computer-readable storage medium” shall also be taken to include any medium that is capable of storing or encoding a set of instructions for execution by the computer and that cause the computer to perform any one or more of the methodologies of the present disclosure. The term “computer-readable storage medium” shall accordingly be taken to include, but not be limited to, solid-state memories, optical media and magnetic media.

[0054] Some portions of the preceding detailed descriptions have been presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the ways used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of operations leading to a desired result. The operations are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0055] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "receiving" or "determining" or "transmitting" or "reflecting" or "specifying" or "identifying" or "providing" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage devices.

[0056] The present disclosure also relates to an apparatus for performing the operations herein. This apparatus may be specially constructed for the intended purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, each coupled to a computer system bus.

[0057] The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the method. The structure for a variety of these systems will appear as set forth in the description below. In addition, the present disclosure is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the disclosure as described herein.

[0058] The present disclosure may be provided as a computer program product, or software, that may include a computer-readable medium having stored thereon instructions, which may be used to program a computer system (or other electronic devices) to perform a process according to the present disclosure. A computer-readable medium includes any mechanism for storing information in a form readable by a machine (e.g., a computer system). For example, a computer-readable (e.g., computer-readable) medium includes a read only memory ("ROM"), random access memory ("RAM"), magnetic disk storage media, optical storage media, flash memory devices, etc.

[0059] In the foregoing specification, implementations of the disclosure have been described with reference to specific example implementations thereof. It will be evident that various modifications may be made thereto without departing from the broader spirit and scope of implementations of the disclosure as set forth in the following claims. The specification and drawings are, accordingly, to be regarded in an illustrative sense rather than a restrictive sense.

CLAIMS

What is claimed is:

1. A method, comprising:
 - receiving, by a controller, a write command specifying a data item and an identifier of a data stream comprising the data item;
 - determining, by parsing the identifier of the data stream, a data attribute shared by data items comprised by the data stream;
 - determining, using the data attribute, a storage operation parameter; and
 - transmitting, to a memory device, an instruction specifying the data item and the storage operation parameter.
2. The method of claim 1, wherein the storage operation parameter identifies a section of the memory device to be utilized for storing the data item.
3. The method of claim 1, wherein the data attribute comprises a value reflecting a retention time of data items comprised by the data stream.
4. The method of claim 1, wherein the data attribute specifies a workload type of data items comprised by the data stream.
5. The method of claim 1, wherein parsing the identifier of the data stream further comprises:
 - identifying a bit string of a pre-defined size starting from a pre-defined position within the identifier of the data stream.
6. The method of claim 1, wherein determining the storage operation parameter comprises:
 - identifying a section of the memory device to be utilized for storing the data item, in order to uniformly distribute programming cycles across storage media of the memory device.
7. The method of claim 1, wherein determining the storage operation parameter comprises:

identifying a first section of the memory device to be utilized for storing the data item in a physical proximity of a second section of the memory device, wherein the second section stores one or more data items of a second data stream characterized by the data attribute.

8. A system, comprising:
 - a memory device;
 - a controller operatively coupled to the memory devices, the controller to:
 - receive a write command specifying a data item and an identifier of a data stream comprising the data item;
 - determine, by parsing the identifier of the data stream, a data attribute shared by data items comprised by the data stream;
 - determine, using the data attribute, a storage operation parameter; and
 - transmit, to a memory device, an instruction specifying the data item and the storage operation parameter.
9. The system of claim 8, wherein the storage operation parameter identifies a section of the memory device to be utilized for storing the data item.
10. The system of claim 8, wherein the data attribute comprises a value reflecting a retention time of data items comprised by the data stream.
11. The system of claim 8, wherein the data attribute specifies a workload type of data items comprised by the data stream.
12. The system of claim 8, wherein the part of the identifier of the data stream is provided by a bit string of a pre-defined size starting from a pre-defined position within the identifier of the data stream.
13. The system of claim 8, wherein determining the storage operation parameter comprises:
 - identifying a section of the memory device to be utilized for storing the data item, in order to uniformly distribute programming cycles across storage media of the memory device.

14. The system of claim 8, wherein determining the storage operation parameter comprises:

identifying a section of the memory device to be utilized for storing the data item in a physical proximity of a second section of the memory device, wherein the second section stores one or more data items of a second data stream characterized by the data attribute.

15. A method, comprising:

receiving, by a processor, a plurality of data items to be written to a storage device;

identifying, among the plurality of data items, a first data item and a second data item sharing a data attribute;

generate a data stream identifier comprising an encoded form of the data attribute; and

transmitting, to a controller of the storage device, one or more write commands specifying data comprised by the first data item and the second data item, wherein each write command further specifies the data stream identifier.

16. The method of claim 15, wherein the storage device is a solid state drive (SSD).

17. The method of claim 15, wherein the data attribute specifies a retention time of the first data item and the second data item.

18. The method of claim 17, wherein identifying the first data item and the second data item further comprises:

estimating the retention time of the first data item and the second data item based on an average frequency of requested overwrite operations.

19. The method of claim 15, wherein the data attribute specifies a workload type of the first data item and the second data item.

20. The method of claim 15, wherein the encoded form of the data attribute is provided by a bit string of a pre-defined size starting from a pre-defined position within the data stream identifier.

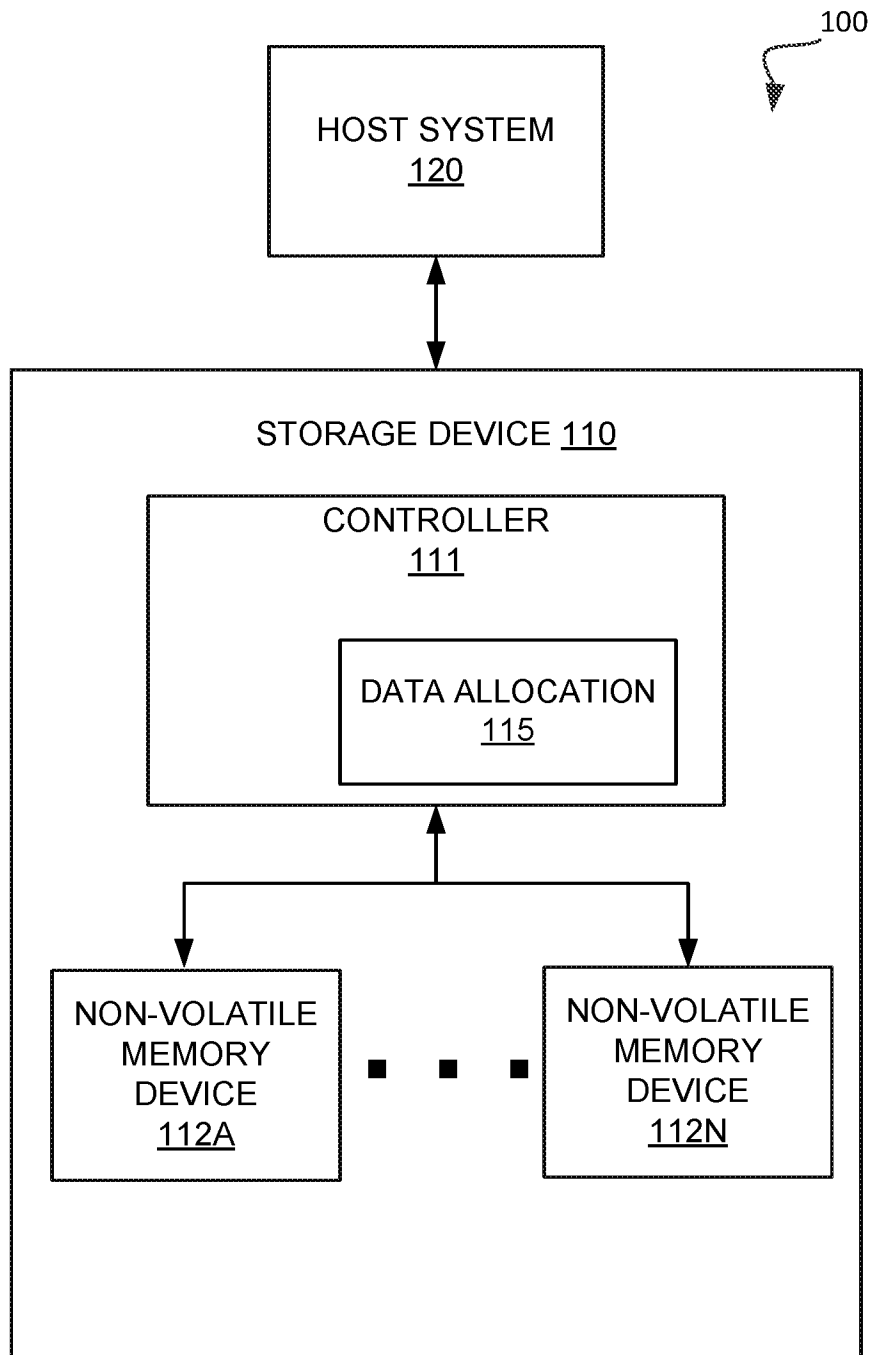


FIG. 1

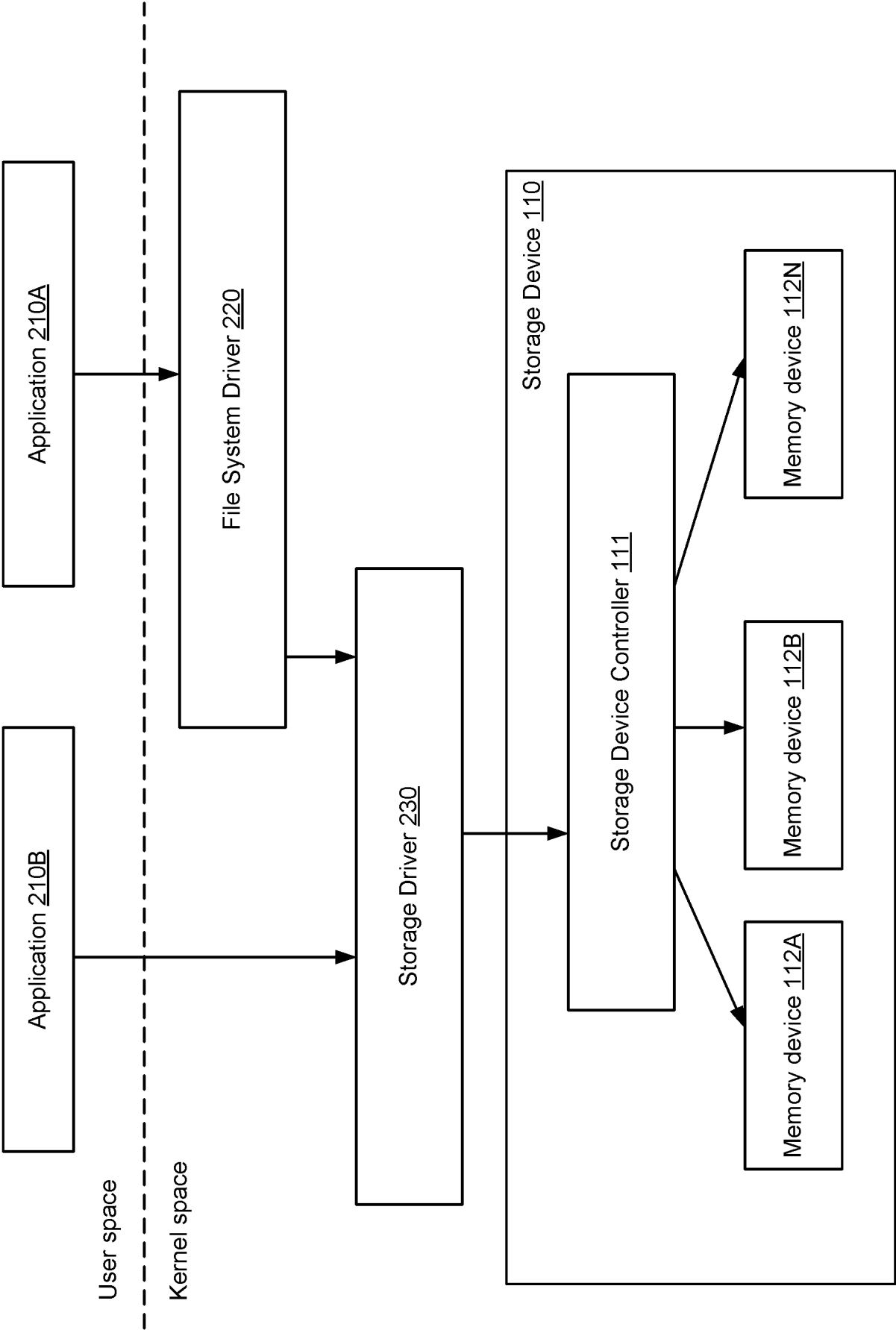


FIG. 2

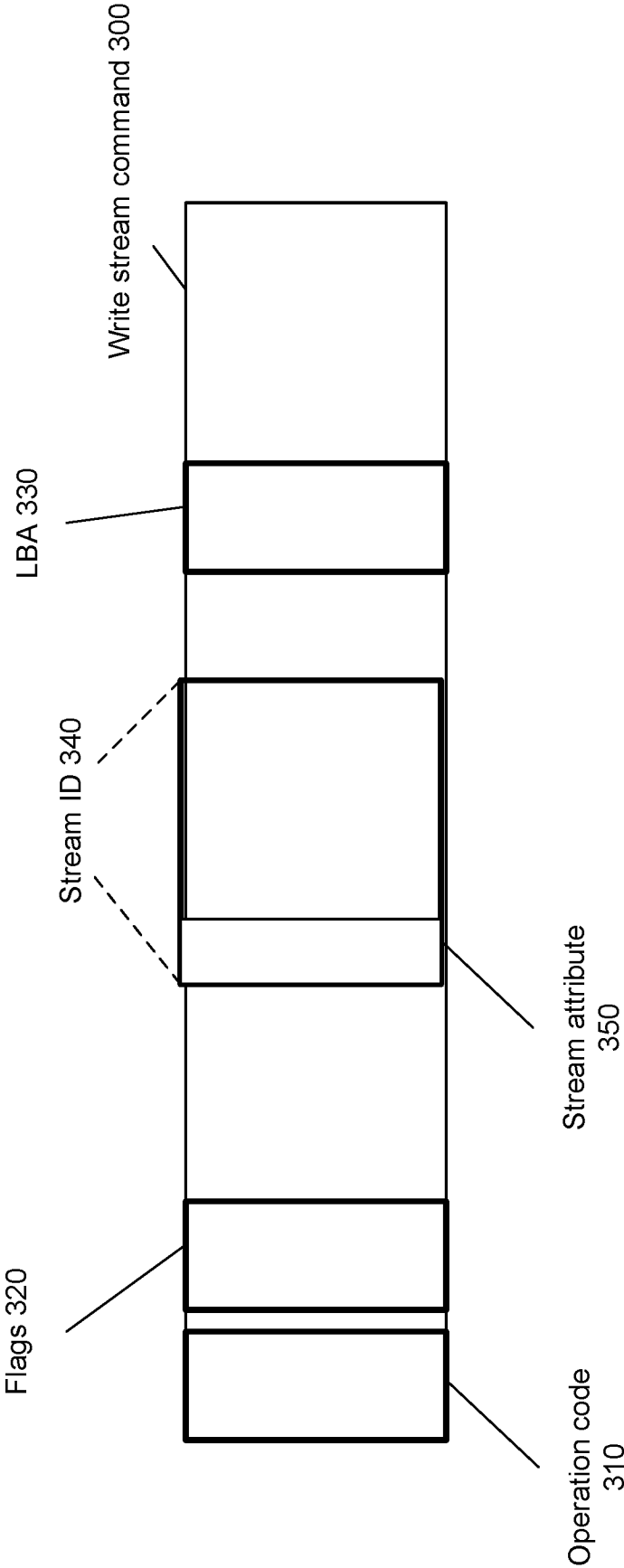


FIG. 3

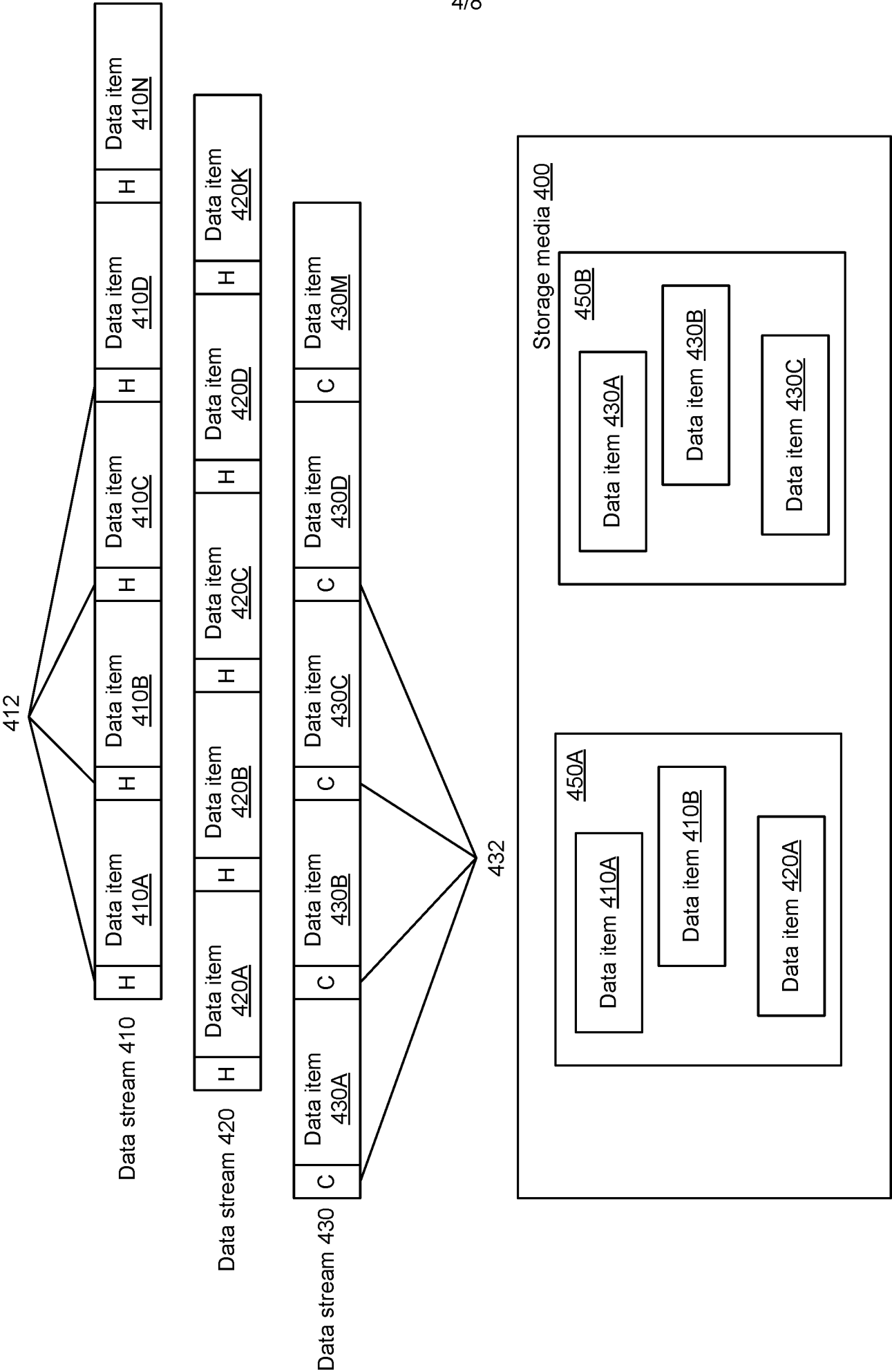


FIG. 4

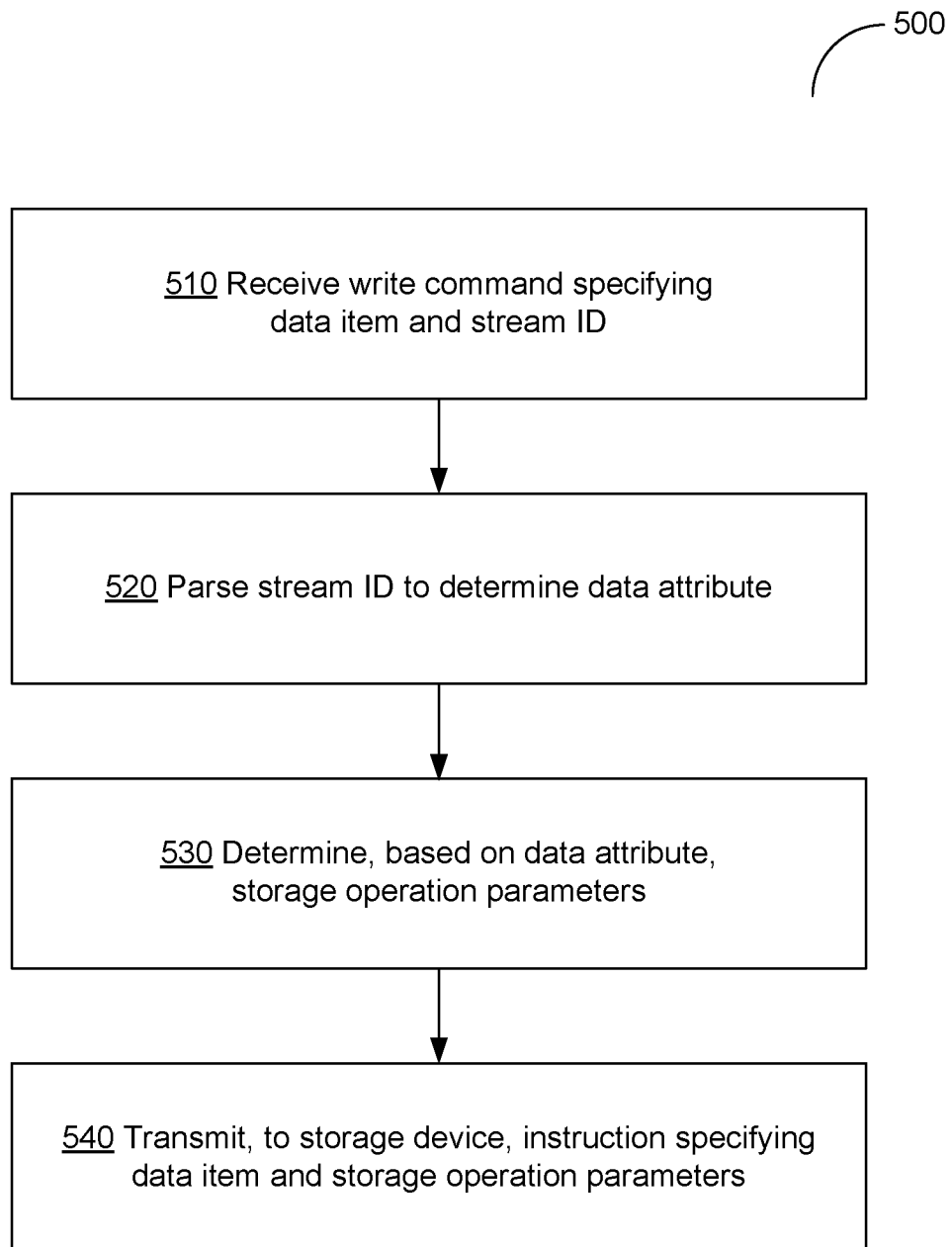


FIG. 5

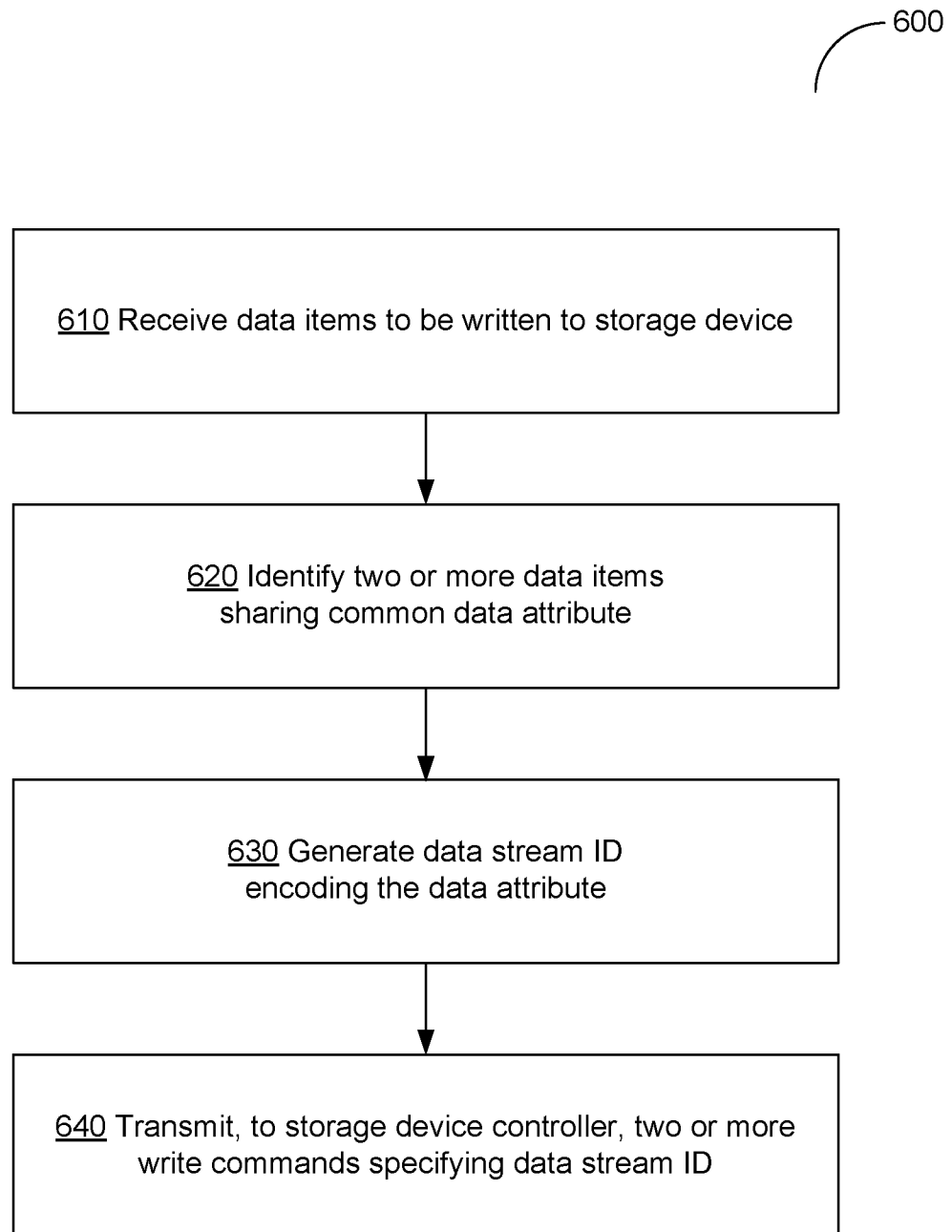


FIG. 6

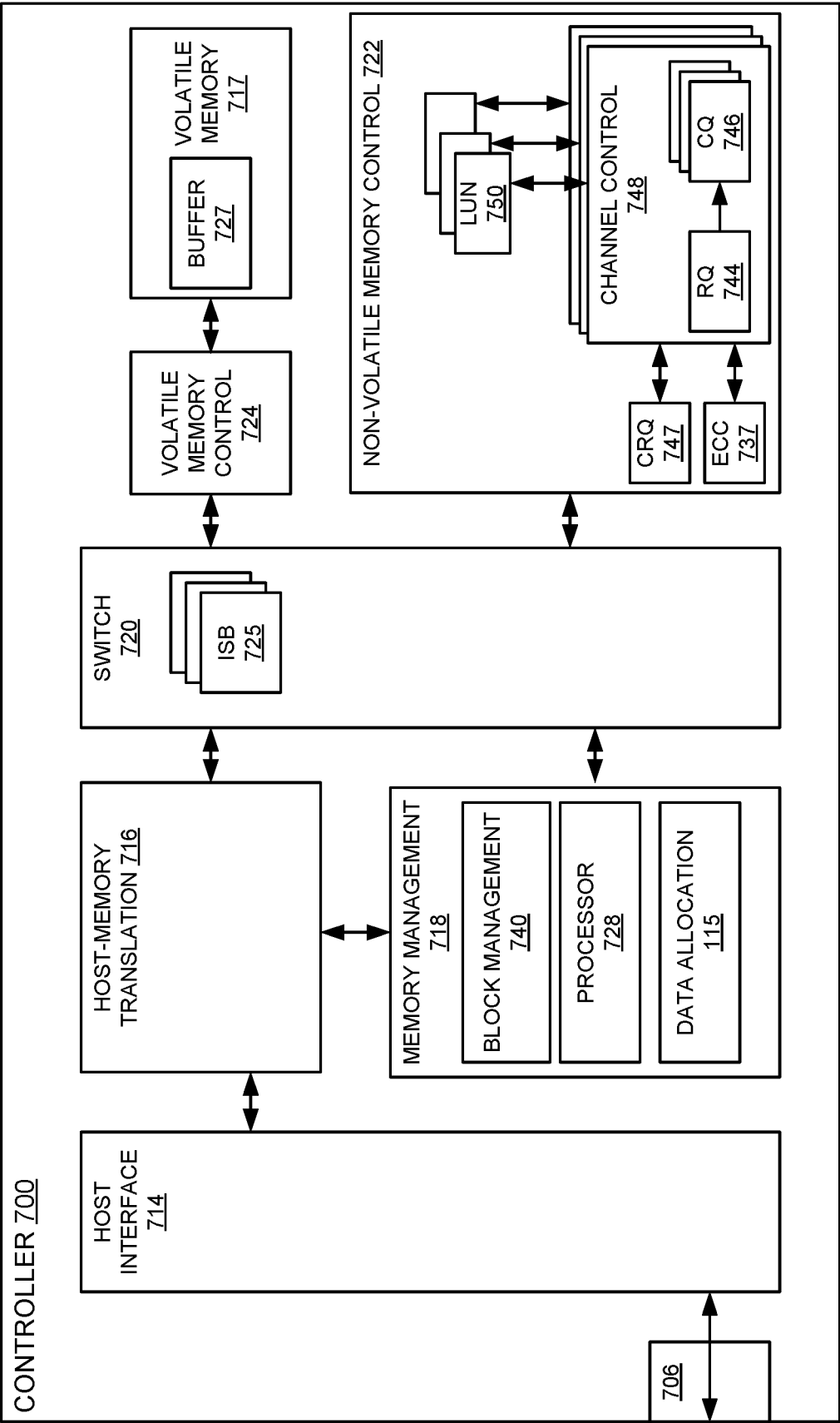


FIG. 7

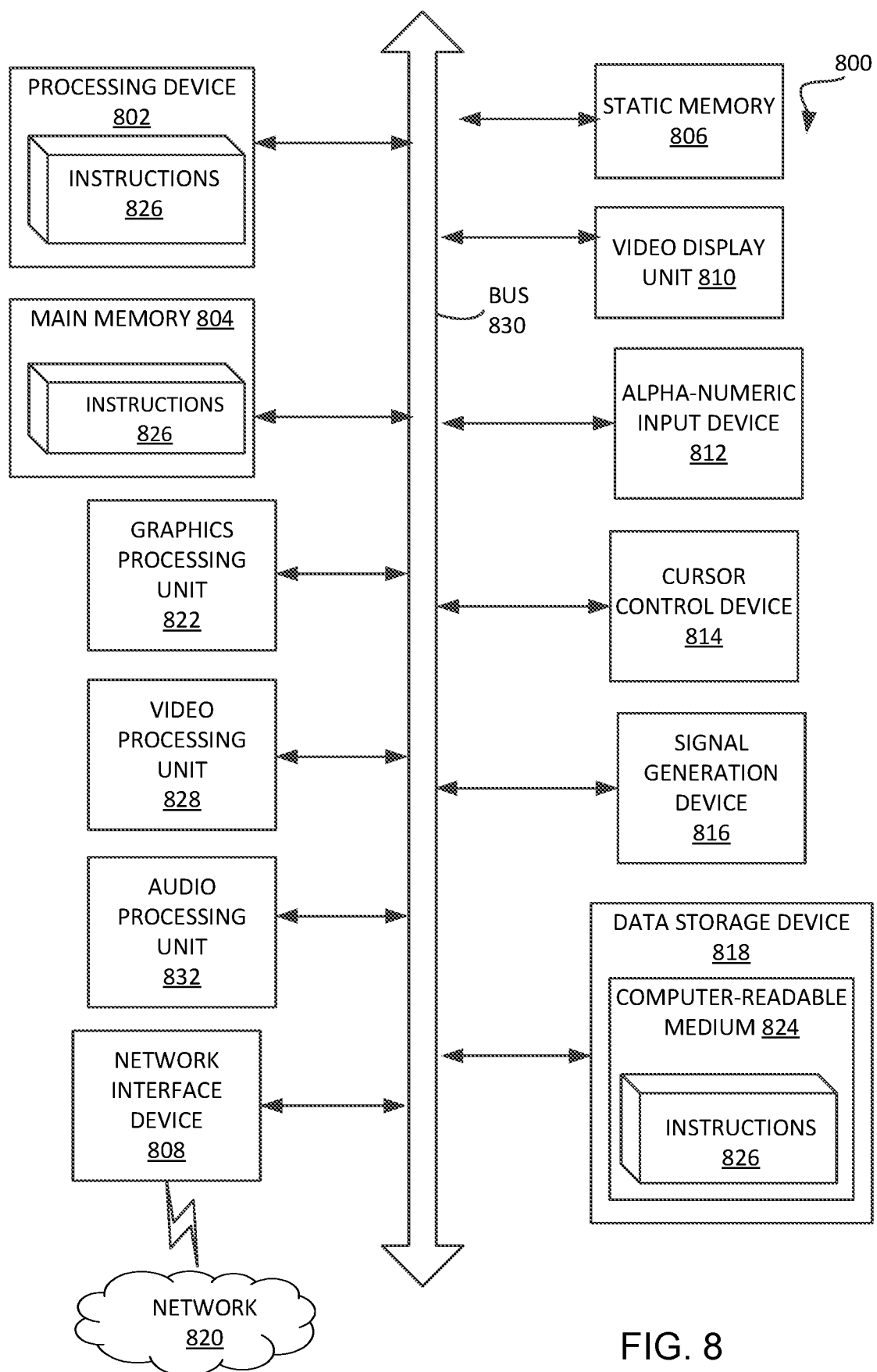


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/065813**A. CLASSIFICATION OF SUBJECT MATTER****G06F 3/06(2006.01)i**

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 3/06; G06F 12/00; G06F 12/02; G06F 12/06; G06F 12/08

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: SSD, controller, non-volatile memory, data item, stream, identifier, attribute

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2017-0017411 A1 (SAMSUNG ELECTRONICS CO., LTD.) 19 January 2017 See paragraphs [0011]-[0012], [0027]-[0028], [0032]-[0033], [0040]-[0048], [0053], [0057], [0069]; and figures 1, 3.	1-4, 6-11, 13-14
Y		5, 12, 15-20
Y	US 2013-0103891 A1 (ERAN SHARON et al.) 25 April 2013 See paragraphs [0034], [0047], [0076], [0089]; and figure 4.	5, 12, 15-20
A	US 2016-0147454 A1 (APPLE INC.) 26 May 2016 See paragraphs [0055]-[0056], [0070]-[0076]; and figures 1-2.	1-20
A	US 2013-0159626 A1 (SHACHAR KATZ et al.) 20 June 2013 See paragraphs [0020]-[0037], [0041]; and figures 1-2.	1-20
A	KR 10-2015-0020136 A (LSI CORPORATION) 25 February 2015 See paragraphs [0012]-[0014], [0023]-[0026]; and figure 2.	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

"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

19 April 2019 (19.04.2019)

Date of mailing of the international search report

22 April 2019 (22.04.2019)

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

KANG, Hee Gok

Telephone No. +82-42-481-8264



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2018/065813

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017-0017411 A1	19/01/2017	CN 106354425 A CN 106354745 A EP 3118745 A1 JP 2017-021804 A JP 2017-021805 A JP 2018-037088 A KR 10-2017-0008152 A KR 10-2017-0008153 A KR 10-2017-0124444 A KR 10-2018-0026327 A TW 201710881 A US 2017-0017663 A1 US 2017-0075832 A1 US 2017-0228157 A1	25/01/2017 25/01/2017 18/01/2017 26/01/2017 26/01/2017 08/03/2018 23/01/2017 23/01/2017 10/11/2017 12/03/2018 16/03/2017 19/01/2017 16/03/2017 10/08/2017
US 2013-0103891 A1	25/04/2013	US 8799559 B2	05/08/2014
US 2016-0147454 A1	26/05/2016	US 2014-0143482 A1 US 2014-0143483 A1 US 8677054 B1 US 9280299 B2 US 9753644 B2	22/05/2014 22/05/2014 18/03/2014 08/03/2016 05/09/2017
US 2013-0159626 A1	20/06/2013	US 10203881 B2	12/02/2019
KR 10-2015-0020136 A	25/02/2015	CN 103348330 A CN 103348330 B CN 103392207 A CN 103392207 B CN 103415844 A CN 103415844 B CN 103902403 A CN 103902403 B CN 103942114 A CN 103942114 B CN 104375956 A CN 104375956 B CN 104679446 A CN 104679446 B DE 102014111668 A1 DE 102014111670 A1 EP 2646922 A2 EP 2646922 A4 EP 2666091 A2 EP 2750018 A2 EP 2750018 A3 EP 2757480 A1 EP 2764516 A1 EP 2764516 A4	09/10/2013 24/05/2017 13/11/2013 04/08/2017 27/11/2013 20/04/2016 02/07/2014 03/05/2017 23/07/2014 10/08/2018 25/02/2015 14/11/2017 03/06/2015 03/10/2017 19/02/2015 19/02/2015 09/10/2013 25/11/2015 27/11/2013 02/07/2014 20/05/2015 23/07/2014 13/08/2014 08/07/2015

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2018/065813

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		GB 2519211 A	15/04/2015
		GB 2519629 A	29/04/2015
		JP 2013-545202 A	19/12/2013
		JP 2014-130587 A	10/07/2014
		JP 2014-142931 A	07/08/2014
		JP 2014-507717 A	27/03/2014
		JP 2014-534503 A	18/12/2014
		JP 2015-036981 A	23/02/2015
		JP 2015-036982 A	23/02/2015
		JP 5978259 B2	24/08/2016
		JP 5980798 B2	31/08/2016
		JP 6265746 B2	24/01/2018
		JP 6285709 B2	28/02/2018
		JP 6443794 B2	26/12/2018
		KR 10-1454807 B1	04/11/2014
		KR 10-1522848 B1	26/05/2015
		KR 10-1562781 B1	23/10/2015
		KR 10-1564569 B1	03/11/2015
		KR 10-1636785 B1	06/07/2016
		KR 10-2013-0114681 A	17/10/2013
		KR 10-2013-0118370 A	29/10/2013
		KR 10-2013-0132662 A	04/12/2013
		KR 10-2014-0084337 A	04/07/2014
		KR 10-2014-0092750 A	24/07/2014
		KR 10-2014-0094467 A	30/07/2014
		KR 10-2014-0094661 A	30/07/2014
		KR 10-2015-0020137 A	25/02/2015
		TW 201232264 A	01/08/2012
		TW 201241615 A	16/10/2012
		TW 201324150 A	16/06/2013
		TW 201415225 A	16/04/2014
		TW 201432702 A	16/08/2014
		TW 201443640 A	16/11/2014
		TW 201518944 A	16/05/2015
		TW 201523249 A	16/06/2015
		TW I516931 B	11/01/2016
		TW I529734 B	11/04/2016
		TW I531902 B	01/05/2016
		TW I534617 B	21/05/2016
		TW I609263 B	21/12/2017
		TW I637262 B	01/10/2018
		US 2013-073895 A1	21/03/2013
		US 2013-246839 A1	19/09/2013
		US 2013-290618 A1	31/10/2013
		US 2014-082261 A1	20/03/2014
		US 2014-108703 A1	17/04/2014
		US 2014-189421 A1	03/07/2014
		US 2014-208003 A1	24/07/2014
		US 2014-208004 A1	24/07/2014
		US 2014-208061 A1	24/07/2014

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2018/065813Patent document
cited in search reportPublication
datePatent family
member(s)Publication
date

US 2014-208062 A1	24/07/2014
US 2014-237166 A1	21/08/2014
US 2015-234599 A1	20/08/2015
US 2016-070496 A1	10/03/2016
US 2017-046273 A1	16/02/2017
US 2017-123733 A1	04/05/2017
US 2017-147435 A1	25/05/2017
US 2017-161191 A1	08/06/2017
US 2018-143874 A1	24/05/2018
US 2018-232179 A1	16/08/2018
US 8656101 B2	18/02/2014
US 8949517 B2	03/02/2015
US 9105305 B2	11/08/2015
US 9183140 B2	10/11/2015
US 9189385 B2	17/11/2015
US 9195594 B2	24/11/2015
US 9329991 B2	03/05/2016
US 9396104 B1	19/07/2016
US 9477406 B2	25/10/2016
US 9495288 B2	15/11/2016
US 9569320 B2	14/02/2017
US 9582431 B2	28/02/2017
US 9727414 B2	08/08/2017
US 9851910 B2	26/12/2017
US 9886383 B2	06/02/2018
US 9971547 B2	15/05/2018
WO 2012-075200 A2	07/06/2012
WO 2012-075200 A3	16/08/2012
WO 2012-099937 A2	26/07/2012
WO 2012-099937 A3	13/09/2012
WO 2013-052562 A1	11/04/2013
WO 2014-028183 A1	20/02/2014