

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4905511号
(P4905511)

(45) 発行日 平成24年3月28日(2012.3.28)

(24) 登録日 平成24年1月20日(2012.1.20)

(51) Int.Cl.

G06F 12/00 (2006.01)

F 1

G06F 12/00 501B
G06F 12/00 514E

請求項の数 6 (全 13 頁)

(21) 出願番号 特願2009-155599 (P2009-155599)
 (22) 出願日 平成21年6月30日 (2009.6.30)
 (65) 公開番号 特開2011-13800 (P2011-13800A)
 (43) 公開日 平成23年1月20日 (2011.1.20)
 審査請求日 平成22年7月15日 (2010.7.15)

(73) 特許権者 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番
 1号
 (74) 代理人 100090516
 弁理士 松倉 秀実
 (74) 代理人 100113608
 弁理士 平川 明
 (74) 代理人 100105407
 弁理士 高田 大輔
 (74) 代理人 100089244
 弁理士 遠山 勉
 (72) 発明者 和田 美穂子
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】記憶装置の制御部及び制御方法

(57) 【特許請求の範囲】

【請求項 1】

物理的な記憶装置をそれぞれ管理する複数の制御部を備えたストレージ装置の制御部において、

前記制御部は、

他の制御部との接続を介して制御部間通信を行う通信部と、

論理的に定義した記憶領域としての仮想論理ボリュームの論理アドレスを物理的な記憶装置の記憶領域としての実論理ボリュームの論理参照番号及び論理アドレスと対応付ける変換テーブル及びそれぞれの制御部の識別情報と前記物理的な記憶装置とを対応付けるストレージ管理テーブル、を記憶する記憶部と、

仮想論理ボリュームに書き込み要求が発生したときに、前記変換テーブルにおいて前記書き込み要求先の論理的に定義した記憶領域と前記物理的な記憶装置の記憶領域とが対応付けられているかを判定し、対応付けられていないときに物理的な記憶領域を対応付ける、前記仮想論理ボリュームを管理する管理部と、

前記仮想論理ボリュームの使用率が閾値を超えたとき、前記仮想論理ボリュームに対応する前記物理的な記憶装置が前記他の制御部に対応付けられているかどうか判定する検知部と、

前記検知部による判定結果に基づいて、前記仮想論理ボリュームに割り当てられた実論理ボリュームに記憶されているデータを、前記仮想論理ボリュームを管理する管理部を含む制御部が対応付けられている実論理ボリュームにマイグレーションするマイグレーション

10

20

ン部と、
を備えた記憶装置の制御部。

【請求項 2】

前記マイグレーションを行う際、マイグレーション先の実論理ボリュームを前記変換テーブルによって管理する管理部が、前記ストレージ管理テーブルによって対応付けられている物理的な記憶領域と前記マイグレーション先の実論理ボリュームの記憶領域とを対応付ける請求項 1 に記載の記憶装置の制御部。

【請求項 3】

物理的な記憶装置をそれぞれ管理する複数の制御部を備えたストレージ装置の制御部において、

論理的に定義した記憶領域としての仮想論理ボリュームの論理アドレスを物理的な記憶装置の記憶領域としての実論理ボリュームの論理参照番号及び論理アドレスと対応付ける変換テーブル及びそれぞれの制御部の識別情報と前記物理的な記憶装置とを対応付けるストレージ管理テーブル、を記憶するための記憶部を有する制御部が、

他の制御部との接続を介して制御部間通信を行うステップと、

仮想論理ボリュームに書き込み要求が発生したときに、前記変換テーブルにおいて前記書き込み要求先の論理的に定義した記憶領域と前記物理的な記憶装置の記憶領域とが対応付けられているかを判定し、対応付けられていないときに物理的な記憶領域を対応付ける、前記仮想論理ボリュームを管理するステップと、

前記仮想論理ボリュームの使用率が閾値を超えたとき、前記仮想論理ボリュームに対応する前記物理的な記憶装置が前記他の制御部に対応付けられているかどうか判定するステップと、

前記判定するステップの結果に基づいて、前記仮想論理ボリュームに割り当てられた実論理ボリュームに記憶されているデータを、前記仮想論理ボリュームを管理する管理部を含む制御部が対応付けられている実論理ボリュームにマイグレーションするステップと、を実行する制御方法。

【請求項 4】

前記マイグレーションを行う際、マイグレーション先の実論理ボリュームを前記変換テーブルによって管理する管理部が前記ストレージ管理テーブルによって対応付けられている物理的な記憶領域と前記マイグレーション先の実論理ボリュームの記憶領域とを対応付ける請求項 3 に記載の制御方法。

【請求項 5】

物理的な記憶装置をそれぞれ管理する複数の制御部を備えたストレージ装置のコンピュータにおいて、

論理的に定義した記憶領域としての仮想論理ボリュームの論理アドレスを物理的な記憶装置の記憶領域としての実論理ボリュームの論理参照番号及び論理アドレスと対応付ける変換テーブル及びそれぞれの制御部の識別情報と前記物理的な記憶装置とを対応付けるストレージ管理テーブル、を記憶するための記憶部を備えるコンピュータに、

他の制御部との接続を介して制御部間通信を行うステップと、

仮想論理ボリュームに書き込み要求が発生したときに、前記変換テーブルにおいて前記書き込み要求先の論理的に定義した記憶領域と前記物理的な記憶装置の記憶領域とが対応付けられているかを判定し、対応付けられていないときに物理的な記憶領域を対応付ける、前記仮想論理ボリュームを管理するステップと、

前記仮想論理ボリュームの使用率が閾値を超えたとき、前記仮想論理ボリュームに対応する前記物理的な記憶装置が前記他の制御部に対応付けられているかどうか判定するステップと、

前記判定するステップの結果に基づいて、前記仮想論理ボリュームに割り当てられた実論理ボリュームに記憶されているデータを、前記仮想論理ボリュームを管理する管理部を含む制御部が対応付けられている実論理ボリュームにマイグレーションするステップと、を実行させるための制御プログラム。

【請求項 6】

前記マイグレーションを行う際、マイグレーション先の実論理ボリューム前記変換テーブルによって管理する管理部が前記ストレージ管理テーブルによって対応付けられている物理的な記憶領域と前記マイグレーション先の実論理ボリュームの記憶領域とを対応付ける請求項5に記載の制御プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本願明細書で開示される技術は、記憶装置のボリューム管理に関する。

【背景技術】

【0002】

記憶装置の仮想化技術のひとつに、書き込み要求の発生時に動的に記憶領域を割り当てる仮想論理ボリューム（以下、単にシンプロビジョニングボリューム或はシンプロボリュームとも記載）がある。

【0003】

通常の論理ボリューム（以下、実論理ボリュームとも記載）では、作成時に定義された容量分の記憶領域が全て物理ディスク又はアレイグループ等の実ストレージ上に予め確保される。

【0004】

これに対し、シンプロボリュームでは、作成時に容量が定義されるだけで実ストレージ上の記憶領域は確保されず、当該シンプロボリュームに対する書き込み要求が発生したときに初めて記憶領域が必要量だけ割り当てられる。

【0005】

シンプロボリュームでは、定義された容量分の記憶領域をあらかじめすべて用意する必要がなく、書き込み要求発生時に実際に使用するサイズの領域が動的に確保される。このため、仮想論理ボリュームの利用によって、記憶装置の導入時は必要最小限の物理ディスクを搭載し、その後の使用状況に応じて、容量が不足すれば物理ディスクを追加するという運用が可能になる。このようにディスクの利用効率を高めることによって、記憶装置の導入コスト及び運用コストを削減することができる。

【先行技術文献】

30

【特許文献】

【0006】

【特許文献1】特開2008-310734号公報

【特許文献2】特開2007-156815号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

シンプロボリュームを採用する記憶装置は、シンプロボリュームの仮想的な記憶領域に対するアクセスを実ストレージ上の実際の記憶領域に対するアクセスに読み替えてアクセスを行う制御部（CM：Centralized Module）を備えている。

40

【0008】

記憶装置は、CMを複数備えることができ、各CMには実ストレージがRAIDグループ単位で接続される。ここで、実ストレージを接続しているCMが、実ストレージの担当CMである。また、シンプロボリュームから実ストレージへ読み替えてアクセスするため、シンプロボリュームと実ストレージとの対応関係を示すマッピング情報をキャッシュメモリに保持したCMがシンプロボリュームの担当CMである。

【0009】

シンプロボリュームの記憶領域と対応する実ストレージの記憶領域は、複数のCMのうち、何れのCMが担当する実ボリュームについても可能である。即ち、シンプロボリュームの記憶領域と対応する実ストレージの記憶領域は、複数の実ストレージに散在すること

50

になる。

【0010】

このため、シンプロボリュームに対してホストコンピュータからデータアクセスがあった場合、記憶装置は、先ずアクセス対象であるシンプロボリュームの担当 CM にてマッピング情報から実ストレージを求める。次に記憶装置は、シンプロボリュームの担当 CM と実ストレージの担当 CM とで CM 間通信を行いRead/Write処理といったアクセスを行う。

【0011】

つまりシンプロボリュームに対するデータアクセスの処理時間は、CM 間通信を行うために、従来の実論理ボリュームに対するデータアクセスの処理時間よりも長くなる。

【0012】

また、シンプロボリュームは、定義された容量よりも使用容量が少なければ、上述のようにメリットがあるが、使用容量が定義された容量とほぼ同等になった場合、定義された容量と同じ記憶容量の実ストレージが必要になるので、メリットがなくなる。

【0013】

このため、シンプロボリュームの使用容量が定義された容量と同等となった段階においては、シンプロボリュームよりもデータアクセス性能のよい実論理ボリュームの方が使用に適していることになる。

【0014】

そこで、開示の一実施形態は、仮想論理ボリュームの使用容量が増加して、仮想論理ボリュームのメリットが失われた場合に、仮想論理ボリュームを実論理ボリュームへマイグレーションしてアクセス性能の向上を図る技術を提供する。

【課題を解決するための手段】

【0015】

上記課題を解決するため、一実施形態としての制御部は、
他の制御部との接続を介して制御部間通信を行う通信部と、

論理的に定義した記憶領域を物理的な記憶領域と予め対応づけて実論理ボリュームとして管理し、データを書き加えるときに論理的に定義した記憶領域を物理的な記憶領域と対応づけて仮想論理ボリュームとして管理する管理部と、

前記仮想論理ボリュームの使用率が閾値を超えること且つ仮想論理ボリュームの記憶領域に対応する物理的な記憶領域が他の制御部で管理されていることを条件とし、当該条件を満たしているか否かを検知する検知部と、

前記条件を満たした場合に、当該仮想論理ボリュームを実論理ボリュームにマイグレーションするマイグレーション部と、
を備えた。

【発明の効果】

【0016】

開示の一実施形態は、仮想論理ボリュームの使用容量が増加して、仮想論理ボリュームのメリットが失われた場合に、仮想論理ボリュームを実論理ボリュームへマイグレーションしてアクセス性能の向上を図る技術を提供できる。

【図面の簡単な説明】

【0017】

【図1】実施形態1としての記憶装置の概略図

【図2】変換テーブルの説明図

【図3】シンプロボリュームを設定する手順の説明図

【図4】シンプロボリュームに対する読み込み手順の説明図

【図5】シンプロボリュームに対する書き込み手順の説明図

【図6】実施形態2としての記憶装置のブロック図

【発明を実施するための形態】

【0018】

以下、図面を参照して本発明を実施するための形態について説明する。以下の実施の形

10

20

30

40

50

態の構成は例示であり、本発明は実施の形態の構成に限定されない。

【0019】

実施形態1

図1は、本発明の一実施形態としての記憶装置の概略図である。

【0020】

図1に示すように、記憶装置10は、CA(Channel Adapter)11、制御部(CM:Centralized Module)12、実ストレージ13を備える。

【0021】

CA11は、ホストコンピュータ90と接続して当該ホストコンピュータ90との通信を制御するインターフェイスである。

10

【0022】

CM12は、ホストコンピュータ90からの要求に応じた実ストレージ13に対するアクセスを制御する。

【0023】

実ストレージ13は、ホストコンピュータ90から要求されるデータを記憶する大容量の記憶部であり、例えばハードディスクドライブである。

【0024】

本実施形態1の記憶装置10は、大容量化や冗長化のために複数の実ストレージ13をRAID(Redundant Arrays of Inexpensive Disks)によりグループ化して備え、実ストレージ13のRAIDグループ毎にCM12を設け、複数のCM12を備えている。

20

【0025】

また、CM12は、CPU21や、メモリ22、DA(Device Adapter)23、FRT(Front-end Router)24、記憶部25を備える。

【0026】

メモリ22は、記憶部13に対するアクセスを行う際のキャッシュや、CPU21が演算処理を行う際のメインメモリとして機能する。

【0027】

DA23は、実ストレージ13と接続して実ストレージ13に対するアクセスを制御するインターフェイスである。

【0028】

30

FRT24は、通信経路4Aを介してCM12間で行う通信を制御するインターフェイスである。

【0029】

記憶部25は、ハードディスクドライブやメモリディスクドライブ、ROM等、OSやファームウェア(プログラム)を記憶するユニットである。

【0030】

CPU21は、前記OSやファームウェアを記憶部25から適宜読み出して実行し、CA11やDA23、FRT24、記憶部13から読み出した情報を演算処理することにより、管理部C1や検知部C2、マイグレーション部C3、読み書き部C4としても機能する。

40

【0031】

管理部C1としてCPU21は、論理的に定義した記憶領域を予め物理的な記憶領域と対応づけて実論理ボリュームとして管理する。また、管理部C1としてCPU21は、論理的に定義した記憶領域をデータを書き加えるときに物理的な記憶領域と対応付けてシンプロボリュームとして管理する。

【0032】

なお、管理部C1としてのCPU21は、論理的な記憶領域と物理的な記憶領域との対応付けをメモリ22内の変換テーブルに登録している。図2は、当該変換テーブルの説明図である。図2に示すように、変換テーブルは、論理ボリュームのLBA(Logical Block Addressing)と実ストレージのLUN(Logical Unit Number)、LBAとを対応付けてい

50

る。

【0033】

また、管理部C1としてのCPU21は、DA23に実ストレージ13が接続されると、当該実ストレージ13から識別情報を読み出し、自身のCM12が担当する実ストレージ13であることを示す情報として、メモリ22内の担当ストレージテーブルに加える。更にCPU21は、自身のCM12の識別情報と、担当する実ストレージ13の識別情報とを対応つけて、他のCM12へ送信し、各CM12の担当ストレージテーブルを更新させる。この担当ストレージテーブルは、どのCM12が、どの実ストレージを担当しているのかを示すテーブルであり、担当するCM12の識別情報と、担当される実ストレージ13の識別情報を対応付けている。

10

【0034】

検知部C2としてCPU21は、シンプロボリュームの使用率が閾値を超えること且つシンプロボリュームの記憶領域に対応する物理的な記憶領域が他の制御部で管理（担当）されていることを条件とし、当該条件を満たしているか否かを検知する。例えば、CPU21は、変換テーブルを参照し、シンプロボリュームの論理的な記憶領域と対応付けられている実ストレージ（記憶部13）の記憶領域の容量Qzと、シンプロボリュームとして定義した記憶領域の全容量Qaとの比Qa/Qzを使用率として求める。そして、CPU21は、変換テーブルを参照し、シンプロボリュームの論理的な記憶領域が、他のCM12の管理下にある実ストレージ（記憶部13）の記憶領域と対応付けられているか否かを判定する。

20

【0035】

マイグレーション部C3としてCPU21は、検知部C2で条件を満たしたと判定した場合に、当該シンプロボリュームを実論理ボリュームにマイグレーションする。即ち、CPU21は、実ストレージ13の空き領域のうち、当該シンプロボリュームと同じ容量の物理的な記憶領域を確保して、当該記憶領域にシンプロボリューム内のデータを移動する。このとき、シンプロボリュームで未割当だった領域についてはデータの移動は必要ないため行わない。

【0036】

CM12は、CPU21を備え、プログラムを実行することで各部C1～C4の機能を実現するコンピュータでもある。なお、プログラム（ファームウェア）は、コンピュータが読み取り可能な記録媒体に記録されていても良い。ここで、コンピュータが読み取り可能な記録媒体とは、データやプログラム等の情報を電気的、磁気的、光学的、機械的、または化学的作用によって蓄積し、コンピュータから読み取ることができる記録媒体をいう。このような記録媒体の内コンピュータから取り外し可能なものとしては、例えばフレキシブルディスク、光磁気ディスク、CD-ROM、CD-R/W、DVD、DAT、8mmテープ、メモリカード等がある。

30

【0037】

また、コンピュータに固定された記録媒体としてハードディスクやROM（リードオンリーメモリ）等がある。

【0038】

図3は、記憶装置10において、シンプロボリュームを設定する手順の説明図である。

40

【0039】

シンプロボリュームを新たに作成する場合、オペレータは、管理端末（不図示）等の記憶装置10に接続した端末や記憶装置10の操作部（不図示）から記憶装置10にアクセスして、シンプロボリュームを作成するCM12を選択し、記憶容量を指定してシンプロボリュームの作成を指示する。当該指示を受信した記憶装置10は、選択されたCM12のCPU21が、ボリューム名や、記憶容量、シンプロボリュームであることを示す種別といった定義をメモリ22へ書き込むことでシンプロボリュームを作成する（S1）。

【0040】

そして、オペレータがマイグレーションを行うシンプロボリュームの使用率を入力し、

50

当該使用率を記憶装置 10 が受信すると (S2)、CM12 の CPU21 が当該使用率を閾値としてメモリ 22 へ書き込む (S3)。

【0041】

この定義に従って CM12 がシンプロボリュームを管理するので、例えば OS は、当該シンプロボリュームを定義された記憶容量のデバイスとして認識できる。

なお、ボリュームを作成した場合、CM12 の CPU21 は、ボリュームを特定する識別情報と自身の CM12 を他の CM12 と識別する識別情報とを対応付けたボリューム位置情報を他の全ての CM12 へ送信する。即ち、どのボリュームが、どの CM12 の管理しているのか (担当しているのか) を示すボリューム位置情報を全ての CM12 で共有する。

10

【0042】

図4は、シンプロボリュームに対する読み込み手順の説明図である。

【0043】

ホストコンピュータ90がシンプロボリュームに対するリード要求を記憶装置10へ送信すると、CA11を介して当該リード要求をCM12が受信する (S21)。

【0044】

リード要求を受信した CM12 は、読み込み対象のシンプロボリュームが自身の管理下か否かを判定する。例えば、CM12 は、メモリ 22 を参照し、読み込み対象のシンプロボリュームが定義されているか否かを判定する (S22)。

【0045】

リード対象のシンプロボリュームが、リード要求を受信した CM12 の管理下で無かった場合 (S22, No)、リード要求を受信した CM12 は、ボリューム位置情報を参照してリード対象のシンプロボリュームを管理する CM12 を求める。そして、リード要求を受信した CM12 は、FRT24 を介する CM 間通信により、リード対象のシンプロボリュームを管理する CM12 にリード要求を伝達する (S23)。

20

【0046】

ステップ S23 でリード要求を伝達された CM12 或いは、ステップ S22 で読み込み対象が自身の管理下であった CM12 は、変換テーブルを参照し、読み込み対象の実ストレージ 13 を検索する (S24)。

【0047】

CM12 は、担当ストレージテーブルを参照し、ステップ S24 で索出した読み込み対象の実ストレージ 13 が、自身の担当か否か、即ち他の CM12 を介さずに直接接続している実ストレージ 13 か否かを判定する (S25)。

30

【0048】

ステップ S25 において、読み込み対象の実ストレージが自身の担当で無い場合 (S25, No)、CM12 は、FRT24 を介する CM 間通信により、読み込み対象の実ストレージ 13 を担当する CM12 へリード要求を伝達する (S26)。

【0049】

ステップ S26 でリード要求を伝達された CM12 或いは、ステップ S25 で読み込み対象の実ストレージが自身の担当であった CM12 は、リード要求に従い実ストレージ 13 からデータを読み出す。そして、CM12 は、当該データをホストコンピュータ90からリード要求を受信した CM12 へ送信してメモリ 22 へ展開させる (S27)。

40

【0050】

当該データを受信した CM12 は、メモリ 22 に展開されたデータをホストコンピュータ90へ送信して読み込み処理を完了する (S28)。

【0051】

図5はシンプロボリュームに対する書き込み手順の説明図である。

【0052】

ホストコンピュータ90がシンプロボリュームに対するライト要求を記憶装置10に送信すると、CA11を介して当該ライト要求をCM12が受信する (S31)。

50

【0053】

ライト要求を受信したCM12は、変換テーブルを参照し、当該ライト要求によって書き込む論理ボリュームの記憶領域が実ストレージ13の記憶領域と対応付けられているか否かを判定する(S32)。

【0054】

ステップS32で実ストレージが割り当てられている場合(S32, Yes)、読書き部C4としてのCPU21は、ライト要求に従って書き込み処理を行う(S33)。

【0055】

また、ステップS32で実ストレージが割り当てられていない場合(S32, No)、管理部C1としてのCPU21は、書き込み対象のシンプロボリュームの記憶領域に、実ボリューム13の記憶領域を割り当てる(S34)。10

【0056】

ステップS34の割り当て後、読書き部C4としてのCPU21は、ライト要求に従って書き込み処理を行う(S35)。

【0057】

ステップS34, 35で新たなデータが書き加えられると、検知部C2としてのCPU21は、シンプロボリュームの使用率が閾値を超えたか否かを検知する(S36)。

【0058】

ステップS36で使用率が閾値を超えていなければ、CPU21は処理を終了し、使用率が閾値を超えていれば、シンプロボリュームの記憶領域と対応する物理的な記憶領域が複数のCM12が担当する実ストレージに散在しているか否かを判定する。即ち、CPU21は、使用率が閾値を超えたシンプロボリュームの記憶領域が、他のCM12で担当する実ストレージの記憶領域と対応しているか否かを判定する(S37)。20

【0059】

ステップS36及びS37の条件を満たした場合、当該シンプロボリュームを実論理ボリュームへマイグレーションする(S38)。

【0060】

シンプロボリュームは、初期投資を抑えることができるというメリットを有するが、条件によっては、アクセス性能の低下を招くことがある。例えば、シンプロボリュームの記憶領域が、複数のCM12でそれぞれ管理する物理的な記憶領域と対応している場合、図3に示すようにCM間通信でリード要求を伝達する必要があり、アクセス性能が低下する。30

【0061】

本実施形態1の記憶装置10は、シンプロボリュームの使用率が閾値を超えてシンプロボリュームのメリットが失われた場合には、シンプロボリュームを実論理ボリュームへマイグレーションして、アクセス性能の向上を図ることができる。なお、本実施形態1の記憶装置10は、ステップS37でアクセス性能の低下の有無を判定し、アクセス性能が低下する場合にマイグレーションを行うので、アクセス性能が低下していないときに、無駄にマイグレーションすることがない。

【0062】

また、本実施形態1の記憶装置10は、マイグレーションを行う際、マイグレーション先の実論理ボリュームを管理するCM12と同じCM12が管理する物理的な記憶領域と当該マイグレーション先の実論理ボリュームの記憶領域とを対応付けても良い。これにより、少なくとも図3のステップS26のCM間通信が不要となり、アクセス性能が向上する。

【0063】

更に、シンプロボリュームの記憶領域を管理するCM12が、ホストコンピュータ90からアクセス要求を受信したCM12と同じであれば、図3のステップ23に示すCM間通信が不要である。このためシンプロボリュームを作成する場合、当該シンプロボリュームの記憶領域を管理するCM12は、ホストコンピュータ90からアクセス要求を受信す4050

るCM12と同じものが選択されることが多い。そこで、本実施形態の記憶装置10は、マイグレーションを行う際、マイグレーション対象の仮想論理ボリュームを管理する管理部と同じ管理部が管理する実論理ボリュームをマイグレーション先とする。これにより、適切なマイグレーション先へマイグレーションすることができる。

【0064】

実施形態2

図6は、本発明の実施形態2としての記憶装置のブロック図である。

【0065】

図6に示すように、記憶装置10は、複数の制御部(CM:Centralized Module)12と、複数の実ストレージ13を備える。

10

【0066】

CM12は、ホストコンピュータ90からの要求に応じた実ストレージ13に対するアクセスを制御する制御部である。

【0067】

実ストレージ13は、ホストコンピュータ90から要求されるデータを記憶する大容量の記憶部であり、例えばハードディスクドライブである。

【0068】

CM12は、メモリ22、通信部(FRT:Front-end Router)24、記憶部25、管理部C1や、検知部C2、マイグレーション部C3、読み書き部C4を備える。

【0069】

メモリ22は、記憶部13に対するアクセスを行う際のキャッシング等として機能する。

20

【0070】

FRT24は、通信経路4Aを介してCM12間で行う通信を制御するインターフェイスである。

【0071】

記憶部25は、ハードディスクドライブやメモリディスクドライブ、ROM等、設定情報やファームウェア(プログラム)を記憶するユニットである。

【0072】

管理部C1は、論理的に定義した記憶領域を予め物理的な記憶領域と対応づけて実論理ボリュームとして管理する。また、管理部C1は、論理的に定義した記憶領域をデータを書き加えるときに物理的な記憶領域と対応付けてシンプロボリュームとして管理する。

30

【0073】

なお、管理部C1は、論理的な記憶領域と物理的な記憶領域との対応付けをメモリ22内の変換テーブルに登録している。

【0074】

また、管理部C1は、実ストレージ13が接続されると、当該実ストレージ13から識別情報を読み出し、自身の属するCM12が担当する実ストレージ13であることを示す情報として、メモリ22内の担当ストレージテーブルに加える。更に管理部C1は、自身のCM12の識別情報と、担当する実ストレージ13の識別情報とを対応づけて、他のCM12へ送信し、各CM12の担当ストレージテーブルを更新させる。この担当ストレージテーブルは、どのCM12が、どの実ストレージを担当しているのかを示すテーブルであり、担当するCM12の識別情報と、担当される実ストレージ13の識別情報とを対応付けている。

40

【0075】

検知部C2は、シンプロボリュームの使用率が閾値を超えること且つシンプロボリュームの記憶領域に対応する物理的な記憶領域が他の制御部で管理されていることを条件とし、当該条件を満たしているか否かを検知する。例えば、検知部C2は、変換テーブルを参照し、シンプロボリュームの論理的な記憶領域と対応付けられている実ストレージ(記憶部13)の記憶領域の容量Qzと、シンプロボリュームとして定義した記憶領域の全容量Qaとの比Qa/Qzを使用率として求める。そして、検知部C2は、変換テーブルを参

50

照し、シンプロボリュームの論理的な記憶領域が、他の CM12 の管理下にある実ストレージ（記憶部 13）の記憶領域と対応付けられているか否かを判定する。

【0076】

マイグレーション部 C3 は、検知部 C2 で条件を満たしたと判定した場合に、当該シンプロボリュームを実論理ボリュームにマイグレーションする。即ち、マイグレーション部 C3 は、実ストレージ 13 の空き領域のうち、当該シンプロボリュームと同じ容量の物理的な記憶領域を確保し、当該記憶領域にシンプロボリューム内のデータを移動する。

【0077】

これらの構成により、本実施形態 2 の記憶装置 10 は、シンプロボリュームの使用率が閾値を超えてシンプロボリュームのメリットが失われた場合には、シンプロボリュームを実論理ボリュームへマイグレーションして、アクセス性能の向上を図ることができる。なお、本実施形態 1 の記憶装置 10 は、ステップ S37 でアクセス性能の低下の有無を判定し、アクセス性能が低下する場合にマイグレーションを行うので、アクセス性能が低下していないときに、無駄にマイグレーションすることがない。

【0078】

また、本実施形態 2 の記憶装置 10 は、マイグレーションを行う際、マイグレーション先の実論理ボリュームを管理する CM12 と同じ CM12 が管理する物理的な記憶領域と当該マイグレーション先の実論理ボリュームの記憶領域とを対応付けても良い。これにより、少なくとも図 3 のステップ S26 の CM 間通信が不要となり、アクセス性能が向上する。

【0079】

更に本実施形態の記憶装置 10 は、マイグレーションを行う際、マイグレーション対象の仮想論理ボリュームを管理する管理部と同じ管理部が管理する実論理ボリュームをマイグレーション先としても良い。これにより、適切なマイグレーション先へマイグレーションすることができる。

【0080】

本例の記憶装置 10 の構成要素のうち、情報を処理する要素は、特定の動作を行う基本的な回路を組み合わせて各々の機能を実現したハードウェアである。また、これら情報を処理する要素の一部或いは全部は、汎用のプロセッサがソフトウェアとしてのプログラムを実行することによって各々の機能を実現するものでも良い。

【0081】

記憶装置 10 において、管理部 C1 や、検知部 C2、マイグレーション部 C3、読み書き部 C4 は、前記情報を処理する要素である。

【0082】

これら情報を処理する要素であるハードウェアは、例えば、FPGA [Field Programmable Gate Array]、ASIC [Application Specific Integrated Circuit]、LSI [Large Scale Integration] といった基本的な回路を備える。また、当該ハードウェアは、I C [Integrated Circuit]、ゲートアレイ、論理回路、信号処理回路、アナログ回路といった基本的な回路を備えても良い。

【0083】

論理回路としては、例えば、AND、OR、NOT、NAND、NOR、フリップフロップ、カウンタ回路がある。信号処理回路には、信号値に対し、例えば、加算、乗算、除算、反転、積和演算、微分、積分を実行する回路が、含まれていてもよい。アナログ回路には、例えば、信号値に対して、增幅、加算、乗算、微分、積分を実行する回路が、含まれていてもよい。

【符号の説明】

【0084】

10 記憶装置

11 C A

12 制御部 (CM)

10

20

30

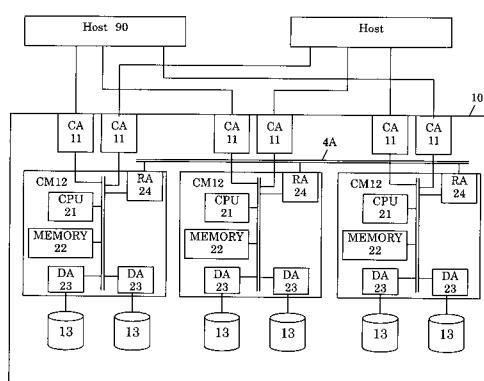
40

50

1 3 実ストレージ
 9 0 ホストコンピュータ
 2 1 C P U
 2 2 メモリ 2 2
 2 3 D A
 2 4 通信部 (F R T)
 2 5 記憶部
 C 1 管理部
 C 2 検知部
 C 3 マイグレーション部
 C 4 読書き部

10

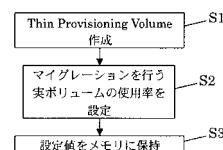
【図 1】



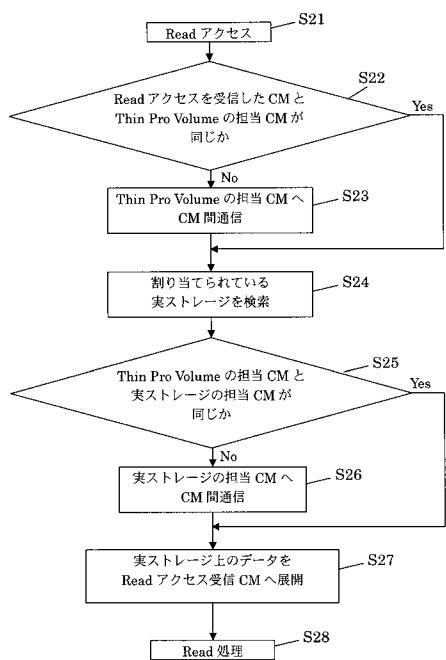
【図 2】

LBA変換テーブル(例)		
論理ボリューム	物理LUN	物理LBA
32MB	0x0000	0x0000000000001000
32MB	0x0003	0x000000000040000
32MB	0xFFFF	0xFFFFFFF0FFFFFFF
32MB	0x00A0	0x0000000070800030
32MB	0xFFFF	0xFFFFFFF0FFFFFFF

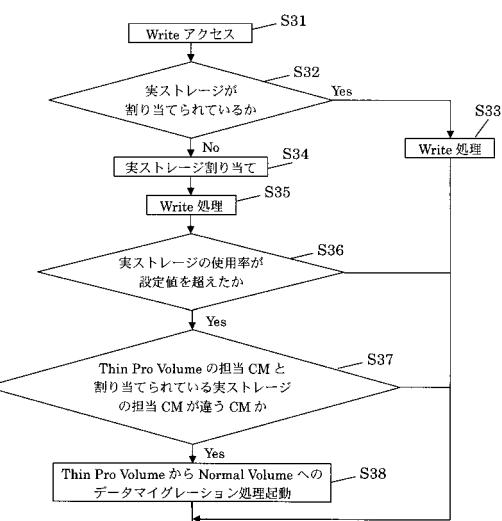
【図 3】



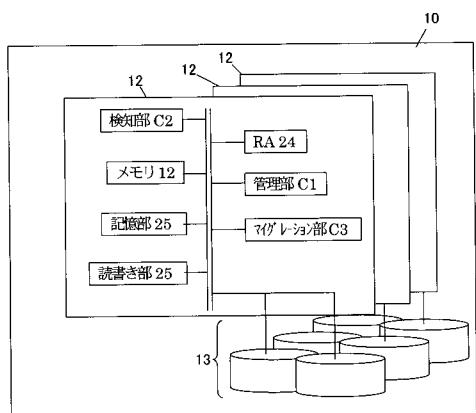
【図4】



【図5】



【図6】



フロントページの続き

審査官 池田 聰史

(56)参考文献 特開2008-310734 (JP, A)

特開2007-156815 (JP, A)

特開2005-275525 (JP, A)

特開2004-199551 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 12 / 00