

(21) Application No: 1317108.7  
 (22) Date of Filing: 26.09.2013  
 (30) Priority Data:  
 (31) 13647273 (32) 08.10.2012 (33) US

(51) INT CL:  
 G06F 12/06 (2006.01) G06F 12/02 (2006.01)  
 G06F 13/16 (2006.01)

(56) Documents Cited:  
 WO 2013/165386 A1 US 6336174 B1  
 US 20120137055 A1 US 20110167197 A1  
 US 20090248965 A1 US 20080126716 A1

(71) Applicant(s):  
**HGST Netherlands B.V.**  
 (Incorporated in the Netherlands)  
 Herikerbergweg 38, Luna Arena,  
 Amsterdam 1101 CM, Netherlands, Netherlands

(58) Field of Search:  
 INT CL G06F  
 Other: EPODOC, WPI, TXTEN, INSPEC, XPESP,  
 XPESP2, XPIEE, XPIPCOM, XPI3E, XPMISC, XPLNCS,  
 XPRD, XPSRNG, TDB

(72) Inventor(s):  
**Frank R Chu**  
**Luiz M Franca-Neto**  
**Timothy Tsai**  
**Qingbo Wang**

(74) Agent and/or Address for Service:  
**Mewburn Ellis LLP**  
 33 Gutter Lane, LONDON, EC2V 8AS, United Kingdom

(54) Title of the Invention: **Apparatus and Method for Low Power Low Latency High Capacity Storage Class Memory**  
 Abstract Title: **Storage class memory having low power, low latency, and high capacity**

(57) Direct attached dual in line memory (DIMM) card 102 contains dynamic random access memory (DRAM) 104, at least one of non-volatile memory 105-108 and a processor 110 selectively allocating data among the DRAM and the non-volatile memories based on the data set size. Such hybrid arrangement has low power and high capacity because of the non-volatile memories, such as solid-state memories, while low latency is obtained by distributing data-sets according to their size. Smaller data sets are placed in DRAM, medium sized data in phase change memory (PCM) 105, and large data-sets in NAND flash 108. Data-sets can also be straddled by partitioning them across different memory technologies. Data writes can involve buffering to DRAM, sending the data to the NAND flash 108, and then a selective migration among the memories. Other non volatile memories can include resistive RAM (ReRAM) 106 and spin transfer torque RAM (STT-RAM) 107.

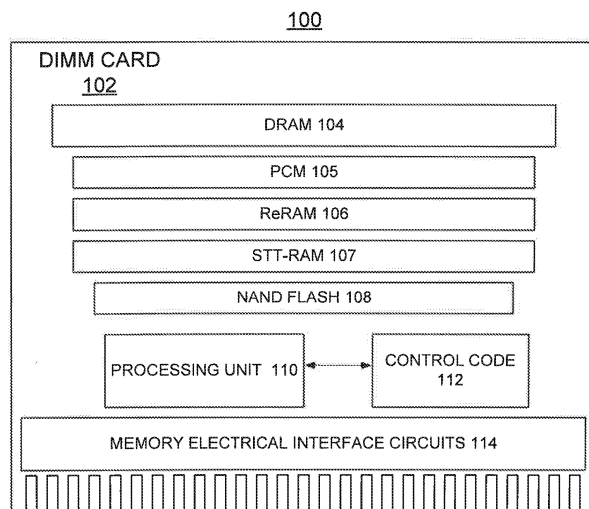


FIG. 1

100

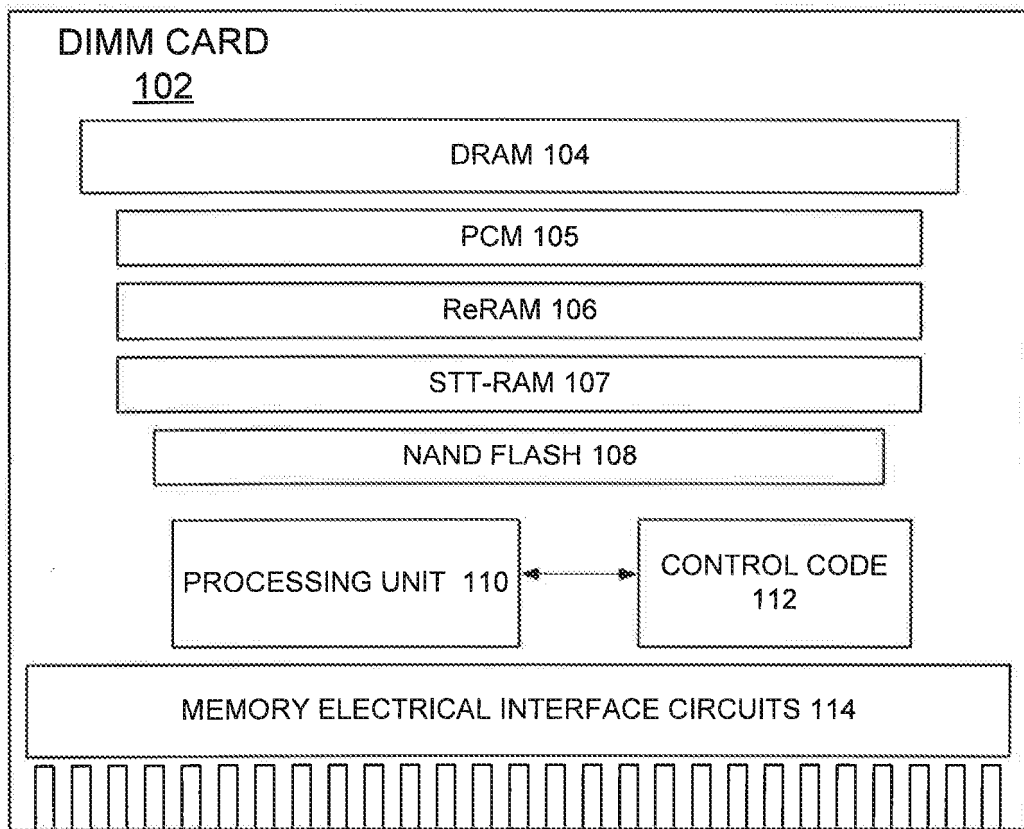
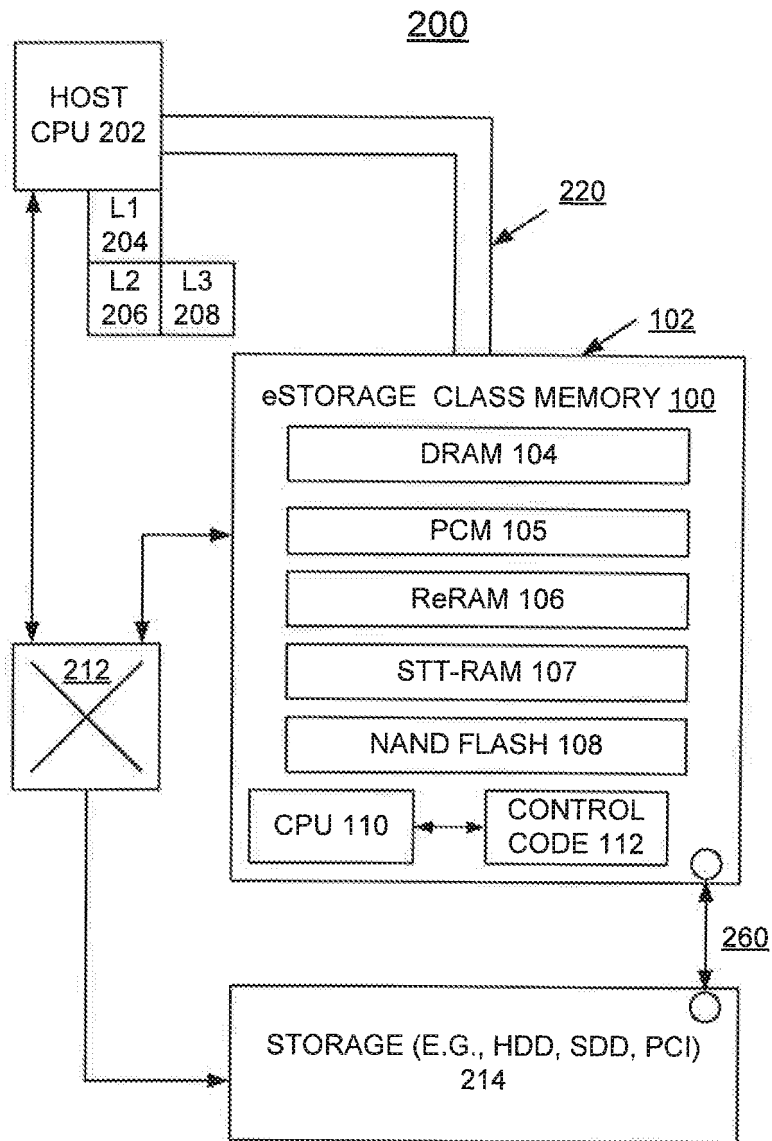
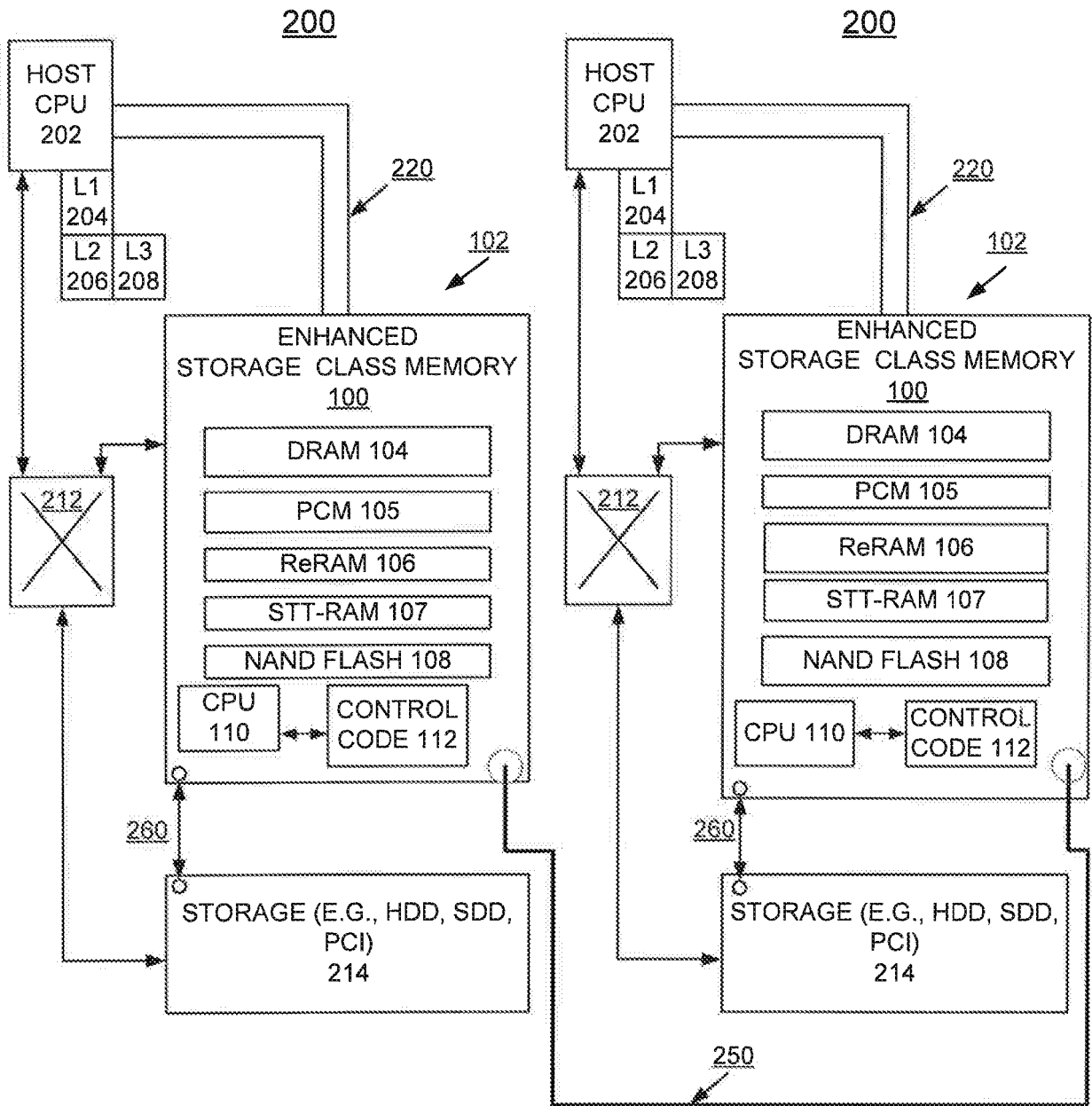


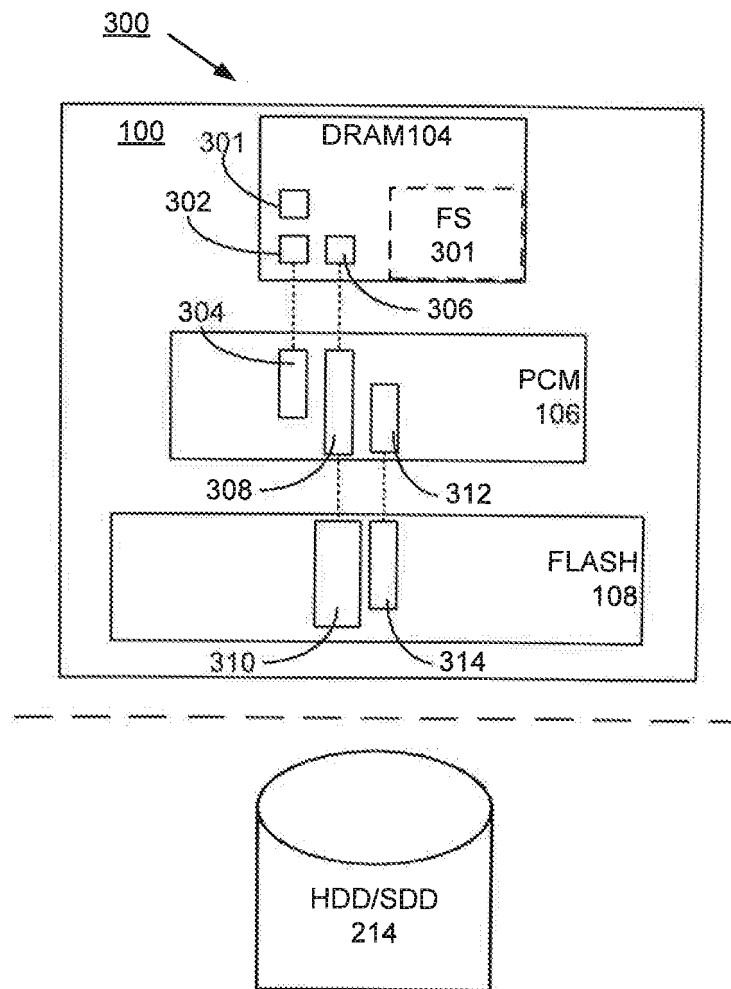
FIG. 1



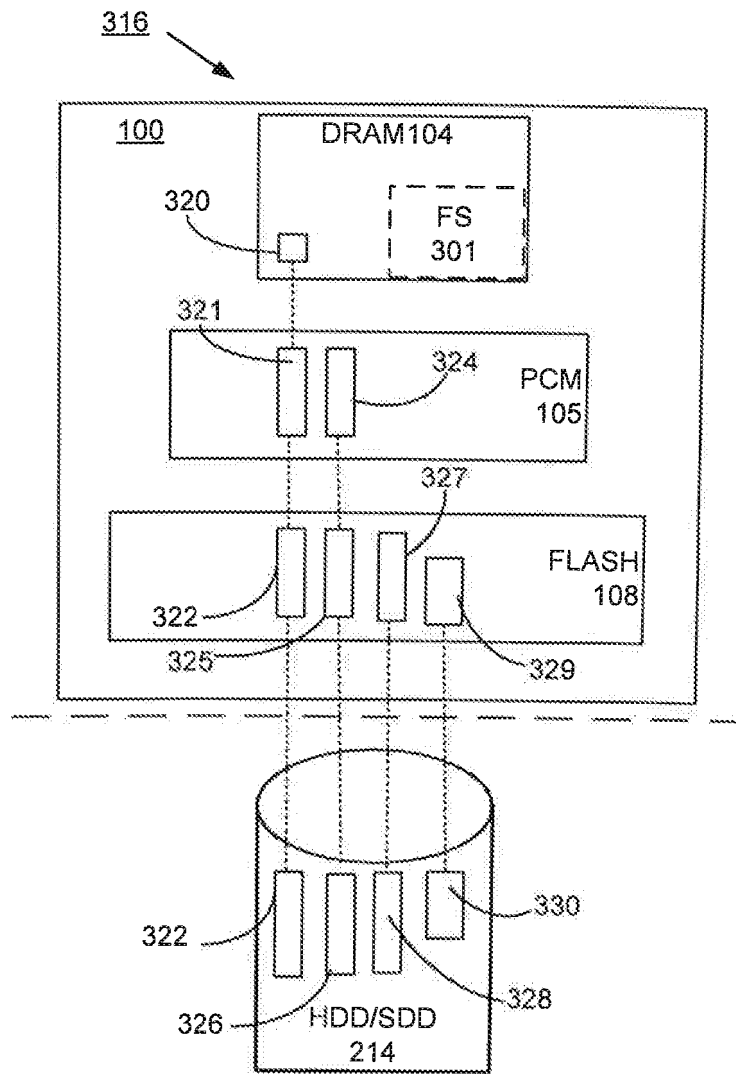
**FIG. 2A**



**FIG. 2B**



**FIG. 3A**



**FIG. 3B**

FIG. 4A

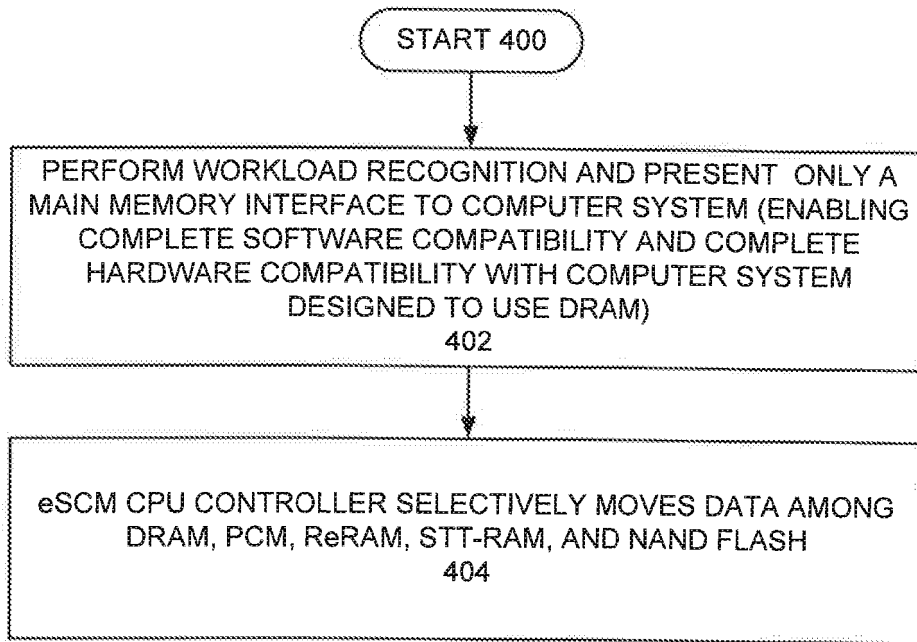
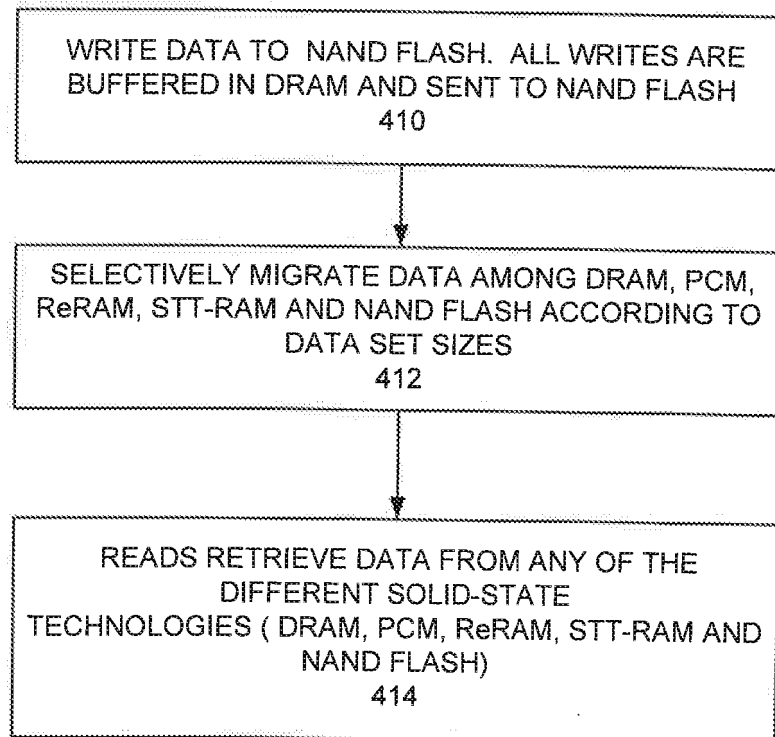
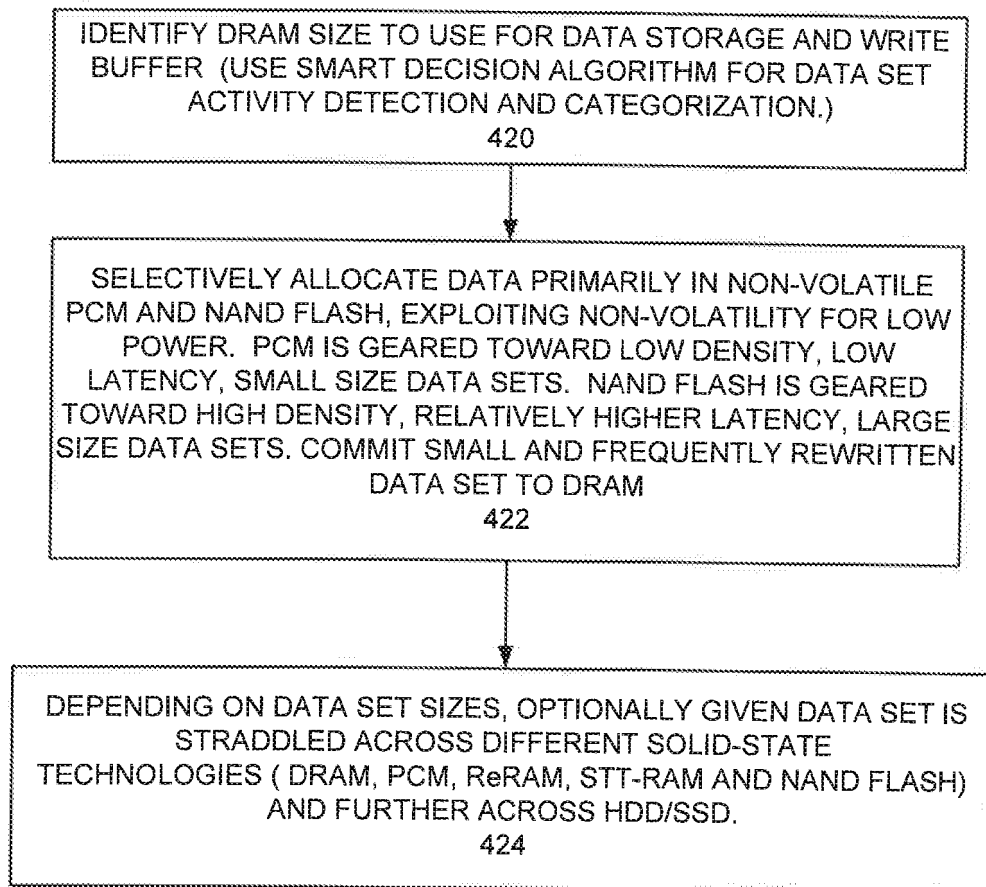


FIG. 4B





## FIG. 4C



500

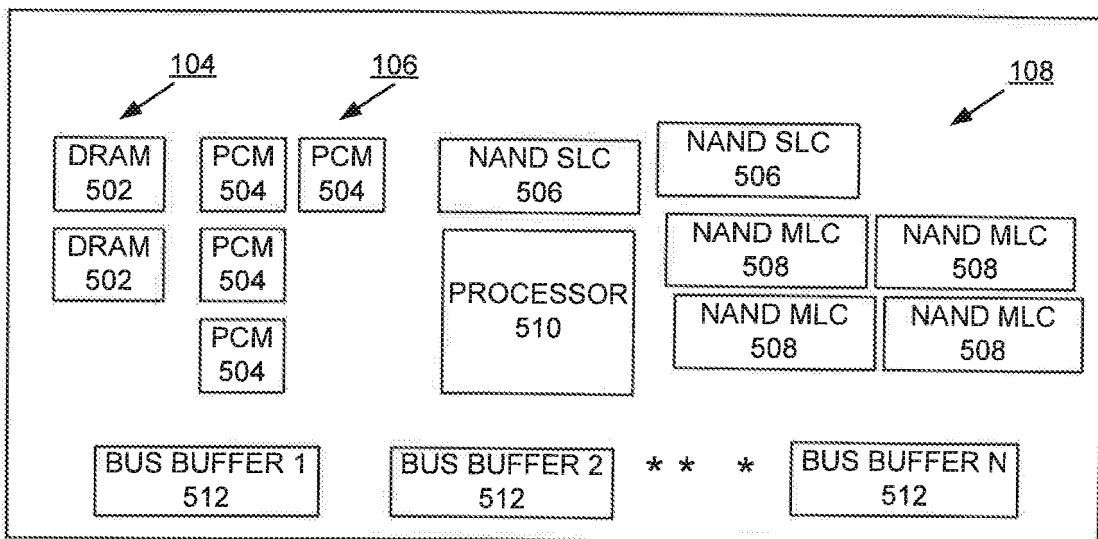


FIG. 5

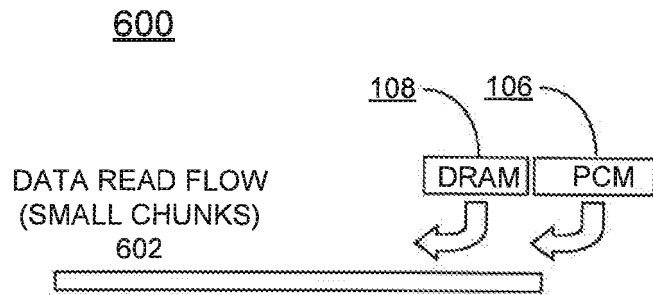


FIG. 6A

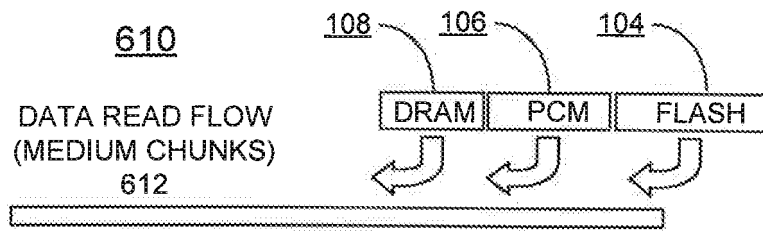


FIG. 6B



FIG. 6C

## **APPARATUS AND METHOD FOR LOW POWER LOW LATENCY HIGH CAPACITY STORAGE CLASS MEMORY**

**[0001]** The present invention relates generally to the data storage field, and more particularly, relates to a method and a storage system for implementing storage class memory with large size, low power and low latency in data accesses. This storage class memory can be attached directly to the memory bus or to peripheral interfaces in computer systems such as peripheral component interconnect (PCI), or PCIe or common storage interfaces such as Serial ATA or SATA, or Serial Attached SCSI (SAS).

**[0002]** Non-volatile solid state memory technologies, such as NAND Flash, have been used for data storage in computer systems. Solid State Drives (SSDs) used in computer systems can take both the form factors and interfaces of hard disk drives (HDDs). SSDs nevertheless provide for faster data access solution than HDDs. SSDs have recently evolved to provide alternative form factor and access through a PCIe interface. In the interest of providing even faster access to stored data, it has been proposed to use direct attachment to the memory bus in a computer system for those solid state storage solutions.

**[0003]** On the memory bus in computer systems, due to the performance requirement in bandwidth and low latency, volatile dynamic random access memory (DRAM) is typically used. Moreover, since data in memory is frequently accessed, non-volatile memory technologies might be exposed to early failure given the relatively low endurance of current non-volatile solid state technology.

[0004] Recently, given the significant gap in bandwidth and latency between memory and storage in computer systems, a new hierarchy called Storage Class Memory (SCM) has been proposed. A SCM would have attributes of low latency and high bandwidth closer to memory requirements than common storage hierarchy, and SCM would have also the attribute of non-volatility associated with storage technologies.

[0005] Unfortunately, the Storage Class Memory concept has found only partial realization. In some instances, SCM is basically a typical NAND Flash-based solid state storage where some improvements were gained at latency in data access. In other realization, SCM is mostly a memory solution where non-volatility was added to the realization. In this latter case, capacity of the SCM was compromised or the SCM cost became relatively unattractive.

[0006] An aspect of the present invention is to provide an apparatus and method for a Storage Class Memory (SCM) that provides low power, high performance, low latency and non-volatility, without sacrificing capacity thus realizing the required attributes for a SCM.

Aspects of the present invention are to provide a method and a storage system for implementing enhanced solid-state storage usage. Other important aspects of the present invention are to provide such method and storage system substantially without negative effect and to overcome some of the disadvantages of prior art arrangements.

[0007] In brief, a method and a storage system are provided for implementing enhanced solid-state storage class memory (eSCM) including a direct attached dual in line memory (DIMM) card containing dynamic random access memory (DRAM), and at least one non-volatile memory, for example, Phase Change memory (PCM), Resistive RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash chips. An eSCM processor controls selectively moving data among the DRAM, and the at least one non-volatile memory based upon a data set size.

[0008] The present invention together with the above and other objects and advantages may best be understood from the following detailed description of the embodiments of the invention illustrated in the drawings, wherein:

[0009] FIG. 1 is a block diagram representation of an enhanced solid-state storage class memory (eSCM) for implementing enhanced solid-state storage performance in accordance with an embodiment of the invention;

[0010] FIG. 2A is a block diagram representation of a computer system including the enhanced solid-state Storage Class Memory of FIG. 1 in accordance with an embodiment of the invention;

[0011] FIG. 2B is a block diagram representation of computer systems, each including the enhanced solid-state Storage Class Memory of FIG. 1 in accordance with an embodiment of the invention where the SCM in the different computer system are capable of exchange data without interference of the host CPU and this embodiment supports cloud applications;

[0012] FIGS. 3A, and 3B schematically illustrates example data location based on data set sizes of the enhanced solid-state Storage Class Memory of FIG. 1 and HDD/SSD of FIG. 2 for implementing enhanced solid-state storage usage performance in accordance with an embodiment of the invention;

[0013] FIGS. 4A, 4B, 4C are flow charts illustrating example operations of the enhanced solid-state Storage Class Memory of FIG. 1 for implementing enhanced solid-state storage usage performance in accordance with embodiments of the invention;

[0014] FIG. 5 schematically illustrates another more detailed example enhanced solid-state Storage Class Memory for implementing enhanced solid-state storage performance in accordance with embodiments of the invention; and

[0015] FIGS. 6A, 6B, 6C are charts schematically illustrating example operations of the enhanced solid-state Storage Class Memory of FIG. 1 for implementing a process of latency hiding in a Storage Class Memory in accordance with embodiments of the invention.

[0016] In many computer systems main memory typically includes dynamic random access memory (DRAM). DRAM is generally expensive and has generally high power dissipation resulting from required memory refreshing.

[0017] A need exists for an effective and efficient method and a storage system for implementing enhanced solid-state storage performance including a low cost, low power and high capacity storage system.

[0018] In the following detailed description of embodiments of the invention, reference is made to the accompanying drawings, which illustrate example embodiments by which the invention may be practiced. It is to be understood that other embodiments may be utilized and structural changes may be made without departing from the scope of the invention.

[0019] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention. As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms



"comprises" and/or "comprising," when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

**[0020]** In accordance with features of the embodiments of the invention, a method and a storage system are provided for implementing an enhanced solid-state Storage Class Memory including a direct attached dual in line memory (DIMM) card containing dynamic random access memory (DRAM), and at least one non-volatile memory, such as Phase Change Memory (PCM), Resistive RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash chips.

**[0021]** The apparatus and method for a low power low latency high capacity enhanced Storage Class Memory disclosed in one embodiment uses the direct attached dual in line memory (DIMM) card containing a multiplicity of solid state memory technologies and a method to manage storage data with the objective of providing data protection against power disruption, low power operation and low latency in data access. In such enhanced storage class memory, for illustration only dynamic random access memory (DRAM), phase-change-memory (PCM), Resistive RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash chips provide an example of implementation. The skilled in the art will readily find variations on the example using different memory technologies without departing from the spirit of this invention.

[0022] In another embodiment, the enhanced Storage Class Memory may use other interfaces to the computer system different from the used above in the illustration of an eSCM used in direct attachment to the memory bus.

[0023] Different solid state memory technologies offer different benefits for the final eSCM solution. The eSCM embodiments of the present invention exploit in a hybrid arrangement those different technologies to improve the final solution. In one illustrative embodiment, large capacity and low cost are achieved by using NAND Flash. Other solid state memory technologies like Phase Change Memory are added to the hybrid solution to provide low latency access and non-volatility. Very frequently overwriting of data is supported by substantial presence of DRAM in the eSCM.

[0024] Low power is achieved by the non-volatility attribute of the eSCM disclosed, since relative to a purely DRAM solution there is no need to refresh data in the non-DRAM SCM memory cells.

[0025] Low latency is achieved by a specific algorithm in the eSCM by distributing data among the different solid state technologies according to data set size committed to the memory. This is a dynamic strategy that takes advantage of statistics of the eSCM data traffic.

[0026] Those skilled in the art will recognize that this dynamic strategy of the present invention provided by such method and storage system achieves low latency

objectives substantially without negative effect and that overcomes some of the disadvantages of prior art arrangements.

[0027] In accordance with features of the embodiments of the invention, a method and a storage system are provided for implementing an enhanced solid-state Storage Class Memory including a direct attached dual in line memory (DIMM) card, for example, containing dynamic random access memory (DRAM), Phase Change memory (PCM), Resistive RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash chips.

[0028] Having reference now to the drawings, in FIG. 1, there is shown an example solid-state storage system generally designated by the reference character 100 for implementing enhanced solid-state Storage Class Memory in accordance with an embodiment of the invention. Solid-state storage system 100 includes solid-state storage devices contained on a direct attached dual in line memory (DIMM) card 102. Enhanced solid-state Storage Class Memory (eSCM) system 100 enables a low power, low cost, large memory space, for example, a memory space in hundreds of GBs.

[0029] Enhanced solid-state Storage Class Memory (eSCM) system 100, for example, includes volatile data storage dynamic random access memory (DRAM) 104, and non-volatile data storage devices including phase-change-memory (PCM) 105, Resistive RAM (ReRAM) 106, Spin-Transfer-Torque RAM (STT-RAM) 107 and NAND flash memory 108 contained on the DIMM card 102. An eSCM processing unit 110,

such as an embedded processing unit, is provided with the DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND flash memory 108 on the DIMM card 102. The eSCM processing unit or eSCM controller 110 selectively moves data among the DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND flash memory 108 enabling enhanced latency and throughput performance. eSCM system 100 includes control code 112 for implementing smart decision algorithms for data set activity detection and categorization. eSCM system 100 includes memory electrical interface circuits 114 coupled to the eSCM processor unit 110.

**[0030]** Referring also to FIG. 2A, there is shown an example processor or computer system including the eSCM system 100 of FIG. 1 in accordance with an embodiment of the invention. System 200 includes a central processor unit (CPU) 202 and a plurality of cache memory L1, 204, L2, 206, L3, 208. System 200 includes a memory controller 212, and storage 214, such as, a direct access storage devices (DASD), such as Solid State drive (SSD), or hard disk drive (HDD) including a Shingled Disk Drive (SDD), or a peripheral component interconnect (PCI) computer bus for attaching hardware devices (not shown) in the system 200. For generality purposes, CPU 202 is depicted as also connected to the eSCM 100 by an interface 220, such as a system bus 220. In system 200, hierarchy of DRAM 104 is encompassed by the eSCM 100 and management of data movements among the hybrid collection of solid state memory technologies present in the eSCM 100 is driven by specific algorithms housed in the eSCM processor 110 itself, for example, as described below. eSCM 100 has an interface

260 without interference of the host CPU 202. Those skilled in the art will recognize that the eSCM capability to transfer data between eSCM 100 and storage 214 without the host CPUs intervention.

[0031] Referring also to FIG. 2B, there is shown an example pair of computer systems 200 of FIG. 2A, each including the enhanced Storage Class Memory 100 of FIG. 1 in accordance with an embodiment of the invention where the eSCM 100 in the different computer systems 200 are capable of exchange data as indicated at an interface 250 without interference of the host CPU 202. Those skilled in the art will recognize that the eSCM capability to transfer data between computer systems without the host CPUs intervention can be extended to many more than two computer systems and be the used to support efficient data movement for a large assembly of computer systems as used in cloud applications.

[0032] In accordance with features of the embodiments of the invention, eSCM processor 110 communicates with the memory controller or CPU 202 as a standard main memory DRAM module in the Dual Inline Memory Module (DIMM) socket. The memory bus 220 can be standard DRAM bus with 240 lines or narrower high speed Fully-Buffered DRAM bus. In both cases all signals in the bus are routed to the eSCM processor 110, which will according to predefined algorithms decide to commit the data to DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, or NAND Flash 108.

[0033] It should be understood that principles of the present invention are not limited to a particular bus arrangement, and many other bus configurations are possible without departing from the spirit of this invention.

[0034] In accordance with features of the embodiments of the invention, control code 112 enables eSCM processor 110 of the eSCM system 100 to use its own intelligent data detection algorithms to determine when data should be committed to DRAM 104, PCM 105 or NAND Flash 108. Optionally, the eSCM processor 110 can coordinate with the host CPU 202 and learn from this CPU 202 specific data requirements that recommend a particular data set to be committed to one of the technologies or memory tier available of DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, or NAND Flash 108.

[0035] In accordance with features of the embodiments of the invention, in another innovation, data sets are committed to the different solid state memory technologies according to data set sizes. It is a departure from typical hierarchical memory concepts where data is committed to different memory (or storage) hierarchy according to frequency of reuse and spatial and location proximity correlation. Memory control code 112 of the eSCM system 100 allows for coordination, detection and categorization of features with host CPU 202. For example, control code 112 of the invention optionally allows the CPU 202 of the host system 200 to determine the sizes of DRAM 104 for cache or for write buffer, what data set should be immediately committed

to PCM 105 or NAND Flash 108, and what addresses should be fetched directly from PCM 105 or NAND Flash 108 in a read operation, among combination of these features.

[0036] eSCM system 100 and system 200 are shown in simplified form sufficient for understanding the present invention. It should be understood that principles of the present invention are not limited to the illustrated eSCM system 100 and the illustrated system 200. The illustrated system 200 is not intended to imply architectural or functional limitations. The present invention can be used with various hardware implementations and systems and various other internal hardware devices in accordance with an embodiment of the invention.

[0037] In accordance with features of the embodiments of the invention, the eSCM processor 110 selectively moves data among the DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND flash memory 108 enabling enhanced latency and throughput performance. Using the three technology direct attached DIMM card 102, for example, including DRAM 104, PCM 105 and NAND Flash 108 of the invention provides enhanced latency and throughput performance as compared to the latency incurred if a large data set were to be only available in storage 214, such as HDD or SSD. eSCM 100 is a low latency storage, which has main memory class.

[0038] In accordance with features of the embodiments of the invention, the cost of the eSCM system 100 is diminished by extensive use of low cost NAND Flash memory 108. Low power is achieved by both extensively use of non-volatile memory

space including PCM 105 and NAND flash memory 108 and selective power down of unused memory chips including DRAM 104. An extremely large memory space advantageously is defined by PCM 105 and NAND Flash 108 enabling DRAM tier 104 to work more as a write buffer than as a cache for both other tiers. Data in a read operation can be retrieved directly from PCM 105 or NAND Flash 108, when not available in DRAM 104. Hence, in an embodiment, there could be only one copy of the data in the eSCM 100; hence none of the solid state technologies is used as cache.

[0039] Referring to FIG. 3A, in another innovation in this invention, example operations generally designated by the reference character 300 of the eSCM 100, including straddling data sets across different memory technologies. In FIG. 3A, the eSCM 100 is shown together with the storage 214 of system 200. In FIG. 3A, the example data locations based on data set sizes are illustrated with DRAM 104, PCM 105 and NAND flash 108 of the of the eSCM 100. Smaller data sets, as in a first data set indicated by 301, are completely placed in DRAM. Progressive larger data set, which are expected to be later read as a single set, are stored completely on other solid state memory technologies or stored across different solid state memory technologies. A second data set indicated by 302, 304 is respectively stored in the DRAM 104 and PCM 105. That is, the second data set 302,304 has part of its data stored in DRAM 104 and part of its data stored in PCM 105; hence this is a data set that straddles along two different solid state memory technologies. A third data set indicated by 306, 308, 310 is respectively stored part in the DRAM 104, part in PCM 105 and part in NAND flash 108.



A fourth data set indicated by 312, 314 is respectively stored part in the PCM 105 and part in NAND flash 108.

[0040] Referring also to FIG. 3B, in another innovation in this invention, example operations generally designated by the reference character 318 of the eSCM 100, including straddling data sets across different memory technologies of the eSCM 100 and the HDD/SSD 214 of system 200. A fifth data set indicated by 320, 321, 322, 323 is respectively stored part in DRAM 104, part in PCM 106, part in NAND flash 108 and part in HDD/SSD 214. A sixth data set indicated by 324, 325, 326 is respectively stored part in PCM 106, part in NAND flash 108 and part in HDD/SSD 214. A seventh data set indicated by 327, 328 is respectively stored part in NAND flash 108 and part in HDD/SSD 214. A further data set indicated by 329 is stored in the NAND flash 108 and data set indicated by 330 is stored in the HDD/SSD 214. It should be understood that this innovation will be used to support another innovation in this invention, where the higher read latency of a given solid state memory technology is partially or completely hidden by the operation of another solid state memory technology with lower read latency.

[0041] Referring now to FIGS. 4A, 4B, 4C are flow charts illustrating example operations of the eSCM 100 for implementing enhanced solid-state storage usage performance in accordance with embodiments of the invention.

[0042] In FIG. 4A, example operations, for example, performed by CPU 110, start as indicated at a block 400. eSCM CPU or eSCM controller 110 performs workload

recognition and presents only a memory interface to the computer system 200 as indicated at a block 402, which allows not only complete software compatibility but also complete hardware compatibility with computer systems using only DRAM. Hence, existing DIMMs in substantially all existing systems can be swapped out for the new eSCM 100 in accordance with embodiments of the invention. As indicated at a block 404, eSCM controller 110 selectively moves data among DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND Flash 108 with data types used to achieve improved latency, and throughput performance characteristics.

[0043] In FIG. 4B, example operations, for example, performed by eSCM controller 110 continue with writing data to the NAND Flash 108 and never are initial writes to PCM 105, with all writes buffered in DRAM and sent to the NAND Flash as indicated at a block 410 in accordance with embodiments of the invention. This strategy exploits both the lower write process time in NAND Flash as opposed to PCM, and also the possibility of a design decision to further parallelize access to the much larger capacity available in NAND Flash relative to PCM in an embodiment of this invention. As indicated at a block 412, data is selectively migrated among DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND Flash 108 according to data set sizes. As indicated at a block 414, reads retrieve data from any of the memory technologies including DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND Flash 108. This is another detail that indicates none of these solid state memory technologies are being used as cache of another in an embodiment of the invention. Nevertheless, those

skilled in the art will readily recognize that adding a cache strategy in addition to the strategies described in this invention is straightforward without departing from the spirit of this invention.

[0044] In FIG. 4C, example operations, for example, performed by eSCM controller 110 include identifying DRAM size to use for data storage and write buffer, a smart decision algorithm is used for data set activity detection and categorization as indicated at a block 420. As indicated at a block 422 data is selectively allocated primarily in non-volatile PCM 105, and NAND Flash 108, exploiting non-volatility for low power instead of refreshing large DRAM sets. PCM 105 by array design is geared toward low density, low latency, and smaller sized data sets. NAND Flash 108 by array design is geared toward high density, relatively higher latency, and larger sized data sets. The smaller data sets with high frequency of writes are preferably committed to DRAM 104 itself, from which they can be retrieved with minimal latency.

[0045] Another important innovation, as indicated at a block 424, depending on data set sizes optionally a given data set is straddled across different solid-state technologies including DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND Flash 108, and optionally further across HDD/SSD 204. This allows for hiding latencies of PCM 105 or NAND Flash 108 in some data sets as detailed below.

[0046] Referring now to FIG. 5, there is schematically shown another more detailed example solid-state storage system generally designated by the reference

character 500 for implementing enhanced solid-state Enhanced Storage Class Memory (eSCM) in accordance with embodiments of the invention. In the example eSCM 500, ReRAM, and STT-RAM are not shown. In this embodiment, NAND Flash is further partitioned in single-level (SLC) and multi-level cell (MLC) technologies. Thus, Solid-state Enhanced Storage Class Memory (eSCM) system 500 includes DRAM 104 including DRAM chips 502, PCM 105 including PCM chips 504, and NAND Flash 108 including a combination of NAND Flash single-level cell (SLC) chips 506 and NAND Flash multi-level cell (MLC) chips 508. Solid-state storage system eSCM 500 includes a processor 510 and a plurality of bus buffers 1-N, 512, together with the DRAM chips 502, PCM chips 504, and NAND Flash SLC chips 506 and NAND Flash MLC chips 508.

**[0047]** In accordance with features of the embodiments of the invention, bandwidth is handled by eSCM processor 510 by buffering and parallelization, using bus buffers 1-N, 512 with the DRAM chips 502, PCM chips 504, and NAND Flash SLC chips 506 and NAND Flash MLC chips 508.

**[0048]** Recalling that according to size, data sets can straddle different solid state memory technologies, latency from one solid state memory technology can be hidden or partially hidden by another lower latency solid state technology. Referring now to FIGS. 6A, 6B, 6C charts are shown schematically illustrating example read operations of the eSCM system 100 or solid-state storage system eSCM 500 for implementing enhanced solid-state storage usage performance in accordance with embodiments of the invention.

As described above to implement enhanced solid-state storage latency performance, data is migrated among DRAM 108, PCM 105 and NAND Flash 104 and DRAM chips 502, PCM chips 504, and NAND Flash SLC chips 506 and NAND Flash MLC chips 508 depending on data set sizes.

[0049] In FIG. 6A, example read operations generally designated by the reference character 600, for example performed by eSCM controller 110, with data read flow 602 of small chunks from DRAM 104 and PCM 105. For example, small requests are sometimes as small as 32 B or 64 B, but average main memory accesses tend to get chunks of 16 KB average. In this example, a data set straddling DRAM 104 and PCM 105 technologies is to be read. The SCM responds to the read request by sending first the part of the data requested which resides in DRAM 104. In parallel, and hidden from the host, the SCM starts fetching the remaining data from PCM 105. Hence, the latency from the PCM 105 technology is hidden or partially hidden by the latency in DRAM 104 access and the time taken to transfer that data from DRAM 104 to the host 202. Depending on the size of the partitions of the data set in DRAM 104 and PCM 105, the latency of these two technologies and the speed of the bus, the higher latency from the PCM access can be completely hidden and unnoticed by the host 202. Hence, this solution behaves as if all the data set were in DRAM, but the cost of this solution will be proportionally lowered by the relative amount of PCM and DRAM. Moreover, in an event of a power loss, only the part of the data residing in DRAM needs to be saved to a non-volatile memory in the eSCM.

[0050] In FIG. 6B, example read operations generally designated by the reference character 610, for example performed by eSCM controller 110, with data read flow 612 of medium sized chunks from DRAM 108, PCM 105 and NAND Flash 104. In this case the data set is large enough to straddle three different technologies. As part of the data is sequentially read and sent to the host from the lowest latency memory technologies, data from the remaining memory technologies also requested is being fetched. Depending on the size of the partitions of the requested data set allocated to each of the memory technologies, the actual latencies of the different solid state memory technologies, and the speed of the bus, the latency from the PCM 105 and NAND Flash 108 accesses can be completely hidden and unnoticed by the host.

[0051] Those skilled in the art will readily recognize other memory technologies can be used in the eSCM and benefit from the same invention described here. Those skilled in the art will also recognize that the size of the data set partitions in each memory technology the data set straddles is a function of the actual latencies of the solid state memory technologies used and the speed of the bus. In an embodiment, careful design might offer partial or total hidden latencies according to how critical a data set is.

[0052] In FIG. 6C, example read operations generally designated by the reference character 620, for example performed by eSCM controller 110, with data read flow 622 of very long chunks from NAND Flash 108. For example, any request for more than 320 KB will allow NAND Flash 108 to engage reading. In such a very large size data set, the

latency from NAND Flash 108 may itself be of less importance, and the SCM could allocate the entire data set in NAND Flash 108.

[0053] Those skilled in the art will readily recognize that the strategy of allocating data primarily according to data set size can be used in conjunction with ancillary strategies for the case where a large amount of data of a particular size might not fit the memory space available at a particular solid state memory technology. In such a case, a secondary criteria based on frequency of use of a data set can be used to decide which data set will be placed in total or in part (in case it straddles more than one solid state technology) in the lower latency position in the storage.

[0054] Those skilled in the art will readily recognize that the strategy of allocating data of the invention includes that a given data set optionally is straddled across different solid-state technologies including DRAM 104, PCM 105, ReRAM 106, STT-RAM 107, and NAND Flash 108, and optionally further across HDD/SSD 204.

[0055] While the present invention has been described with reference to the details of the embodiments of the invention shown in the drawing, these details are not intended to limit the scope of the invention as claimed in the appended claims, as interpreted by the description and drawings.

## Claims

1           1.       A method for implementing enhanced solid-state storage performance  
2 comprising:

3                 providing a direct attached dual in line memory (DIMM) card containing  
4 dynamic random access memory (DRAM), and at least one non-volatile memory; and  
5                 selectively allocating data among the DRAM, and the at least one non-volatile  
6 memory based upon a data set size.

1           2.       The method as recited in claim 1 wherein selectively allocating data  
2 among the DRAM, and the at least one non-volatile memory based upon a data set size  
3 includes selectively partitioning data among the DRAM, and the at least one non-  
4 volatile memory.

1           3.       The method as recited in claim 1 includes performing a read operation  
2 where part of the requested data is read from DRAM with simultaneously fetching parts  
3 of requested data from the at least one non-volatile memory.

1           4.       The method as recited in claim 1 wherein providing the DRAM and the  
2 at least one non-volatile memory includes providing Phase Change memory (PCM) and  
3 NAND flash memory.



1           5.       The method as recited in claim 4 includes providing an eSCM processor  
2 on said DIMM card with said DRAM, said PCM, and said NAND flash memory, and  
3           using said eSCM processor, selectively moving data among said DRAM, said  
4 PCM, and said NAND flash memory for enabling enhanced latency and throughput  
5 performance.

1           6.       The method as recited in claim 5 wherein selectively moving data among  
2 said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3 and throughput performance includes selectively migrating data among said DRAM,  
4 said PCM, and said NAND flash memory based upon data set sizes.

1           7.       The method as recited in claim 6 further includes selectively migrating  
2 data among said DRAM, said PCM, and said NAND flash memory based upon  
3 frequency of use.

1           8.       The method as recited in claim 5 wherein selectively moving data among  
2 said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3 and throughput performance includes writing first to said NAND flash memory and  
4 selectively moving data to at least one of said PCM, and said DRAM.

1           9.       The method as recited in claim 5 wherein selectively moving data among  
2 said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3 and throughput performance includes buffering data being written to said DRAM,  
4 before committing to said NAND flash memory; and selectively maintaining said data  
5 committed in said NAND flash memory or selectively moving data committed to said  
6 NAND flash memory to at least one of said DRAM and said PCM.

1           10.      The method as recited in claim 5 wherein selectively moving data among  
2 said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3 and throughput performance includes presenting a memory interface to a second  
4 computer system, said second computer system including a direct attached dual in line  
5 memory (DIMM) card containing dynamic random access memory (DRAM), and at  
6 least one non-volatile memory.

1           11.      The method as recited in claim 5 wherein selectively moving data among  
2 said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3 and throughput performance includes reading data from any of said DRAM, said PCM,  
4 and said NAND flash memory.

1           12.     The method as recited in claim 5 wherein selectively moving data among  
2     said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3     and throughput performance includes selectively allocating data primarily in said non-  
4     volatile memory including said PCM, and said NAND flash memory.

1           13.     The method as recited in claim 5 wherein selectively moving data among  
2     said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3     and throughput performance includes using said PCM primarily for storing relatively  
4     middle sized data sets.

1           14.     The method as recited in claim 5 wherein selectively moving data among  
2     said DRAM, said PCM, and said NAND flash memory for enabling enhanced latency  
3     and throughput performance includes using said NAND flash memory primarily for  
4     storing relatively large data sets.

1           15.    An apparatus for implementing enhanced solid-state storage performance  
2 comprising:  
3           a direct attached dual in line memory (DIMM) card, said DIMM card containing  
4 dynamic random access memory (DRAM), and at least one non-volatile memory; Phase  
5 Change memory (PCM) and NAND flash memory;  
6           an eSCM processor coupled to said DRAM, and said at least one non-volatile  
7 memory on said DIMM card, and  
8           said eSCM processor, selectively allocating data among the DRAM, and the at  
9 least one non-volatile memory based upon a data set size.

1           16.    The apparatus as recited in claim 15 includes control code stored on a  
2 computer readable medium, and wherein said eSCM processor uses said control code  
3 for implementing enhanced solid-state storage performance.

1           17.    The apparatus as recited in claim 15 wherein said wherein said at least  
2 one non-volatile memory includes at least one of a Phase Change memory (PCM),  
3 Resistive RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash  
4 memory.

1           18.    The apparatus as recited in claim 15 wherein said at least one non-  
2 volatile memory includes Phase Change memory (PCM) and NAND flash memory.

1            19.     The apparatus as recited in claim 18 includes said eSCM processor, using  
2     an intelligent data size detection algorithm for selectively moving data among said  
3     DRAM, said PCM, and said NAND flash memory.

1            20.     The apparatus as recited in claim 19 wherein said eSCM processor,  
2     selectively moving data among said DRAM, said PCM, and said NAND flash memory  
3     includes said eSCM processor writing to said NAND flash memory and selectively  
4     moving data to said PCM.

1            21.     The apparatus as recited in claim 19 wherein said eSCM processor,  
2     selectively moving data among said DRAM, said PCM, and said NAND flash memory  
3     includes said eSCM processor selectively migrating data among said DRAM, said PCM,  
4     and said NAND flash memory according to data set sizes.

1            22.     The apparatus as recited in claim 19 wherein said eSCM processor,  
2     selectively moving data among said DRAM, said PCM, and said NAND flash memory  
3     includes said eSCM processor selectively allocating data primarily in non-volatile  
4     memory including said PCM, and said NAND flash memory.

1            23.     The apparatus as recited in claim 19 wherein said eSCM processor,  
2     selectively moving data among said DRAM, said PCM, and said NAND flash memory  
3     includes said eSCM processor using said PCM primarily for storing relatively medium  
4     sized data sets.

1           24.     The apparatus as recited in claim 19 wherein said eSCM processor,  
2 selectively moving data among said DRAM, said PCM, and said NAND flash memory  
3 includes said eSCM processor using said NAND flash memory primarily for storing  
4 high density large data sets.

1           25.     The apparatus as recited in claim 15 includes a memory interface to a  
2 second computer system, said second computer system including a direct attached dual  
3 in line memory (DIMM) card containing dynamic random access memory (DRAM),  
4 and at least one non-volatile memory coupled to said memory interface.

1           26.     An enhanced solid-state storage class memory (eSCM) system  
2 comprising:  
3           a dynamic random access memory (DRAM),  
4           at least one non-volatile memory;  
5           a processor coupled to said DRAM, and said at least one non-volatile memory;  
6 said processor allocating data among the DRAM, and the at least one non-volatile  
7 memory based upon a data set size.

1           27.     The eSCM system as recited in claim 26 wherein said at least one non-  
2 volatile memory includes Phase Change memory (PCM) and NAND flash memory.

1           28.     The eSCM system as recited in claim 27 includes a direct attached dual  
2 in line memory (DIMM) card, said DIMM card containing said dynamic random access  
3 memory (DRAM), Phase Change memory (PCM) and NAND flash memory.

1           29.     The eSCM system as recited in claim 26 includes a memory interface to a  
2 second computer system, said second computer system including a direct attached dual  
3 in line memory (DIMM) card containing dynamic random access memory (DRAM),  
4 and at least one non-volatile memory coupled to said memory interface.

1           30.     The eSCM system as recited in claim 26 includes a memory interface to  
2 another storage system including at least one of a Solid State drive (SSD), and a hard  
3 disk drive (HDD).

1           31.     The eSCM system as recited in claim 26 wherein said at least one non-  
2 volatile memory includes at least one of a Phase Change memory (PCM), Resistive  
3 RAM (ReRAM), Spin-Transfer-Torque RAM (STT-RAM), and NAND flash memory.

1           32.     The eSCM system as recited in claim 26 includes a memory interface to a  
2 second storage system including at least one of a Solid State drive (SSD), and a hard  
3 disk drive (HDD) and further includes said processor selectively migrating data among  
4 said DRAM, said PCM, said ReRAM, said STT-RAM and said NAND flash memory;  
5 and said second storage system.

6 33. A method for implementing enhanced solid state performance substantially as  
7 any one herein described with reference to the accompanying drawings.

8 34. An apparatus for implementing enhanced solid state performance substantially as  
9 herein described with reference to an as illustrated in Figs. 1 to 4 and 6 or Fig. 5 of the  
10 accompanying drawings.

11 35. An enhanced solid-state storage class memory system substantially as herein  
12 described with reference to an as illustrated in Figs. 1 to 4 and 6 or Fig. 5 of the  
13 accompanying drawings.





Application No: GB1317108.7

Examiner: Alessandro Potenza

Claims searched: 1-35

Date of search: 21 February 2014

## Patents Act 1977: Search Report under Section 17

### Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1-35	US 2008/0126716 A1 (IBM) see figure 1 and paragraphs 4, 7-8, 26-27, 36-37
X	1-35	US 2012/0137055 A1 (SAMSUNG) see figure 1 and paragraphs 71 and 79
X,E	1, 15, 26	WO 2013/165386 A1 (HEWLETT PACKARD) see figure 3 and paragraphs 11, 14, and 26
A	-	US 6336174 B1 (MAXTOR)
A	-	US 2011/0167197 A1 (LEINWANDER)
A	-	US 2009/0248965 A1 (ASMEDIA TECHNOLOGY)

### Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

### Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC<sup>X</sup> :

Worldwide search of patent documents classified in the following areas of the IPC

G06F

The following online and other databases have been used in the preparation of this search report

EPODOC, WPI, TXTEN, INSPEC, XPESP, XPESP2, XPIEE, XPIPCOM, XPI3E, XPMISC, XPLNCS, XPRD, XPSRNG, TDB



**International Classification:**

<b>Subclass</b>	<b>Subgroup</b>	<b>Valid From</b>
G06F	0012/06	01/01/2006
G06F	0012/02	01/01/2006
G06F	0013/16	01/01/2006