(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 May 2014 (22.05.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/077962 A1

(51) **International Patent Classification**:
G06F 12/00 (2006.01) *G06F 12/08* (2006.01)

(21) **International Application Number**:
PCT/US2013/062725

(22) **International Filing Date**:
30 September 2013 (30.09.2013)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
13/675,913 13 November 2012 (13.11.2012) US

(71) **Applicants**: **WESTERN DIGITAL TECHNOLOGIES, INC.** [US/US]; 3355 Michelson Drive, Suite 100, Irvine, CA 92612 (US). **SKYERA, INC.** [US/US]; 1704 Automation Parkway, San Jose, CA 95131 (US).

(72) **Inventors**: **DANILAK, Radoslav**; 10425 Glencore Drive, Cupertino, CA 95014 (US). **MULLENDORE, Rodney, N.**; 5813 Killarney Circle, San Jose, CA 95138 (US). **TOMLIN, Andrew, J.**; C/o Western Digital Technologies, Inc., 3355 Michelson Drive, Suite 100, Irvine, CA 92612 (US). **JONES, Justin**; C/o Western Digital Technologies, Inc., 3355 Michelson Drive, Suite 100, Irvine, CA 92612 (US). **YANG, Jui-Yao**; C/o Western Digital Technologies, Inc., 3355 Michelson Drive, Suite 100, Irvine, CA 92612 (US).

(74) **Agent**: **DELANEY, Karoline, A.**; KNOBBE MARTENS OLSON & BEAR, LLP, 2040 Main Street, 14th Floor, Irvine, CA 92614 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, 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, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) **Title**: METHODS AND DEVICES FOR AVOIDING LOWER PAGE CORRUPTION IN DATA STORAGE DEVICES

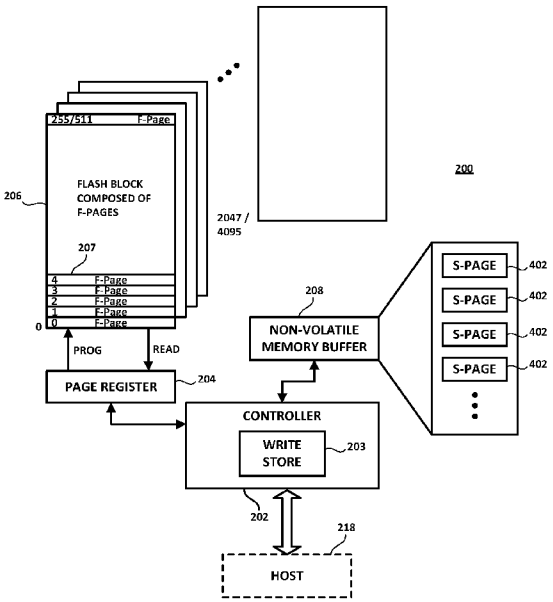

FIG. 2

(57) **Abstract**: A data storage device may comprise a plurality of Multi-Level Ceil (MLC) non- volatile memory devices comprising a plurality of lower pages and a corresponding plurality of higher-order pages. A controller may be configured to write data to and read data from the plurality of lower pages and the corresponding plurality of higher- order pages. A buffer may be coupled to the controller, which may be configured to accumulate data to be written to the MLC non-volatile memory devices, allocate space in the buffer and write the accumulated data to the allocated space, At least a portion of the accumulated data may be written in a lower page of the MLC non- volatile memory devices and the space in the buffer that stores data written to the lower page may be de-allocated when ail higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

WO 2014/077962 A1

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).

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))*

METHODS AND DEVICES FOR AVOIDING LOWER PAGE CORRUPTION IN DATA STORAGE DEVICES

BACKGROUND

[0001] Flash memory is a non-volatile computer storage technology that can be electrically erased and reprogrammed. Flash memory is typically written in blocks and allocated, garbage collected and erased in larger super blocks or S-Blocks.

[0002] Flash memory comprises a plurality of cells, with each cell being configured to store one, two or more bits per cell. SLC is an abbreviation of "Single-Level Cell", which denotes a configuration in which each cell stores one bit. SLC is characterized not only by fast transfer speeds, low power consumption and high cell endurance, but also by relatively high cost. MLC is an abbreviation of "Multi-level Cell", which denotes a configuration in which each cell stores two or more bits per cell. The acronym MLC is often used to denote a Flash memory having cells that store two bits per cell. That same acronym MLC is also used, however, to designated Flash memory having cells configured to store three bits per cell (also called "TLC" or Triple or Three Level Cell) or even a greater number of bits per cell. When MLC is used to designate a memory that stores two bits in each cell, such an MLC Flash memory may be characterized by somewhat slower transfer speeds, higher power consumption and lower cell endurance than a Single-Level Cell memory. Such MLC memories, however, enjoy a comparatively lower manufacturing cost per bit than do SLC memories.

[0003] In MLC NAND Flash memory, the same physical page of memory cells may be used to store two or more logical pages of data, with each cell being configured to store 2 or more bits. When two bits per cell are stored, a first bit of a lower page is stored first, and then the next bit or bits of one or more higher-order pages are stored. The lower page is programmed first, followed by the higher-order page or pages. When programming the upper page, programming voltages are applied to the same cells that already store valid data in the lower page. Should power fail during the programming of the higher-order page or pages, the stored data in the lower page may be irrecoverably corrupted, as may be the data intended to be stored in the higher-order page or pages. This problem is compounded by the fact that the host may have already received an acknowledgment from the data storage

device indicating that the data stored in the lower page has already been saved to the Flash memory.

[0004] Fig. 1 is a block diagram of aspects of conventional Flash data storage device. As shown therein, the Flash data storage device 100 comprises a controller 104. The controller 104 is coupled to an array of non-volatile memory (e.g., Flash memory devices), collectively referenced at numeral 102. Conventionally, to provide power-fail protection, conventional Flash data storage devices include a backup power source, as shown at 106 in Fig. 1. As indicated at 106, super-capacitors or an array of discrete capacitors are conventionally used to maintain the controller 104 and the non-volatile memory 102 powered-up during a power loss, typically only long enough to finish programming the data to the Flash memory devices 102. Indeed, these super-capacitors or array of discrete capacitors are configured to store a sufficiently large amount of energy to enable the controller 104 to complete any firmware operation (such as a write operation) upon power loss. This is not optimal, however, because super-capacitors are large, unreliable, prone to problems and expensive.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Fig. 1 is a block diagram of aspects of a conventional Flash data storage device.

[0006] Fig. 2 is a diagram showing aspects of the physical and logical data organization of a data storage device according to one embodiment.

[0007] Fig. 3 is a block diagram of an S-Block, according to one embodiment.

[0008] Fig. 4 is a block diagram of an S-Page, according to one embodiment.

[0009] Fig. 5 is a block diagram of a data storage device according to one embodiment.

[0010] Fig. 6 is a flowchart of a method of controlling a data storage device according to one embodiment.

DETAILED DESCRIPTION

[0011] Within the scope of the present disclosure, the acronym “MLC”

expressly denotes a Flash memory comprising cells that store two or more bits per cell. In the case wherein such MLC Flash memory is configured to store two bits per cell, data is stored in a lower page and a corresponding upper page. In the case wherein such MLC Flash memory is configured to store three or more bits per cell, data is stored in a lower page and one or more corresponding higher-order pages. The phrase "higher-order pages" is expressly intended to cover the upper page and/or the upper page and one or more pages of higher order.

[0012] Fig. 2 is a diagram showing aspects of the physical and logical data organization of a data storage device 200 according to one embodiment. In one embodiment, the data storage device is an SSD. In another embodiment, the data storage device is a hybrid drive including Flash memory and rotating magnetic storage media. The disclosure is applicable to both SSD and hybrid implementations, but for the sake of simplicity the various embodiments are described with reference to SSD-based implementations. A data storage device controller 202 according to one embodiment may be configured to be coupled to a host, as shown at reference numeral 218. The host 218 may utilize a logical block addressing (LBA) scheme. While the LBA size is normally fixed, the host can vary the size of the LBA dynamically. For example, the physical data storage device may be logically portioned to support partitions configured for LBAs of different sizes. However, such partitions are not required for the physical device to support LBAs of different sizes at the same time. For example, the LBA size may vary by interface and interface mode. Indeed, while 512 bytes is most common, 4 KB is also becoming more common, as are 512+ (520, 528 etc.) and 4 KB+ (4 KB+8, 4K+16 etc.) formats. As shown therein, the data storage device controller 202 may comprise or be coupled to a page register 204. The page register 204 may be configured to enable the controller 202 to read data from and store data to the data storage device 200. The controller 202 may be configured to program and read data from an array of Flash memory devices responsive to data access commands from the host 218. While the description herein refers to Flash memory generally, it is understood that the array of memory devices may comprise one or more of various types of non-volatile memory devices such as Flash integrated circuits, Chalcogenide RAM (C-RAM), Phase Change Memory (PC-RAM or PRAM), Programmable Metallization Cell RAM (PMC-RAM or PMCm), Ovonic Unified

Memory (OUM), Resistance RAM (RRAM), NAND memory (e.g., single-level cell (SLC) memory, multi-level cell (MLC) memory, or any combination thereof), NOR memory, EEPROM, Ferroelectric Memory (FeRAM), Magnetoresistive RAM (MRAM), other discrete NVM (non-volatile memory) chips, or any combination thereof.

[0013] The page register 204 may be configured to enable the controller 202 to read data from and store data to the array. According to one embodiment, the array of Flash memory devices may comprise a plurality of non-volatile memory devices in die (e.g., 128 dies), each of which comprises a plurality of blocks, such as shown at 206 in Fig. 2. Other page registers 204 (not shown), may be coupled to blocks on other die. A combination of Flash Blocks, grouped together, may be called a Superblock or S-Block. In some embodiments, the individual blocks that form an S-Block may be chosen from one or more dies, planes or other levels of granularity. An S-Block, therefore, may comprise a plurality of Flash Blocks, spread across one or more die, that are combined together. In this manner, the S-Block may form a unit on which the Flash Management System (FMS) operates. In some embodiments, the individual blocks that form an S-Block may be chosen according to a different granularity than at the die level, such as the case when the memory devices include dies that are sub-divided into structures such as planes (i.e., blocks may be taken from individual planes). According to one embodiment, allocation, erasure and garbage collection may be carried out at the S-Block level. In other embodiments, the FMS may perform data operations according to other logical groupings such as pages, blocks, planes, dies, etc.

[0014] According to one embodiment, the array of MLC non-volatile memory devices may comprise a plurality of lower pages and a corresponding plurality of higher-order pages. That is, each lower page may be associated with a corresponding single upper page or with a corresponding plurality (i.e., two or more) of higher-order pages. The MLC non-volatile memory devices may be organized in Flash Blocks 206, with each Flash Block comprising a plurality of Flash Pages (F-Pages) 207, as shown in Fig. 2. Alternatively, a different physical organization may be employed. An F-Page, according to one embodiment, may be the size of the minimum unit of program of the non-volatile memory devices. The controller 202 may be coupled to the plurality of non-volatile memory devices and may be

configured to write data to and read data from the plurality of lower pages and to one or more corresponding higher-order pages. To address the lower-page corruption problem afflicting conventional data storage devices, one embodiment of a data storage device 200 comprises a buffer 208. In one embodiment, the buffer comprises a non-volatile memory. As write commands are received from the host 218 and executed by the controller 202, the data to be written to the MLC non-volatile memory devices (i.e., Flash Blocks 206) may be accumulated. According to one embodiment, the data to be written may be accumulated within a write store 203. The write store 203 may be coupled to the controller 202. In one embodiment, the write store 203 is a store that is internal to the controller 202. Contemporaneously with the accumulation of the data in the write store 203 (or at least shortly before or after such accumulation), the controller 202 may allocate memory space in the buffer 208 and may write the accumulated data to the allocated space in the buffer 208. According to one embodiment, the data to be written to the MLC non-volatile memory devices may be stored in both the write store 203 and the buffer 208, such that the buffer 208 and the write store 203 mirror or substantially mirror each other. Indeed, according to one embodiment, the controller 202 may be configured to accumulate data to be written in the write store 203 and in the buffer 208 until a complete F-page is constructed. According to one embodiment, a partial F-Page may be packed with a predetermined coded value and considered to be complete. Completed F-Pages may be written to the MLC non-volatile memory devices. That is, at least a portion of the accumulated data may be written to a lower page of the MLC non-volatile memory devices. According to one embodiment, previously-allocated space in the buffer 208 may be de-allocated when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices. That is, the controller 202 may be configured to keep lower page data in the buffer 208 until the upper page or high-order pages corresponding to the lower page have been programmed. This ensures that a power-safe copy of the write data is maintained in the buffer 208 until both the lower and higher-order page or pages have been programed, after which the data may be considered to be power-fail and corruption safe.

[0015] According to one embodiment, in the normal course of operation of the data storage device 200, as complete F-Pages are constructed (and

accumulated in the write store 203 and written out to the buffer 208), they may be written out to the MLC non-volatile memory devices. According to one embodiment, it is the completed F-Pages stored in the write store 203 that are written out to the MLC non-volatile memory devices. Indeed, it may be preferable to write the completed F-Pages out from the write store 203 rather than from the buffer 208, as the write store 203 may be able to support a greater bandwidth than the buffer 208.

[0016] According to one embodiment, the buffer 208 may comprise non-volatile memory. For example, the buffer 208 may comprise memory that is non-volatile and that is characterized as having high access, read and write speeds. According to one embodiment, the buffer 208 may comprise Magnetic Random Access Memory (MRAM). Other memory types may also be used. MRAM may exhibit performance that is similar to that of SRAM, a density comparable to DRAM and low power consumption. Moreover, MRAM is not known to degrade over time. Although relatively costly, MRAM is well suited to the task of the buffer 208; namely, to store a power-safe copy of lower page data at least until the corresponding higher-order page(s) have been safely stored in the Flash Blocks 206. It is to be noted, however that, as of this writing, the cost of implementing the buffer 208 in MRAM is still less costly than the conventional use of super-capacitors or the use of an array of discrete capacitors. In the implementation in which the buffer 208 comprises non-volatile memory such as MRAM, the data written thereto that has not yet been safely stored in the Flash Blocks 206 (such as lower page data whose corresponding higher-order page(s) were not stored to Flash before a power-fail event) may be read out from the buffer 208 by the controller 202 and stored in the Flash Blocks 206 upon restoration of the power to the data storage device 200. That is, the controller 202 may be further configured to read data from the buffer 208 and write at least a portion of the read data to the non-volatile memory devices of the Flash Blocks 206, after power is restored to the data storage device 200 subsequent to a loss of power thereto.

[0017] According to one embodiment, the buffer 208 may be configured to be at least sufficiently large to enable recovery from lower page corruption after a power loss to the data storage device. The size of the buffer 208, therefore, may vary with, for example, the number of pages between lower and higher-order pages, the size of the F-Pages, the number of planes and dies of the data storage device.

For example, the size of the buffer 208 may vary from a few MB to a few hundreds of MB, although other implementations may utilize other sizes to good effect. According to one embodiment, the buffer 208 may comprise a plurality of buffers for each die, as each die completes its programming at a different time. This enables multiple pages per die to be managed independently, leading to an efficient configuration of the buffer 208. The buffer 208 may, for example, be implemented as a plurality of buffers from which the controller 202 allocates space, stores data and de-allocates space, as the higher-order page(s) of corresponding lower pages are stored in the Flash Blocks 206. Such a buffer configuration is well suited to buffering the stream of write data from host write commands until the probability of lower page corruption upon power fail is acceptably small or zero. According to one embodiment, the controller 202 may be configured to generate and send a write acknowledgement to the host 218 after (e.g., as soon as) the accumulated data is written to the allocated space in the buffer. That is, from the host's perspective, the data may be considered to have been safely stored in Flash as soon as it is stored in the buffer 208. According to one embodiment, the MLC non-volatile devices may be run in lower-page only mode or in "SLC" mode. In that case, since there are no higher-order pages to contend with, de-allocation of space in the buffer 208 may be carried out as soon as the page in MLC lower-page only mode or the page in SLC mode is programmed and need not be delayed while waiting for any higher-ordered pages to be programmed.

[0018] Fig. 3 is a block diagram of an S-Block, according to one embodiment. As shown therein, an S-Block 302 may comprise one Flash block (F-Block) 206 per die. An S-Block, therefore, may be thought of as a collection of F-Blocks, one F-Block per die, that are combined together to form a unit of the Flash Management System (FMS) of the data storage device. According to one embodiment, allocation, erasure and GC may be managed at the S-Block level. Each F-Block 206, as shown in Fig. 3, may comprise a plurality of Flash pages (F-Page) such as, for example, 256 or 512 F-Pages. An F-Page, according to one embodiment, may be the size of the minimum unit of program for a given non-volatile memory device. Fig. 4 shows a super page (S-page), according to one embodiment. As shown therein, an S-page 402 may comprise one F-Page per F-Block of an S-Block, meaning that an S-page spans across an entire S-Block. According to one

embodiment shown in Fig. 2, the data may be accumulated, written and stored in units of S-Pages 402. The buffer 208 may be configured to store write data organized differently than S-Pages 402, depending upon the specific implementation.

[0019] Fig. 5 is a block diagram of another data storage device 500, according to one embodiment. With reference to Fig. 2, like reference numerals denote like elements and the description of such like elements omitted for brevity. In this embodiment, the controller 502 need not (but may) comprise a write store, such as write store 203 in Fig. 2. A volatile memory buffer 504 (or a plurality of such volatile memory buffers) may be coupled to the controller 502. For example, the volatile memory buffer 504 may comprise or be configured in a Dynamic Random Access Memory (DRAM). The volatile memory buffer 504 is shown in Fig. 5 as being external to the controller 502. However, the volatile memory buffer 504 may also be internal to the controller 502, which may translate into faster access times by the controller 502. However, size and/or other considerations may recommend that the volatile memory buffer 504 be configured as an external memory buffer coupled to the controller 502. The data storage device 500 of Fig. 5 may also comprise a backup source of power such as shown at 506 and a non-volatile memory, as shown at 508. The backup source of power 506 may be coupled to the volatile memory buffer 504, to the controller 502 and to the non-volatile memory 508. The backup source of power 506 may be configured to keep at least the volatile memory buffer 504, the controller 502 and/or the non-volatile memory 508 powered-up for a period of time upon a power fail event, without loss of data in the volatile memory 504 (at least until the contents thereof can be saved to the non-volatile memory 508). According to one embodiment, the backup power source 506 may be configured to power at least portions of the data storage device of Fig. 5 at least as long as necessary for the controller 502 to write the data from the volatile buffer 504 to the non-volatile memory 508. The backup power source 506 may comprise capacitors, super-capacitors and/or any energy storage elements. In one embodiment where the data storage device is a hybrid disk drive or a solid state drive coupled with a hard disk drive, the power source 506 may be provided by the BEMF (back electromotive force) generated from the spindle motor of the hard disk drive.

[0020] Indeed, during normal operation, as write commands are

received from the host 218 and executed by the controller 502, the data to be written to the MLC non-volatile memory devices (i.e., Flash Blocks 206) may be written to both the volatile memory buffer 504 and to the MLC non-volatile memory devices. According to one embodiment, the controller 502 may be configured to accumulate data to be written (as directed by write commands issued by the host 218, for example) internally (in a write store 203, for example) until a complete F-page is constructed. Alternatively and according to one embodiment, a partial F-Page may be packed with a predetermined coded value and considered to be complete. Completed F-Pages may then be written both to the volatile memory buffer 504 and to the MLC non-volatile memory devices. According to one embodiment, the data may also be accumulated, written and stored in units of S-Pages 402 or any other data organization unit. Indeed, the volatile memory buffer 504 may be configured to store write data organized differently than S-Pages 402, depending upon the specific implementation. According to one embodiment, therefore, in addition to writing the accumulated data (or a portion thereof) to volatile memory buffer 504, the accumulated data (or a portion thereof) may be written to one or more lower and/or upper pages of the MLC non-volatile memory devices (the Flash Blocks 206). According to one embodiment, previously-allocated space in the volatile memory buffer 504 may be de-allocated and the de-allocated space therein reused for new write data when all higher-order pages corresponding to a previously-programmed lower page have been written in the MLC non-volatile memory devices and thus may be considered to be effectively corruption-safe. That is, the controller 202 may be configured to keep lower page data in the volatile memory buffer 504 at least until the upper page or high-order pages corresponding to the lower page have been programmed in the MLC non-volatile memory devices.

[0021] In the event of a power loss, the backup power source 506 may supply power at least to the controller 502, the volatile memory buffer 504 and/or the non-volatile memory 508. During the time the controller 502, the volatile memory buffer 504 and/or the non-volatile memory 508 are powered by the backup power source 506, the controller 502 may cause data stored in the volatile memory buffer 504 to be copied to the non-volatile memory 508, thereby saving the data that has not yet been saved to the MLC non-volatile memory devices in a corruption safe manner, thereby enabling the controller to acknowledge the write to the host 218.

When power to the MLC non-volatile memory devices is restored, the data saved in the non-volatile memory 508 may be programmed into the MLC non-volatile memory devices.

[0022] It is to be noted that the non-volatile memory 508 may, according to one embodiment, be written to only in the event of a power failure. Moreover, by powering only the controller 502, the volatile memory buffer 504 and the non-volatile memory 508 in the event of a power failure, comparatively less power is required than would be required to also power the dies of the MLC non-volatile memory devices (e.g., Flash Blocks 206). It is to be noted that the non-volatile memory 508 may draw its power from the controller 502 that is powered by the backup power source 506. According to one embodiment, the non-volatile memory 508 may comprise MRAM. According to one embodiment, the backup power source 506 need only be coupled to the controller 502 and to the volatile memory buffer 504. In addition, by using the volatile memory buffer 504 as the primary write location in the data path and the non-volatile memory 508 in the event of a power failure, the wear on the non-volatile memory 508 is reduced.

[0023] Fig. 6 is a flowchart of a method of controlling a data storage device according to one embodiment. The data storage device may comprise a buffer and a plurality of Multi-Level Cell (MLC) non-volatile memory devices (comprising, e.g., the Flash Blocks 206). The MLC non-volatile memory devices may comprise, as described herein, a plurality of lower pages and a corresponding plurality of higher-order pages. A controller, such as shown at 502, may be configured to write and read data to and from the plurality of lower pages and the corresponding plurality of higher-order pages. According to one embodiment and as shown in Fig. 6, the method may comprise accumulating data to be written to the MLC non-volatile memory devices, as shown at Block B61. Block B62 calls for allocating space in a buffer 208, 504 and writing the accumulated data to the allocated space in the buffer 208, 504. As shown at B63, at least a portion of the accumulated data may be written to one or more lower pages of the MLC non-volatile memory devices. It is understood that a portion of the accumulated data may also be written to one or more upper pages of the MLC non-volatile memory devices. When all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices, space in the buffer that stores data written to

the lower page may be safely de-allocated, as called for in Block B64.

[0024] According to one embodiment and as shown in Fig. 2, the buffer may comprise or be configured as a non-volatile memory such as, for example, Magnetic Random Access Memory (MRAM). As described relative to Fig. 2 and as shown at B65 in Fig. 5, the method may also comprise reading data from the buffer 208 and writing at least a portion of the read data to the MLC non-volatile memory devices, after power is restored to the data storage device subsequent to a loss thereof. As described relative to Fig. 5, the method may also comprise reading data from the volatile memory buffer 504 and writing at least a portion of the read data to the non-volatile memory 508, after power is restored (e.g., by the backup power source 506) to the data storage device after a power loss, as shown at B65 in Fig. 6. According to one embodiment, the buffer 208, 504 may be at least sufficiently large to enable recovery from a possible lower page corruption after a power loss to the data storage device. According to one embodiment, the buffer 208, 504 may be configured as one or more buffers. A write acknowledgment to the host 218 may be generated and sent after the accumulated data is written to the allocated space in the buffer 208, 504, as the data may be considered, from that point in time forward, to be corruption-safe. According to one embodiment, the MLC non-volatile memory devices may be configured to comprise a plurality of blocks, each of which comprising a plurality of physical pages. A collection of such blocks may define a superblock (S-Block). A collection of physical pages with one physical page per block in an S-Block may define a superpage (S-Page), which may be the unit by which the controller 202 accumulates, writes and stores host and/or other data. Other data organizations and units may be implemented within the present context.

[0025] As shown in and described relative to Fig. 5, the method may also comprise providing a backup source of power 506 and a non-volatile memory, such as a non-volatile memory, such as shown at 508. The buffer 504, in this embodiment, may comprise volatile memory. The method may, according to one embodiment, further comprise powering at least a portion of the data storage device 500 for at least as long as necessary for the controller 502 to write the data from the buffer 504 to the non-volatile memory 508. According to one embodiment, after power is restored to the data storage device after a loss thereof, at least a portion of the data in the non-volatile memory 508 may be written to the MLC non-volatile

memory devices. A write store (203 in Figs. 2 and 5) may be provided, and the controller 202, 502 may be configured to write the accumulated data to the write store 203 as the accumulated data is written to the allocated space in the buffer 208, 504. According to one embodiment, the controller 202, 502 may also be configured to de-allocate space, in the buffer 208, 504, that stores data written to the higher-order pages when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices, as such data may be considered to be corruption-safe.

[0026] While certain embodiments have been described, these embodiments have been presented by way of example only, and are not intended to limit the scope of the present disclosure. Indeed, the novel methods, devices and systems described herein may be embodied in a variety of other forms. Furthermore, various omissions, substitutions and changes in the form of the methods and systems described herein may be made without departing from the spirit of the present disclosure. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope and spirit of the present disclosure. For example, those skilled in the art will appreciate that in various embodiments, the actual structures may differ from those shown in the figures. Depending on the embodiment, certain of the steps described in the example above may be removed, others may be added. Also, the features and attributes of the specific embodiments disclosed above may be combined in different ways to form additional embodiments, all of which fall within the scope of the present disclosure. Although the present disclosure provides certain preferred embodiments and applications, other embodiments that are apparent to those of ordinary skill in the art, including embodiments which do not provide all of the features and advantages set forth herein, are also within the scope of this disclosure. Accordingly, the scope of the present disclosure is intended to be defined only by reference to the appended claims.

CLAIMS:

1. A data storage device, comprising:

a plurality of Multi-Level Cell (MLC) non-volatile memory devices comprising a plurality of lower pages and a corresponding plurality of higher-order pages;

a controller coupled to the plurality of MLC non-volatile memory devices and configured to write data to and read data from the plurality of lower pages and the corresponding plurality of higher-order pages; and

a buffer coupled to the controller;

wherein the controller is configured to:

accumulate data to be written to the MLC non-volatile memory devices;

allocate space in the buffer and write the accumulated data to the allocated space in the buffer;

write at least a portion of the accumulated data in a lower page of the MLC non-volatile memory devices; and

de-allocate space in the buffer that stores data written to the lower page when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

2. The data storage device of claim 1, wherein the buffer comprises non-volatile memory.

3. The data storage device of claim 2, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

4. The data storage device of claim 1, wherein the controller is further configured to read data from the buffer and write at least a portion of the read data to the MLC non-volatile memory devices, after power is restored to the data storage device after a loss of power.

5. The data storage device of claim 1, wherein the buffer is at least sufficiently large to enable recovery from lower page corruption after a power loss to the data storage device.

6. The data storage device of claim 1, wherein the MLC non-volatile devices are configured to operate in lower page only mode or in Single Level Cell (SLC) mode.

7. The data storage device of claim 1, wherein the controller is further configured to generate and send a write acknowledgement to a host after the accumulated data is written to the allocated space in the buffer.

8. The data storage device of claim 1, wherein the MLC non-volatile memory devices comprise a plurality of blocks, each of the plurality of blocks comprising a plurality of physical pages, a collection of blocks defining a superblock (S-Block), a collection of physical pages with one physical page per block in an S-Block defining a superpage (S-Page), and wherein the data is accumulated, written and stored in units of S-Pages.

9. The data storage device of claim 1, further comprising a backup source of power and a non-volatile memory, wherein the buffer comprises volatile memory and wherein the backup source of power is configured to power at least a portion of the data storage device at least as long as necessary for the controller to write the data from the buffer to the non-volatile memory.

10. The data storage device of claim 9, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

11. The data storage device of claim 9, wherein the controller is further configured to write at least a portion of the data in the non-volatile memory to the MLC non-volatile memory devices, after power is restored to the data storage device after a loss of power.

12. The data storage device of claim 1, further comprising a write store and wherein the controller is further configured to write the accumulated data to the write store as the accumulated data is written to the allocated space in the buffer.

13. The data storage device of claim 1, wherein controller is further configured to also de-allocate space in the buffer that stores data written to the higher-order pages when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

14. A data storage device controller, comprising:

a processor configured to couple to a buffer and to a plurality of Multi-Level Cell (MLC) non-volatile memory devices that comprise a plurality of lower pages and a corresponding plurality of higher-order pages, the processor being further configured to:

read data from the plurality of lower pages and the corresponding plurality of higher-order pages; and

write data to the plurality of lower pages and the corresponding plurality of higher-order pages by at least:

accumulating data to be written to the MLC non-volatile memory devices;

allocating space in the buffer and writing the accumulated data to the allocated space in the buffer;

writing at least a portion of the accumulated data in a lower page of the MLC non-volatile memory devices; and

de-allocating space in the buffer that stores data written to the lower page when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

15. The data storage device controller of claim 14, wherein the buffer comprises non-volatile memory.

16. The data storage device controller of claim 15, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

17. The data storage device controller of claim 14, wherein the processor is further configured to read data from the buffer and write at least a portion of the read data to the MLC non-volatile memory devices, after power is restored to the

data storage device after a loss of power.

18. The data storage device controller of claim 14, wherein the buffer is at least sufficiently large to enable recovery from lower page corruption after a power loss to the data storage device.

19. The data storage device controller of claim 14, wherein the MLC non-volatile devices are configured to operate in lower page only mode or in Single Level Cell (SLC) mode.

20. The data storage device controller of claim 14, wherein the processor is further configured to generate and send a write acknowledgement to a host after the accumulated data is written to the allocated space in the buffer.

21. The data storage device controller of claim 14, wherein the MLC non-volatile memory devices comprise a plurality of blocks, each of the plurality of blocks comprising a plurality of physical pages, a collection of blocks defining a superblock (S-Block), a collection of physical pages with one physical page per block in an S-Block defining a superpage (S-Page), and wherein the data is accumulated, written and stored in units of S-Pages.

22. The data storage device controller of claim 14, further comprising a backup source of power and a non-volatile memory, wherein the buffer comprises volatile memory and wherein the backup source of power is configured to power at least a portion of the data storage device at least as long as necessary for the controller to write the data from the buffer to the non-volatile memory.

23. The data storage device controller of claim 22, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

24. The data storage device controller of claim 22, wherein the processor is further configured to write at least a portion of the data in the non-volatile memory to the MLC non-volatile memory devices, after power is restored to the data storage

device after a loss of power.

25. The data storage device controller of claim 14, further comprising a write store and wherein the processor is further configured to write the accumulated data to the write store as the accumulated data is written to the allocated space in the buffer.

26. The data storage device controller of claim 14, wherein processor is further configured to also de-allocate space in the buffer that stores data written to the higher-order pages when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

27. A method of controlling a data storage device, the data storage device comprising a buffer and a plurality of Multi-Level Cell (MLC) non-volatile memory devices that comprise a plurality of lower pages and a corresponding plurality of higher-order pages, the method comprising:

reading data from the plurality of lower pages and the corresponding plurality of higher-order pages; and

writing data to the plurality of lower pages and the corresponding plurality of higher-order pages by at least:

accumulating data to be written to the MLC non-volatile memory devices;

allocating space in the buffer and writing the accumulated data to the allocated space in the buffer;

writing at least a portion of the accumulated data in a lower page of the MLC non-volatile memory devices; and

de-allocating space in the buffer that stores data written to the lower page when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

28. The method of claim 27, wherein the buffer comprises non-volatile memory.

29. The method of claim 28, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

30. The method of claim 27, further comprising reading data from the buffer and writing at least a portion of the read data to the MLC non-volatile memory devices, after power is restored to the data storage device after a loss of power.

31. The method of claim 27, wherein the buffer is at least sufficiently large to enable recovery from lower page corruption after a power loss to the data storage device.

32. The method of claim 27, wherein the MLC non-volatile devices are configured to operate in lower page only mode or in Single Level Cell (SLC) mode.

33. The method of claim 27, further comprising generating and sending a write acknowledgement to a host after the accumulated data is written to the allocated space in the buffer.

34. The method of claim 27, wherein the MLC non-volatile memory devices comprise a plurality of blocks, each of the plurality of blocks comprising a plurality of physical pages, a collection of blocks defining a superblock (S-Block), a collection of physical pages with one physical page per block in an S-Block defining a superpage (S-Page), and accumulating, writing and storing is carried out in units of S-Pages.

35. The method of claim 27, further comprising a backup source of power and a non-volatile memory, wherein the buffer comprises volatile memory and wherein the method further comprises the backup source of power powering at least a portion of the data storage device at least as long as necessary for the controller to write the data from the buffer to the non-volatile memory.

36. The method of claim 35, wherein the non-volatile memory comprises Magnetic Random Access Memory (MRAM).

37. The method of claim 35, further comprising writing at least a portion of the data in the non-volatile memory to the MLC non-volatile memory devices, after power is restored to the data storage device after a loss of power.

38. The method of claim 27, wherein the data storage device further comprises a write store and wherein writing data further comprises writing the accumulated data to the write store as the accumulated data is written to the allocated space in the buffer.

39. The method of claim 27, further comprising de-allocating space in the buffer that stores data written to the higher-order pages when all higher-order pages corresponding to the lower page have been written in the MLC non-volatile memory devices.

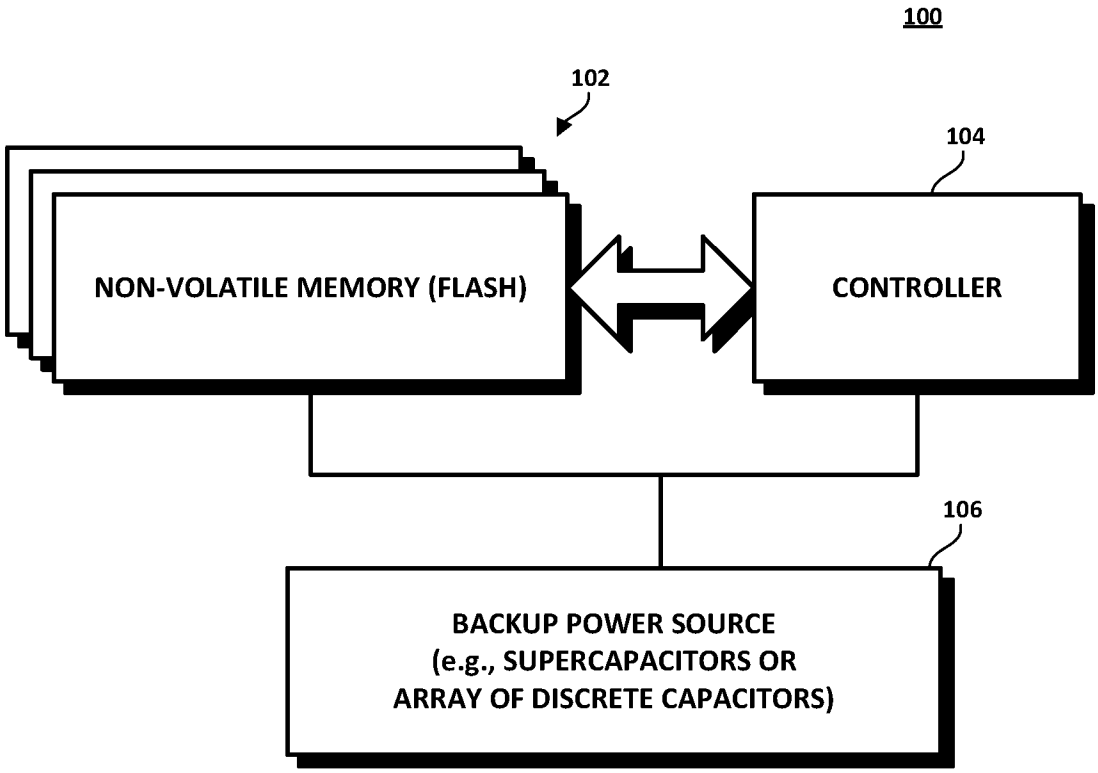

FIG. 1

(Prior Art)

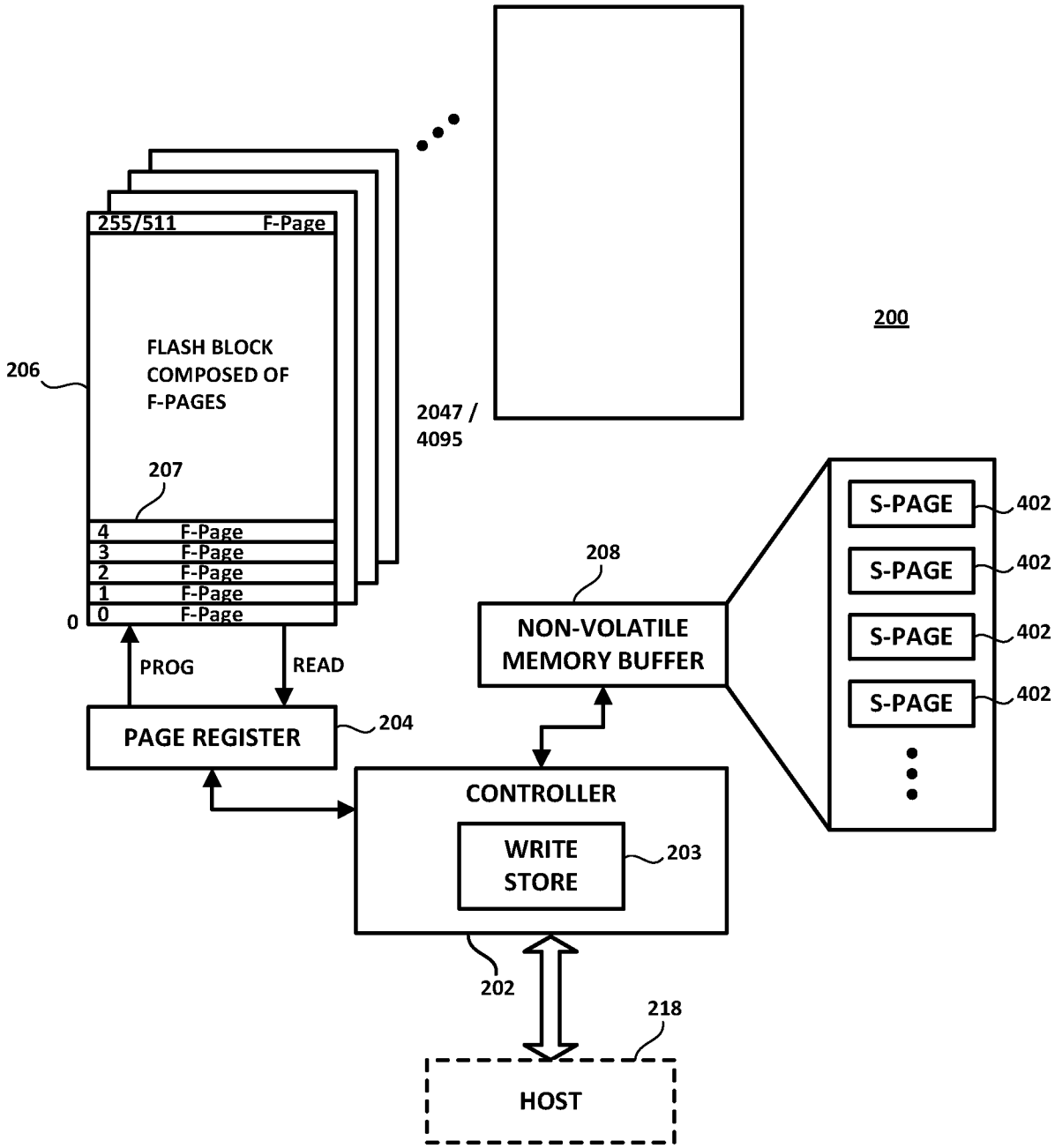
200
2047 / 4095
255/511 F-Page
FLASH BLOCK COMPOSED OF F-PAGES
206
207
4 F-Page
3 F-Page
2 F-Page
1 F-Page
0 F-Page
0
PROG
READ
PAGE REGISTER
204
208
NON-VOLATILE MEMORY BUFFER
S-PAGE
S-PAGE
S-PAGE
S-PAGE
402
402
402
402
CONTROLLER
WRITE STORE
203
202
218
HOST

FIG. 2

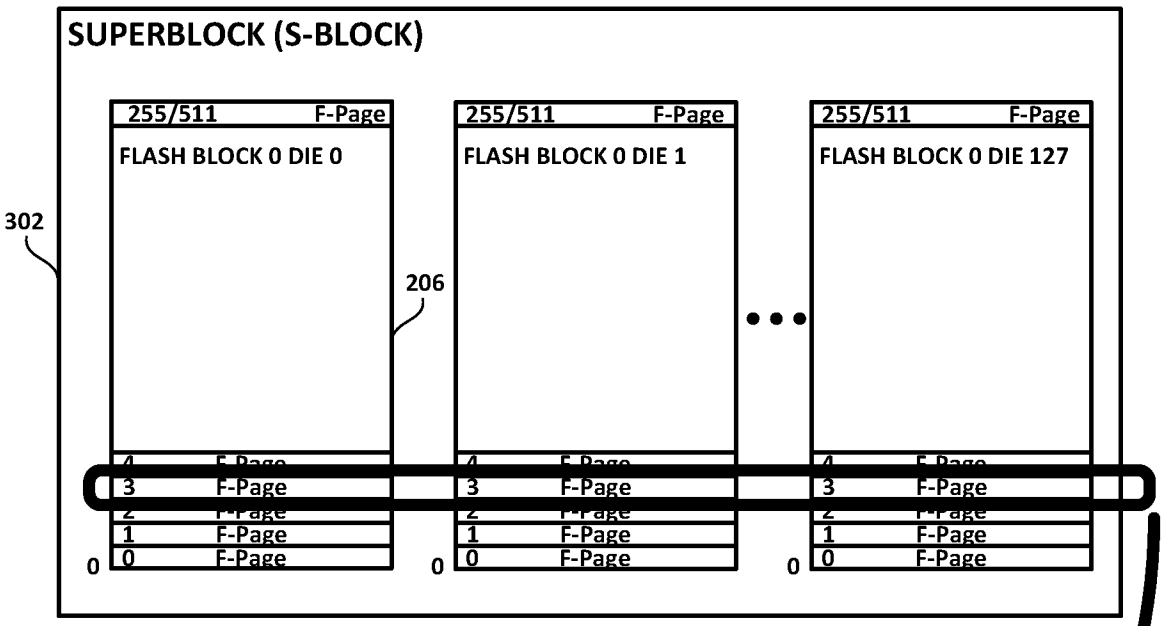
SUPERBLOCK (S-BLOCK)
255/511 F-Page
FLASH BLOCK 0 DIE 0
FLASH BLOCK 0 DIE 1
FLASH BLOCK 0 DIE 127
302
206
3 F-Page
1 F-Page
0 F-Page
0
0
0

FIG. 3

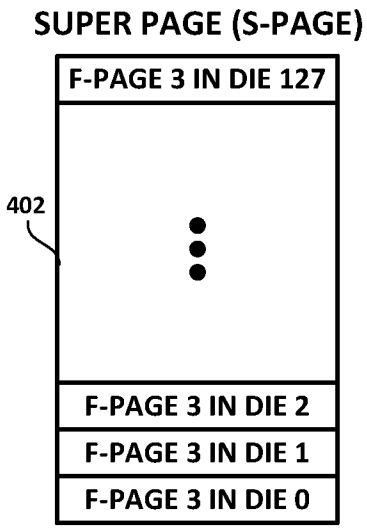
SUPER PAGE (S-PAGE)
F-PAGE 3 IN DIE 127
402
F-PAGE 3 IN DIE 2
F-PAGE 3 IN DIE 1
F-PAGE 3 IN DIE 0

FIG. 4

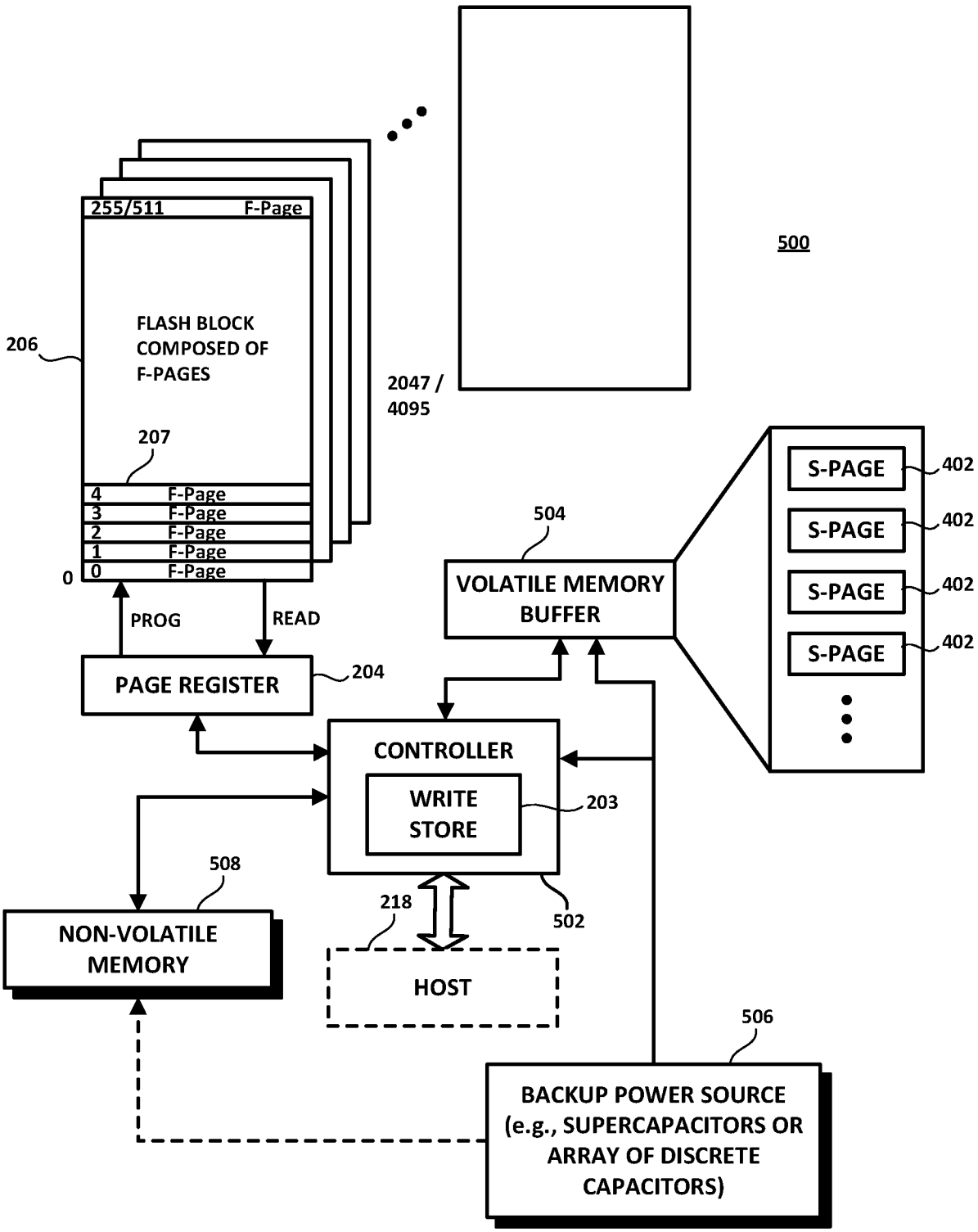
500
2047 / 4095
255/511 F-Page
FLASH BLOCK COMPOSED OF F-PAGES
206
207
4 F-Page
3 F-Page
2 F-Page
1 F-Page
0 F-Page
0
PROG
READ
PAGE REGISTER
204
504
VOLATILE MEMORY BUFFER
S-PAGE
S-PAGE
S-PAGE
S-PAGE
402
402
402
402
CONTROLLER
WRITE STORE
203
508
NON-VOLATILE MEMORY
218
502
HOST
506
BACKUP POWER SOURCE (e.g., SUPERCAPACITORS OR ARRAY OF DISCRETE CAPACITORS)

FIG. 5

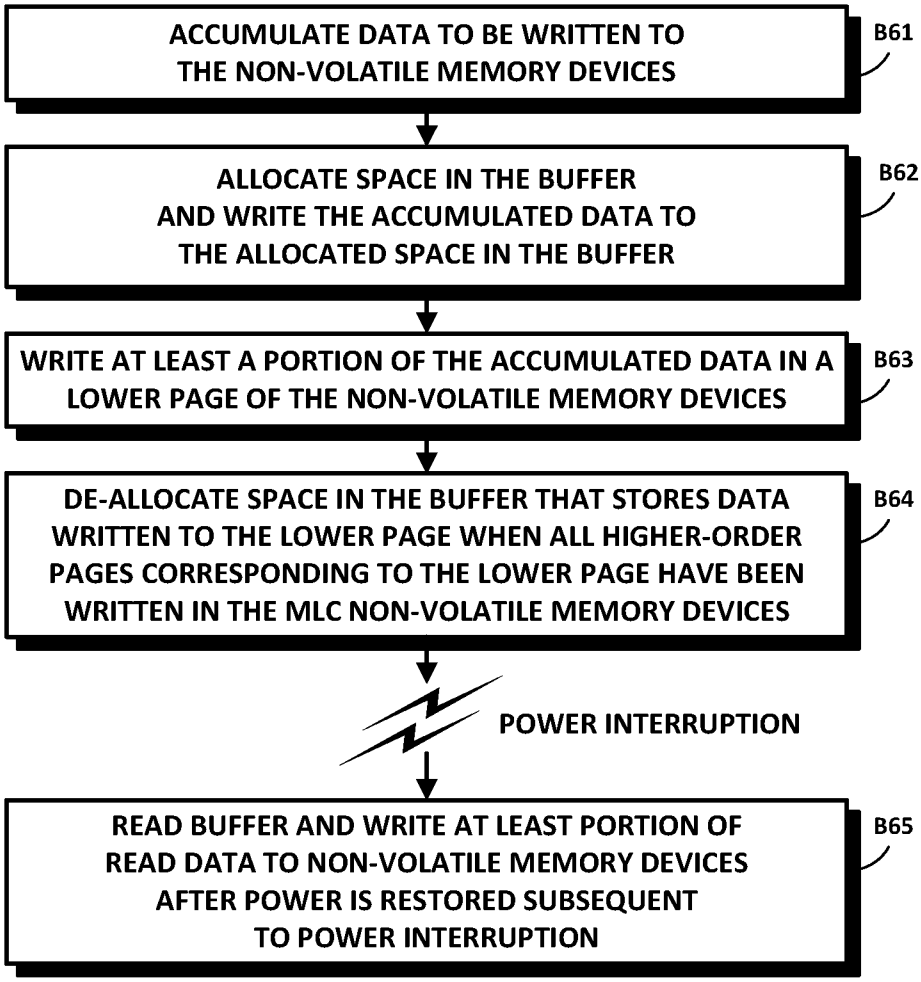
ACCUMULATE DATA TO BE WRITTEN TO THE NON-VOLATILE MEMORY DEVICES
B61
ALLOCATE SPACE IN THE BUFFER AND WRITE THE ACCUMULATED DATA TO THE ALLOCATED SPACE IN THE BUFFER
B62
WRITE AT LEAST A PORTION OF THE ACCUMULATED DATA IN A LOWER PAGE OF THE NON-VOLATILE MEMORY DEVICES
B63
DE-ALLOCATE SPACE IN THE BUFFER THAT STORES DATA WRITTEN TO THE LOWER PAGE WHEN ALL HIGHER-ORDER PAGES CORRESPONDING TO THE LOWER PAGE HAVE BEEN WRITTEN IN THE MLC NON-VOLATILE MEMORY DEVICES
B64
POWER INTERRUPTION
READ BUFFER AND WRITE AT LEAST PORTION OF READ DATA TO NON-VOLATILE MEMORY DEVICES AFTER POWER IS RESTORED SUBSEQUENT TO POWER INTERRUPTION
B65

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/062725

A. CLASSIFICATION OF SUBJECT MATTER

G06F 12/00(2006.01)i, G06F 12/08(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 12/00; G06F 12/02; H03M 13/05; G06F 11/10; G06F 11/14; 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: multi level cell, buffer, memory, power, loss, corrupt, interrupt, allocate, lower, higher, page, read, write, space, accumulate, and similar terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2012-0054582 A1 (MATTHEW BYOM et al.) 01 March 2012 See paragraphs [0021]-[0047] and [0061]-[0076]; claims 9 and 18; and figures 1-2 and 6.	1-39
A	US 2011-0202710 A1 (QUN ZHAO et al.) 18 August 2011 See paragraphs [0018]-[0023]; claims 1-2; and figures 1-2.	1-39
A	US 2010-0318839 A1 (CHRIS NGA YEE AVILA et al.) 16 December 2010 See paragraphs [0044]-[0046]; claim 12; and figures 6A-6C.	1-39
A	US 2008-0189473 A1 (MICHAEL MURRAY) 07 August 2008 See paragraphs [0038]-[0039]; claim 26; and figures 6-7.	1-39
A	US 2012-0166720 A1 (FRANKIE F. ROOHPARVAR) 28 July 2012 See paragraphs [0015]-[0018] and figure 1.	1-39

☐ 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	Date of mailing of the international search report
27 January 2014 (27.01.2014)	**28 January 2014 (28.01.2014)**

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea
Facsimile No. +82-42-472-7140

Authorized officer
NHO, Ji Myong
Telephone No. +82-42-481-8528

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2013/062725

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012-0054582 A1	01/03/2012	US 8341500 B2	25/12/2012
US 2011-0202710 A1	18/08/2011	CN 102792381 A	21/11/2012
		US 8549214 B2	01/10/2013
		WO 2011-103168 A1	25/08/2011
US 2010-0318839 A1	16/12/2010	US 8307241 B2	06/11/2012
US 2008-0189473 A1	07/08/2008	CN 101641679 A	03/02/2010
		KR 10-2009-0117787 A	12/11/2009
		TW 200847170 A	01/12/2008
		TW I395223 B	01/05/2013
		WO 2008-097617 A2	14/08/2008
		WO 2008-097617 A3	19/03/2009
US 2012-0166720 A1	28/06/2012	US 2007-288702 A1	13/12/2007
		US 2009-327594 A1	31/12/2009
		US 7586784 B2	08/09/2009
		US 8134872 B2	13/03/2012