

(51) International Patent Classification:
G06F 12/16 (2006.01)(21) International Application Number:
PCT/US2014/041786(22) International Filing Date:
10 June 2014 (10.06.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/833,381 10 June 2013 (10.06.2013) US
14/010,350 26 August 2013 (26.08.2013) US

(71) Applicant: WESTERN DIGITAL TECHNOLOGIES, INC. [US/US]; 3355 Michelson Drive, Suite 100, Irvine, California 92612 (US).

(72) Inventors: DEFOREST, Marvin R.; 3355 Michelson Drive, Suite 100, Irvine, California 92612 (US). HORN, Robert L.; 3355 Michelson Drive, Suite 100, Irvine, California 92612 (US).

(74) Agents: CHAN, David H. et al.; Western Digital Technologies, Inc., 3355 Michelson Drive, Suite 100, Irvine, California 92612 (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, IR, 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, 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))

[Continued on next page]

(54) Title: MIGRATION OF ENCRYPTED DATA FOR DATA STORAGE SYSTEMS

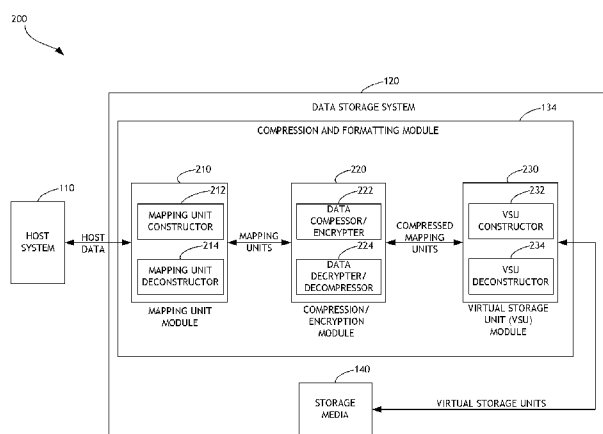


FIGURE 2

(57) Abstract: Systems and methods for compression, formatting, and migration of data for data storage systems are disclosed. In some embodiments, data repacking can be used in any situation where embedded metadata needs to be accessed, such as during data migration, and where the underlying data is encrypted. In some embodiments, performance is increased because encrypted data is repacked without first performing decryption. In addition, data may also be compressed and repacking can be performed without performing decompression. Advantageously, there is no need to retrieve or wait for the availability of encryption key (or keys) or expand resources in decrypting (and decompressing) data before repacking it and encrypting repacked data. Available capacity for storing user data, reliability, and performance of the data storage system can be increased.



-
- *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))*

MIGRATION OF ENCRYPTED DATA FOR DATA STORAGE SYSTEMS

BACKGROUND

Technical Field

[0001] This disclosure relates to data storage systems for computer systems. More particularly, the disclosure relates to migration of encrypted data for storage systems.

Description of the Related Art

[0002] Modern computer systems continually require increases in data storage system capacity, reliability, and performance. In 1981, a typical hard disk drive used in a personal computer had capacity of about 10 megabytes (MB). Today, capacities of hard disk drives are measured in terabytes (TB). Shingled magnetic recording is utilized in order to increase capacity of magnetic storage. In addition, modern disk drives often include solid state memory to increase performance. It is desirable to provide mechanisms and architectures for increasing capacity, reliability, and performance of data storage systems.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] Systems and methods that embody the various features of the invention will now be described with reference to the following drawings, in which:

[0004] Fig. 1 illustrates a combination of a host system and a data storage system according to some embodiments of the invention.

[0005] Fig. 2 illustrates compression and formatting utilized by a data storage system according to some embodiments of the invention.

[0006] Fig. 3 illustrates compression and formatting data flow utilized by a data storage system according to some embodiments of the invention.

[0007] Fig. 4 illustrates migration of compressed and encrypted data according to some embodiments of the invention.

[0008] Fig. 5 is a flow diagram illustrating a process 500 of migrating encrypted data according to some embodiments of the invention.

DETAILED DESCRIPTION

[0009] While certain embodiments are described, these embodiments are presented by way of example only, and are not intended to limit the scope of protection. Indeed, the novel methods 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 scope of protection.

Overview

[0010] Data storage systems can utilize compression techniques in order to meet the demand for increased capacity, reliability, and performance. Compressing host system data before storing it in a data storage system typically causes compressed data to have smaller size than the original, non-compressed data provided by a host system. Accordingly, fixed sized host data units provided by the host system become variable sized data units after compression. Numerous challenges exist with storing, moving, and retrieving such variable sized data. For example, storing variable sized data in a data storage system media may necessitate making considerable changes to the way the data storage system stores data and keeps track of the stored data. As another example, because the host system usually expects fixed size data to be returned during data retrieval, considerable changes to data retrieval mechanisms may also be necessitated.

[0011] In some embodiments, mechanisms for compression, formatting, and migration of host data (or user data) for storing in data storage systems are provided. A data storage system can encrypt and compress fixed sized data before storing it on a media and format obtained variable sized compressed data for storing on the media that typically has fixed size storage granularity, such as sectors for magnetic media and pages for non-volatile solid-state memory. One or more modules can compress incoming host data and create an output stream of fixed size “virtual” storage units (or storage units) containing one or more compressed host data units, error correction and detection code (EDC) information, and metadata for

decoding the logical contents of the storage units. Virtual storage units serve as fixed sized containers for compressed host data.

[0012] In some embodiments, during writing of data to a data storage system, uncompressed, fixed sized host data units are grouped by adjacent logical addresses, stripped of EDC data, and encrypted and compressed into one or more variable sized compressed data units. In order to maintain data protection and error detection and recovery, uncompressed data of a compressed data unit is used to determine new EDC that is appended to the compressed data unit. The compression logic also creates compressed data unit metadata that may include host logical addresses and length of compressed data. Compressed payload can include the metadata, determined EDC, and compressed host data. The metadata may be uncompressed and unencrypted for subsequent repacking and migration. One or more compressed payloads, which can include one or more full or partial payloads, can be appended to create a fixed size virtual storage unit. The virtual storage unit can also include metadata, which may be uncompressed and unencrypted, for decoding the logical contents of the virtual storage unit. For example, the virtual storage unit metadata can include an offset of the first compressed mapping unit, a virtual data unit block address field, and a sequence number. Additional EDC data can be appended to the virtual storage unit. The virtual storage unit can be moved from a volatile memory buffer to the non-volatile media.

[0013] In some embodiments, one or more virtual storage units can be moved or migrated between different types of media in a data storage system. Migration of stored data can involve utilizing metadata to repack and/or reformat compressed (and optionally encrypted) data unit(s) for storage on a different media type. The metadata can be uncompressed and unencrypted so that migration is advantageously performed without decrypting or decompressing the data.

System Overview

[0014] Fig. 1 illustrates a combination 100 of a host system and a data storage system according to some embodiments of the invention. As is shown, the data storage system 120 (e.g., a hybrid hard drive) includes a controller 130 and

storage media 140. Media 140 includes a non-volatile solid-state memory array 150. The solid-state memory array 150 may comprise 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. Media 140 also includes magnetic storage 160, which comprises magnetic recording media 164. In one embodiment, magnetic storage 160 utilizes shingled magnetic recording. Storage media 140 can further comprise other types of storage. In some embodiments, the storage media 140 can comprise only one of solid-state memory array 150 or the magnetic storage 160.

[0015] The controller 130 can be configured to receive data and/or storage access commands from a storage interface module 112 (e.g., a device driver) of a host system 110. Storage access commands communicated by the storage interface module 112 can include write data and read data commands issued by the host system 110. Read and write commands can specify a logical address (e.g., logical block addresses or LBAs) used to access the data storage system 120. The controller 130 can execute the received commands in the media 140.

[0016] Data storage system 120 can store data communicated by the host system 110. In other words, the data storage system 120 can act as memory storage for the host system 110. To facilitate this function, the controller 130 can implement a logical interface. The logical interface can present to the host system 110 data storage system's memory as a set of logical addresses (e.g., contiguous address) where user data can be stored. Internally, the controller 130 can map logical addresses to various physical locations or addresses in the media 140. The controller 130 includes a buffer 132, which can comprise volatile memory, such as random access memory (RAM) (e.g., DRAM, SRAM, SDRAM, DDR, etc.). In one embodiment, the buffer 132 is external to the controller 130. The controller 130 also includes a compression, formatting, and data migration module 134 configured to compress, optionally encrypt, and format host data for storing in the media 140,

repack and migrate compressed and encrypted data, and decompress, decrypt, and format stored data for providing to the host system 110. The compression, formatting, and data migration module 134 can utilize the buffer 132 for storing and retrieving data while performing compression, optional encryption, migration, repacking, and/or formatting tasks. In one embodiment, the compression, formatting, and data migration module 134 is implemented in hardware, such as a system on-chip (SoC) hardware module. Hardware implementation can be realized using an application specific integrated circuit (ASIC), field-programmable gate array (FPGA), or a combination thereof. In another embodiment, the compression, formatting, and data migration module 134 can be implemented in firmware or as a combination of hardware and firmware.

[0017] In some embodiments, the host system 110 stores in the data storage system 120 and the data storage system returns to the host system data formatted as one or more logical host data units. Logical host data units can span ranges of logical addresses, with each logical address corresponding to a physical location in the media 140. Host data units can be fixed sized data units of any suitable size, such as 512 bytes, 1024 bytes (1 KB), 2048 bytes (2 KB), 4096 bytes (4 KB), and the like. In one embodiment, 512 byte data units or sectors are utilized by the host system 110. In some embodiments, the media 140 is divided into fixed sized physical media storage units (or physical storage units or memory data units or memory data unit containers) of a suitable size, such as 512 bytes, 1024 bytes (1 KB), 2048 bytes (2 KB), 4096 bytes (4 KB), and the like. Physical media storage units can correspond to the smallest programmable and readable portion of memory and can span physical addresses of individual memory locations. In some embodiments, different media types include media storage units of different sizes. For example, the solid-state memory array 150 can be divided into dies, blocks, and pages, such as 2 KB pages, and the magnetic recording media 164 can be divided into tracks, clusters, and sectors, such as 512 byte sectors. In certain embodiments, a particular media type utilizes one or more media storage units of different sizes. For instance, the solid-state memory array 150 can utilize media storage units of various sizes due to a variable size of error detection and correction data (e.g., ECC). For example, a physical page have fixed 2 KB size, but because ECC data

varies in size depending on the desired error detection and correction strength, storage space for storing host data in a page is also variable (e.g., 2 KB – ECC data size).

Compression and Formatting

[0018] Fig. 2 illustrates compression and formatting 200 utilized by the data storage system 120 according to some embodiments of the invention. The host system 110 provides host data (or user data) for storing in the data storage system 120. The host system 110 can provide host data in form of fixed sized host data units, such as 512 byte logical sectors. The host system 110 can provide host data as part of one or more write data commands. On the data storage system 120 side, the compression, formatting, and data migration module 134 compresses and formats host data units for storage in the media 140.

[0019] In some embodiments, host data units are processed by a mapping unit module 210. A mapping unit can be selected as an integer multiple of the host data unit size, such as $N \times s_l$ or s_l/N , where s_l is the host data unit size (e.g., 512 bytes) and N is an integer. The size of the mapping unit can be selected so as to achieve an optimal or substantially optimal balance between reduction of storage space gained from compression and latency associated with accessing compressed and formatted data stored in the media. In some embodiments, a mapping unit constructor 212 assembles one or more host data units into one or more mapping units. In some embodiments, the mapping unit constructor 212 assembles consecutive or sequential host data units into one or more mapping units. In one embodiment, consecutive or sequential host data units are associated with sequential logical addresses. In addition, aligned host data units of the size equal to the integer multiple size of the host data unit size can be stored into a mapping unit. For example, suppose that the host system 110 sends write data commands with host data units corresponding to 512-byte logical sectors 0-7, 8-15, and 16-23. Suppose that the mapping unit constructor 212 is configured to assemble aligned 4 KB of host data (or eight aligned 512-byte logical sectors) into a mapping unit, where “aligned” refers to alignments of the starting (and ending) logical addresses as well as the size. For the above example, the mapping unit constructor 212

generates three mapping units containing host data units for logical sectors 0-7, 8-15, and 16-23. Continuing the previous example, suppose that the host system 110 sends write data commands with host data units corresponding to logical sectors 1-8. These host data units are associated with eight logical sectors and have 4 KB aggregate size. However, these host data units are not aligned with chunks of eight logical sectors (e.g., a media storage unit boundary line is crossed). That is, the starting logical sector of these host data units is 1, not 0. Accordingly, the mapping unit constructor 212 generates two separate partial mapping units as discussed below: the first corresponding to logical sectors 1-7 and the second corresponding to logical sector 8. In some embodiments, the mapping unit constructor 212 assembles non-aligned host data units into mapping units. In some embodiments, the mapping unit constructor 212 assembles non-consecutive or consecutive and non-consecutive host data units into mapping units.

[0020] In some embodiments, the mapping unit constructor 212 generates partial mapping units (or PMUs) that comprise less host data than the selected or target mapping unit size. For example, suppose that the host system sends write data commands with host data units corresponding to 512-byte logical sectors 0-10, and that the mapping unit constructor 212 is configured to assemble aligned 4 KB of host data (or eight aligned 512-byte logical sectors) into a mapping unit. The mapping unit constructor 212 generates a first full or complete mapping unit corresponding to host data units for logical sectors 0-7 and a second partial mapping unit corresponding to host data units for logical sectors 8-10. Partial mapping unit corresponding to data associated with logical sector 8-10 can be later integrated into a full mapping unit when the host system 110 writes data into logical sectors 11-15.

[0021] In some embodiments, mapping units (full or partial) constructed by the mapping unit module 210 are compressed, and optionally encrypted, by the compression and encryption module 220. In some embodiments, a data compression module 222 first compresses and then, optionally, encrypts the compressed data. Encryption is performed after compression as encrypted data is randomized and thus may not compress well. Encryption does not change the size of the data. In some embodiments, lossless compression is used, such as Lempel-Ziv-Welch (LZW) compression. Compressed mapping unit(s) generated by the data

compression module 222 can be referred to as compressed payload. In some embodiments, partial mapping units are not compressed in order to make later integration into full mapping units more efficient. In some embodiments, full mapping units are not compressed. For example, one uncompressed mapping unit can be stored as a virtual storage unit, as explained below.

[0022] In some embodiments, compressed (and optionally encrypted) mapping unit data has a size that is smaller (or equal) than that of an uncompressed mapping unit. Accordingly, compressed mapping units can have variable size. However, as explained above, the media 140 can have fixed size storage granularity, such as fixed sized media storage units (or physical storage units or memory data units). For example, the solid-state memory array 150 can comprise memory pages, with a page being the smallest programmable and/or retrievable memory unit. As another example, the magnetic recording media 164 can comprise sectors, with a sector being the smallest programmable and/or retrievable memory unit. Compressed mapping units are provided to the virtual storage unit module 230 that assembles one or more compressed mapping units (or, in some embodiments, uncompressed mapping units) into fixed sized storage units for storing in the media.

[0023] In some embodiments, a virtual storage unit constructor 232 is configured to assemble compressed mapping units into virtual storage units having a size that is aligned with the size of the media storage unit. For example, the size of the virtual storage unit can be an integer multiple of the media storage unit, such as $N \times s_2$ or s_2/N , where s_2 is the media storage unit size and N is an integer. In one embodiment, the size of the virtual storage unit is the same as the size of the media storage unit, such as 4 KB. In another embodiment, the size of the virtual storage unit is selected such that one or more compressed mapping units are assembled into a virtual storage unit. In some embodiments, different sizes of media storage units can be utilized for different types of media or different sizes of media storage units can be utilized for a media type. As is illustrated in Fig. 1, the data storage system 120 includes solid-state memory 150 and magnetic storage 160, which can have different sizes of media storage units. For example, solid-state memory 150 can include memory pages that are 8 KB in size, while the magnetic recording media 164 can include 4 KB sectors. The virtual storage unit constructor 232 can assemble two

types of virtual storage units: 8 KB storage unit for being programmed in the solid-state memory 150 and 4 KB storage unit for being written in the magnetic recording media 164. As another example, solid-state memory 150 can include memory pages that are 8 KB in size, but variable size error detection and correction information can be generated and stored in the page. Thereby, variable size virtual storage units can be utilized for storing data in the solid-state memory 150.

[0024] In some embodiments, a virtual storage unit that includes one or more compressed (or uncompressed) data units is padded with filler data so that the size of the virtual storage unit is aligned with the size of a corresponding media storage unit (e.g., corresponding to the media where the virtual storage unit will be stored). For example, suppose that the size of a compressed mapping unit is 3.5 KB and 4 KB virtual storage units are used. The virtual storage unit constructor 232 can assemble 3.5 KB compressed mapping unit along with 0.5 KB (or 512 bytes) of filler data into a virtual storage unit. As another example, suppose that the combined size of two compressed mapping units is 3.7 KB and 4 KB virtual storage units are used. The virtual storage unit constructor 232 can assemble the two compressed mapping units having combined size of 3.7 KB along with 0.3 KB of filler data into a virtual storage unit. Filler data can comprise one or more suitable values, such as all zero values. As yet another example, suppose that compressed data units are formatted for storage in the magnetic storage 160, and the magnetic storage provides support for only limited mapping between logical and physical addresses. For instance, suppose that the host system 110 “understands” or utilizes 1 GB of logical host data units, such as logical sectors, and the mapping supports indirection for each logical host data unit but no additional data units. Such mapping may be utilized by a traditional hard disk drive where, unlike shingled magnetic disk recording, there is a one-to-one correspondence between logical addresses and physical addresses where host data is stored. In such data storage system, because address indirection is not supported, only one compressed mapping unit is stored in a virtual storage unit or in a media storage unit. The compressed mapping unit is padded to fill the entire storage unit (e.g., storage unit 440 in Fig. 4).

[0025] In some embodiments, compressed mapping units can be split across multiple virtual storage units. Compressed mapping units can include

compressed payload or uncompressed payload (e.g., partial mapping unit payload may not be compressed). For example, suppose that a first compressed mapping unit is sized 1 KB, a second compressed mapping unit is sized 3.5 KB, a third compressed mapping unit is sized 3 KB, and 4 KB virtual storage units are used. The virtual storage unit constructor 232 can assemble the first compressed mapping unit (sized 1 KB) and 3 KB of the second compressed mapping unit (or split compressed mapping unit) into a first virtual storage unit. Further, the virtual storage unit constructor 232 can assemble the remaining 0.5 KB of the second compressed mapping unit (or remainder compressed mapping unit), third compressed mapping unit (sized 3 KB), and 0.5 KB of filler data into a second virtual storage unit. In some embodiments, compressed mapping units are split across sequential virtual storage units. In other embodiments, compressed mapping units are split across non-sequential virtual storage units.

[0026] In some embodiments, multiple compressed mapping units assembled into a virtual storage unit include only sequential host data units. For example, a first compressed mapping unit that includes host data units corresponding to logical 512-byte sectors 0-7 and a second compressed mapping unit that includes host data units corresponding to logical sectors 8-15 are combined into a virtual storage unit. In some embodiments, virtual storage units can include non-sequential host data units. For example, a first compressed mapping unit that includes host data units corresponding to logical sectors 0-7 and a second compressed mapping unit that includes host data units corresponding to logical sectors 32-39 are combined into a virtual storage unit.

[0027] In some embodiments, the mapping unit module 210, compression and encryption module 220, and virtual storage unit module 230 operate on data stored in the buffer 132. That is, mapping units, compressed mapping units, and virtual storage units are created and manipulated in the buffer 132. Assembled virtual storage units can be stored in the non-volatile media 140.

[0028] In some embodiments, reading data comprises performing the inverse of the operations associated with writing data. In response to a read data command received from the host system, the command including logical addresses for host data to be retrieved, the virtual storage unit deconstructor 234 identifies or

locates in the media 140 (and/or in the buffer 132) one or more virtual storage units with the host data to be retrieved. The virtual storage unit deconstructor 234 can use mapping data for performing the identification. The virtual storage unit deconstructor 234 identifies one or more compressed mapping units from the retrieved one or more virtual storage units, the one or more compressed mapping units containing host data to be retrieved. Data decrypter and decompressor 224 decrypts (if the data is encrypted) and decompresses (if the data is compressed) the payloads of the one or more compressed mapping units to obtain one or more mapping units containing host data to be retrieved. Mapping unit deconstructor 214 unpacks host data to be retrieved from the one or more mapping units. Host data to be retrieved is returned to the host system 110.

[0029] In some embodiments, one or more of the mapping unit module 210, compression and encryption module 220, and virtual storage unit module 230 are implemented in hardware, such as in one or more a system on-chip (SoC) hardware modules. Hardware implementation can be realized using an application specific integrated circuit (ASIC), field-programmable gate array (FPGA), or a combination thereof. In another embodiment, modules 210, 220, and 230 can be implemented in firmware or as a combination of hardware and firmware.

[0030] Fig. 3 illustrates compression and formatting data flow 300 utilized by a data storage system according to some embodiments of the invention. The data flow 300 can be implemented by the controller 130 and/or the compression, formatting, and data migration module 134. Host data 310 is received from the host system 110 as a set of host data units 312. In one embodiment, a host data unit 312 is a logical sector, such as a 512-byte logical sector. A host data unit 312 includes host (or user) data and error correction and detection (EDC) data that enables error detection and/or reconstruction of the host data. In one embodiment, EDC data can be parity data checksum, cyclic redundancy check (CRC) data, or a suitable error correction code. In one embodiment, a 512-byte logical sector of host data includes 4 bytes of EDC data.

[0031] In some embodiments, host data 310 is assembled into one or more mapping units 320 as discussed above. This can be performed by the mapping unit module 210. As explained above, a mapping unit 320 can be selected

as an integer multiple of the host data unit size. In some embodiments, a mapping unit 320 includes one or more host data units 324 (which collectively make up a mapping unit payload), mapping unit metadata 322 which identifies the host data units, and mapping unit EDC 326. Mapping unit metadata 322 can include the following information: payload type (e.g., compressed or uncompressed), count (or number) of host data units in the payload, logical addresses of one or more host data units in the payload, and payload length. In other embodiments, metadata 322 can include additional or alternative fields. In some embodiments, when the mapping unit module 210 assembles sequential host data units into a mapping unit, the logical address of the first host data unit in the payload and the host data unit size are sufficient to identify and locate all host data units in the payload. For example, if host data units corresponding to 512-byte logical sectors 0-7 are assembled into a 4 KB mapping unit, starting logical address associated with logical sector 0 and logical sector size (512 bytes) are sufficient to describe all logical sectors. The count of the host units in the payload in this case is set to 8. As another example, if host data units corresponding to 512-byte logical sectors 0-4 are assembled into a partial mapping unit, the count of the host units in the payload is set to 5.

[0032] In some embodiments, mapping unit EDC 326 is calculated for the entire payload of the mapping unit. The size of the mapping unit EDC 326 is smaller than the aggregate size of EDC data associated with the host data units 324 in the payload. For example, suppose that 512-byte logical sector of host data includes 4 bytes of EDC data and 8 logical sectors (or 4 KB) are assembled into a mapping unit 320. A 4-byte mapping unit EDC 326 can be generated, which results in a 28 byte savings as compared to the aggregate size of EDC data (32 bytes) for the host data units in the payload. In some embodiments, this saving of space allows for including metadata 322 into the mapping unit 320 without taking up any additional space. In some embodiments, mapping unit EDC 326 is seeded by the starting logical address of the payload.

[0033] In some embodiments, mapping units 320 are compressed, and optionally encrypted, to produce one or more compressed mapping units 330 as discussed above. This can be performed by the compression and encryption module 220. A compressed mapping unit, such as mapping units 331, 337, or 338,

includes compressed payload 334 (corresponding to the compressed mapping unit), compressed mapping unit metadata 332 which identifies host data that has been compressed, and mapping unit EDC 336, which can be the same as EDC 326. Mapping unit metadata 332 can include the following information: payload type (e.g., compressed or uncompressed), count (or number) of host data units in the compressed payload, logical addresses of one or more host data units in the compressed payload, and compressed payload length. In other embodiments, metadata 332 can include additional or alternative fields. Because compression produces variable sized data, compressed payload length field can be different for two or more different compressed mapping units, such as compressed mapping units 331, 337, and 338. In some embodiments, when sequential host data units are assembled into a mapping unit and into the compressed payload, the logical address of the first host data unit in the compressed payload and the host data unit size are sufficient to describe all host data in the compressed payload. In some embodiments, metadata 322 is reprocessed in order to generate metadata 332. For example, metadata 332 includes metadata 322 along with a field indicating length of the compressed data in the compressed mapping unit.

[0034] In some embodiments, when encryption is performed after compression, the compressed payload 334 and the mapping unit EDC 336 are encrypted. In certain embodiments, the metadata 332 is not encrypted or compressed. The metadata 332 can describe the contents of compressed payload, including describing logical addresses associated with host data in the compressed payload. Because in some embodiments, the host system 110 provides the decryption and/or encryption key(s) to the data storage system 120, not encrypting the metadata 332 allows the data storage system to manipulate compressed (or uncompressed) encrypted payload for various operations, such as background or housekeeping tasks, without having to perform decryption (or decompression) first. Even if the encryption key is available, it may be more efficient to manipulate encrypted payload without first decrypting (or decompressing) it. For example, data may need to be moved between different media types, such as the non-volatile solid-state memory array 150 and the magnetic media 164. As explained above, different media types can utilize media storage units of different sizes. For instance, the

solid-state memory array 150 can be divided into dies, blocks, and pages, such as 2 KB pages, and the magnetic recording media 164 can be divided into tracks, clusters, and sectors, such as 512 byte sectors. When compressed and encrypted data is moved between different media types, data may need to be reformatted and/or repacked to be stored in media storage units of a different type (or types).

[0035] As another example, the data storage system 120 can perform garbage collection without having to obtain the decryption key(s) from the host system 110 or decrypting stored data. Garbage collection can include omitting or removing from a virtual storage unit one or more mapping units having invalid data. For example, suppose that a virtual storage unit includes mapping units for host data units corresponding to 512-byte logical sectors 0-7. Suppose further that as a result of host system activity, host data stored in logical sector 3 is updated, thereby making data in the virtual storage unit corresponding to logical sector 3 invalid. Garbage collection operation can create a new virtual storage unit that includes mapping units corresponding to logical sectors 0-2 and 4-7. As another example, the data storage system 120 can repack compressed payload without the decryption key(s). Repacking can include defragmentation of compressed payload, such as assembling partial mapping units into full mapping units, format conversion, such as storing virtual storage unit payload in a virtual storage unit of a different size for storage on different type of media (or on same media), and the like. Defragmentation can include reordering the mapping units for host data units in a virtual storage unit. For example, suppose that a virtual storage unit includes mapping units for host data units corresponding to 512-byte logical sectors 0, 1, 3, 2, and 4-7. Defragmentation operation can reorder the mappings units in the virtual sector so that the mapping units are arranged in the order corresponding to host data units for logical sectors 0-7. In some embodiments, defragmentation can be performed in conjunction with garbage collection. For example, suppose that a virtual storage unit includes mapping units for host data units corresponding to 512-byte logical sectors 0, 1, 2, 4, 3, and 5-7. Suppose further that as a result of host system activity, host data stored in logical sector 2 is updated, thereby making data in the virtual storage unit corresponding to logical sector 2 invalid. Performing garbage collection in conjunction with defragmentation can create a new virtual

storage unit that includes mapping units corresponding to logical sectors 0, 1, 3, 4, 5, 6, and 7 in that order. In some embodiments, the metadata 332 is encrypted and/or compressed. In addition to garbage collection, the data storage system 120 may perform other housekeeping operations without decrypting the data, including wear leveling, data refresh, etc.

[0036] In some embodiments, partial, full, and/or multiple compressed/uncompressed mapping units are assembled into a virtual storage unit 340. This can be performed by the virtual storage unit module 230. In some embodiments, virtual storage units can include one or more of the following: uncompressed mapping unit(s) (full and/or partial), parts (split/remainder parts) of one or more compressed mapping units split across several virtual storage units, compressed mapping unit(s), filler data, and so on. In the illustrated example, the virtual storage unit 340 comprises one or more full or partial compressed mapping units 344 (collectively making up compressed payload), optional filler data 346 so that the size of the virtual storage unit is aligned with the size of a corresponding media storage unit, virtual storage unit metadata 342, and virtual storage unit EDC 348 determined for the contents of the virtual storage unit. Virtual storage unit metadata 342 can include the following information: virtual storage unit logical address for identifying the physical location of the virtual storage unit in the media 140, offset within the virtual storage unit of metadata field of the first compressed mapping unit in the compressed payload, and a sequence number of the virtual storage unit (or another form of a timestamp). In other embodiments, metadata 342 can include additional or alternative fields. Compressed mapping units 344 can include compressed and, optionally, encrypted data along with uncompressed and unencrypted metadata.

[0037] In some embodiments, returning stored data to the host system 110 comprises performing the inverse of the operations associated with writing data. In response to a read data command received from the host system, the command including logical addresses for the host data to be retrieved, the virtual storage unit module 230 identifies or locates in the media 140 (and/or in the buffer 132) one or more virtual storage units 340 with the host data to be retrieved. In some embodiments, the virtual unit storage unit module 230 utilizes a mapping table (or

another suitable data structure) that reflects associations between the logical addresses used by the host system 110 for addressing data stored in the data storage system 120 and virtual storage units storing data associated with the logical addresses. The virtual storage unit module 230 can also use virtual storage unit metadata 342 to identify and retrieve from the media (and/or from the buffer) virtual storage units with the host data to be retrieved.

[0038] Using the metadata 332, the virtual storage unit module 230 identifies one or more compressed mapping units from the retrieved one or more virtual storage units, the one or more compressed mapping units containing host data to be retrieved. In some embodiments, a retrieved virtual storage unit can include one or more compressed mapping units that contain host data to be retrieved as well as one or more compressed mapping units that do not contain sought after host data. This classification can be performed using the metadata fields of the compressed mapping units. In such cases, one or more mapping units that do not contain host data to be retrieved are skipped, thereby improving efficiency of the data transfer operation. For example, suppose that the host system 110 requested host data corresponding to logical sector 3, but a retrieved mapping unit contains host data corresponding to logical sectors 0-7. The data storage system can return host data corresponding only to logical sector 3. As another example, suppose that the host system 110 requested host data corresponding to logical sectors 0-7, and a first retrieved mapping unit contains host data corresponding to logical sectors 0-7. Suppose further that a second mapping unit contains a “fresh” (or latest) copy of host data corresponding to logical sector 3. The data storage system can return host data contained in the first mapping unit for logical sectors 0-2 and 4-7 (thereby skipping host data for logical sector 3 contained in the first mapping unit) and return host data contained in the second mapping unit for logical sector 3. Data compression and encryption module 220 decrypts (if the data is encrypted) and decompresses (if the data is compressed) one or more compressed mapping units that contain host data to be retrieved. Mapping unit module 210 unpacks host data to be retrieved from the obtained one or more mapping units. Host data to be retrieved is returned to the host system 110. In some embodiments, mapping unit metadata 322 is used to ensure that correct data is returned to the host system 110.

[0039] The above described packing schemes used in one or more embodiments are useful in a mixed media environment where several types of media are present in the data storage system (e.g., a hybrid hard disk with solid state memory). As the media types have various physical sector/memory unit sizes, the packing schemes enable packing characteristics to be easily adjusted across the media sizes. Further, even in data storage systems having single media type, such as a solid-state memory array, the above described packing schemes are useful when the various types of media storage units (or physical storage units or memory data units) are utilized. Thus, a controller employing the packing schemes can service the mixed media storage components and handle migration of data from one media component to another or between different types of media storage unit of a particular media storage component, while fulfilling the requirements of a compression and encryption-based system. For example, the virtual storage unit size (and format) can be adjusted depending on a particular media storage unit where the virtual storage unit will be stored. In one embodiment, migration can be performed seamlessly, with repacking, if any, being handled automatically inline. Also, in one embodiment the packing/unpacking processing (including the necessary metadata handling, data mapping, compression, decompression, encryption, decryption) is automated in hardware, further enhancing the performance of the overall system.

Examples of Repacking and Migration

[0040] Fig. 4 illustrates migration 400 of compressed and encrypted data according to some embodiments of the invention. The illustrated migration can be performed by the controller 130 and/or the compression, formatting, and data migration module 134. Block 410 depicts user data units received from the host system 110, which may be organized as logical host data units (e.g., addressed by LBAs). User data 410 is compressed and stored in a compressed mapping unit 420 which includes compressed user data and metadata that identifies user data that has been compressed. Metadata can include one or more of the following fields: length after compression, starting LBA, physical location information, sequence number, etc. As is illustrated, compressed data is smaller in size than incoming user data.

Compressed data is encrypted, and a CMU 430 with compressed and encrypted data is generated. The metadata is not compressed or encrypted. The CMU 430 can be stored in a virtual storage unit (not shown), and written to the media. In some embodiments, user data may only be encrypted and not compressed.

[0041] In one embodiment, the encrypted CMU 430 (which may be stored in a virtual storage unit) is stored in magnetic media (such as magnetic media 164) which does not support mapping that allows for storage of multiple CMUs in media storage units (or physical storage units). As explained above, such magnetic media may, as a traditional hard disk drive, support one-to-one correspondence between logical addresses and physical addresses where user data is stored. As is illustrated, CMU 430 with compressed and encrypted data is padded with filler data before being stored in a magnetic media physical storage unit 440 on the magnetic media.

[0042] Suppose that CMU 430 stored in the physical storage unit 440 is migrated (along with other stored CMUs) from the magnetic storage to solid-state memory array 150. In one embodiment, such migration may be done to improve performance associated with accessing user data (e.g., data stored in the solid-state memory array may be accessed faster than data stored in the magnetic storage). This can be achieved without decrypting (and decompressing) user data by using CMU metadata, which is not encrypted (or compressed). Because the metadata identifies the bounds of a CMU (e.g., start and end) and the user data stored in the CMU, encrypted (and compressed) data can be repacked into physical storage units utilized by solid-state memory 150 without decryption (or decompression). In one embodiment, solid-state memory 150 supports storage of virtual storage units that comprise multiple CMUs (such as VSU 340 illustrated in Fig. 3). Multiple CMUs stored in separate physical storage units in the magnetic media may be repacked into a single solid-state memory array physical storage unit, such as storage unit 460. In one embodiment, because a physical storage unit of the solid-state memory array may be larger in size than a physical storage unit of the magnetic storage, migration of data from the magnetic storage to the solid-state memory array involves formatting the data to be stored in a physical storage unit of a different size.

[0043] In one embodiment, physical storage unit 460 includes multiple CMUs 466 (CMUs 1, 2, . . . , N) which contain compressed, and optionally, encrypted data. These CMUs are complete or full CMUs in the sense that they are not split up among different physical storage units. Storage unit 460 may include a partial CMU 464 which is part of CMU that is split among two (or two or more) physical storage units. As is illustrated, physical storage unit 450 includes the other part of the partial CMU 464. In one embodiment, split CMUs are stored in sequential physical storage units, such as in storage units numbered M-1 and M. Storage unit 460 may also include field 462, which is an offset to the first full CMU (e.g., CMU 1) and allows the partial CMU 464 to be skipped. In some embodiments, the storage unit 460 may not include any partial CMUs, and the offset field 462 may be set to a zero value or may be omitted altogether. Storage unit 460 may include a partial CMU 468 (CMU N+1) as the last CMU.

[0044] Although migration of compressed and encrypted data from magnetic media to solid-state media is illustrated in Fig. 4, migration of data from solid-state media to magnetic media is the reverse of the illustrated sequence. Compressed and encrypted data can be migrated to magnetic media without decryption or decompression. Depending on whether magnetic media where data is migrated supports only one or multiple CMUs per physical storage unit, encrypted (and compressed) CMUs are repacked and stored into physical storage units utilized by the magnetic media. Further, the migration sequence illustrated in Fig. 4 is applicable to migration of encrypted (and compressed) data between different types of physical storage units on a single media, such as solid-state memory that stores variable types and sizes of error correction data. Repacking and migration of encrypted (and compressed) data can thus be performed without needing or utilizing the encryption key (or encryption keys).

[0045] Fig. 5 is a flow diagram illustrating a process 500 of migrating encrypted data according to some embodiments of the invention. The illustrated process can be executed by the controller 130 and/or the compression, formatting, and data migration module 134. In block 502, the process 500 uses unencrypted metadata to determine format of encrypted data (which may also be compressed). The process 500 can use metadata to determine the boundaries (e.g., start and end)

of various encrypted data units, such as mapping units and/or compressed mapping units. In block 504, the process 500 repacks the encrypted data without decrypting it. The process 500 can perform this step because the format of the data has been determined in block 504. For example, the process 500 can repack encrypted data into different type of virtual storage units. In block 506, the process 500 stores repacked encrypted data in the storage media.

Conclusion

[0046] Disclosed approaches for compression, formatting, and migration of data for data storage systems can increase available capacity for storing user data, reliability, and performance of a data storage system. Disclosed approaches for repacking data can be applied in any situation where embedded metadata needs to be accessed, such as during data migration, and where the underlying data is encrypted. In some embodiments, performance is increased because encrypted data is repacked without first performing decryption. In addition, data may also be compressed and repacking can be performed without performing decompression. Advantageously, there is no need to retrieve or wait for the availability of encryption key (or keys) or expand resources in decrypting (and decompressing) data before repacking it and encrypting repacked data.

Other Variations

[0047] Those skilled in the art will appreciate that in some embodiments, additional or alternative modules can be used to perform compression, formatting, and data migration. Mapping units, compressed mapping units, and/or virtual storage units can include additional or alternative fields. The actual steps taken in the disclosed processes, such as the processes illustrated in Figs. 4 and 5, may differ from those shown in the figures. Additional system components can be utilized, and disclosed system components can be combined or omitted. Depending on the embodiment, certain of the steps described above may be removed, others may be added.

[0048] 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 protection. Indeed, the novel methods 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 protection. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope and spirit of the protection. For example, the systems and methods disclosed herein can be applied to hard disk drives, hybrid hard drives, and the like. In addition, other forms of storage (e.g., DRAM or SRAM, battery backed-up volatile DRAM or SRAM devices, EPROM, EEPROM memory, etc.) may additionally or alternatively be used. As another example, the various components illustrated in the figures may be implemented as software and/or firmware on a processor, ASIC/FPGA, or dedicated hardware. 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.

WHAT IS CLAIMED IS:

1. A data storage system comprising:
non-volatile memory (NVM) configured to store data in a plurality of memory data units, a memory data unit corresponding to a unit of storage in the NVM, wherein at least some memory data units have a different format;
and
a controller configured to:
retrieve from the NVM data stored in a first memory data unit, the first memory data unit further storing metadata associated with the data, wherein the data is encrypted and the metadata is not encrypted,
and
using the metadata and without decrypting the data, store the data in a second memory data unit, the format of the second memory data unit being different than the format of the first memory data unit.
2. The data storage system of claim 1, wherein the NVM comprises a non-volatile solid-state memory array that utilizes memory data units of a first format type and a magnetic storage that utilizes memory data units of a second format type different from the first format type.
3. The data storage system of claim 2, wherein a memory data unit of the first format type is configured to store a different maximum size of user data than a memory data unit of the second format type, and wherein user data comprises data received from a host system.
4. The data storage system of claim 1, wherein the NVM comprises a non-volatile solid-state memory array that utilizes memory data units of a first format type and memory data units of a second format type different from the first format type.
5. The data storage system of claim 4, wherein a memory data unit of the first format type is configured to store a different maximum size of user data than a memory data unit of the second format type, and wherein user data comprises data received from a host system.
6. The data storage system of claim 1, wherein the data stored in the first memory data unit comprises compressed and encrypted data, and wherein the

controller is further configured to store the compressed and encrypted data in the second memory data unit without decompressing the data.

7. The data storage system of claim 6, wherein the metadata is not compressed.

8. The data storage system of claim 1, wherein the metadata indicates a plurality of logical addresses associated with the data stored in the first memory data unit.

9. The data storage system of claim 1, wherein the controller is configured to store the data in the second memory data unit in response to performing one or more of garbage collection, wear leveling, and data refresh operations.

10. A method of migrating data in a data storage system comprising a non-volatile memory (NVM) configured to store data in a plurality of memory data units, a memory data unit corresponding to a unit of storage in the NVM, wherein at least some memory data units have a different format, the method comprising:

retrieving from the NVM data stored in a first memory data unit, the first memory data unit further storing metadata associated with the data, wherein the data is encrypted and the metadata is not encrypted; and

using the metadata and without decrypting the data, storing the data in a second memory data unit, the format of the second memory data unit being different than the format of the first memory data unit,

wherein the method is performed under control of a controller.

11. The method of claim 10, wherein the NVM comprises a non-volatile solid-state memory array that utilizes memory data units of a first format type and a magnetic storage that utilizes memory data units of a second format type different from the first format type.

12. The method of claim 11, wherein a memory data unit of the first format type is configured to store a different maximum size of user data than a memory data unit of the second format type, and wherein user data comprises data received from a host system.

13. The method of claim 10, wherein the NVM comprises a non-volatile solid-state memory array that utilizes memory data units of a first format type and memory data units of a second format type different from the first format type.

14. The method of claim 13, wherein a memory data unit of the first format type is configured to store a different maximum size of user data than a memory data unit of the second format type, and wherein user data comprises data received from a host system.

15. The method of claim 10, further comprising storing compressed and encrypted data stored in the first memory data unit in the second memory data unit without decompressing the data.

16. The method of claim 15, wherein the metadata is not compressed.

17. The method of claim 10, wherein the metadata indicates a plurality of logical addresses associated with the data stored in the first memory data unit.

18. The method of claim 10, wherein storing the data in the second memory data unit is performed in response to executing one or more of garbage collection, wear leveling, and data refresh operations.

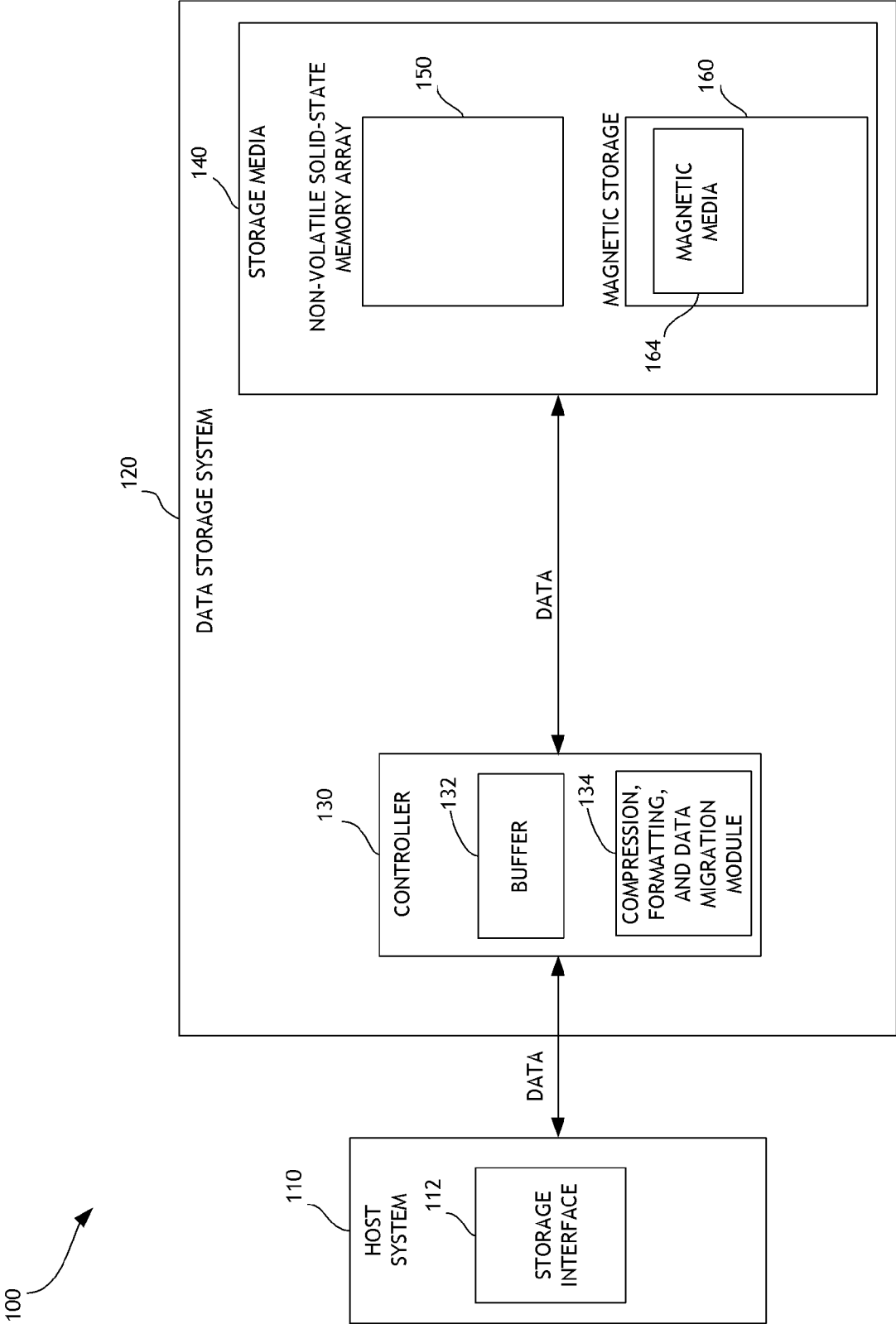


FIGURE 1

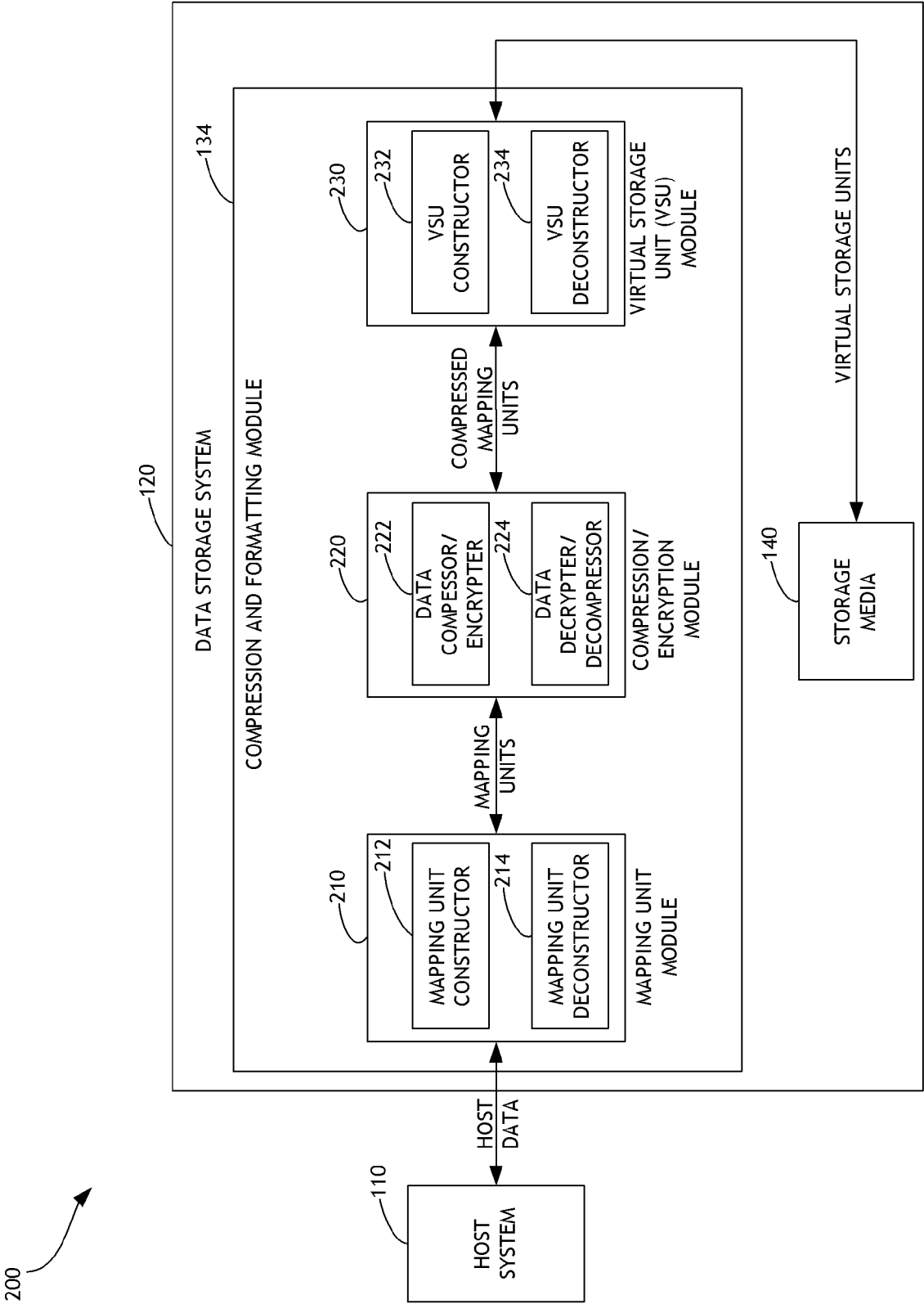


FIGURE 2

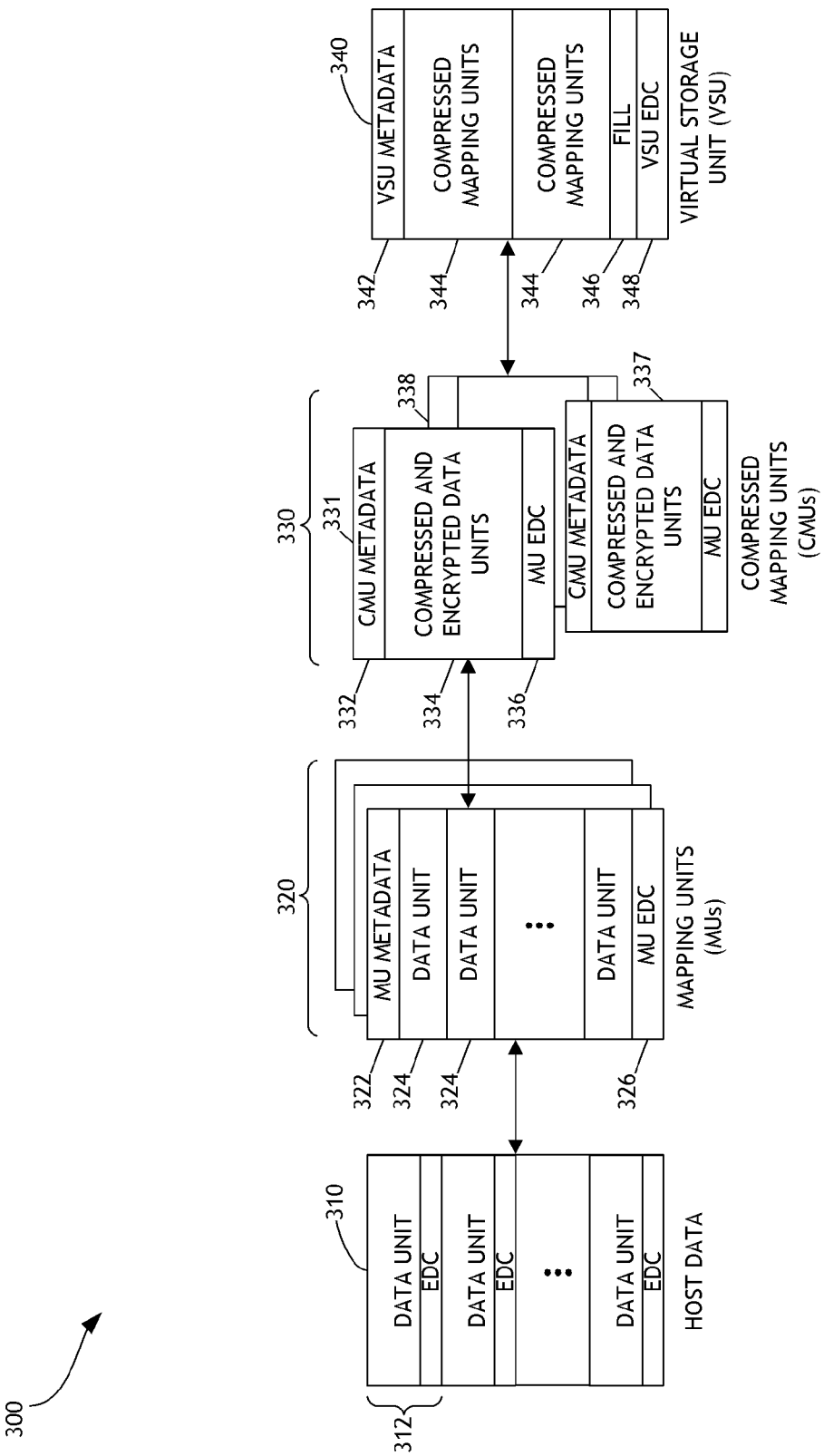
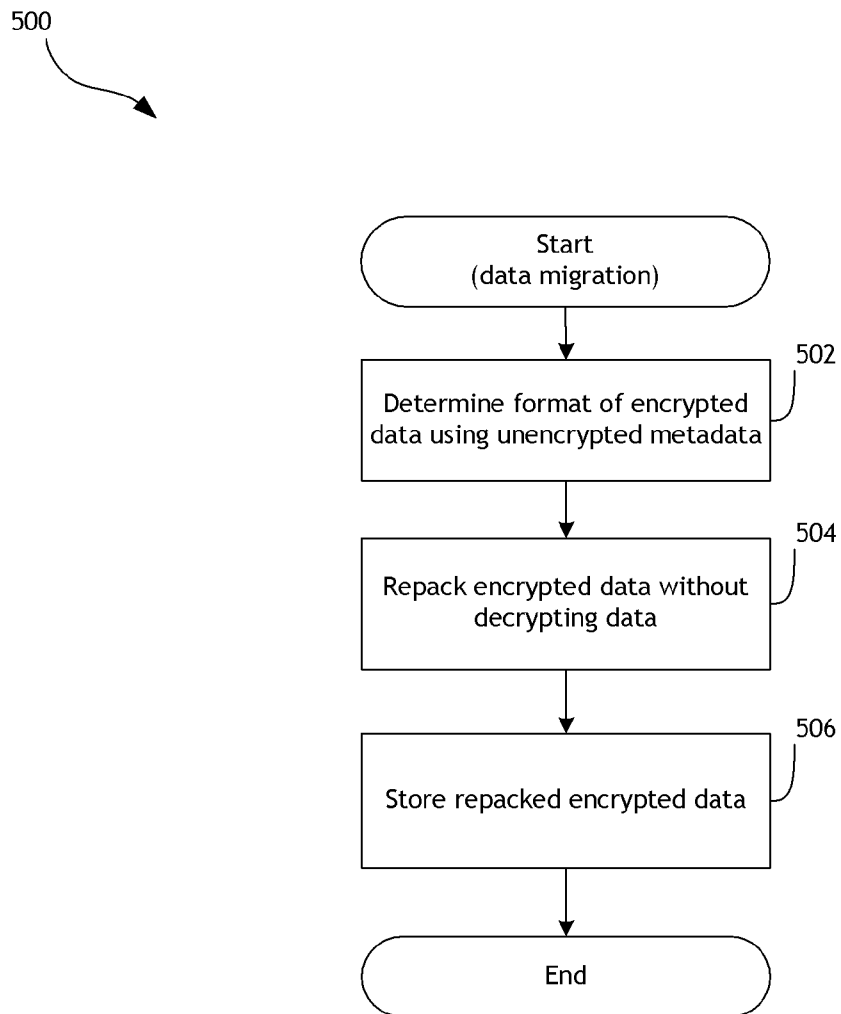


FIGURE 3

**FIGURE 5**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/041786**A. CLASSIFICATION OF SUBJECT MATTER****G06F 12/16(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 12/16; G06F 12/02; G06F 17/00; G06F 12/00; G06F 12/14; G11C 11/00; G06F 11/00; G06F 13/00

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: NVM(non-volatile memory), different, format, encrypt, compress

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1906311 A2 (SAMSUNG ELECTRONICS CO., LTD.) 02 April 2008 See paragraphs 2-3, 5, 8, 11, 21-22, 25, 28; claim 3; and figure 2.	1-18
Y	US 2011-0016091 A1 (ANAND PRAHLAD et al.) 20 January 2011 See paragraphs 9-10, 49, 63, 76-77, 83; and figure 1.	1-18
A	US 2010-0246246 A1 (HO-JUNG KIM et al.) 30 September 2010 See paragraphs 2 and 7; and figure 1.	1-18
A	US 2009-0327759 A1 (ROBERT ROYER et al.) 31 December 2009 See paragraphs 3 and 9; and figure 1.	1-18
A	US 2007-0061511 A1 (ROBERT W. FABER) 15 March 2007 See paragraphs 5-6; and figure 1.	1-18



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

22 October 2014 (22.10.2014)

Date of mailing of the international search report

22 October 2014 (22.10.2014)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

BYUN, Sung Cheal

Telephone No. +82-42-481-8262



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/041786

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1906311 A2	02/04/2008	CN 101154189 A	02/04/2008
		CN 101154189 B	14/12/2011
		DE 602007011549 D1	10/02/2011
		EP 1906311 A3	24/06/2009
		EP 1906311 B1	29/12/2010
		JP 2008-084317 A	10/04/2008
		KR 10-0855467 B1	01/09/2008
		KR 10-2008-0028685 A	01/04/2008
		US 2008-0077729 A1	27/03/2008
		US 8429327 B2	23/04/2013
US 2011-0016091 A1	20/01/2011	US 2009-0319585 A1	24/12/2009
		US 2013-0290280 A1	31/10/2013
		US 8219524 B2	10/07/2012
		US 8484162 B2	09/07/2013
US 2010-0246246 A1	30/09/2010	KR 10-2010-0107609 A	06/10/2010
		US 8111563 B2	07/02/2012
US 2009-0327759 A1	31/12/2009	US 8555086 B2	08/10/2013
US 2007-0061511 A1	15/03/2007	US 7533215 B2	12/05/2009