

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6479639号
(P6479639)

(45) 発行日 平成31年3月6日 (2019.3.6)

(24) 登録日 平成31年2月15日 (2019.2.15)

(51) Int. Cl.

F I

G O 6 F 12/02 (2006.01)

G O 6 F 12/00 (2006.01)

G O 6 F 12/02 5 1 0 A

G O 6 F 12/00 5 9 7 U

請求項の数 10 (全 23 頁)

(21) 出願番号	特願2015-242243 (P2015-242243)	(73) 特許権者	318010018
(22) 出願日	平成27年12月11日 (2015.12.11)		東芝メモリ株式会社
(65) 公開番号	特開2016-115355 (P2016-115355A)		東京都港区芝浦一丁目1番1号
(43) 公開日	平成28年6月23日 (2016.6.23)	(74) 代理人	110001737
審査請求日	平成29年8月4日 (2017.8.4)		特許業務法人スズエ国際特許事務所
(31) 優先権主張番号	62/090,690	(72) 発明者	菅野 伸一
(32) 優先日	平成26年12月11日 (2014.12.11)		東京都港区芝浦一丁目1番1号 株式会社
(33) 優先権主張国	米国 (US)		東芝内
(31) 優先権主張番号	14/656,413	(72) 発明者	橋本 大輔
(32) 優先日	平成27年3月12日 (2015.3.12)		東京都港区芝浦一丁目1番1号 株式会社
(33) 優先権主張国	米国 (US)		東芝内
		審査官	滝谷 亮一
			最終頁に続く

(54) 【発明の名称】 情報処理装置、プログラム、及び、情報処理システム

(57) 【特許請求の範囲】

【請求項1】

メモリシステムと通信する情報処理装置であって、
前記メモリシステムは、複数の第1の消去単位エリアを含む第1のスペースと、複数の第2の消去単位エリアを含む第2のスペースとを具備する不揮発性メモリと、前記不揮発性メモリを制御するコントローラとを具備し、
前記情報処理装置は、前記メモリシステムの外部にプロセッサを具備し、
前記プロセッサは、
第1のオブジェクトに対応する第1の論理アドレス空間に対する第1の書き込み頻度と、第2のオブジェクトに対応する第2の論理アドレス空間に対する第2の書き込み頻度とが異なる書き込み頻度グループに属する場合に、前記第1及び第2の論理アドレス空間を前記第1のスペースに割り当て、第3のオブジェクトに対応する第3の論理アドレス空間に対する第3の書き込み頻度と、第4のオブジェクトに対応する第4の論理アドレス空間に対する第4の書き込み頻度とが異なる書き込み頻度グループに属する場合に、前記第3及び第4の論理アドレス空間を前記第2のスペースに割り当て、
書き込みコマンド、前記書き込みコマンドに対応する論理アドレス空間に割り当てられたスペースの識別データ、書き込みデータ、前記書き込みデータの論理アドレスを前記コントローラへ送信し、
読み出しコマンド、前記識別データ、読み出しデータの論理アドレスを前記コントローラへ送信し、前記コントローラから前記読み出しデータを受信する、

10

20

情報処理装置。

【請求項 2】

前記プロセッサは、前記第 1 乃至第 4 の論理アドレス空間ごとに前記第 1 乃至第 4 の書き込み頻度を算出し、

前記第 1 乃至第 4 の論理アドレス空間ごとの前記第 1 乃至第 4 の書き込み頻度に基づいて、前記第 1 及び第 2 の論理アドレス空間を前記第 1 のスペースに割り当て、前記第 3 及び第 4 の論理アドレス空間を前記第 2 のスペースに割り当てる、

請求項 1 の情報処理装置。

【請求項 3】

前記プロセッサは、前記第 1 のオブジェクトを示す第 1 のオブジェクト識別データ及び前記第 1 の書き込み頻度に基づいて、前記第 1 のオブジェクトに対応する前記第 1 の論理アドレス空間を前記第 1 のスペースに割り当て、前記第 2 のオブジェクトを示す第 2 のオブジェクト識別データ及び前記第 2 の書き込み頻度に基づいて、前記第 2 のオブジェクトに対応する前記第 2 の論理アドレス空間を前記第 1 のスペースに割り当て、前記第 3 のオブジェクトを示す第 3 のオブジェクト識別データ及び前記第 3 の書き込み頻度に基づいて、前記第 3 のオブジェクトに対応する前記第 3 の論理アドレス空間を前記第 2 のスペースに割り当て、前記第 4 のオブジェクトを示す第 4 のオブジェクト識別データ及び前記第 4 の書き込み頻度に基づいて、前記第 4 のオブジェクトに対応する前記第 4 の論理アドレス空間を前記第 2 のスペースに割り当てる、

請求項 1 の情報処理装置。

【請求項 4】

前記スペースの前記識別データは、オブジェクトに対応する論理アドレス空間に割り当てられたスペースを示し、

前記プロセッサは、前記書き込みコマンド、前記スペースの前記識別データ、前記書き込みデータ、前記書き込みデータの論理アドレスに加えて、前記書き込みデータのサイズを送信する、

請求項 1 乃至請求項 3 のうちのいずれか 1 項の情報処理装置。

【請求項 5】

コンピュータを、

割り当て部と発送部として機能させ、

前記コンピュータは、メモリシステムと通信するプロセッサを具備し、

前記メモリシステムは、複数の第 1 の消去単位エリアを含む第 1 のスペースと、複数の第 2 の消去単位エリアを含む第 2 のスペースとを具備する不揮発性メモリと、前記不揮発性メモリを制御するコントローラとを具備し、

前記割り当て部は、前記メモリシステムの外部にあり、

前記割り当て部は、第 1 のオブジェクトに対応する第 1 の論理アドレス空間に対する第 1 の書き込み頻度と、第 2 のオブジェクトに対応する第 2 の論理アドレス空間に対する第 2 の書き込み頻度とが異なる書き込み頻度グループに属する場合に、前記第 1 及び第 2 の論理アドレス空間を前記第 1 のスペースに割り当て、

前記割り当て部は、第 3 のオブジェクトに対応する第 3 の論理アドレス空間に対する第 3 の書き込み頻度と、第 4 のオブジェクトに対応する第 4 の論理アドレス空間に対する第 4 の書き込み頻度とが異なる書き込み頻度グループに属する場合に、前記第 3 及び第 4 の論理アドレス空間を前記第 2 のスペースに割り当て、

前記発送部は、書き込みコマンド、前記書き込みコマンドに対応する論理アドレス空間に割り当てられたスペースの識別データ、書き込みデータ、前記書き込みデータの論理アドレスを前記コントローラへ送信し、

前記発送部は、読み出しコマンド、前記識別データ、読み出しデータの論理アドレスを前記コントローラへ送信し、前記コントローラから前記読み出しデータを受信する、プログラム。

【請求項 6】

前記コンピュータを、さらに、算出部として機能させ、

前記算出部は、前記第1乃至第4の論理アドレス空間ごとに前記第1乃至第4の書き込み頻度を算出し、

前記割り当て部は、前記第1乃至第4の論理アドレス空間ごとの前記第1乃至第4の書き込み頻度に基づいて、前記第1及び第2の論理アドレス空間を前記第1のスペースに割り当て、前記第3及び第4の論理アドレス空間を前記第2のスペースに割り当てる、請求項5のプログラム。

【請求項7】

前記割り当て部は、前記第1のオブジェクトを示す第1のオブジェクト識別データ及び前記第1の書き込み頻度に基づいて、前記第1のオブジェクトに対応する前記第1の論理アドレス空間を前記第1のスペースに割り当て、前記第2のオブジェクトを示す第2のオブジェクト識別データ及び前記第2の書き込み頻度に基づいて、前記第2のオブジェクトに対応する前記第2の論理アドレス空間を前記第1のスペースに割り当て、前記第3のオブジェクトを示す第3のオブジェクト識別データ及び前記第3の書き込み頻度に基づいて、前記第3のオブジェクトに対応する前記第3の論理アドレス空間を前記第2のスペースに割り当て、前記第4のオブジェクトを示す第4のオブジェクト識別データ及び前記第4の書き込み頻度に基づいて、前記第4のオブジェクトに対応する前記第4の論理アドレス空間を前記第2のスペースに割り当てる、請求項5のプログラム。

【請求項8】

前記スペースの前記識別データは、オブジェクトに対応する論理アドレス空間に割り当てられたスペースを示し、

前記発送部は、前記書き込みコマンド、前記スペースの前記識別データ、前記書き込みデータ、前記書き込みデータの前記論理アドレスに加えて、前記書き込みデータのサイズを送信する、請求項5乃至請求項7のうちのいずれか1項のプログラム。

【請求項9】

請求項1乃至請求項4のいずれか1項の前記情報処理装置と前記メモリシステムとを具備する情報処理システムであって、

前記コントローラは、

前記複数の第1の消去単位エリアを前記第1のスペースに割り当て、前記複数の第2の消去単位エリアを前記第2のスペースに割り当て、

前記書き込みコマンド、前記識別データ、前記書き込みデータ、前記論理アドレスを、前記情報処理装置から受信し、

前記論理アドレス及び前記識別データと、前記識別データによって示される前記スペースに対応する物理アドレスとを関連付け、

前記物理アドレスによって示される位置に、前記書き込みデータを書き込む、情報処理システム。

【請求項10】

請求項1乃至請求項4のいずれか1項の前記情報処理装置と前記メモリシステムとを具備する情報処理システムであって、

前記コントローラは、

前記読み出しコマンド、前記識別データ、前記論理アドレスを、前記情報処理装置から受信し、

前記論理アドレス及び前記識別データと、前記識別データによって示される前記スペースに対応する物理アドレスとの関連付けに基づいて、前記論理アドレスを前記物理アドレスへ変換し、

前記読み出しデータを、前記物理アドレスによって示される位置から読み出し、

前記読み出しデータを、前記情報処理装置へ送信する、情報処理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本実施形態は、情報処理装置、プログラム、及び、情報処理システムに関する。

【背景技術】

【0002】

ソリッドステートドライブ（SSD）は、不揮発性半導体メモリを備え、ハードディスクドライブ（HDD）と同様のインタフェースを持つ。例えば、SSDは、データ書き込み時に、書き込みコマンド、書き込み先のLogical Block Addressing（LBA）、書き込みデータを情報処理装置から受け、Look Up Table（LUT）に基づいて、LBAをPhysical Block Addressing（PBA）に変換し、PBAの示す位置に書き込みデータを書き込む。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】 米国特許第7739444号明細書

【特許文献2】 米国特許出願公開第2013/0138867号明細書

【特許文献3】 米国特許出願公開第2014/0195725号明細書

【発明の概要】

【発明が解決しようとする課題】

20

【0004】

本実施形態は、不揮発性メモリへコマンドを送る情報処理装置、プログラム、及び、情報処理システムを提供する。

【課題を解決するための手段】

【0005】

本実施形態によれば、情報処理装置は、メモリシステムと通信する。メモリシステムは、複数の第1の消去単位エリアを含む第1のスペースと、複数の第2の消去単位エリアを含む第2のスペースとを具備する不揮発性メモリと、不揮発性メモリを制御するコントローラとを含む。情報処理装置は、メモリシステムの外部にプロセッサを含む。プロセッサは、第1のオブジェクトに対応する第1の論理アドレス空間に対する第1の書き込み頻度と、第2のオブジェクトに対応する第2の論理アドレス空間に対する第2の書き込み頻度とが異なる書き込み頻度グループに属する場合に、第1及び第2の論理アドレス空間を第1のスペースに割り当て、第3のオブジェクトに対応する第3の論理アドレス空間に対する第3の書き込み頻度と、第4のオブジェクトに対応する第4の論理アドレス空間に対する第4の書き込み頻度とが異なる書き込み頻度グループに属する場合に、第3及び第4の論理アドレス空間を第2のスペースに割り当て、書き込みコマンド、書き込みコマンドに対応する論理アドレス空間に割り当てられたスペースの識別データ、書き込みデータ、書き込みデータの論理アドレスをコントローラへ送信し、読み出しコマンド、識別データ、読み出しデータの論理アドレスをコントローラへ送信し、コントローラから読み出しデータを受信する。

30

40

【図面の簡単な説明】

【0006】

【図1】 第1の実施形態に係る情報処理システムの構成の一例を示すブロック図。

【図2】 LBA空間と、ネームスペースと、アドレス変換テーブルと、ガベージコレクション部と、管理データとの関係の一例を示すブロック図。

【図3】 第1の実施形態に係る受付部及び設定部の処理の一例を示すフローチャート。

【図4】 第1の実施形態に係るガベージコレクション部及びアドレス変換部の処理の一例を示すフローチャート。

【図5】 第1の実施形態に係るネームスペースの割り当て状態を例示するブロック図。

【図6】 第1の実施形態に係る情報処理装置の処理の一例を示すフローチャート。

50

【図 7】第 2 の実施形態に係る情報処理システムの構成の一例を示すブロック図。

【図 8】第 2 の実施形態に係る変換テーブルの一例を示すデータ構造図。

【図 9】第 2 の実施形態に係るメモリシステムの書き込み処理の一例を示すフローチャート。

【図 10】第 2 の実施形態に係るメモリシステムの読み出し処理の一例を示すフローチャート。

【図 11】第 3 の実施形態に係る情報処理システムの構成の一例を示すブロック図。

【図 12】第 3 の実施形態に係るストレージシステムを示す斜視図。

【発明を実施するための形態】

【0007】

10

以下、図面を参照して実施形態について説明する。以下の説明において、略同一の機能及び構成要素については、同一符号を付し、重複説明は必要な場合にのみ行う。以下の実施形態において、アクセスとは、データの読み出し及びデータの書き込みの双方を意味する。

【0008】

[第 1 の実施形態]

図 1 は、本実施形態に係る情報処理システムの構成の一例を示すブロック図である。

【0009】

情報処理システム 1 は、情報処理装置 2 と、メモリシステム 3 とを含む。情報処理システム 1 は、複数の情報処理装置 2 を含むとしてもよい。情報処理システム 1 が複数の情報処理装置を備える場合については、後述の第 2 の実施形態で説明する。

20

【0010】

(メモリシステム 3 の説明)

メモリシステム 3 は、例えば SSD であり、コントローラ 4 と不揮発性メモリ 5 とを含む。メモリシステム 3 は、情報処理装置 2 に内蔵されてもよく、情報処理装置 2 とメモリシステム 3 とは、ネットワークなどによりデータを送受信可能に接続されるとしてもよい。

【0011】

本実施形態において、不揮発性メモリ 5 としては、例えば、少なくとも 1 つの NAND 型フラッシュメモリが用いられる。しかしながら、本実施形態は、例えば、NOR 型フラッシュメモリ、MRAM (Magnetoresistive Random Access Memory)、PRAM (Phase change Random Access Memory)、ReRAM (Resistive Random Access Memory)、又は、FeRAM (Ferroelectric Random Access Memory) などのような各種の不揮発性メモリが複数の書き込み管理エリアを含む場合に、適用可能である。ここで、書き込み管理エリアは、書き込み回数を管理する単位のエリアとする。不揮発性メモリ 5 は、3 次元構造のフラッシュメモリを含むとしてもよい。

30

【0012】

例えば、不揮発性メモリ 5 は、複数のブロック (物理ブロック) を備える。複数のブロックは、ワード線とビット線との交点に配置される複数のメモリセルを含む。不揮発性メモリ 5 では、ブロック単位にデータが一括して消去される。すなわち、ブロックは、データ消去単位の領域である。データ書き込み及びデータ読み出しは、各ブロックにおけるページ (ワード線) 単位で行う。すなわち、ページは、データの書き込み単位又はデータの読み出し単位の領域である。

40

【0013】

本実施形態では、書き込み回数がブロック単位で管理されとする。

【0014】

情報処理装置 2 は、メモリシステム 3 のホスト装置である。情報処理装置 2 は、メモリシステム 3 へ、不揮発性メモリ 5 のブロックと、少なくとも 1 つのブロックを含むスペースとを関係付けるための設定コマンド C1 を送る。

【0015】

50

以下の説明では、スペースは、ネームスペースであるとして説明する。

【0016】

また、情報処理装置2は、書き込みコマンドC2とともに、ネームスペースの識別データ(NSID)6、書き込み先を示すLBA7、書き込みデータのデータサイズ8、及び、書き込みデータ9を、メモリシステム3へ送る。

【0017】

本実施形態において、複数のネームスペース $NS_0 \sim NS_M$ (M は1以上の整数)のそれぞれは、不揮発性メモリ5に含まれる複数のブロック $B_0 \sim B_N$ (N は M 以上の整数)を区分けすることによって得られるスペースである。本実施形態では、ネームスペース NS_0 がブロック $B_0 \sim B_2$ を含み、ネームスペース NS_M がブロック $B_{N-2} \sim B_N$ を含む。他のネームスペース $NS_1 \sim NS_{M-1}$ もネームスペース NS_0 、 NS_M と同様であるとする。なお、本実施形態におけるネームスペース $NS_0 \sim NS_M$ とブロック $B_0 \sim B_N$ との間の割り当て関係は一例であり、1つのネームスペースに割り当てられるブロックの個数などは適宜変更可能である。複数のネームスペースの間でブロックの個数は異なってもよい。

10

【0018】

コントローラ4は、記憶部10、バッファメモリ $F_0 \sim F_M$ 、プロセッサ11を含む。

【0019】

記憶部10は、各ネームスペース $NS_0 \sim NS_M$ に対応するアドレス変換テーブル $T_0 \sim T_M$ を格納する。例えば、記憶部10は、作業用メモリとして使用されてもよい。記憶部10は、例えばDRAM (Dynamic Random Access Memory) 又はSRAM (Static Random Access Memory) などのような揮発性メモリでもよく、又は、不揮発性メモリでもよい。記憶部10は、不揮発性メモリと揮発性メモリとの組み合わせでもよい。

20

【0020】

アドレス変換テーブル $T_0 \sim T_M$ は、それぞれ、ネームスペース $NS_0 \sim NS_M$ に対するデータ書き込みにおいて、LBAとPBAとを関係付けたデータであり、例えばLUTである。なお、アドレス変換テーブル $T_0 \sim T_M$ の一部又は全部は、メモリ12などのような他のメモリに格納されてもよい。

【0021】

バッファメモリ $F_0 \sim F_M$ は、それぞれ、ネームスペース $NS_0 \sim NS_M$ に対するデータ書き込みにおいて、書き込みデータを書き込みに適したデータ量になるまで格納する。

30

【0022】

プロセッサ11は、メモリ12、受付部13、設定部14、アドレス変換部15、書き込み部16、ガベージコレクション部 $G_0 \sim G_M$ を含む。

【0023】

メモリ12は、プログラム17及び管理データ18を格納する。本実施形態において、メモリ12は、プロセッサ11に含まれているが、プロセッサ11の外部に備えられてもよい。メモリ12は、例えば不揮発性のメモリとする。

【0024】

プログラム17は、例えばファームウェアである。プロセッサ11は、プログラム17を実行することにより、受付部13、設定部14、アドレス変換部15、書き込み部16、ガベージコレクション部 $G_0 \sim G_M$ として機能する。なお、プログラム17及び管理データ18の一部又は全部は、記憶部10などのような他のメモリに格納されてもよい。

40

【0025】

管理データ18は、ネームスペース $NS_0 \sim NS_M$ とブロック $B_0 \sim B_N$ との関係を示すデータである。管理データ18を参照することにより、どのブロックがどのネームスペースに所属しているか判断可能である。

【0026】

受付部13は、情報処理装置2から、不揮発性メモリ5の各ブロックと、各ネームスペースとを関係付けるための設定コマンドC1を受け付ける。また、受付部13は、情報処理装置2から、書き込みコマンドC2、NSID6、LBA7、データサイズ8、データ

50

9を受け付ける。

【0027】

以下の説明では、説明を簡略化するため、書き込みコマンドC2に、ネームスペースNS₀を示すNSID6が付されている場合について説明する。しかしながら、書き込みコマンドC2に、他のネームスペースNS₁～NS_Mを示すNSIDを付することも可能である。

【0028】

設定部14は、受付部13によってネームスペースの設定コマンドC1が受け付けられた場合に、設定コマンドC1に基づいて、ネームスペースNS₀～NS_Mに、ブロックB₀～B_Nを割り当てて管理データ18を生成し、管理データ18をメモリ12に格納する。ネームスペースNS₀～NS_MとブロックB₀～B_Nとの間の割り当ては、設定部14がネームスペースNS₀～NS_Mのデータ格納状態を監視し、各ネームスペースNS₀～NS_Mで、データ容量、アクセス頻度、書き込み頻度、アクセス回数、書き込み回数、又は、データ格納率が同レベルとなるように行われてもよく、情報処理装置2からの指示に応じて行われてもよく、メモリシステム3の管理者の指示に応じて行われてもよい。

【0029】

ここで、データ容量とは書き込み可能なデータサイズであり、アクセス頻度又は書き込み頻度とは単位時間当たりのアクセス回数又は書き込み回数であり、データ格納率とは、ある領域サイズに対するデータ格納済みの領域サイズの割合を示す値とする。

【0030】

また、設定部14は、ネームスペースNS₀～NS_Mごとに実行されたガベージコレクションの結果に基づいて、データを格納していない空ブロックを、ガベージコレクション前に属するネームスペースから他のネームスペースに変更し、管理データ18を更新する。これにより、ネームスペースNS₀～NS_M間でウェアレベリング(Wear Leveling)を実行することができる。ネームスペースNS₀～NS_MとブロックB₀～B_Nとの間の割り当て変更は、設定部14がネームスペースNS₀～NS_Mのデータ格納状態を監視し、上記の管理データ18の生成時と同様に監視結果に応じて行われてもよく、情報処理装置2からの指示に応じて行われてもよく、メモリシステム3の管理者の指示に応じて行われてもよい。例えば、ネームスペースNS₀～NS_Mの変更は、データ容量の少ない、アクセス頻度の低い、アクセス回数の少ない、又は、データ格納率の低いネームスペースの空ブロックを、データ容量の多い、アクセス頻度の高い、アクセス回数の多い、又は、データ格納率の高いネームスペースに変換する。

【0031】

さらに、設定部14は、オーバプロビジョニング(Over Provisioning)のために、設定コマンドC1に基づいて、ネームスペースNS₀～NS_Mごとに、不揮発性メモリ5に、普段使用されない予備領域P₀～P_Mを設定する。予備領域P₀～P_Mの設定は、設定部14がネームスペースNS₀～NS_Mの各データ容量などに応じて行うとしてもよく、情報処理装置2からの指示に応じて行われてもよく、メモリシステム3の管理者の指示に応じて行われてもよい。

【0032】

本実施形態において、予備領域P₀～P_Mは、不揮発性メモリ5に確保されているが、不揮発性メモリ5ではないメモリシステム3内の他のメモリに確保されていてもよい。例えば、予備領域P₀～P_Mは、DRAM、SRAMなどのメモリに確保されてもよい。

【0033】

アドレス変換部15は、受付部13によって書き込みコマンドC2が受け付けられた場合に、書き込みコマンドC2に付されていたNSID6の示すネームスペースNS₀に対応するアドレス変換テーブルT₀に対して、書き込みコマンドC2に付されていたLBA7をPBAに変換するための関係付けを行う。

【0034】

本実施形態において、アドレス変換部15は、プロセッサ11によって実現されるとし

10

20

30

40

50

ている。しかしながら、アドレス変換部15は、プロセッサ11と別の構成としてもよい。

【0035】

また、アドレス変換部15は、アドレス変換テーブル $T_0 \sim T_M$ に基づいてアドレス変換を行うが、これに代えてキーバリュ型検索によりアドレス変換を行ってもよい。例えば、LBAをキーとし、PBAをバリュとすることで、キーバリュ型検索によるアドレス変換は実現可能である。

【0036】

書き込み部16は、アドレス変換部15によって得られたPBAの示す位置に、書き込みデータ9を書き込む。本実施形態において、書き込み部16は、書き込みコマンドC2に付されていたNSID6の示すネームスペース NS_0 に対応するバッファメモリ F_0 に、書き込みデータ9を格納する。そして、書き込み部16は、バッファメモリ F_0 がネームスペース NS_0 に適したデータ量になった場合に、PBAの示す位置に、バッファメモリ F_0 のデータを書き込む。

【0037】

ガベージコレクション部 $G_0 \sim G_M$ は、ネームスペース $NS_0 \sim NS_M$ に対応しており、ネームスペース $NS_0 \sim NS_M$ ごとにガベージコレクションを実行する。ガベージコレクションとは、不要になったメモリ領域を解放する処理、又は、隙間のあるメモリ領域に書き込まれているデータを集めて連続した利用可能なメモリ領域を確保する処理である。ガベージコレクション部 $G_0 \sim G_M$ は、並列にガベージコレクションを実行可能としてもよく、順

【0038】

ガベージコレクション部 $G_0 \sim G_M$ のうちガベージコレクション部 G_0 を用いて具体的に説明する。ガベージコレクション部 G_0 は、まず、管理データ18に基づいて、ネームスペース NS_0 に対応するブロック $B_0 \sim B_2$ を選択する。次に、ガベージコレクション部 G_0 は、選択されたブロック $B_0 \sim B_2$ に対するガベージコレクションを実行する。そして、ガベージコレクション部 G_0 によるガベージコレクション結果に基づいて、アドレス変換部15は、アドレス変換テーブル T_0 を更新する。

【0039】

なお、本実施形態においては、LBAとPBAとがアドレス変換テーブル $T_0 \sim T_M$ で関係付けられており、PBAで特定可能なブロックとNSIDとが管理データ18で関係付けられている。したがって、LBAが重複しないユニークなアドレスであり管理データ18が生成されれば、プロセッサ11側で、書き込みコマンドC2に付されているLBA7から書き込み先のネームスペース NS_0 を特定可能である。したがって、LBA7が重複せず管理データ18が生成された後には、書き込みコマンドC2にNSID6が付されることを省略し、LBA7とアドレス変換テーブル $T_0 \sim T_M$ と管理データ18に基づいて、プロセッサ11側でNSID6が求められるとしてもよい。

【0040】

図2は、LBA空間と、ネームスペース $NS_0 \sim NS_M$ と、アドレス変換テーブル $T_0 \sim T_M$ と、ガベージコレクション部 $G_0 \sim G_M$ と、管理データ18との関係の一例を示すブロック図である。

【0041】

情報処理装置2のLBA空間 $A_0 \sim A_M$ は、それぞれネームスペース $NS_0 \sim NS_M$ に割り当てられている。

【0042】

LBA空間 A_0 は論理アドレス $0 \sim E_0$ を含む。LBA空間 A_1 は論理アドレス $0 \sim E_1$ を含む。LBA空間 A_M は論理アドレス $0 \sim E_M$ を含む。他のLBA空間 $A_2 \sim A_{M-1}$ も同様に複数の論理アドレスを含む。

【0043】

以下の説明では、説明を簡略化するために、LBA空間 A_0 とこのLBA空間 A_0 に対し

て割り当てられているネームスペース NS_0 とを代表して説明する。しかしながら、他の LBA 空間 $A_1 \sim A_M$ 及びネームスペース $NS_1 \sim NS_M$ についても同様である。

【0044】

情報処理装置2は、 LBA 空間 A_0 のデータを不揮発性メモリ5に書き込む場合に、書き込みコマンドC2、 LBA 空間 A_0 に対応するネームスペース NS_0 を示す $NSID6$ 、 LBA 空間 A_0 内の $LBA7$ 、データサイズ8、 $LBA7$ に対応する書き込みデータ9をメモリシステム3に送る。

【0045】

管理データ18は、ネームスペース NS_0 とブロック $B_0 \sim B_2$ と関係付けている。

【0046】

ガベージコレクション部 G_0 は、管理データ18に基づいて、ガベージコレクション部 G_0 に対応するネームスペース NS_0 に含まれているブロック $B_0 \sim B_2$ に対して、ガベージコレクションを実行する。

【0047】

ガベージコレクションの結果、ブロック $B_0 \sim B_2$ 内でデータの配置が変化する。このため、ガベージコレクション部 G_0 は、図2において省略されているアドレス変換部15にアドレス変換テーブル T_0 の更新を指示し、アドレス変換部15は、ガベージコレクション後のデータ配置と整合するように、ネームスペース NS_0 に対応するアドレス変換テーブル T_0 を更新する。

【0048】

図3は、本実施形態に係る受付部13及び設定部14の処理の一例を示すフローチャートである。

【0049】

ステップS301において、受付部13は、ネームスペース $NS_0 \sim NS_M$ の設定コマンドC1を受け付ける。

【0050】

ステップS302において、設定部14は、ネームスペース $NS_0 \sim NS_M$ に、不揮発性メモリ5のブロック $B_0 \sim B_N$ を割り当て、管理データ18を生成する。

【0051】

ステップS303において、設定部14は、管理データ18をメモリ12に格納する。

【0052】

図4は、本実施形態に係るガベージコレクション部 G_0 及びアドレス変換部15の処理の一例を示すフローチャートである。なお、他のガベージコレクション部 $G_1 \sim G_M$ でも同様の処理が実行される。この図4の処理は、例えば、情報処理装置2からの指示に応じて行われてもよく、メモリシステム3の管理者の指示に応じて行われてもよい。また、ガベージコレクション部 G_0 は、例えば、ガベージコレクション対象のネームスペース NS_0 のデータ格納状態を監視し、ガベージコレクションの開始を判断するなどのように、自発的に図4の処理を実行してもよい。より具体的に説明すると、例えば、ガベージコレクション部 G_0 は、ネームスペース NS_0 内の空きブロックが所定数以下、又は、ネームスペース NS_0 内の全ブロックに対する空きブロックの割合が所定値以下、の場合に、ネームスペース NS_0 に対するガベージコレクションを実行する。

【0053】

ステップS401において、ガベージコレクション部 G_0 は、管理データ18に基づいて、ガベージコレクション対象のネームスペース NS_0 に対応するブロック $B_0 \sim B_2$ を選択する。

【0054】

ステップS402において、ガベージコレクション部 G_0 は、選択されたネームスペース NS_0 内のブロック $B_0 \sim B_2$ に対するガベージコレクションを実行する。

【0055】

ステップS403において、アドレス変換部15は、ガベージコレクション対象のネー

10

20

30

40

50

ムスペース NS_0 に対応するアドレス変換テーブル T_0 を、ガベージコレクション後のブロック $B_0 \sim B_2$ の状態にそって更新する。

【0056】

以上説明した本実施形態においては、ネームスペース $NS_0 \sim NS_M$ ごとに、予め定められた又は情報処理装置 2 から設定されたブロック量を割り当てることができ、ネームスペース $NS_0 \sim NS_M$ に対応するデータを、そのネームスペース $NS_0 \sim NS_M$ に割り当てられたブロック $B_0 \sim B_M$ に書き込むことができ、ネームスペース $NS_0 \sim NS_M$ ごとに異なるデータ量を設定することができる。

【0057】

本実施形態においては、ガベージコレクションを、ネームスペース $NS_0 \sim NS_M$ ごとに独立して、効率的に、実行することができる。

10

【0058】

本実施形態においては、ガベージコレクションの結果、データを格納していない空ブロックを、ガベージコレクション前のネームスペースから他のネームスペースに変更することができ、他のネームスペース内で空ブロックを確保することができる。これにより、ブロックに割り当てられるネームスペースを変更することができ、ネームスペース $NS_0 \sim NS_M$ 間でウェアレベリングを実行することができ、不揮発性メモリ 5 を長寿命化することができる。

【0059】

本実施形態では、ネームスペース $NS_0 \sim NS_M$ ごとに異なるデータ量の予備領域 $P_0 \sim P_M$ を設定することができ、ネームスペース $NS_0 \sim NS_M$ ごとにオーバープロビジョニングを実現することができる。これにより、書き込み速度の高速化及び性能維持を実現し、信頼性を向上させることができる。

20

【0060】

本実施形態においては、ネームスペース $NS_0 \sim NS_M$ ごとにアドレス変換テーブル $T_0 \sim T_M$ が管理されており、ネームスペース $NS_0 \sim NS_M$ ごとに、アドレス変換及び LBA と PBA との関係の変更を効率的に行うことができる。

【0061】

本実施形態において、アドレス変換をキーバリュ型検索によって行う場合には、不揮発性メモリ 5 のデータ容量が大きくても、効率的にアドレス変換を行うことができる。

30

【0062】

本実施形態においては、ネームスペース $NS_0 \sim NS_M$ ごとに高度なメモリ管理を実現することができ、不揮発性メモリ 5 を長寿命化することができ、コストを削減することができ、ネームスペース $NS_0 \sim NS_M$ で分けられている不揮発性メモリ 5 に対する書き込み及び読み出しを高速化することができる。

【0063】

(情報処理装置 2 の説明)

情報処理装置 2 は、メモリ 21 とプロセッサ 22 とを含む。

【0064】

情報処理装置 2 では、アプリケーションプログラム、オペレーティングシステムなどの各種のプログラム及び各種のデータ（以下、オブジェクトと記す）が、オブジェクト識別データ（以下、オブジェクト ID と記す）により識別可能である。

40

【0065】

情報処理装置 2 は、オブジェクトに対して LBA 空間 $A_0 \sim A_M$ を割り当て、オブジェクト ID ごとに、当該オブジェクト ID の示すオブジェクトに割り当てられた LBA 空間 $A_0 \sim A_M$ を管理する。

【0066】

メモリ 21 は、プログラム 23 を格納する例えば不揮発性のメモリである。

【0067】

プロセッサ 22 は、メモリ 21 に格納されているプログラム 23 を実行することにより

50

、頻度算出部 2 4、割り当て部 2 5、発信部 2 6として機能する。

【0068】

頻度算出部 2 4は、各オブジェクトの L B A空間に対応する書き込み頻度を算出する。

【0069】

割り当て部 2 5は、各オブジェクトの L B A空間に対応する書き込み頻度に基づいて、複数の L B A空間 $A_0 \sim A_M$ を、複数のネームスペース $NS_0 \sim NS_M$ に割り当てる。

【0070】

発信部 2 6は、割り当て部 2 5の割り当て結果に基づいて、設定コマンド C 1を生成し、設定コマンド C 1をメモリシステム 3へ送る。

【0071】

また、発信部 2 6は、割り当て部 2 5の割り当て結果に基づいて、書き込みコマンド C 2を生成し、書き込みコマンド C 2、書き込みコマンド C 2を発信するオブジェクトの L B A空間に割り当てられたネームスペース NS_0 を示す $NSID_6$ 、データサイズ 8、及び、書き込みデータ 9を、メモリシステム 3へ送る。

【0072】

図 5は、本実施形態に係るネームスペースの割り当て状態を例示するブロック図である。

【0073】

この図 5では、オブジェクトの L B A空間を 4つのネームスペース $NS_0 \sim NS_4$ に割り当てる状態を例示しているが、ネームスペースの数は 2以上でよい。

【0074】

オブジェクトの L B A空間に対応する書き込み頻度は、複数の書き込み頻度グループのいずれかに割り当てられる。図 5では、6つの書き込み頻度グループ $L_0 \sim L_5$ が用いられる場合が例示されている。書き込み頻度グループ $L_0 \sim L_5$ は、書き込み頻度の高い方から低い方へ、順に、書き込み頻度グループ L_0 (Extremely Hot)、 L_1 (Hot)、 L_2 (Warm)、 L_3 (Cool)、 L_4 (Cold)、 L_5 (Extremely cold)とする。

【0075】

書き込み頻度グループ $L_0 \sim L_5$ は、それぞれの性質及びオブジェクト IDに基づいて、ネームスペース $NS_0 \sim NS_4$ に割り当てられる。

【0076】

例えば、同一のネームスペース内に異なる性質の書き込み頻度グループが含まれるように、ネームスペース $NS_0 \sim NS_4$ に対して書き込み頻度グループ $L_0 \sim L_5$ の割り当てが行われる。

【0077】

例えば、書き込み頻度の極めて大きい書き込み頻度グループ L_0 は、予備領域を潤沢に確保するため、複数のネームスペース $NS_0 \sim NS_4$ に割り当てられる。これにより、効率的に予備領域を用いることができる。

【0078】

例えば、書き込み頻度の極めて小さい書き込み頻度グループ L_5 は、複数のネームスペース NS_2 、 NS_3 に分割される。このように、書き込み頻度グループ L_5 が複数のネームスペース NS_2 、 NS_3 に割り当てられることにより、ガベージコレクションを効率化することができる。

【0079】

例えば、書き込み頻度の大きい書き込み頻度グループ L_1 は、ユーザアプリケーションなどのようなオブジェクトに依存して、独立に、ガベージコレクションが実行されるように、オブジェクトごとにネームスペース NS_0 、 NS_3 が割り当てられられ、同じオブジェクトの L B A空間は、同じネームスペース NS_0 、 NS_3 に割り当てられる。これにより、あるユーザアプリケーションに対するガベージコレクションが実行された場合に、他のユーザアプリケーションの性能が低下することを防止することができる。

【0080】

10

20

30

40

50

図6は、本実施形態に係る情報処理装置2の処理の一例を示すフローチャートである。

【0081】

ステップS601において、頻度算出部24は、各オブジェクトのLBA空間に対応する書き込み頻度を算出する。

【0082】

ステップS602において、割り当て部25は、各オブジェクトのLBA空間に対応する書き込み頻度に基づいて、各オブジェクトのLBA空間を書き込み頻度グループ $L_0 \sim L_5$ のいずれかに割り当てる。

【0083】

ステップS603において、割り当て部25は、各オブジェクトID及び書き込み頻度グループ $L_0 \sim L_5$ に基づいて、各オブジェクトのLBA空間を、ネームスペース $NS_0 \sim NS_M$ のうちの少なくとも1つに割り当てる。

【0084】

ステップS604において、発信部26は、書き込みコマンドC2と、ネームスペース $NS_0 \sim NS_M$ の割り当て結果を示すNSID6と、LBA7と、データサイズ8と、書き込みデータ9とを、メモリシステム3に送る。

【0085】

以上説明した本実施形態においては、各オブジェクトのLBA空間 $A_0 \sim A_M$ に対応する書き込み頻度に基づいて、各オブジェクトのLBA空間 $A_0 \sim A_M$ を、ネームスペース $NS_0 \sim NS_M$ に割り当てることができ、メモリシステム3のサービス品質及び装置性能を向上させることができ、装置寿命を長くすることができ、メモリシステム3の設定を適切化することができる。

【0086】

例えば、本実施形態においては、同一のネームスペース内に異なる性質のLBA空間を割り当てることにより、ネームスペース $NS_0 \sim NS_4$ に対する書き込み頻度及び書き込み回数を調整することができる。

【0087】

例えば、本実施形態においては、書き込み頻度の極めて大きいLBA空間を、複数のネームスペースに割り当てることができ、これにより予備領域を潤沢に確保することができる。

【0088】

例えば、本実施形態においては、書き込み頻度の極めて小さいLBA空間を、複数のネームスペースに割り当て、これによりガベージコレクションを効率化することができる。

【0089】

例えば、本実施形態においては、書き込み頻度の大きいオブジェクトのLBA空間を、同じネームスペースに割り当てる。これにより、あるユーザアプリケーションに対するガベージコレクションが実行された場合に、他のユーザアプリケーションの性能が低下することを防止することができる。

【0090】

本実施形態においては、情報処理装置2において、情報処理装置2側のLBA空間 $A_0 \sim A_M$ とメモリシステム3側のネームスペース $NS_0 \sim NS_M$ との割り当てを決定することができる。

【0091】

本実施形態においては、書き込み頻度に基づいて、LBA空間 $A_0 \sim A_M$ とネームスペース $NS_0 \sim NS_M$ との割り当てが行われている。しかしながら、これに代えて、書き込み回数、書き込み頻度及び読み出し頻度の組み合わせなど、他の情報に基づいてLBA空間 $A_0 \sim A_M$ とネームスペース $NS_0 \sim NS_M$ との割り当てが行われてもよい。

【0092】

また、割り当て部25は、ユーザ設定に基づいて、複数のLBA空間 $A_0 \sim A_M$ を、複数のネームスペース $NS_0 \sim NS_M$ に割り当てるとしてもよい。

【0093】

本実施形態においては、ガベージコレクション部 $G_0 \sim G_M$ に代えて、又は、ガベージコレクション部 $G_0 \sim G_M$ とともに、ネームスペース $NS_0 \sim NS_M$ ごとのコンパクション部を備えるとしてもよい。ネームスペース $NS_0 \sim NS_M$ のそれぞれに対応するコンパクション部は、管理データ18に基づいて、ネームスペース $NS_0 \sim NS_M$ のそれぞれに対してコンパクションを実行する。

【0094】

本実施形態において、例えば、情報処理装置2とメモリシステム3との間の設定コマンドC1の通信は、省略されてもよい。例えば、アドレス変換部15は、設定部14の一部又は全部の機能を備えるとしてもよい。例えば、アドレス変換部15は、書き込みコマンドC2に付されているNSID6及びLBA7と、LBA7に対応するPBAとを関係付けることで、管理データ18及びネームスペース $NS_0 \sim NS_M$ ごとのアドレス変換テーブル $T_0 \sim T_M$ を生成するとしてもよい。管理データ18とアドレス変換テーブル $T_0 \sim T_M$ とは、適宜組み合わせることができ、又は、分割することができる。このように設定コマンドC1の通信が省略され、設定部14の一部又は全部の機能がアドレス変換部15に備えられる構成は、後述の第2の実施形態で説明する。

10

【0095】

[第2の実施形態]

本実施形態においては、メモリシステムが複数の情報処理装置からの書き込みデータを書き込み、メモリシステムが読み出しデータを複数の情報処理装置へ送る情報処理システムについて説明する。

20

【0096】

図7は、本実施形態に係る情報処理システムの構成の一例を示すブロック図である。

【0097】

情報処理システム1Aは、複数の情報処理装置 $D_0 \sim D_M$ とメモリシステム3Aとを含む。複数の情報処理装置 $D_0 \sim D_M$ のそれぞれは、上記の情報処理装置2と同様の機能を備える。メモリシステム3Aは、主に、アドレス変換テーブル $T_0 \sim T_M$ 及び管理データ18に代えて変換テーブル(変換データ)20を備える点、複数の情報処理装置 $D_0 \sim D_M$ との間でデータ、情報、信号、コマンドなどの送信及び受信を行う点、設定部14の機能がアドレス変換部15に備えられる点で、上記のメモリシステム3と相違している。本実施形態においては、上記第1の実施形態と異なる点を説明し、同じ部分又は実質的に同じ部分については説明を省略するか又は簡単に説明する。

30

【0098】

メモリシステム3Aは、例えばクラウドコンピューティングシステムに備えられる。メモリシステム3Aは、複数の情報処理装置 $D_0 \sim D_M$ によってシェアされる場合を例として説明するが、例えば複数のユーザによってシェアされてもよい。複数の情報処理装置 $D_0 \sim D_M$ の少なくとも1つは、仮想マシンでもよい。

【0099】

本実施形態において、コマンドに付されるNSIDは、ネームスペースのアクセス鍵として用いられる。

40

【0100】

本実施形態においては、複数の情報処理装置 $D_0 \sim D_M$ は、自己に対応するネームスペース $NS_0 \sim NS_M$ に対してアクセス権限を持つとする。しかしながら、1つの情報処理装置が1以上のネームスペースに対してアクセス権限を持つとしてもよく、複数の情報処理装置が共通のネームスペースに対してアクセス権限を持つとしてもよい。

【0101】

情報処理装置 $D_0 \sim D_M$ のそれぞれは、書き込みコマンドC2とともに、例えば自己に対応する書き込み先スペースを示すNSID6W、書き込み先を示すLBA7W、データサイズ8、及び、書き込みデータ9Wを、メモリシステム3Aへ送る。

【0102】

50

情報処理装置 $D_0 \sim D_M$ のそれぞれは、読み出しコマンド C_3 とともに、例えば自己に対応する読み出し先スペースを示す $NSID6R$ 、読み出し先を示す $LBA7R$ を、メモリシステム $3A$ へ送る。

【0103】

情報処理装置 $D_0 \sim D_M$ のそれぞれは、読み出しコマンド C_3 に対応する読み出しデータ $9R$ 、又は、読み出し不可であったことを示す情報を、メモリシステム $3A$ から受ける。

【0104】

メモリシステム $3A$ は、コントローラ $4A$ と不揮発性メモリ 5 とを含む。

【0105】

コントローラ $4A$ は、インタフェース部 19 、記憶部 10 、バッファメモリ $F_0 \sim F_M$ 、プロセッサ 11 を含む。本実施形態においては、コントローラ $4A$ に備えられるプロセッサの数は 1 以上で自由に変更可能である。

【0106】

インタフェース部 19 は、例えば情報処理装置 $D_0 \sim D_M$ などのような外部装置との間で、データ、情報、信号、コマンドなどの送信及び受信を行う。

【0107】

記憶部 10 は、変換テーブル 20 を格納する。なお、変換テーブル 20 の一部又は全部は、メモリ 12 などのような他のメモリに格納されてもよい。

【0108】

変換テーブル 20 は、 LBA 、 PBA 、 $NSID$ を互いに関係付けたデータである。この変換テーブル 20 については、図 8 を用いて後で説明する。

【0109】

バッファメモリ $F_0 \sim F_M$ は、それぞれ、ネームスペース $NS_0 \sim NS_M$ に対する書き込みバッファメモリ及び読み出しバッファメモリとして使用される。

【0110】

プロセッサ 11 は、プログラム 17 を格納するメモリ 12 、受付部 13 、アドレス変換部 15 、書き込み部 16 、読み出し部 21 、ガベージコレクション部 $G_0 \sim G_M$ を含む。プロセッサ 11 は、プログラム 17 を実行することにより、受付部 13 、アドレス変換部 15 、書き込み部 16 、読み出し部 21 、ガベージコレクション部 $G_0 \sim G_M$ として機能する。

【0111】

受付部 13 は、データ書き込み時に、情報処理装置 $D_0 \sim D_M$ からインタフェース部 19 経由で書き込みコマンド C_2 、 $NSID6W$ 、 $LBA7W$ 、データサイズ 8 、書き込みデータ $9W$ を受け付ける。

【0112】

受付部 13 は、データ読み出し時に、情報処理装置 $D_0 \sim D_M$ からインタフェース部 19 経由で読み出しコマンド C_3 、 $NSID6R$ 、 $LBA7R$ を受け付ける。

【0113】

アドレス変換部 15 は、受付部 13 によって書き込みコマンド C_2 が受け付けられた場合に、書き込みコマンド C_2 に付されていた $LBA7W$ と $NSID6W$ に基づいて、 $NSID6W$ の示すネームスペースにおける書き込み先の PBA を決定し、 $LBA7W$ と $NSID6W$ と決定された PBA とを関係付けた状態で変換テーブル 20 を更新する。

【0114】

アドレス変換部 15 は、受付部 13 によって読み出しコマンド C_3 が受け付けられた場合に、読み出しコマンド C_3 に付されていた $LBA7R$ と $NSID6R$ と変換テーブル 20 とに基づいて、 $NSID6R$ の示すネームスペースにおける読み出し先の PBA を決定する。

【0115】

書き込み部 16 は、 $NSID6W$ の示すネームスペースに対応するバッファメモリを経由して、 $NSID6W$ の示すネームスペースに対応する PBA の示す位置に、書き込みデ

10

20

30

40

50

ータ 9 W を書き込む。

【0116】

読み出し部 21 は、NSID6R の示すネームスペースに対応するバッファメモリを経由して、NSID6R の示すネームスペースに対応する PBA の示す位置から、読み出しデータ 9 R を読み出す。そして、読み出し部 21 は、インタフェース部 19 経由で、読み出しコマンド C3 を発行した情報処理装置に、読み出しデータ 9 R を送る。

【0117】

本実施形態において、ガベージコレクション部 $G_0 \sim G_M$ は、変換テーブル 20 に基づいて、ネームスペース $NS_0 \sim NS_M$ ごとにガベージコレクションを実行する。

【0118】

図 8 は、本実施形態に係る変換テーブル 20 の一例を示すデータ構造図である。

【0119】

変換テーブル 20 は、LBA、PBA、NSID を関係付けて管理する。例えば、変換テーブル 20 は、LBA200 と PBA300 と NS_0 とを関係付けており、LBA201 と PBA301 と NS_0 とを関係付けており、LBA200 と PBA399 と NS_M とを関係付けている。

【0120】

アドレス変換部 15 は、例えば、LBA200 とネームスペース NS_0 を示す NSID とに関係する PBA300 と、LBA200 とネームスペース NS_M を示す NSID とに関係する PBA399 とが互いに異なるように、PBA の決定を行う。

【0121】

これにより、アドレス変換部 15 は、LBA200 とともに受け付けられた NSID がネームスペース NS_0 を示す場合には PBA300 を選択することができ、LBA200 とともに受け付けられた NSID がネームスペース NS_M を示す場合には PBA399 を選択することができる。

【0122】

したがって、複数の情報処理装置 $D_0 \sim D_M$ の間で同じ論理アドレスが使用される場合であっても、メモリシステム 3A を複数の情報処理装置 $D_0 \sim D_M$ でシェアすることができる。

【0123】

図 9 は、本実施形態に係るメモリシステム 3A の書き込み処理の一例を示すフローチャートである。

【0124】

この図 9 の説明では、複数の情報処理装置 $D_0 \sim D_M$ のうちの情報処理装置 D_0 から書き込みコマンド C2 が発行され、書き込みコマンド C2 にはネームスペース NS_0 を示す NSID6W が付されている場合を例として説明する。しかしながら、情報処理装置 $D_1 \sim D_M$ から書き込みコマンド C2 が発行された場合も同様である。また、書き込みコマンド C2 に他のネームスペース $NS_1 \sim NS_M$ のいずれかを示す NSID6W が付されている場合も同様である。

【0125】

ステップ S901 において、受付部 13 は、情報処理装置 D_0 からインタフェース部 19 経由で書き込みコマンド C2、NSID6W、LBA7W、データサイズ 8、書き込みデータ 9W を受け付ける。

【0126】

ステップ S902 において、アドレス変換部 15 は、受付部 13 によって書き込みコマンド C2 が受け付けられた場合に、書き込みコマンド C2 に付されていた LBA7W と NSID6W に基づいて、NSID6W の示すネームスペース NS_0 における書き込み先の PBA を決定する。

【0127】

ステップ S903 において、アドレス変換部 15 は、LBA7W と NSID6W と決定

10

20

30

40

50

されたPBAとを関係付けた状態で変換テーブル20を更新する。

【0128】

ステップS904において、書き込み部16は、NSID6Wの示すネームスペースNS₀に対応するバッファメモリF₀を経由して、NSID6Wの示すネームスペースNS₀に対応するPBAの示す位置に、書き込みデータ9Wを書き込む。

【0129】

図10は、本実施形態に係るメモリシステム3Aの読み出し処理の一例を示すフローチャートである。

【0130】

この図10の説明では、複数の情報処理装置D₀～D_Mのうちの情報処理装置D_Mから読み出しコマンドC3が発行され、読み出しコマンドC3にはネームスペースNS_Mを示すNSID6Rが付されている場合を例として説明する。しかしながら、情報処理装置D₀～D_{M-1}から読み出しコマンドC3が発行された場合も同様である。また、読み出しコマンドC3に他のネームスペースNS₀～NS_{M-1}のいずれかを示すNSID6Rが付されている場合も同様である。

10

【0131】

ステップS1001において、受付部13は、情報処理装置D_Mからインタフェース部19経由で読み出しコマンドC3、NSID6R、LBA7Rを受け付ける。

【0132】

ステップS1002において、アドレス変換部15は、受付部13によって読み出しコマンドC3が受け付けられた場合に、読み出しコマンドC3に付されていたLBA7RとNSID6Rと変換テーブル20とに基づいて、読み出し先のPBAを決定する。

20

【0133】

ステップS1003において、読み出し部21は、NSID6Rの示すネームスペースNS_Mに対応するバッファメモリF_Mを経由して、NSID6Rの示すネームスペースNS_Mに対応するPBAの示す位置から、読み出しデータ9Rを読み出し、インタフェース部19経由で、読み出しコマンドC3を発行した情報処理装置D_Mに、読み出しデータ9Rを送る。

【0134】

以上説明した本実施形態においては、不揮発性メモリ5が複数のネームスペースNS₀～NS_Mに区切られる。情報処理装置D₀～D_Mは、複数のネームスペースNS₀～NS_Mのうち自己がアクセス権限を持つネームスペースをアクセス可能である。これにより、データセキュリティを向上させることができる。

30

【0135】

メモリシステム3Aのコントローラ4Aは、ネームスペースNS₀～NS_Mごとに、独立の制御を行う。これにより、ネームスペースNS₀～NS_Mごとに使用条件を切り替えることができる。

【0136】

メモリシステム3Aは、LBAとPBAとNSIDとを関係付けているため、例えば互いに独立の複数の情報処理装置から同じLBAを受信した場合であっても、NSIDによりデータを区別することができる。

40

【0137】

上記各実施形態において、テーブル形式のデータは、例えばリスト形式などの他のデータ構造で実装されてもよい。

【0138】

[第3の実施形態]

本実施形態においては、上記第1及び第2の実施形態で説明した情報処理システム1, 1Aの詳細構成について説明する。

【0139】

図11は、本実施形態に係る情報処理システムの詳細構成の一例を示すブロック図であ

50

る。

【0140】

この図11において、情報処理システム1Bは、情報処理装置2Bとメモリシステム3Bとを含む。情報処理システム1Bは、上記第2の実施形態と同様に、複数の情報処理装置を備えるとしてもよい。すなわち、上記第1及び第2の実施形態の情報処理装置2、 $D_0 \sim D_M$ のいずれかは、情報処理装置2Bに対応する。

【0141】

上記第1及び第2の実施形態のメモリシステム3、3Aは、メモリシステム3Bに対応する。

【0142】

上記第1及び第2の実施形態のプロセッサ11は、CPU43F、43Bに対応する。

【0143】

上記第1の実施形態のアドレス変換テーブル $T_0 \sim T_M$ 、及び、上記第2の実施形態の変換テーブル20は、LUT45に対応する。

【0144】

上記第1及び第2の実施形態の記憶部10は、DRAM47に対応する。

【0145】

上記第2の実施形態のインタフェース部19は、ホストインタフェース41及びホストインタフェースコントローラ42に対応する。

【0146】

上記第1及び第2の実施形態のバッファメモリ $F_0 \sim F_M$ は、ライトバッファWB及びリードバッファRBに対応する。

【0147】

情報処理装置2Bは、ホスト装置として機能する。

【0148】

コントローラ4は、フロントエンド4Fと、バックエンド4Bとを備える。

【0149】

フロントエンド（ホスト通信部）4Fは、ホストインタフェース41、ホストインタフェースコントローラ42、暗号化／復号化部44、及びCPU43Fを備える。

【0150】

ホストインタフェース41は、情報処理装置2Bとの間で、要求（書き込みコマンド、読み出しコマンド、消去コマンドなど）、LBA及びデータなどを通信する。

【0151】

ホストインタフェースコントローラ（制御部）42は、CPU43Fの制御に基づいて、上記ホストインタフェース41の通信を制御する。

【0152】

暗号化／復号化部（Advanced Encryption Standard (AES)）44は、データ書き込み動作において、ホストインタフェースコントローラ42から送信される書き込みデータ（平文）を暗号化する。暗号化／復号化部44は、データ読み出し動作において、バックエンド4BのリードバッファRBから送信される暗号化された読み出しデータを復号化する。なお、この暗号化／復号化部44を介さずに、書き込みデータ及び読み出しデータを送信することも、必要に応じて可能である。

【0153】

CPU43Fは、フロントエンド4Fの上記各構成41、42、44を制御し、フロントエンド4Fの全体の動作を制御する。

【0154】

バックエンド（メモリ通信部）4Bは、ライトバッファWB、リードバッファRB、LUT部45、DDRC46、DRAM47、DMAC48、ECC49、ランダムイザRZ、NANDC50、及びCPU43Bを備える。

【0155】

10

20

30

40

50

ライトバッファ（ライトデータ転送部）WBは、情報処理装置2Bから送信された書き込みデータを一時的に格納する。具体的には、ライトバッファWBは、当該書き込みデータが不揮発性メモリ5に適した所定のデータサイズになるまで、一時的にデータを格納する。

【0156】

リードバッファ（リードデータ転送部）RBは、不揮発性メモリ5から読み出された読み出しデータを一時的に格納する。具体的には、リードバッファRBにおいて、読み出しデータは、情報処理装置2Bに適した順序（情報処理装置2Bが指定した論理アドレスLBAの順序）になるように並び替えられる。

【0157】

LUT45は、論理アドレスLBAを所定の物理アドレスPBAに変換するためのデータである。

【0158】

DDRC46は、DRAM47におけるDDR（Double Data Rate）を制御する。

【0159】

DRAM47は、例えば、LUT45を格納する揮発性のメモリである。

【0160】

DMA C（Direct Memory Access Controller）48は、内部バスIBを介して、書き込みデータや読み出しデータなどを転送する。図11においては1つのDMA C128が図示されているが、コントローラ4は、2以上のDMA C128を備えるとしてもよい。DMA C48は、必要に応じて、コントローラ4内の様々な位置に設定される。

【0161】

ECC（誤り訂正部）49は、ライトバッファWBから送信される書き込みデータにECC（Error Correcting Code）を付加する。ECC49は、リードバッファRBに送信する際に、付加したECCを用いて、不揮発性メモリ5から読み出した読み出しデータを必要に応じて訂正する。

【0162】

ランダムライザRZ（又はスクランブラ）は、データ書き込み動作の際に、書き込みデータが不揮発性メモリ5の特定のページ又はワード線方向などに偏らないように、書き込みデータを分散させる。このように、書き込みデータを分散させることで、書き込み回数を平準化でき、不揮発性メモリ5のメモリセルMCのセル寿命を長期化できる。そのため、不揮発性メモリ5の信頼性を向上できる。また、データ読み出し動作の際に、不揮発性メモリ5から読み出された読み出しデータはランダムライザRZを通過する。

【0163】

NAND C（NAND Controller）50は、所定の速度の要求を満たすため、複数のチャンネル（ここでは、4つのチャンネルCH0～CH3）を用いて、並列に不揮発性メモリ5にアクセスする。

【0164】

CPU43Bは、バックエンド4Bの上記各構成（45～50、RZ）を制御し、バックエンド4Bの全体の動作を制御する。

【0165】

なお、図11に示したコントローラ4の構成は例示であり、この構成に限定されることはない。

【0166】

図12は、本実施形態に係るストレージシステムを示す斜視図である。

【0167】

ストレージシステム100は、SSDとしてのメモリシステム3Bを備える。

【0168】

メモリシステム3Bは、例えば比較的小型のモジュールであり、その外形寸法の一例は、20mm×30mm程度である。なお、メモリシステム3Bの大きさ及び寸法は、これ

10

20

30

40

50

に限られるものではなく、種々の大きさのものに適宜変更可能である。

【0169】

また、メモリシステム3Bは、例えば、企業（エンタープライズ）で運用されるデータセンター又はクラウドコンピューティングシステムにおいて、サーバのような情報処理装置2Bに装着されて使用可能である。そのため、メモリシステム3Bは、エンタープライズ用SSD（eSSD）であってもよい。

【0170】

メモリシステム3Bは、例えば上方に開口した複数のコネクタ（例えばスロット）30を備える。各コネクタ30は、例えばSAS（Serial Attached SCSI）コネクタ等である。このSASコネクタによれば、6GbpsのDual Portにより、情報処理装置3Bと各メモリシステム3Bとが互いに高速通信を行うことが可能である。なお、これに限られず、各コネクタ30は、例えば、PCIe（PCI Express）又はNVMe（NVM Express）等であってもよい。

10

【0171】

複数のメモリシステム3Bは、情報処理装置2Bのコネクタ30にそれぞれ装着され、略鉛直方向に起立した姿勢で互いに並べて支持される。このような構成によれば、複数のメモリシステム3Bをコンパクトに纏めて実装可能であり、メモリシステム3Bの小型化を図ることができる。さらに、本実施形態に係るメモリシステム3Bの各形状は、2.5型のSFF（Small Form Factor）である。このような形状により、メモリシステム3Bは、エンタープライズ用HDD（eHDD）と互換形状（コンパチ形状）を図ることができ、eHDDとの容易なシステム互換性を実現することができる。

20

【0172】

なお、メモリシステム3Bは、エンタープライズ用に限られない。例えば、メモリシステム3Bは、ノートブック型ポータブルコンピュータ又はタブレット型端末のようなコンピュータ用の電子機器の記憶媒体としても適用可能である。

【0173】

以上説明したように、本実施形態で説明した構成を持つ情報処理システム1B及びストレージシステム100においては、大容量の記憶に、上記第1及び第2の実施形態と同様の効果を得ることができる。

【0174】

30

本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。

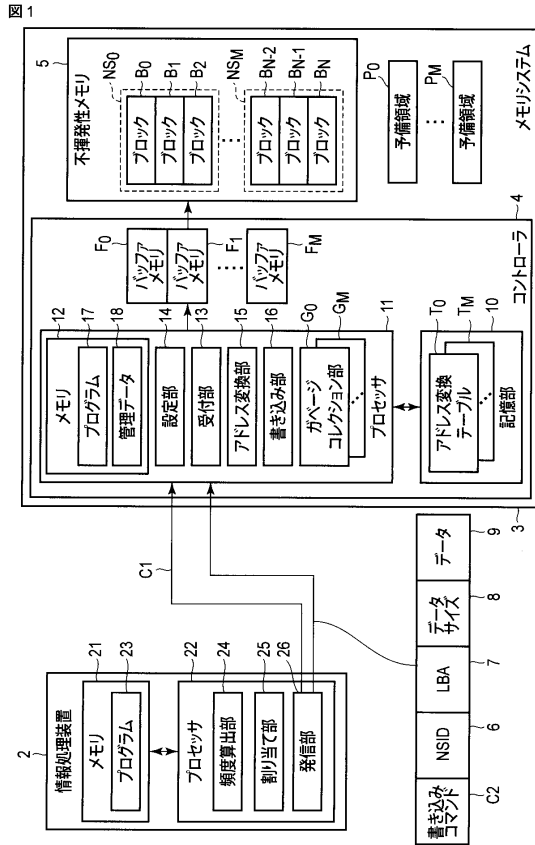
【符号の説明】

【0175】

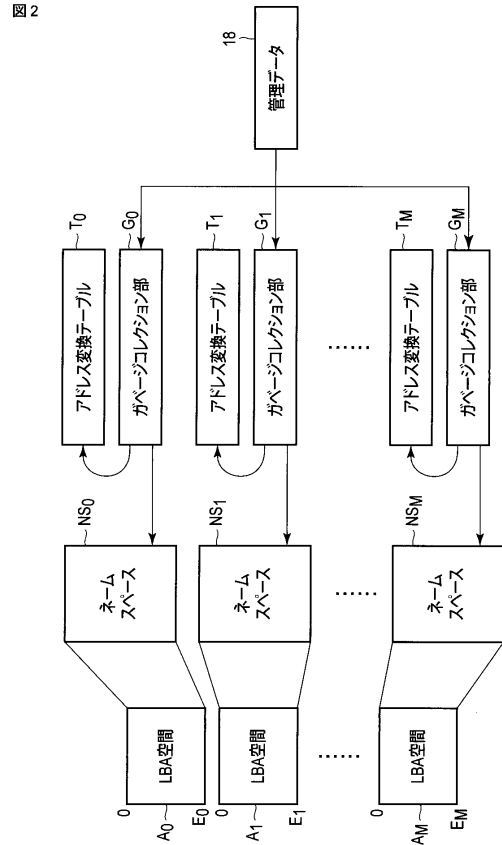
2, $D_0 \sim D_M$ …情報処理装置、21…メモリ、22…プロセッサ、23…プログラム、24…頻度算出部、25…割り当て部、26…発信部、3…メモリシステム、4…コントローラ、5…不揮発性メモリ、C2…書き込みコマンド、6…NSID、7…LBA、8…データサイズ、9…データ、10…記憶部、12…メモリ、13…受付部、14…設定部、15…アドレス変換部、16…書き込み部、17…プログラム、18…管理データ、 $G_0 \sim G_M$ …ガベージコレクション部、 $T_0 \sim T_M$ …アドレス変換テーブル、 $F_0 \sim F_M$ …バッファメモリ、 $NS_0 \sim NS_M$ …ネームスペース、 $B_0 \sim B_N$ …ブロック、 $P_0 \sim P_M$ …予備領域、 $A_0 \sim A_M$ …LBA空間、 $L_0 \sim L_5$ …書き込み頻度グループ、20…変換テーブル。

40

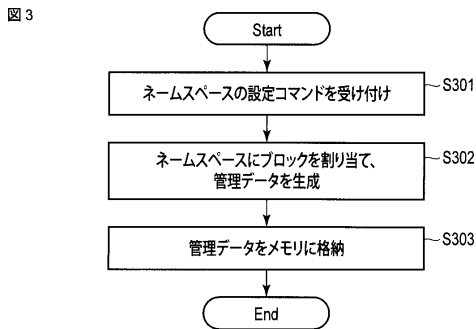
【図 1】



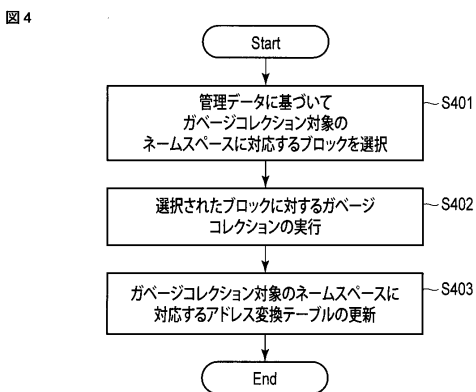
【図 2】



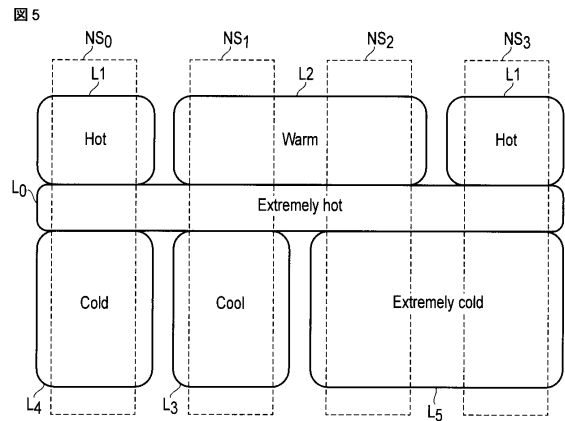
【図 3】



【図 4】

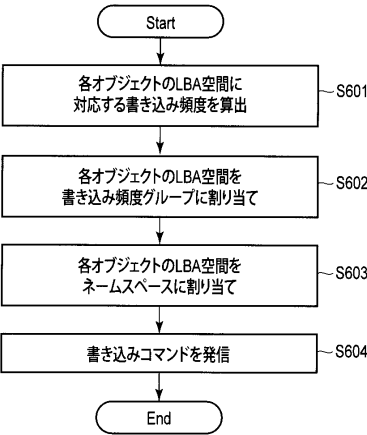


【図 5】



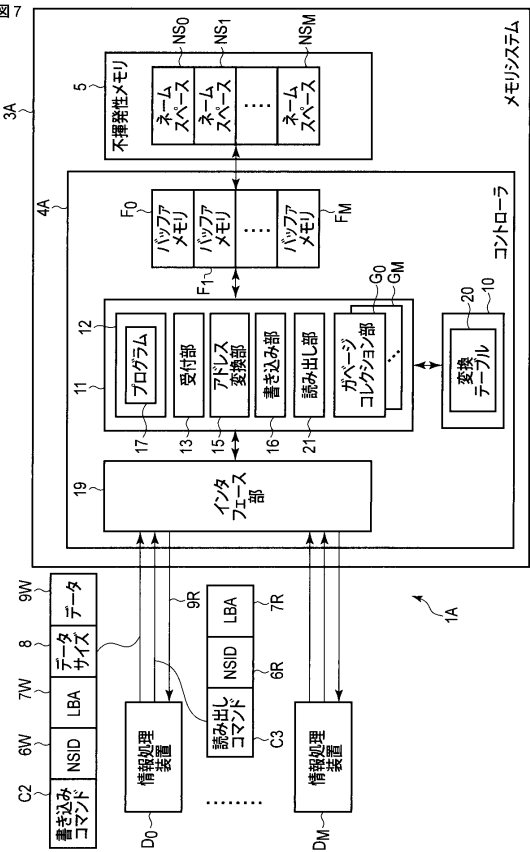
【図 6】

図 6



【図 7】

図 7



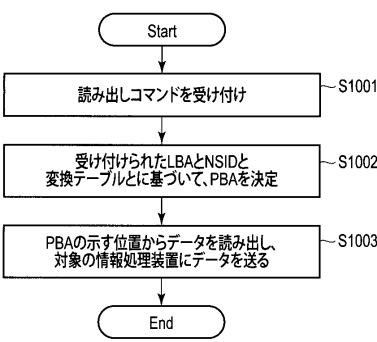
【図 8】

図 8

LBA	PBA	ネームスペース
LBA200	PBA300	NS ₀
LBA201	PBA301	NS ₀
⋮	⋮	⋮
LBA200	PBA399	NS _M

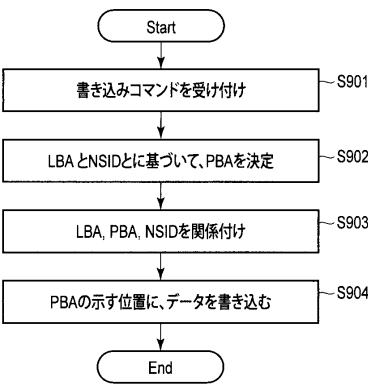
【図 10】

図 10

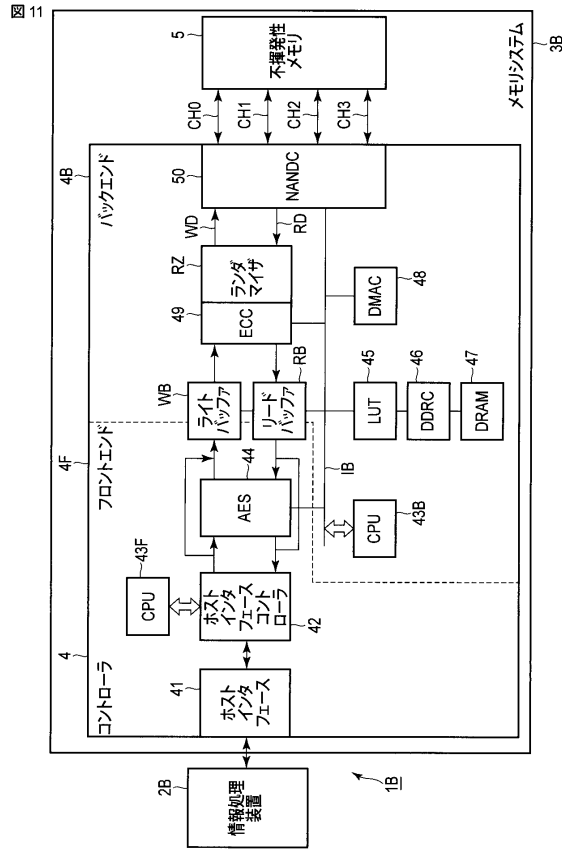


【図 9】

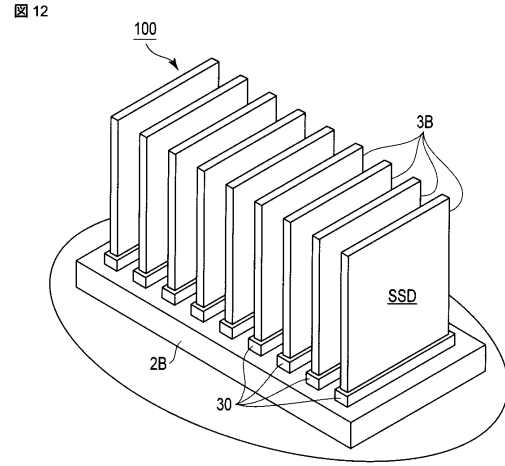
図 9



【図 1 1】



【図 1 2】



フロントページの続き

(56)参考文献 特表2009-503629 (JP, A)

特開2013-161222 (JP, A)

NVM Express Revision 1.1b, NVM Express, Inc, 2014年 7月 2日, 第8-10, 48-50, 118, 135-137ページ, <URL:http://nvmexpress.org/wp-content/uploads/NVM-Express-1_0e.pdf>

(58)調査した分野(Int.Cl., DB名)

G06F 12/02

G06F 12/00