

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-3694
(P2012-3694A)

(43) 公開日 平成24年1月5日(2012.1.5)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 3/06 (2006.01)	G06F 3/06 301Z	5B014
G06F 13/10 (2006.01)	G06F 3/06 302J	5B065
G06F 12/00 (2006.01)	G06F 13/10 340A	5B082
	G06F 12/00 514E	
	G06F 12/00 501B	

審査請求 未請求 請求項の数 10 O L (全 30 頁)

(21) 出願番号 特願2010-140601 (P2010-140601)
(22) 出願日 平成22年6月21日 (2010.6.21)

(71) 出願人 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 100108187
弁理士 横山 淳一
(72) 発明者 松村 忠
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
(72) 発明者 吉田 雅裕
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
(72) 発明者 内山 賢治
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

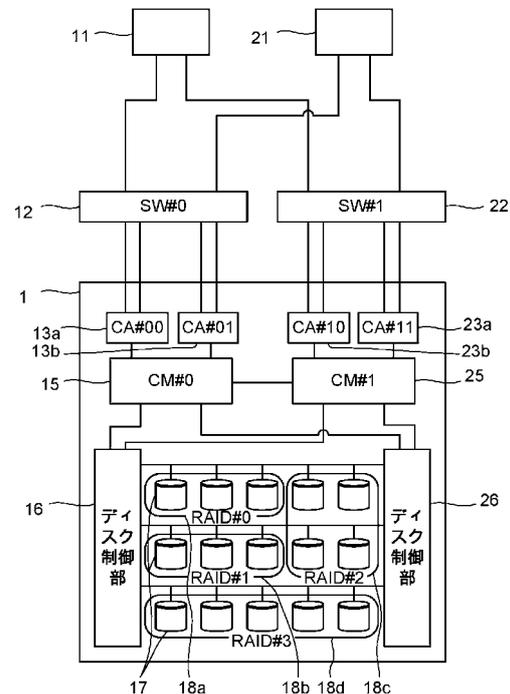
(54) 【発明の名称】 ストレージ制御装置、ストレージ制御方法、ストレージシステム

(57) 【要約】

【課題】低消費電力で、仮想的ボリュームを設定した複数の記憶装置を駆動させるストレージ制御装置を提供する。

【解決手段】各々が複数の記憶装置を有する複数のRAID装置に複数の論理ボリュームを設定し、その少なくとも一つを仮想ボリュームとして割り当てると共に、記憶されたデータの容量に応じて、未割当ての論理ボリュームを仮想ボリュームに追加して割り当てる仮想ボリューム制御部と、仮想ボリューム - 論理ボリューム間のアクセス変換情報を格納する記憶部と、アクセス変換情報に基づき仮想ボリュームへのアクセスを変換するアクセス制御部と、論理ボリュームの各々へのアクセス頻度を監視するアクセス監視部と、アクセス頻度に応じて決定された論理ボリュームのデータを他のRAID装置に移動する移動制御部と、データ移動先の論理ボリュームへのアクセスのためにアクセス変換情報を更新するアクセス変換情報更新部とを有するストレージ制御装置。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

各々が複数の記憶装置を有する複数の R A I D グループを制御するストレージ制御装置において、

各々の前記 R A I D グループに複数の論理ボリュームを設定し、前記複数の論理ボリュームの少なくとも一つを仮想ボリュームに対して割り当てると共に、前記仮想ボリュームに記憶されるデータ量に応じて、前記複数の論理ボリュームのうち少なくとも一つの未割当ての論理ボリュームを前記仮想ボリュームに追加して割り当てる仮想ボリューム制御部と、

前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス変換情報を格納する記憶部と、

前記アクセス変換情報に基づき、前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス制御部と、

前記論理ボリュームの各々へのアクセス頻度を監視するアクセス監視部と、

前記アクセス監視部が監視したアクセス頻度に応じて論理ボリュームを決定し、決定された論理ボリュームのデータを当該論理ボリュームが設定された R A I D グループとは異なる他の R A I D グループに移動する移動制御部と、

データ移動先の論理ボリュームへのアクセスのために前記アクセス変換情報を更新するアクセス変換情報更新部と

を有するストレージ制御装置。

10

20

【請求項 2】

前記移動制御部は、アクセス頻度が所定値より高い論理ボリュームのデータを所定の R A I D グループに移動することを特徴とする請求項 1 に記載のストレージ制御装置。

【請求項 3】

前記移動制御部は、アクセス頻度が所定値より低い論理ボリュームのデータを所定の R A I D グループに移動することを特徴とする請求項 1 又は 2 に記載のストレージ制御装置。

【請求項 4】

更に、前記論理ボリュームへのアクセス頻度に応じて前記 R A I D グループの各々の省電力制御を行う省電力制御部を有する請求項 1 乃至 3 のいずれか 1 項に記載のストレージ制御装置。

30

【請求項 5】

前記複数の記憶装置のうち少なくとも一つは、ディスクと、該ディスクを回転させるスピンドルモータとを有するハードディスクドライブであり、

前記省電力制御部は、アクセス頻度が所定値より低いディスクのスピンドルモータの回転を停止することを特徴とする請求項 4 に記載のストレージ制御装置。

【請求項 6】

前記複数の記憶装置のうち少なくとも一つは、ソリッドステートディスクであり、

前記省電力制御部は、アクセス頻度が所定値より低いソリッドステートディスクへの電源供給を停止することを特徴とする請求項 4 又は 5 に記載のストレージ制御装置。

40

【請求項 7】

前記アクセス監視部は、アクセスされた論理ボリュームの情報とアクセス時刻の情報をモニタし、

更に、前記アクセスされた論理ボリュームの情報と前記アクセス時刻の情報を解析して、前記複数の論理ボリュームのうち前記 R A I D グループが異なりアクセス時刻が重複する 2 つの論理ボリュームのデータが同一の R A I D グループに配置されるように、前記複数の論理ボリュームのうち、データの移動元となる R A I D グループの論理ボリュームと、前記複数の論理ボリュームのうち、データの移動先となる R A I D グループの論理ボリュームとを決定するアクセス解析部を有することを特徴とする請求項 1 乃至 6 のいずれか 1 項に記載のストレージ制御装置。

50

【請求項 8】

前記アクセス解析部は、前記複数の論理ボリュームのうちアクセス頻度が所定の閾値未満の論理ボリュームを前記データの移設元となる論理ボリュームに決定することを特徴とする請求項 7 に記載のストレージ制御装置。

【請求項 9】

各々が複数の記憶装置を有する複数の R A I D グループにそれぞれ複数の論理ボリュームを設定し、前記複数の論理ボリュームの少なくとも一つを仮想ボリュームに対して割り当てるとともに、前記仮想ボリュームに記憶されるデータ量に応じて、前記複数の論理ボリュームのうち少なくとも一つの未割り当ての論理ボリュームを前記仮想ボリュームに追加して割り当て、前記仮想ボリュームへのアクセスを前記複数の論理ボリュームへのアクセスに変換する工程を介して前記複数の記憶装置へアクセスが行われるストレージシステムを制御するストレージ制御方法において、

前記論理ボリュームの各々へのアクセス頻度を監視する工程と、

前記監視したアクセス頻度に応じて論理ボリュームを決定し、決定された論理ボリュームのデータを、当該論理ボリュームが設定された R A I D グループとは異なる他の R A I D グループに移動する工程と、

データ移動先の論理ボリュームへのアクセスのために前記アクセス変換情報を更新する工程と

を有することを特徴とするストレージ制御方法。

【請求項 10】

各々が複数の記憶装置を有する複数の R A I D グループと、

各々の前記 R A I D グループに複数の論理ボリュームを設定し、前記複数の論理ボリュームの少なくとも一つを仮想ボリュームとして割り当てると共に、前記仮想ボリュームに記憶されるデータ量に応じて、前記複数の論理ボリュームのうち少なくとも一つの未割り当ての論理ボリュームを前記仮想ボリュームに追加して割り当てるとともに、

前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス変換情報を格納する記憶部と、

前記アクセス変換情報に基づき、前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス制御部と、

前記論理ボリュームの各々へのアクセス頻度を監視するアクセス監視部と、

前記アクセス監視部が監視したアクセス頻度に応じて論理ボリュームを決定し、決定された論理ボリュームに記憶されたデータを当該論理ボリュームが設定された R A I D グループとは異なる他の R A I D グループに移動する移動制御部と、

移動先の論理ボリュームへのアクセスのために前記アクセス変換情報を更新するアクセス変換情報更新部と

を有するストレージシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ制御装置及びストレージ制御方法に関するものである。

【背景技術】

【0002】

ストレージ装置の制御方法として、例えば、M A I D (M a s s i v e A r r a y o f I d l e D i s k s) 技術がある。M A I D 技術は、R A I D (R e d u n d a n t A r r a y s o f I n e x p e n s i v e D i s k s) 装置 (R A I D グループ) において、アクセスがないストレージ装置のディスクをアイドル状態に遷移させることで、ストレージ装置の消費電力を抑制する技術である。この M A I D 技術を使用して、R A I D グループ単位にディスクモータを制御する方式がある。R A I D グループ内のディスクのモータオフ時間が長くなるほど、ストレージ装置の消費電力量は減る。そこで

、例えば、同一ホスト又はサーバが、一つのRAIDグループ内の論理ボリュームにアクセスし、他のRAIDグループ内のディスクのモータをオフとするように運用することが開示されている。また、例えば、一つのRAIDグループ内のボリュームが複数のサーバで用いられる場合、各サーバは、同じような時間帯にそのRAIDグループ内の論理ボリュームにアクセスするように運用することが開示されている。

【0003】

ストレージ装置の消費電力の低減に関連する技術が特許文献1、2に開示されている。

【0004】

シンプロビジョニング(Thin Provisioning)による仮想ボリュームを構築したストレージ装置群において、仮想ボリュームの容量が不足した場合に論理ボリュームがプール領域からアサインされる。プール領域には、通常、各レベルのRAIDグループがあり、各RAIDグループを構成する一部の論理ボリュームが、仮想ボリュームの一領域として獲得される。ファイルシステムが仮想ボリュームにおいてデータの追加及び削除を繰り返し行くと、仮想ボリュームに割り振られるRAIDグループの数が増加する傾向にある。割り振られるRAIDグループの数が増加すると、アクセスが頻繁に行われるデータの量に対してストレージ装置の数が多くなり、消費電力が増加する。

10

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2009-110451号公報

20

【特許文献2】国際公開第2003/050707号パンフレット

【発明の概要】

【発明が解決しようとする課題】

【0006】

本発明は、少ない消費電力で仮想ボリュームを設定した複数のストレージ装置を駆動させるストレージ制御装置を提供する。

【課題を解決するための手段】

【0007】

本発明の一側面によると、

各々が複数の記憶装置を有する複数のRAIDグループを制御するストレージ制御装置において、

30

各々の前記RAIDグループに複数の論理ボリュームを設定し、前記複数の論理ボリュームの少なくとも一つを仮想ボリュームに対して割り当てると共に、前記仮想ボリュームに記憶されるデータ量に応じて、前記複数の論理ボリュームのうち少なくとも一つの未割当ての論理ボリュームを前記仮想ボリュームに追加して割り当てる仮想ボリューム制御部と、

前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス変換情報を格納する記憶部と、

前記アクセス変換情報に基づき、前記仮想ボリュームへのアクセスを前記割り当てられた論理ボリュームへのアクセスに変換するアクセス制御部と、

40

前記論理ボリュームの各々へのアクセス頻度を監視するアクセス監視部と、

前記アクセス監視部が監視したアクセス頻度に応じて論理ボリュームを決定し、決定された論理ボリュームのデータを当該論理ボリュームが設定されたRAIDグループとは異なる他のRAIDグループに移動する移動制御部と、

データ移動先の論理ボリュームへのアクセスのために前記アクセス変換情報を更新するアクセス変換情報更新部と

を有するストレージ制御装置が提供される。

【発明の効果】

【0008】

本発明のストレージ制御装置によれば、少ない消費電力で仮想ボリュームを設定した複

50

数のストレージ装置を駆動させることができる。

【図面の簡単な説明】

【0009】

【図1】図1は、第1実施形態のストレージ制御装置を備えるストレージシステムの物理的構成の一例を示す図である。

【図2】図2A～2Cは、サーバから認識できる仮想ボリュームと、RAIDグループに設定される論理ボリュームとの対応関係を示す図である。

【図3】図3は、第1実施形態のストレージ制御装置の物理的構成を示す図である。

【図4】図4は、第1実施形態のストレージ制御装置のストレージコントローラモジュールの機能的構成を示す処理ブロック図である。

【図5】図5は、LBA変換テーブルの一例である。

【図6】図6は、アクセス解析部134が作成するアクセス管理テーブルの一例である。

【図7】図7は、プールボリュームに格納されたデータの再配置を説明する図である。

【図8】図8は、プールボリュームに格納されたデータの再配置を説明する図である。

【図9】図9は、再配置制御部135によって書き換えられたLBA変換テーブル77である。

【図10】図10は、第1実施形態のストレージ制御装置の制御フローである。

【図11】図11は、アクセス解析部が行う解析と、再配置制御部が行う再配置の一例である制御フローである。

【図12】図12は、アクセス解析部が行う解析と、再配置制御部が行う再配置の一例である制御フローである。

【図13】図13は、アクセスモニタ部がモニタリングした情報に基づき、アクセス解析部が作成したアクセスフラグマップである。

【図14】図14は、アクセス解析部による解析の経過を示すアクセスフラグマップである。

【図15】図15は、アクセス解析部による解析の経過を示すアクセスフラグマップである。

【図16】図16は、第2実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析のフローチャートの一例である。

【図17】図17は、第2実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析のフローチャートの一例である。

【図18】図18は、第3実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析のフローチャートの一例である。

【図19】図19は、第3実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析のフローチャートの一例である。

【図20】図20は、第3実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析のフローチャートの一例である。

【図21】図21Aは、図14Aに示されるアクセスフラグマップの解析において、プールボリューム#0-0をRAIDグループ#1へ移動すると仮定した場合の仮重複アクセス特性である。図21Bは、図14Aに示されるアクセスフラグマップの解析において、プールボリューム#1-1をRAIDグループ#0へ移動すると仮定した場合の仮重複アクセス特性である。

【発明を実施するための形態】

【0010】

図1～図3は、第1実施形態のストレージ制御装置を示す図である。

【0011】

図1は、第1実施形態のストレージ制御装置15及び25を備えるストレージシステム1の物理的構成の一例を示す図である。ストレージシステム1は、ストレージ制御装置15、25、チャンネルアダプタ13a、13b、23a、23b、ディスク制御部16、26、RAIDグループ#0～3で示されるRAIDグループ18a～18dを含む。

10

20

30

40

50

【 0 0 1 2 】

ストレージ制御装置 1 5 (C M # 0)、2 5 (C M # 1) は、ディスク制御部 1 6、2 6 を介して R A I D グループ # 0 ~ 3 で示される 4 つの R A I D 装置 (R A I D グループ) 1 8 a ~ 1 8 d に接続されている。R A I D 装置 1 8 a ~ 1 8 d は、それぞれ複数の記憶装置 1 7 を含む。各々の記憶装置 1 7 は、記憶媒体を駆動して情報の記録、再生を行う装置であり、例えばハードディスクドライブ (H D D) である。H D D は、ディスクと、そのディスクを回転させるディスクモータを備える。

【 0 0 1 3 】

記憶装置 1 7 内のディスクを回転させるディスクモータは、ストレージ装置に接続されたホストやサーバから認識できるボリューム単位ではなく、R A I D グループ単位の I / O (I n p u t / O u t p u t) アクセスに基づき制御される。R A I D グループ内のいずれかのボリュームにアクセスがあれば、アクセスがあったボリュームを含む R A I D グループのディスクのモータはオンとなる。また、R A I D グループ内にある全てのボリュームに対して所定の条件でアクセスがなくなれば、その R A I D グループ内のディスクのモータはオフとなる。

10

【 0 0 1 4 】

本実施形態においては、冗長化のため、2 台のストレージ制御装置 1 5、2 5 によって R A I D グループ 1 8 a ~ 1 8 d を管理しているが、これによりストレージ制御装置の数が限定されるものではない。

【 0 0 1 5 】

ストレージ制御装置 1 5 は、チャンネルアダプタ 1 3 a、1 3 b と接続されている。チャンネルアダプタ 1 3 a、1 3 b は、ストレージ制御装置 1 5 とサーバ 1 1 及びサーバ 2 1 との間でデータの授受を制御する。サーバ 1 1 及びサーバ 2 1 とストレージ制御装置 1 5 との間のデータ授受は、スイッチ 1 2 (S W # 0) を経由して行われる。

20

【 0 0 1 6 】

ストレージ制御装置 2 5 は、チャンネルアダプタ 2 3 a、2 3 b と接続されている。チャンネルアダプタ 2 3 a、2 3 b は、ストレージ制御装置 2 5 とサーバ 1 1 及びサーバ 2 1 との間でデータの授受を制御する。サーバ 1 1 及びサーバ 2 1 とストレージ制御装置 1 5 との間のデータの授受は、スイッチ 2 2 (S W # 1) を経由して行われてもよい。

【 0 0 1 7 】

ストレージ制御装置 1 5、2 5 は、2 台のサーバ 1 1、2 1 とデータの授受を行うが、これによりストレージ制御装置とデータを行うサーバの数が限定されるものではない。スイッチ 1 2、2 2 は、ネームサーバ機能やルーティング機能等を備えていてもよい。

30

【 0 0 1 8 】

R A I D グループ 1 8 a ~ 1 8 d は、それぞれ、サーバ 1 1、2 1 がアクセスする対象のデータを格納する。ストレージ制御装置 1 5、2 5 は、R A I D グループ 1 8 a ~ 1 8 d それぞれに少なくとも一つの論理ボリューム (図示せず) を設定する。

【 0 0 1 9 】

ここで、論理ボリュームは、R A I D 装置 1 8 a ~ 1 8 d が有する複数の記憶装置にまたがって、R A I D の規格に基づき設定される論理的な記憶領域である。各記憶装置に実際に割り振られている物理ブロックアドレスは、この論理ボリュームに割り当てられている。

40

【 0 0 2 0 】

論理ボリュームは、サーバ 1 1、2 1 が一つのまとまったデータ格納領域として認識できる仮想ボリュームに、変換表を利用して割り当てられる。変換表には、仮想ボリュームの論理ブロックアドレスと、論理ボリュームの論理ブロックアドレスとが割り付けられている。一つの仮想ボリュームは、一つの R A I D グループに設定された論理ボリュームから割り当てられてもよいし、複数の R A I D グループにわたって設定された論理ボリュームが割り当てられてもよい。

【 0 0 2 1 】

50

仮想ボリュームの設定（定義）は、例えば、システム管理者が、ストレージ制御装置 15 に対して図示されないインターフェースを介してストレージ管理ツールを接続することにより行われる。

【0022】

なお、例えば、論理ボリュームの論理ブロックアドレス（LBA）と各記憶装置に実際に割り振られている物理ブロックアドレスとの間に更に他のLBAが設定されていてもよい。また、例えば、論理ボリュームのLBAと仮想ボリュームのLBAとの間に更に他のLBAが設定されていてもよい。

【0023】

実施形態において、仮想ボリュームは、シンプロビジョニングによる仮想ボリュームである。複数の記憶装置にまたがって設定された複数の論理ボリュームは、プール領域（ストレージプール）を構成するプールボリュームとして利用される。仮想ボリュームにアクセスが有ると、プール領域の論理ボリュームのうち仮想ボリュームに割り当てられていないものが割り当てられる。プール領域には、通常、各レベルのRAIDグループがある。各RAIDグループを構成する少なくとも一つの論理ボリュームのうち、仮想ボリュームに未割当の論理ボリュームが、仮想ボリュームの一領域として獲得される。この新たな論理ボリュームの獲得により、一つの仮想ボリュームは複数のRAIDグループに設定された論理ボリュームを含んでいることがある。仮想ボリュームに対する論理ボリュームの割り当ては、後述のLBA変換テーブル77により、仮想ボリュームに設定された論理ブロックアドレスと、論理ボリュームに設定された論理ブロックアドレスとが対応付けられることにより実現される。

【0024】

サーバ11、21がアクセスする対象のデータは、アクセスするデータの記憶位置、読み込み命令か書き込み命令かを識別する情報、および書込む対象のデータ情報を含む。

【0025】

図2A～2Cは、サーバから認識できる仮想ボリュームと、RAIDグループに設定される論理ボリュームとの対応関係を示す図である。

【0026】

図2Aは、ストレージプールとして準備された複数のRAIDグループを示す模式図である。図2Aにおいて、ストレージプール19の中にRAIDグループとしてRAID18a（#0）、RAID18b（#1）が表示されている。各RAIDグループ18はそれぞれ複数の記憶装置17を含む。

【0027】

図2Bは、RAIDグループ#0に設定されたプールボリューム群51（#0）と、RAIDグループ#1に設定されたプールボリューム群52（#1）とを示す模式図である。プールボリューム群51（#0）は、プールボリューム（論理ボリューム）61（#0-0～#0-n）を有する。プールボリューム群52（#1）は、プールボリューム62（#1-0～#1-n）を有する。プールボリューム群51（#0）の各プールボリューム61は、RAID18aに含まれる複数の記憶装置17aにまたがって設定される。プールボリューム群52の各プールボリューム62は、RAID18bに含まれる複数の記憶装置17bにまたがって設定される。

【0028】

図2Cは、プールボリュームと仮想ボリュームとの対応関係の一例を示す模式図である。図2Bのプールボリューム群51のプールボリューム#0-0、#0-1、及びプールボリューム群52のプールボリューム#1-0が、サーバ11が認識する仮想ボリューム81の領域74（#0-0、#0-1、及び#0-2）に割り当てられている様子が示されている。

【0029】

なお、仮想ボリューム81内の領域74（#0-0、#0-1、及び#0-2）内の境界は、説明の便宜上示されたものである。実際には、サーバ側からは領域#0-0、#0

10

20

30

40

50

- 1、及び# 0 - 2の境界は識別できない。サーバ11は、LBA変換テーブル77における仮想LBAにアクセスすることでプールボリュームへアクセスし、更に論理ブロックアドレス-物理ブロックアドレス変換テーブル(図示せず)により、各記憶装置17(17a、17b)に設定された物理ボリュームにアクセスする。図2Cにおいて、図2Bに示されるプールボリュームのうちサーバに割り当てられていない領域は、「未割当」と表示されている。図2Cの仮想ボリューム81は、割り当てられたプールボリューム及び未割当のプールボリュームを含む全体の領域が、サーバ側に認識されている。

【0030】

図3は、第1実施形態のストレージ制御装置15、25の物理的構成を示す図である。ストレージ制御装置15は、制御部2、メモリ3を備え、ストレージ制御装置25は、制御部4、メモリ5を備える。ストレージ制御装置15の制御部2、メモリ3とストレージ制御装置25の制御部4、メモリ5とは、それぞれ同じ機能である。よって、ストレージ制御装置15の制御部2、メモリ3について主に説明する。

10

【0031】

メモリ3は、制御部2が実行するプログラム、制御部2が処理する途中結果の情報、制御部2が読み書きする変換表等を記憶する。メモリ3は、例えば、ランダムアクセスメモリ(RAM)、リードオンリーメモリ(ROM)等である。メモリ3には、後述のI/O制御部、シンプロ制御部(仮想ボリューム制御部)、アクセスモニタ部、アクセス解析部、再配置制御部(移動制御部及びアクセス変換情報更新部)、ディスクモータ制御部、構成管理部、及びRAID制御部の機能を実行するための各種プログラムが格納されている。

20

【0032】

制御部2は、メモリ3に格納されるプログラムを実行することにより、図4で説明するコントローラモジュール115として機能する。同様に制御部4は、図4で説明するコントローラモジュール125として機能する。制御部2は、例えば、CPU(Central Processing Unit)である。

【0033】

図4は、第1実施形態のストレージ制御装置のストレージコントローラモジュールの機能的構成を示す処理ブロック図である。

【0034】

コントローラモジュール115、125は、サーバからのアクセス特性が類似するプールボリュームのデータ全てが同一RAIDグループに含まれるように、複数のRAIDグループのうち一つの論理ボリューム(プールボリューム)へデータの移動を行い、その移動に伴いデータを記憶した論理ボリュームと仮想ボリュームに含まれる領域との対応を書換える。

30

【0035】

コントローラモジュール115は、I/O制御部131、アクセスモニタ部133、アクセス解析部134、再配置制御部(移動制御部及びアクセス変換情報更新部)135、ディスクモータ制御部136、構成管理部137を含む。コントローラモジュール125は、I/O制御部141、アクセスモニタ部143、アクセス解析部144、再配置制御部145、ディスクモータ制御部146、及び構成管理部147を含む。

40

【0036】

コントローラモジュール115の、I/O制御部131、アクセスモニタ部133、アクセス解析部134、再配置制御部135、ディスクモータ制御部136、構成管理部137と、コントローラモジュール125の、I/O制御部141、アクセスモニタ部143、アクセス解析部144、再配置制御部145、ディスクモータ制御部146、構成管理部147とは、それぞれ同じ機能である。よって、コントローラモジュール115の各機能的構成について主に説明する。

【0037】

I/O制御部131は、RAIDグループ18a~18dにそれぞれ設定されるプール

50

ボリュームに記録されているデータに対する、サーバ11、21のアクセスを制御する。

【0038】

I/O制御部131は、RAID制御部138を備える。RAID制御部138は、複数の記憶装置にまたがって設定される少なくとも一つのプールボリュームを備えるRAIDグループ18a~18dを設定する。

【0039】

I/O制御部131は、シンプロ(Thin Provisioning)制御部(仮想ボリューム制御部)132を備える。

【0040】

シンプロ制御部132は、サーバ11、21が認識する仮想ボリュームの論理ブロックアドレス(LBA)に対して、RAIDグループ18a~18dそれぞれに設定されるプールボリュームのLBAを割り当てる。この割り当てに基づき、I/O制御部131は、仮想ボリュームのLBAへのアクセスを、プールボリュームのLBAへのアクセスに変換する。シンプロ制御部132が生成する、後述のLBA変換テーブル77を参照して、I/O制御部131は仮想ボリュームのLBAへのアクセスをプールボリュームのLBAへのアクセスに変換する。

10

【0041】

図5は、LBA変換テーブルの一例である。図5のLBA変換テーブル77は、仮想ボリュームのLBAに対するプールボリュームのLBAの割り当てを定義する。LBA変換テーブル77は、「プールボリューム」、「プールボリュームLBA」、「仮想ボリューム」、「仮想ボリュームLBA」という4つの項目で構成される。プールボリュームの欄の値について、例えばPool Volume#0-0は、RAIDグループ#0の0番目のプールボリュームを示す。(なお、本明細書において、例えば、RAIDグループ#0の0番目のプールボリュームを「プールボリューム#0-0」と表記する場合がある。)プールボリュームLBAの欄の値について、例えばPV LBA#0-0は、RAIDグループ#0の0番目のプールボリュームのLBAを示す。仮想ボリュームの欄の値について、例えば、RAIDグループ#0の0番目のプールボリュームは、Virtual Volume(仮想ボリューム)#0に割り当てられ、RAIDグループ#0の2番目のプールボリュームは、未割当であることを示す。仮想ボリュームの欄の値について、例えば、RAIDグループ#0の0番目のプールボリュームのLBAが、仮想ボリューム#0の0番目のVirtual(仮想)LBAに対応付けられていることを示す。

20

30

【0042】

また、シンプロ制御部132は、仮想ボリュームに実際に割り当てられている記憶容量が不足する場合に、プールボリュームのうち未割当のものを仮想ボリュームに割り当てる。

【0043】

アクセスモタ部133は、I/O制御部131において、サーバ11、21のプールボリュームへのアクセスを監視する。例えば、アクセスモタ部133は、サーバがアクセスした仮想ボリュームと、アクセスされた仮想ボリュームの領域に対応するプールボリュームと、そのアクセスが行われた最新の時刻とを監視する。

40

【0044】

アクセス解析部134は、アクセスモタ部133が監視したアクセス頻度を解析し、プールボリュームの中から、データの移動元となるプールボリュームと、そのプールボリュームが設定されたRAID装置とは異なるRAID装置に含まれ、データの移動先となるプールボリュームとを決定する。

【0045】

アクセス解析部134は、例えば、サーバ11、21がアクセスした仮想ボリュームと、アクセスされた仮想ボリュームの領域に対応するプールボリュームと、そのアクセスが行われた最新の時刻の情報をアクセス管理テーブルに記録する。なお、アクセスモタ部133は、アクセス管理テーブルの代わりに、後述のアクセスフラグマップを作成しても

50

よい。

【 0 0 4 6 】

図 6 は、アクセス解析部 1 3 4 が作成するアクセス管理テーブルの一例である。図 6 のアクセス管理テーブルは、「プールボリューム」「仮想ボリューム」「最新アクセス時刻」の 3 つの項目で構成される。例えば、RAID グループ # 0 の 0 番目のプールボリュームは、仮想ボリューム # 0 に割り当てられ、最新アクセス時刻は A A A A A A A であることを示す。RAID グループ # 0 の 2 番目のプールボリュームは、割り当てられていない (Non) ことを示す。RAID グループ # 1 の 1 番目のプールボリュームは、サーバから一度アクセスがあり仮想ボリューム # 2 に割り当てられている又はサーバからのアクセスはないが仮想ボリューム # 2 に割り当てられているが、サーバからのアクセスがない (0 0 0 0 0 0 0 0) ことを示す。

10

【 0 0 4 7 】

アクセス解析部 1 3 4 は、記録したプールボリューム、仮想ボリューム、アクセス時刻を解析する。解析の方法としては、例えば以下の方法が挙げられる。(I) アクセス解析部 1 3 4 は、所定の期間内にアクセスがあったかを解析し、互いに異なる RAID グループに属しサーバに割り当てられたプールボリュームのうち、所定の期間内(例えば、1 時間、1 日、1 週間など)にアクセスが存在する 2 以上のプールボリュームを選択する。(II) アクセス解析部 1 3 4 は、所定の期間内にアクセスがあったかを解析し、互いに異なる RAID グループに属しサーバに割り当てられたプールボリュームのうち、アクセスが行われていない 2 以上のプールボリュームを選択する。(I) (II) の方法が続けて行われてもよい。

20

【 0 0 4 8 】

再配置制御部 1 3 5 は、アクセス解析部 1 3 4 が選択したプールのうちのプールボリュームのデータ全てを、他のプールボリュームを含む RAID グループにおいて未割当のプールボリュームに移動(コピー)する。また、再配置制御部 1 3 5 は、データの移動に応じて、シンプロ制御部 1 3 2 の L B A 変換テーブル中の、データの移動に関係したプールボリュームの L B A と仮想ボリュームの L B A 変換テーブルを書き換える。すなわち、再配置制御部 1 3 5 は、データの移動先のプールボリュームの L B A をデータの移動元のプールボリュームに対応する仮想ボリュームの L B A に関連付ける。また、再配置制御部 1 3 5 は、データの移動元のボリュームの L B A をデータの移動先のボリュームに対応する仮想ボリュームの L B A に関連付ける。

30

【 0 0 4 9 】

図 7 ~ 8 は、プールボリュームに格納されたデータの再配置を説明する図である。

【 0 0 5 0 】

図 7 は、RAID グループに設定された複数の論理ボリュームと、それを割り当てることにより設定された仮想ボリュームとの対応関係の一例を示す模式図である。

【 0 0 5 1 】

図 7 において、仮想ボリューム # 0 ~ # n が複数の RAID グループに構築されている。例えば、仮想ボリューム # 0 に、RAID グループ # 0 の 0 番目のプールボリュームと、RAID グループ # 1 の 0 番目のプールボリュームと、RAID # 2 の 0 番目のプールボリュームとが割り当てられている。仮想ボリューム # 1 に、RAID グループ # 0 の 1 番目のプールボリュームが割り当てられている。仮想ボリューム # 2 に、RAID グループ # 1 の 1 番目のプールボリュームと、RAID # 2 の 1 番目のプールボリュームとが割り当てられている。仮想ボリューム # n に、RAID グループ # 0 の n 番目のプールボリュームと、RAID グループ # 1 の n 番目のプールボリュームとが割り当てられている。RAID グループ # 0 の 0 番目、RAID グループ # 1 の 0 番目、RAID グループ # 0 の 1 番目、及び RAID グループ # 0 の n 番目のプールボリュームに対してサーバからのアクセスがある。一方、RAID # 2 の 0 番目、RAID グループ # 1 の 1 番目、RAID # 2 の 1 番目、及び RAID グループ # 1 の n 番目のプールボリュームに対してサーバからのアクセスはない。従って、RAID グループ # 0 ~ # 2 に含まれる記憶装置に対し

40

50

てそれぞれアクセスがあるため、RAIDグループ#0～#2に含まれる記憶装置はいずれも駆動させている(Motor On)状態である。

【0052】

図8は、図7に示される複数のプールボリュームとそれに対応する仮想ボリュームに関し、プールボリュームへのアクセスの有無に基づいてデータを移動した3つのボリュームと、仮想ボリュームとの対応関係の一例を示す模式図である。この例ではRAIDグループ#0にアクセスがあるプールボリュームのデータを集め、RAID#2にアクセスのないプールボリュームのデータを集め、RAIDグループ#1に未割当のプールボリュームを集める。すなわち、RAIDグループ#1の0番目のプールボリュームのデータをRAIDグループ#0の2番目のプールボリュームに移動(コピー)させ、RAIDグループ#1の1番目のプールボリュームのデータをRAID#2の2番目のプールボリュームに移動(コピー)させ、RAIDグループ#1のn番目のプールボリュームのデータをRAID#2の3番目のプールボリュームに移動(コピー)させる。

10

【0053】

図9は、再配置制御部135によって書き換えられたLBA変換テーブル77である。上記データの移動に応じて、LBA変換テーブルを書き換え、仮想ボリュームに対するプールボリュームの割り当てを変更する。

【0054】

データの移動と、LBA変換テーブル77の書き換えにより、RAIDグループ#1、#2に含まれる記憶装置に対するアクセスが行われなくなるため、RAIDグループ#1、#2の駆動を停止(Motor Off)することができる。図8に表示される示されるRAIDグループのなかでは、アクセスがあるのはRAIDグループ#0のみであり、ディスク#0-0～#0-3のみを駆動させればよいため、ストレージ装置群の消費電力が抑制されうる。なお、仮想ボリュームに割り当てられているが所定の期間内にアクセスがないRAIDグループ#1を構成するディスク#1-0～#1-3のうちRAIDの冗長性の分だけディスクのモータの駆動を停止させてもよい。このとき、サーバからアクセスがあっても遅滞なく読み出し及び書き込み動作がなされる。

20

【0055】

再び図4を用いて、コントローラモジュール115の説明を続ける。

【0056】

ディスクモータ制御部136は、アクセスモニタ部133が記録した、サーバ11、21のプールボリュームに対するアクセス状況(アクセスされたプールボリューム、アクセスされた時刻)に基づき、RAIDグループ18a～18dに含まれる複数の記憶装置17のディスクモータを制御することで、記憶媒体の駆動(例えばディスクの回転動作)を制御する。ディスクモータ制御部136は、アクセス頻度(アクセスフラグ)が所定値よりも低い記憶媒体の駆動を停止させることで、消費電力を抑制する。例えば、ディスクモータ制御部136は、ある時間帯においてアクセス予定がないHDDのディスクモータの回転を停止させることで、消費電力を抑制する。

30

【0057】

構成管理部137は、ストレージシステムの構成を管理する。例えば、各記憶装置17の記憶媒体の駆動(ディスクの回転)の状態を保持する。

40

【0058】

図10は、第1実施形態のストレージ制御装置を用いたストレージ制御方法のフローである。

【0059】

まず、アクセスモニタ部133は、I/O制御部131において、サーバ11、21のプールボリュームへのアクセスを監視する(S101)。例えば、アクセスモニタ部133は、サーバがアクセスした仮想ボリュームと、アクセスされた仮想ボリュームの領域に対応するプールボリュームと、そのアクセスが行われた最新の時刻とを監視する。

【0060】

50

次いで、アクセス解析部 134 は、アクセスモニタ部 133 が監視したサーバがアクセスした仮想ボリュームと、アクセスされた仮想ボリュームの領域に対応するプールボリュームと、そのアクセスが行われた最新の時刻の情報を解析し、プールボリュームの中から、移動元のプールボリュームと、該移動元のプールボリュームが設定された RAID 装置とは異なる RAID 装置に設定された移動先のプールボリュームとを決定する (S102)。

【0061】

次いで、再配置制御部 135 は、アクセス解析部 134 によって決定された移動元のプールボリュームのデータを移動先のプールボリュームへ移動する (S103)。

【0062】

次いで、再配置制御部 135 は、データの再配置に応じて、シンプロ制御部 132 の LBA 変換テーブル中の、データの移動に関係したプールボリュームの LBA と仮想ボリュームの LBA 変換テーブルを書き換える (S104)。

【0063】

図 11、12 は、アクセス解析部が行う解析 (S102) と、再配置制御部が行う再配置 (S103) の一例である制御フローである。

【0064】

図 11 は、ある時点において、一定時間アクセスがない状態で駆動している一つの RAID グループへ、一定時間内にアクセスがある他の RAID グループに含まれるプールボリュームのうちアクセスがないもののデータを移動する制御フローである。

【0065】

アクセス解析部 134 は、アクセスモニタ部 133 から取得したアクセスの情報 (サーバがアクセスした仮想ボリューム、アクセスされた仮想ボリュームの領域に対応するプールボリューム、そのアクセスが行われた最新の時刻) から、アクセス管理テーブルを作成する。また、アクセス解析部 134 は、作成したアクセス管理テーブルから、記憶媒体の駆動が認識されている RAID グループのうち、一定時間アクセスがないプールボリュームと、そのプールボリュームが属する RAID グループとの 2 項目で構成されるリスト (1) を作成する。例えば、図 6 のアクセス管理テーブルにおいて、リスト (1) にリストアップされる一定時間アクセスがないプールボリュームは、Pool Volume # 1 - 1、Pool Volume # 1 - n、Pool Volume # 2 - 0、Pool Volume # 2 - 1 である。

【0066】

アクセス解析部 134 は、構成管理部 137 によって保持された、記憶媒体の駆動の状態の情報を取得する。更に、アクセス解析部 134 は、アクセスモニタ部 133 が作成したアクセス管理テーブルから、アクセスがあるプールボリュームと、そのプールボリュームが属する RAID グループとの 2 項目で構成されるリスト (2) を作成する (S201)。例えば、図 6 のアクセス管理テーブルにおいて、リスト (2) にリストアップされるアクセスがあるプールボリュームは、Pool Volume # 0 - 0、Pool Volume # 0 - 1、Pool Volume # 0 - n、Pool Volume # 1 - 0 である。

【0067】

次いで、アクセス解析部 134 は、リスト (1) とリスト (2) の両方に属する RAID グループを検索する。検索された RAID グループは、一定時間アクセスが有るプールボリュームとアクセスが無いプールボリュームとが混在しており、移動元の RAID グループの候補となる。アクセス解析部 134 は、この検索結果と、リスト (1) とから、移動元のプールボリュームと、そのプールボリュームが属する RAID グループとの 2 項目で構成されるリスト (3) を作成する (S202)。例えば、図 6 のアクセス管理テーブルから作成したリスト (1) (2) によれば、リスト (3) にリストアップされるのは、Pool Volume # 1 - 1 (RAID グループ # 1)、Pool Volume # 1 - n (RAID グループ # 1) である。

10

20

30

40

50

【0068】

リスト(3)にプールボリュームがリストアップされているとき(S203 Yes判定)、アクセス解析部134は、リスト(3)において、アクセスの無いプールボリュームの数が最も小さいRAIDグループを検索し、それに属するプールボリュームを移動元のプールボリュームに決定する(S204)。こうして決定されたプールボリュームのデータを移動すると、データの移動量が最も少ない点から好ましい。

【0069】

次いで、アクセス解析部134は、リスト(2)にリストアップされておらず、且つ、シンプロ制御部132が作成したLBA変換テーブルを参照して、移動元に記憶されたデータを移動でき仮想ボリュームに未割当のプールボリュームが存在するRAIDグループが存在するかどうか判断する(S205)。リスト(2)にリストアップされておらず、且つ、移動元に記憶されたデータを移動できる未割当のプールボリュームが存在するRAIDグループに属する未割当のプールボリュームは、移動先のプールボリュームの候補となる。一定時間内にアクセスのないデータを特定のRAIDグループに集約するためである。

10

【0070】

リスト(2)にリストアップされておらず、且つ、移動元に記憶されたデータを移動できる未割当のプールボリュームが存在するRAIDグループが存在する場合(S205 Yes判定)、再配置制御部135は、ステップ204で決定した移動元のプールボリュームのデータを、S205で判断した未割当のプールボリューム(移動先)へコピーする(S206)。次いで、再配置制御部135は、データの移動が完了した移動元プールボリュームと、それが属するRAIDグループのリストを削除する(S207)。

20

【0071】

リスト(2)にリストアップされておらず、且つ、移動元に記憶されたデータを移動できる未割当のプールボリュームが存在するRAIDグループが存在しない場合(S205 No判定)、再配置制御部135は、移動を行うことなく、ステップ204で移動元に決定されたプールボリュームと、それが属するRAIDグループのリストを削除する(S207)。

【0072】

次いで、リスト(3)にリストアップされたプールボリュームが残っている場合は(S203 Yes判定)、アクセス解析部134及び再配置制御部135は上記ステップ204~207を繰り返し行い、残っていなければ(S203 No判定)、解析及び再配置処理が終了する。

30

【0073】

なお、図11に示される再配置において、ステップ205のYes判定の後、再配置制御部135が、再配置処理を行うため、及び再配置処理後のプールボリュームのLBAの仮想ボリュームのLBAに対する割り当てを変更するため、予めLBA変換テーブルを作成してもよい。再配置処理が完了した後、シンプロ制御部132は、既存のLBA変換テーブルを、再配置制御部135が作成したLBA変換テーブルと置換する。

【0074】

図12は、ある時点において、一定時間内にアクセスが有るRAIDグループの未割当のプールボリュームへ、一定時間内にアクセスがある他のRAIDグループのプールボリュームを移動する制御フローである。

40

【0075】

アクセス解析部134は、アクセスモニタ部133から取得したアクセスの情報(サーバがアクセスした仮想ボリューム、アクセスされた仮想ボリュームの領域に対応するプールボリューム、そのアクセスが行われた最新の時刻)から、アクセス管理テーブルを作成する。また、アクセス解析部134は、作成したアクセス管理テーブルから、記憶媒体の駆動が認識されているRAIDグループのうち、一定時間アクセスがないプールボリュームと、そのプールボリュームが属するRAIDグループとの2項目で構成されるリスト(

50

1) を作成する。アクセス解析部 134 は、構成管理部 137 によって保持された、記憶媒体の駆動の状態の情報を取得する。更に、アクセス解析部 134 は、アクセスモニタ部 133 が作成したアクセス管理テーブルから、アクセスがあるプールボリュームと、そのプールボリュームが属する RAID グループとの 2 項目で構成されるリスト (2) を作成する。(S301)

次いで、アクセス解析部 134 は、リスト (1) とリスト (2) の両方に属する RAID グループを検索する。検索された RAID グループは、一定時間アクセスが有るプールボリュームとアクセスが無いプールボリュームとが混在しており、移動元の RAID グループの候補となる。アクセス解析部 134 は、この検索結果と、リスト (2) とから、移動元のプールボリュームと、そのプールボリュームが属する RAID グループとの 2 項目で構成されるリスト (4) を作成する (S302)。例えば、図 6 のアクセス管理テーブルから作成したリスト (1) (2) によれば、リスト (4) にリストアップされるのは、Pool Volume # 1 - 0 (RAID グループ # 1) である。

【0076】

リスト (4) にプールボリュームがリストアップされているとき (S303 Yes 判定)、アクセス解析部 134 は、リスト (4) において、アクセスの有るプールボリュームの数が最も小さい RAID グループを検索し、それに属するプールボリュームを移動元のプールボリュームに決定する (S304)。決定されたプールボリュームのデータを移動すると、データの移動量が最も少ないからである。

【0077】

アクセス解析部 134 は、リスト (4) にリストアップされており、且つ、シンプロ制御部 132 が作成した LBA 変換テーブルを参照して、移動元のデータを移動できる未割当のプールボリュームが存在する RAID グループが存在するかどうか判断する (S305)。

【0078】

リスト (4) にリストアップされており、且つ、シンプロ制御部 132 が作成した LBA 変換テーブルを参照して、移動元のデータを移動できる未割当のプールボリュームが存在する RAID グループに属する未割当のプールボリュームは、アクセスのある移動元のデータをコピーする移動先のプールボリュームの候補となる。一定時間内にアクセスのあるデータを特定の RAID グループに集約するためである。

【0079】

リスト (4) にリストアップされており、且つ、シンプロ制御部 132 が作成した LBA 変換テーブルを参照して、移動元のデータを移動できる未割当のプールボリュームが存在する RAID グループが存在する場合 (S305 Yes 判定)、再配置制御部 135 は、ステップ 304 で決定した移動元のプールボリュームのデータを、ステップ 304 で判断した未割当のプールボリューム (移動先) へコピーする (S306)。次いで、再配置制御部 135 は、データの移動が完了した移動元プールボリュームと、それが属する RAID グループのリストを削除する (S307)。

【0080】

リスト (4) にリストアップされており、且つ、シンプロ制御部 132 が作成した LBA 変換テーブルを参照して、移動元のデータを移動できる未割当のプールボリュームが存在する RAID グループが存在しない場合 (S305 No 判定)、再配置制御部 135 は、移動を行うことなく、ステップ 304 で移動元に決定されたプールボリュームをと、それが属する RAID グループのリストを削除する (S307)。

【0081】

次いで、リスト (4) にリストアップされたプールボリュームが残っている場合は (S303 Yes 判定)、アクセス解析部 134 及び再配置制御部 135 は上記ステップ 304 ~ 307 を繰り返し行い、残っていなければ (S303 No 判定)、解析及び再配置処理が終了する。

【0082】

10

20

30

40

50

なお、図 1 2 に示される再配置において、ステップ 2 0 4 の Y e s 判定の後、再配置制御部 1 3 5 は、再配置処理を行い、再配置処理後のプールボリュームの L B A の仮想ボリュームの L B A に対する割り当てを変更するため、予め L B A 変換テーブルを作成してもよい。再配置処理が完了した後、シプロ制御部 1 3 2 は、既存の L B A 変換テーブルを、再配置制御部 1 3 5 が作成した L B A 変換テーブルと置換する。

【 0 0 8 3 】

第 1 実施形態のストレージシステムのストレージ制御装置、及びストレージ制御方法において、R A I D グループを構成する複数の記憶装置にまたがって設定される複数のプールボリュームへのアクセスが監視され、それが解析され、アクセス頻度が類似するプールボリュームのデータが複数の R A I D グループのうち一つの R A I D グループへ移動され、移動に応じて L B A 変換テーブルが書き換えられる。これによりアクセス頻度が高いプールボリュームを有する R A I D グループを構成する複数の記憶装置を駆動させ、アクセス頻度が低いプールボリュームを有する R A I D グループを構成する複数の記憶装置を停止させることができる。ゆえに、ストレージ装置群の消費電力を抑制できる。シプロビジョニングにおいて、仮想ボリュームに対するプールボリュームの割り当て量が増加したときに、第 1 実施形態のストレージ制御装置及びストレージ制御方法によれば、アクセスが頻繁に行われるデータの量に対してモータオン状態のストレージ装置の数が減少するため、ストレージ装置群の消費電力を効果的に抑制できる。また、R A I D グループ間のデータの移動はプールボリューム単位で行われるため、サーバ側に移動のための特別な機能を準備しなくてもよい。また、データの移動に応じて L B A 変換テーブルを書き換えるため、データの移動はサーバ側が認識する仮想ボリュームに影響を与えない。

10

20

【 0 0 8 4 】

なお、第 1 実施形態において、記憶装置 1 7 が、H D D の代わりにソリッドステートディスク (S S D) であってもよい。この場合、例えば、ディスクモータ制御部 1 3 6 は、ある時間帯においてアクセスがない S S D に対する電源の供給を停止する。

【 0 0 8 5 】

以下、アクセス解析部による解析方法及び再配置制御部による再配置方法に関する第 2 実施形態について図 1 3 ~ 1 7 を用いて説明する。第 2 実施形態において、第 1 実施形態と共通するストレージシステムの構成については同一の番号を付し、説明を省略する。

【 0 0 8 6 】

第 2 実施形態において、アクセス解析部はアクセスフラグマップを作成し、それを用いてデータの移動元のプールボリュームを決定する。

30

【 0 0 8 7 】

図 1 3 は、アクセスモニタ部 1 3 3 がモニタリングした情報に基づき、アクセス解析部 1 3 4 が作成したあるプールボリュームのアクセスフラグマップである。図 1 3 において、横方向は時間を示し、所定の時間 T 内にサーバからプールボリュームにアクセスがあった場合にフラグ値を 1 に、アクセスが無かった場合にフラグ値を 0 にするように設定されている。所定の時間 T は、例えば 5 分間、1 時間、2 4 時間、1 週間など、任意に設定されうる。

【 0 0 8 8 】

図 1 4 A は、R A I D グループ # 0 及び # 1 の各プールボリュームに対してアクセスが有るか否かを表示するアクセスフラグマップの一例である。このアクセスフラグマップにおいて、1 時間単位で、現在 (1 7 時) からさかのぼって 2 3 時間前までの各プールボリュームへのアクセスの有無が記されている。アクセスがあれば 1、アクセスが無ければ 0 が記されている。各アクセスフラグマップの左側の項目は、左から順に、プールボリューム、割り当てられた仮想ボリューム、及びグルーピングである。仮想ボリューム及びグルーピングの項目において # N / A は未割当を意味する。

40

【 0 0 8 9 】

また、アクセスフラグマップの上には、各 R A I D グループ単位で重複度が記されている。重複度は、各時間における、各 R A I D グループに属する全てのプールボリュームの

50

フラグの合計値である。重複度が高い時間帯は、RAIDを構成する記憶装置へのアクセス頻度が高いことを示唆する。

【0090】

アクセス解析部134は、このアクセスフラグマップから、異なるRAIDグループに属し、アクセス特性が類似するプールボリュームを選択する。アクセス特性が類似するプールボリュームは、例えば、互いに異なるRAIDグループに属し、ある時間帯にいずれもアクセスがあるプールボリュームである。また、アクセス特性が類似するプールボリュームは、例えば、互いに異なるRAIDグループに属し、ある時間帯にいずれもアクセスがないプールボリュームである。アクセス解析部134は、仮に同一のRAIDグループに属することになれば現状よりも記憶装置の駆動時間が短くなるようにデータの移動元のプールボリュームを選択する。

10

【0091】

図14Aにおいて、9～11時の3時間、プールボリューム#0-0及びプールボリューム#1-1にアクセスがある。アクセス解析部134は、プールボリューム#0-0及びプールボリューム#1-1をデータの移動元候補とする(図14A参照)。

【0092】

続いて、データの移動元候補の2つのプールボリュームのうち一方に含まれる全データを、そのプールボリュームの属さない他のRAIDグループに含まれる未割当のプールボリュームへ移動する。アクセス解析部134は、どちらのプールボリュームをデータの移動元とするかを決定する。移設元のデータを他のRAIDグループに含まれる未割当のプールボリュームへ移動する際に、仮に一つのRAIDグループに属するプールボリュームを移設元のプールボリュームとする場合と、仮に前記他のRAIDグループに属するプールボリュームのデータを移設元のプールボリュームとする場合とを比較して、消費電力がより抑制される場合のプールボリュームを移設元のプールボリュームとして決定する。

20

【0093】

仮に、プールボリューム#0-0の全データをRAIDグループ#1の未割当のプールボリュームに移動した場合、RAIDグループ#0のモータオフ時間は4時間から11時間へと7時間増加する一方、RAIDグループ#1のモータオフ時間は4時間から1時間へと5時間減少する。また、仮に、プールボリューム#1-1の全データをRAIDグループ#1からRAIDグループ#0の未割当のプールボリュームに移動した場合、RAIDグループ#0のモータオフ時間は4時間のままである一方、RAIDグループ#1のモータオフ時間は6時間から10時間へと4時間増加する。RAIDグループ#0を構成する記憶装置の消費電力と、RAIDグループ#1を構成する記憶装置のそれとが同じであるとき、RAIDグループ#1に属するプールボリューム#1-1のデータをRAIDグループ#0の未割当のプールボリュームのデータに移動するほうが、より消費電力を抑えられる。

30

【0094】

データの移動元が決定されたら、アクセス制御部134はデータの移動先のRAIDグループに未割当のプールボリュームがあるか否かを判断する。未割当のプールボリュームがあれば、再配置制御部135は、プールボリューム#1-1のデータをRAIDグループ#0の未割当のプールボリューム#0-3に移動し、その後、プールボリューム#0-0及びプールボリューム#0-3のグルーピングの値をグループ#0としてグルーピングする(図14B参照)。グルーピングされたプールボリュームは、アクセスの頻度、アクセスがあった時間帯などが類似している点からグルーピングされる。既にグルーピングされたプールボリュームは、原則、再度のデータの移動元候補として選択されないようにする。データの移動元であるプールボリューム毎に移動先のプールボリュームを決定し、データの移動を行うという1回の処理の中で、移動先に移動させたデータを再び移動させるという、無駄な処理を行わないようにするためである。但し、時間が経過するにつれてアクセス特性が変化した場合、グルーピングしている意味がなくなることもある。その場合においては、プールボリュームのグルーピングを解除し、新たにプールボリュームのグル

40

50

ーピングを組みなおしてもよい。

【 0 0 9 5 】

また更に、18～24時の6時間、プールボリューム#0-2及びプールボリューム#1-2にアクセスがある。アクセス解析部134は、プールボリューム#0-2及びプールボリューム#1-2をデータの移動元候補とする(図15A参照)。

【 0 0 9 6 】

続いて、データの移動元候補の2つのプールボリュームのうち一方に含まれる全データを、そのプールボリュームの属さない他のRAIDグループに含まれる未割当のプールボリュームへ移動する。アクセス解析部134は、どちらのプールボリュームをデータの移動元とするかを判断する。仮に、プールボリューム#0-2の全データをRAIDグループ#1の未割当のプールボリュームに移動した場合、RAIDグループ#0のモータオフ時間は4時間から10時間へと6時間増加する。一方、RAIDグループ#1のモータオフ時間は10時間のままである。また、仮に、プールボリューム#1-2の全データをRAIDグループ#1からRAIDグループ#0の図示されない未割当のプールボリュームに移動した場合、RAIDグループ#0のモータオフ時間は4時間から2時間へと2時間減少する。一方、RAIDグループ#1のモータオフ時間は10時間から16時間へと6時間増加する。RAIDグループ#0を構成する記憶装置の消費電力と、RAIDグループ#1を構成する記憶装置のそれとが同じであるとき、RAIDグループ#1に属するプールボリューム#0-2をRAIDグループ#1の未割当のプールボリュームに移動するほうが、より消費電力を抑えられる。

10

20

【 0 0 9 7 】

再配置制御部135は、プールボリューム#0-2をRAIDグループ#1の未割当のプールボリューム#1-3に移動し、その後、プールボリューム#1-2及びプールボリューム#1-3のグルーピングの値をグループ#1としてグルーピングする(図15B参照)。

【 0 0 9 8 】

図16～17は、第2実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析、及び解析結果に基づく再配置のフローチャートである。

【 0 0 9 9 】

まず、アクセス解析部134は、現在(17時)において、アクセスが完了したプールボリュームのリスト(1)と、アクセスを開始するプールボリュームのリスト(2)を作成する(S401)。図14Aのアクセスフラグマップによれば、リスト(1)にプールボリューム#0-0がリストアップされ、リスト(2)にプールボリューム#0-2、#1-2がリストアップされる。本工程で作成されるリスト(1)(2)に挙げられるプールボリュームは、データの移動元候補のプールボリュームとなるが、データの移動元候補の組み合わせの選定の方法はこれに限られるものではない。

30

【 0 1 0 0 】

次いで、アクセス解析部134は、リスト(1)にプールボリュームがあるか否かを判定する(S402)。リスト(1)にプールボリュームがあれば(S402 Yes判定)、アクセス解析部134は、リスト(1)のプールボリュームと同じRAIDグループに属すれば、連続モータオフ時間が長くなる、他のRAIDに属するプールボリュームを検索する(S403)。具体的には、リスト(1)のプールボリュームのアクセスフラグと他のプールボリュームのアクセスフラグとで論理和(OR)をとる。論理和をとった結果、どちらの値も0となる連続時間が最も長いプールボリュームを検索結果とする。

40

【 0 1 0 1 】

ステップ403の検索結果のプールボリュームがあれば(S404 Yes判定)、アクセス解析部134はデータの移動元のプールボリュームを決定する(S405)。リスト(1)のプールボリュームのうちステップ403で検索したもの及び検索結果のプールボリュームが、図14Aを用いて説明したデータの移動元候補のプールボリュームである。本工程において、アクセス解析部134は、データの移動元の候補のプールボリュームの

50

うち、どちらをデータの移動元とするかを判断する。

【0102】

次いで、アクセス解析部134は、データの移動先のプールボリュームの有無を判断する。データの移動先のプールボリュームがあれば(S406 Yes判定)、再配置制御部135はデータの移動元のプールボリュームの全データを移動先のプールボリュームへと移動(コピー)する(S407)。その後、アクセス解析部134は、リスト(1)のプールボリュームと検索結果のプールボリュームとをグルーピングする(S408)。また、アクセス解析部134は、データの移動に関わったプールボリュームをリスト(1)から削除する(S409)。図14A、14Bで説明したアクセスフラグマップを用いた解析例においては、プールボリューム#0-2及び#1-2が削除される。

10

【0103】

ステップ403の検索結果のプールボリュームがなければ(S404 No判定)、又はステップ406のデータの移動先のプールボリュームがなければ(S406 No判定)、アクセス解析部134は、データの移動に関わったプールボリュームをリスト(1)から削除する(S409)。

【0104】

ステップ409の後、リスト(1)にプールボリュームが残っていれば(S402 Yes判定)、ステップ403~409を繰り返し行う。リスト(1)にプールボリュームが残っていなければ(S402 No判定)、リスト(2)のプールボリュームに関してデータの移動の検討を行う(S501へ)。

20

【0105】

アクセス解析部134は、リスト(2)にプールボリュームがあるか否かを判定する(S501)。リスト(2)にプールボリュームがあれば(S501 Yes判定)、ステップ403と同様、アクセス解析部134は、リスト(2)のプールボリュームと同じRAIDに属すれば連続モータオフ時間が長くなる、他のRAIDに属するプールボリュームを検索する(S502)。

【0106】

ステップ502の検索結果のプールボリュームがあれば(S503 Yes判定)、アクセス解析部134はデータの移動元のプールボリュームを決定する(S504)。リスト(2)のプールボリュームのうちステップ502で検索したもの、及び検索結果のプールボリュームが、図15Aを用いて説明したデータの移動元候補のプールボリュームである。アクセス解析部134は、どちらのプールボリュームをデータの移動元とするかを判断する。

30

【0107】

次いで、アクセス解析部134は、データの移動先のプールボリュームの有無を判断する。データの移動先のプールボリュームがあれば(S505 Yes判定)、再配置制御部135はデータの移動元のプールボリュームの全データを移動先のプールボリュームへと移動(コピー)する(S506)。その後、アクセス解析部134は、リスト(2)のプールボリュームと検索結果のプールボリュームとをグルーピングする(S507)。また、アクセス解析部134は、データの移動に関わったプールボリュームをリスト(2)から削除する(S508)。図14A~15Bで説明したアクセスフラグマップを用いた解析例においては、プールボリューム#0-2及び#1-2が削除される。

40

【0108】

ステップ502の検索結果のプールボリュームがなければ(S503 No判定)、又はステップ505のデータの移動先のプールボリュームがなければ(S505 No判定)、アクセス解析部134は、データの移動に関わったプールボリュームをリスト(2)から削除する(S508)。

【0109】

ステップ508の後、リスト(2)にプールボリュームが残っていれば(S501 Yes判定)、ステップ502~508を繰り返し行う。リスト(2)にプールボリュームが

50

残っていなければ (S 5 0 1 N o 判定)、アクセスの解析及び再配置は終了する。

【 0 1 1 0 】

なお、上記実施形態において、アクセスフラグマップを作成することにより、過去のアクセス特性に基づきデータの移動が行われているが、サーバからプールボリュームに対して予期しないアクセスがある場合もある。その場合には、例えば、I/O制御部がサーバに対して busy を返す。或いは例えば、I/O制御部が記憶装置のディスクが回転してアクセス可能な状態になるまで待ってから、記憶装置の記憶領域に対してアクセスする。

【 0 1 1 1 】

また、データの移動 (コピー) 中にサーバから移動元のプールボリュームに対して読み込みのアクセスがあるとき、サーバによる記憶領域に記録されたデータの読み込みは、例えば、I/O制御部が移動前のプールボリュームに対して直接アクセスすることによって行われる。また、データの移動 (コピー) 中にサーバから移動元のプールボリュームに対して書き込みのアクセスがあるとき、サーバによる記憶領域へのデータの書き込みは次のように行われる。例えば、I/O制御部がメモリ 2 に対して一時的に書き込むデータを保存し、データの移動が終わった後で、一時的に保存されたデータを移動先のプールボリュームに対して書き込む。

10

【 0 1 1 2 】

第 2 実施形態の解析方法を用いたストレージ制御方法によれば、第 1 実施形態のストレージ制御方法と同様に、ストレージ装置の消費電力を抑制できる。シンプロビジョニングにおいて、仮想ボリュームに対するプールボリュームの割り当て量が増加したときに、ストレージ装置群の消費電力を効果的に抑制できる。また、サーバ側にデータの移動のための特別な機能を準備しなくてもよい。また、データの移動はサーバ側が認識するボリュームに影響を与えない。更に、第 2 実施形態においては、時間ごとにプールボリュームへのアクセスの有無を監視することにより、消費電力を抑制するためのデータの移動をより精密に行うことができる。ゆえに、更に消費電力を抑制することができる。

20

【 0 1 1 3 】

以下、アクセス解析部による解析方法及び再配置制御部による再配置方法に関する第 3 実施形態について図 1 8 ~ 2 1 を用いて説明する。第 3 実施形態において、第 2 実施形態と共通するストレージシステムの構成については同一の番号を付し、説明を省略する。

【 0 1 1 4 】

第 3 実施形態は、第 2 実施形態のアクセス解析部による解析方法において、データの移動元のプールボリュームを決定する方法 (S 4 0 5 及び S 5 0 5) が異なる点を除き、第 2 実施形態の解析及び再配置の方法と同様である。以下、データの移動元のプールボリュームを決定する方法を説明する。

30

【 0 1 1 5 】

図 1 8 ~ 2 0 は、第 3 実施形態に関し、アクセス解析部が作成したアクセスフラグマップを用いて行う解析、及び解析結果に基づく再配置のフローチャートの一例である。

【 0 1 1 6 】

第 3 実施形態において、第 2 実施形態におけるリスト (1) 及び (2) に挙げられた各プールボリュームを「プールボリューム 1」と、リスト (1) 及び (2) のプールボリュームがそれぞれ属する RAID を「RAID (1)」と総称する。また、この「RAID (1)」以外の RAID グループを「RAID (2)」と総称し、RAID (2) に属する各プールボリュームを「プールボリューム 2」と総称する。

40

【 0 1 1 7 】

図 1 4 A に示されるアクセスフラグマップの解析において、リスト (1) にはプールボリューム # 0 - 0、リスト (2) にはプールボリューム # 0 - 2、# 1 - 2 がリストアップされ、これらはプールボリューム 1 と総称される。

【 0 1 1 8 】

例えば、プールボリューム # 0 - 0 が属する RAID グループ # 0 が RAID (1) であるとき、この RAID (1) に対応する RAID (2) は、RAID グループ # 0 以外

50

の R A I D グループ、すなわち R A I D グループ # 1 である。また、R A I D グループ # 1 に属する各プールボリュームがプールボリューム 2 である。

【 0 1 1 9 】

また、例えば、プールボリューム # 1 - 2 が属する R A I D グループ # 1 が R A I D (1) であるとき、この R A I D (1) に対応する R A I D (2) は、R A I D グループ # 1 以外の R A I D グループ、すなわち R A I D グループ # 0 である。また、R A I D グループ # 0 に属する各プールボリュームがプールボリューム 2 である。

【 0 1 2 0 】

まず、アクセス解析部 1 3 4 は、リスト (1) 又は (2) に挙げられたプールボリューム群の中から選択される 1 のプールボリューム 1 のアクセスフラグ値と、R A I D (1) 以外の R A I D に属する各プールボリュームから選択される 1 つのプールボリューム 2 のアクセスフラグ値との各時間における論理和 (O R) を算出する。算出した論理和を仮グルーピングアクセス特性と呼ぶ。アクセス解析部 1 3 4 は、更に、仮グルーピングアクセス特性から、仮にプールボリューム 1 と 2 とが同一 R A I D に属すると仮定するならば、連続してモータをオフにすることができる (すなわち、プールボリューム 1 にも 2 にも連続してアクセスがない) 時間の最大値を算出する (S 6 0 1) 。

【 0 1 2 1 】

次いで、アクセス解析部 1 3 4 は、ステップ 6 0 1 で算出した連続モータオフ時間の最大値が、グルーピング最大連続モータオフ時間よりも長いかなかを判断する (S 6 0 2)

。

【 0 1 2 2 】

グルーピング最大連続モータオフ時間は、プールボリューム 1 と各プールボリューム 2 との間で既に算出された連続モータオフ時間の最大値のうち、最大のものを意味する。グルーピング最大連続モータオフ時間の初期値は 0 である。

【 0 1 2 3 】

最大連続モータオフ時間のほうが長くないとき (S 6 0 2 N o 判定)、プールボリューム 1 と他のプールボリューム 2 との間での仮グルーピングアクセス特性と、モータオフ時間の最大値の算出とが繰り返し行われる (S 6 0 1 ~ 6 0 2) 。

【 0 1 2 4 】

最大連続モータオフ時間のほうが長いとき (S 6 0 2 Y e s 判定)、アクセス解析部 1 3 4 は、もしプールボリューム 1 を R A I D (2) へ移動するならば、R A I D (1) と R A I D (2) それぞれのアクセスフラグがどのように変化するかを算出する。すなわち、アクセス解析部 1 3 4 は、プールボリューム 1 のアクセスフラグを除く R A I D (1) の重複度を、仮重複アクセス特性 1 - 1 として算出し、また、その重複度の最大値を最大重複度 1 - 1 とする。また、アクセス解析部 1 3 4 は、プールボリューム 1 のアクセスフラグを加えた R A I D (2) の重複度を仮重複アクセス特性 1 - 2 として算出し、また、その重複度の最大値を最大重複度 1 - 2 とする (S 6 0 3) 。

【 0 1 2 5 】

図 2 1 A は、図 1 4 A に示されるアクセスフラグマップの解析において、プールボリューム # 0 - 0 を R A I D グループ # 1 へ移動すると仮定した場合の R A I D グループ # 0 と R A I D グループ # 1 の重複度 (仮重複アクセス特性) である。最大重複度 1 - 1 は、5 ~ 8 時における値の 2 であり、最大重複度 1 - 2 は、1 ~ 2 時及び 9 ~ 1 1 時における値の 2 である。

【 0 1 2 6 】

次いで、アクセス解析部 1 3 4 は、もしプールボリューム 2 を R A I D (1) へ移動するならば、R A I D (1) と R A I D (2) それぞれのアクセスフラグがどのように変化するかを算出する。すなわち、アクセス解析部 1 3 4 は、プールボリューム 2 のアクセスフラグを加えた R A I D (1) の重複度を仮重複アクセス特性 2 - 1 として算出し、また、その重複度の最大値を最大重複度 2 - 1 とする。また、アクセス解析部 1 3 4 は、プールボリューム 2 のアクセスフラグを除いた R A I D (2) の重複度を仮重複アクセス特性

10

20

30

40

50

2 - 2として算出し、また、その重複度の最大値を最大重複度2 - 2とする(S 6 0 4)。

【0 1 2 7】

図2 1 Bは、図1 4 Aに示されるアクセスフラグマップの解析において、プールボリューム# 1 - 1をRAIDグループ# 0へ移動すると仮定した場合のRAIDグループ# 0とRAIDグループ# 1の重複度(仮重複アクセス特性)である。最大重複度2 - 1は、9時における値の4であり、最大重複度2 - 2は、1 ~ 8時及び1 8 ~ 2 4時における値の1である。

【0 1 2 8】

次いで、アクセス解析部1 3 4は、(A)最大重複度1 - 1又は1 - 2が閾値a以上である時間帯が存在するか否か、或いは(B)仮重複アクセス特性1 - 1又は1 - 2において閾値b以上の時間帯が閾値c時間以上連続して存在するか否かを判定する(S 6 0 5)。但し、閾値 $a > b$ である。最大重複度の閾値a、仮重複アクセス特性の閾値b、及び時間の閾値cは、RAID装置の構成、プールボリュームの設定、ストレージシステムの使用態様等によって適宜設定される。

10

【0 1 2 9】

ステップ6 0 5において(A)及び(B)のいずれも満たさない場合(S 6 0 5 N o 判定)、アクセス解析部1 3 4は、仮重複アクセス特性1 - 1からモータオフ時間を仮モータオフ時間1 - 1として算出し、仮重複アクセス特性1 - 2からモータオフ期間を仮モータオフ時間1 - 2として算出する(S 6 0 6)。

20

【0 1 3 0】

ステップ6 0 5において(A)又は(B)を満たす場合(S 6 0 5 Y e s 判定)、アクセス解析部1 3 4は、仮モータオフ時間1 - 1及び1 - 2をゼロ時間(無効)とする(S 6 0 7)。

【0 1 3 1】

次いで、アクセス解析部1 3 4は、(A)最大重複度2 - 1又は2 - 2が閾値a以上である時間帯が存在するか否か、或いは(B)仮重複アクセス特性2 - 1又は2 - 2において閾値b以上の時間帯が閾値c時間以上連続して存在するか否かを判定する(S 6 0 8)。但し、 $a > b$ である。

【0 1 3 2】

ステップ6 0 8において(A)及び(B)のいずれも満たさない場合(S 6 0 8 N o 判定)、アクセス解析部1 3 4は、仮重複アクセス特性2 - 1からモータオフ時間を仮モータオフ時間2 - 1として算出し、仮重複アクセス特性2 - 2からモータオフ期間を仮モータオフ時間2 - 2として算出する(S 6 0 9)。

30

【0 1 3 3】

ステップ6 0 8において(A)又は(B)を満たす場合(S 6 0 8 Y e s 判定)、アクセス解析部1 3 4は、仮モータオフ時間2 - 1及び2 - 2をゼロ時間(無効)とする(S 6 1 0)。

【0 1 3 4】

次いで、アクセス解析部1 3 4は、仮モータオフ時間1 - 1、1 - 2、2 - 1、2 - 2はゼロ以外か否かを判断する(S 7 0 1)。仮モータオフ時間がいずれもゼロの場合(S 7 0 1 Y e s 判定)、ステップ6 0 1に戻る。

40

【0 1 3 5】

仮モータオフ時間のいずれか一つでもゼロではない場合(S 7 0 1 Y e s 判定)、アクセス解析部1 3 4は、プールボリューム1とともにデータの移動元候補となるプールボリュームの値である「グルーピング候補プールボリューム」をプールボリューム2に設定する(S 7 0 2)。

【0 1 3 6】

次いで、アクセス解析部1 3 4は、グルーピング連続最大モータオフ時間を、ステップ6 0 1で算出された連続モータオフ時間の最大値に更新する(S 7 0 3)。

50

【 0 1 3 7 】

次に、アクセス解析部 1 3 4 は、R A I D (1) と R A I D (2) のハード構成が異なるか否かを判断する (S 7 0 4) 。比較するハード構成は、R A I D を構成する記憶装置に含まれるディスク単体の消費電力、記憶装置に含まれるディスク数などである。

【 0 1 3 8 】

ハード構成が異なる場合 (S 7 0 4 Y e s) 、アクセス解析部 1 3 4 は、仮モータオフ時間だけモータオン状態であると仮定した場合に消費する電力量 (削減電力量) を算出する (S 7 0 5 、 S 7 0 6) 。

【 0 1 3 9 】

すなわち、アクセス解析部 1 3 4 は、例えば、仮モータオフ時間 1 - 1 と、R A I D (1) を構成するディスク数と、ディスク単体の消費電力の積から削減電力量 1 - 1 を算出する。また、アクセス解析部 1 3 4 は、例えば、仮モータオフ時間 1 - 2 、R A I D (2) を構成するディスク数、ディスク単体の消費電力から削減電力量 1 - 2 を算出する。また、アクセス解析部は、削減電力量 1 - 1 と削減電力量 1 - 2 の合計を削減電力量 1 として算出する (S 7 0 5) 。

10

【 0 1 4 0 】

次いで、例えば、アクセス解析部 1 3 4 は、仮モータオフ時間 1 - 1 と、R A I D (1) を構成するディスク数と、ディスク単体の消費電力の積から削減電力量 1 - 1 を算出する。また、アクセス解析部 1 3 4 は、例えば、仮モータオフ時間 1 - 2 、R A I D (2) を構成するディスク数、ディスク単体の消費電力から削減電力量 1 - 2 を算出する。また、アクセス解析部は、削減電力量 1 - 1 と削減電力量 1 - 2 の合計を削減電力量 1 として算出する (S 7 0 6) 。

20

【 0 1 4 1 】

次いで、アクセス解析部 1 3 4 は、削減電力量 1 が削減電力量 2 より大きいかな否かを判断する (S 7 0 7) 。大きい場合 (S 7 0 7 Y e s 判定) 、アクセス解析部 1 3 4 は、データの移動元をプールボリューム 1 に、データの移動先を R A I D (2) に設定する (S 7 0 8) 。

【 0 1 4 2 】

削減電力量 1 が削減電力量 2 以下の場合 (S 7 0 7 N o 判定) 、アクセス解析部 1 3 4 は、削減電力量 2 が削減電力量 1 より大きいかな否かを判断する (S 7 0 9) 。削減電力量 2 が削減電力量 1 より大きい場合 (S 7 0 9 Y e s 判定) 、アクセス解析部 1 3 4 は、データの移動元をプールボリューム 2 に、データの移動先を R A I D (1) に設定する (S 7 1 0) 。ステップ 7 0 8 及び 7 1 0 の後、図 2 0 の “ D ” へ進む。

30

【 0 1 4 3 】

削減電力量 2 が削減電力量 1 以下の場合 (S 7 0 9 N o 判定) 、ステップ 8 0 1 へ進む。

【 0 1 4 4 】

アクセス解析部 1 3 4 は、仮モータオフ時間 1 が仮モータオフ時間 2 より長いかな否かを判断する (S 8 0 1) 。長い場合 (S 8 0 1 Y e s 判定) 、アクセス解析部 1 3 4 は、データの移動元をプールボリューム 1 に、データの移動先を R A I D (2) に設定する (S 8 0 2) 。

40

【 0 1 4 5 】

仮モータオフ時間 1 が仮モータオフ時間 2 以下の場合 (S 8 0 1 N o 判定) 、アクセス解析部 1 3 4 は、仮モータオフ時間 2 が仮モータオフ時間 1 より長いかな否かを判断する (S 8 0 3) 。長い場合 (S 8 0 3 Y e s 判定) 、アクセス解析部 1 3 4 は、データの移動元をプールボリューム 2 に、データの移動先を R A I D (1) に設定する (S 8 0 4) 。ステップ 8 0 2 及び 8 0 4 の後、図 2 0 の “ D ” へ進む。

【 0 1 4 6 】

仮モータオフ時間 2 が仮モータオフ時間 1 以下の場合 (S 8 0 3 N o 判定) 、アクセス解析部 1 3 4 は、最大重複度 1 - 2 は最大重複度 2 - 1 より小さいかな否かを判断する (S

50

805)。小さい場合(S805 Yes 判定)、アクセス解析部134は、データの移動元をプールボリューム1に、データの移動先をRAID(2)に設定する。その後、図20の“D”へ進む。

【0147】

最大重複度1-2が最大重複度2-1より小さくない場合(S805 No 判定)、アクセス解析部134は、データの移動元をプールボリューム2に、データの移動先をRAID(1)に設定する(S807)。その後、図20の“D”へ進む。

【0148】

次いで、アクセス解析部134はデータの移動元候補となりうる全てのプールボリューム2について、ステップ601~808の工程を実施したか否か判断する(S808)。未実施のプールボリュームが存在すれば(S808 No 判定)、図18の“A”に戻ってステップ601~808の工程を実施する。全て検索済みであれば(S808 Yes 判定)、一通りの解析を終了する。

【0149】

リスト(1)及び(2)にリストアップされた各プールボリュームについて、それぞれ、ステップ601~808の工程を実施する。

【0150】

第3実施形態の解析方法を用いたストレージ装置の制御方法によれば、第2実施形態のストレージ制御方法と同様に、消費電力を抑制するためのデータの移動をより精密に行うことができるため、ストレージシステムの消費電力を高度に抑制できる。シミュレーションにおいて、仮想ボリュームに対するプールボリュームの割り当て量が増加したときに、ストレージ装置群の消費電力を効果的に抑制できる。また、サーバ側に、データの移動のための特別な機能を準備しなくてもよい。また、データの移動はサーバ側が認識する仮想ボリュームに影響を与えない。更に、第3実施形態の制御方法によれば、アクセス頻度が所定の閾値以上のRAIDグループに対して、更にアクセス頻度を高めるようなデータの移動を防止することができる。アクセス頻度を所定の閾値以内とすることで、特定のRAID装置に対するアクセスの集中を防止し、アクセス速度を維持することができる。

【0151】

尚、本発明は上記実施形態に限定されるものではない。矛盾のない限りにおいて、複数の実施例を組み合わせても構わない。上記実施形態は例示であり、本発明の特許請求の範囲に記載された技術的思想と実質的に同一な構成を有し、同様な作用効果を奏するものは、いかなるものであっても本発明の技術的範囲に包含される。

【符号の説明】

【0152】

- 1 ストレージシステム
- 2、4 制御部
- 3、5 メモリ
- 11、21 サーバ
- 12、22 スイッチ
- 13a、13b、23a、23b チャンnelアダプタ
- 15、25 ストレージ制御装置
- 16、26 ディスク制御部
- 18a~18d RAIDグループ
- 17 記憶装置
- 51、52 プールボリューム群
- 61、62 プールボリューム
- 74 仮想ボリューム内の領域
- 77 LBA変換テーブル
- 81 仮想ボリューム
- 115、125 コントローラモジュール

10

20

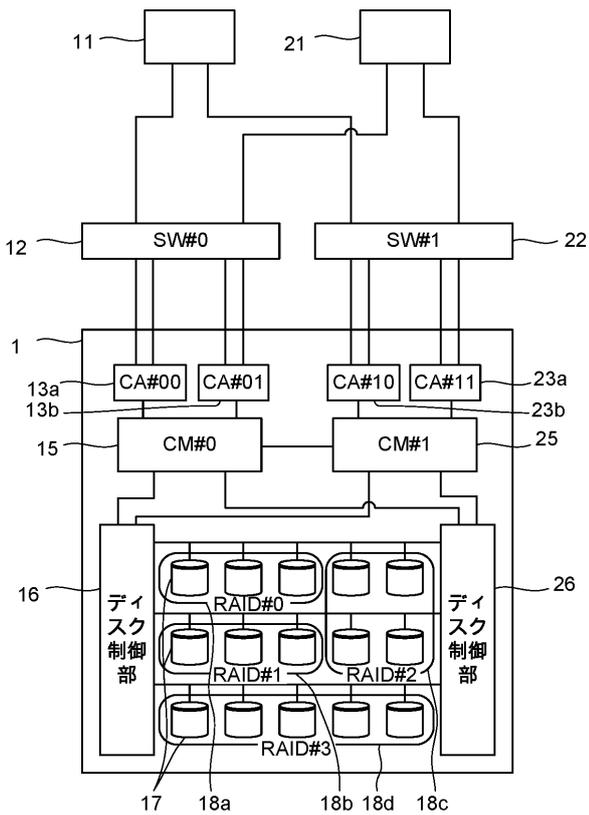
30

40

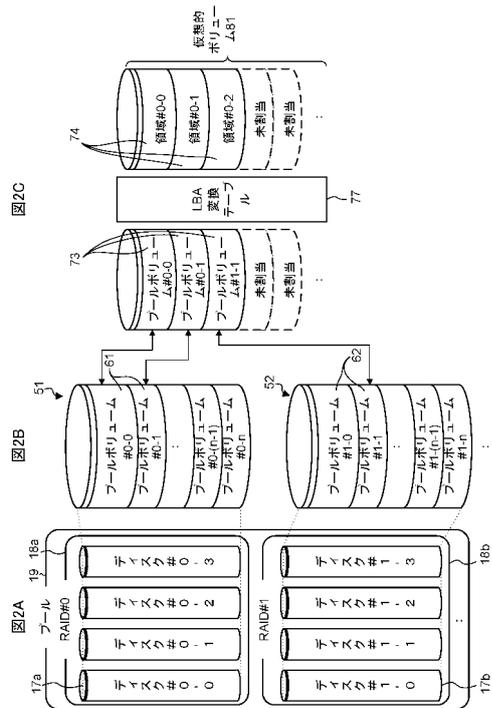
50

- 1 3 1、1 4 1 I / O 制御部
- 1 3 2、1 4 2 シンプロ制御部
- 1 3 3、1 4 3 アクセスモニタ部
- 1 3 4、1 4 4 アクセス解析部
- 1 3 5、1 4 5 再配置制御部
- 1 3 6、1 4 6 ディスクモータ制御部
- 1 3 7、1 4 7 構成管理部
- 1 3 8、1 4 8 R A I D 制御部

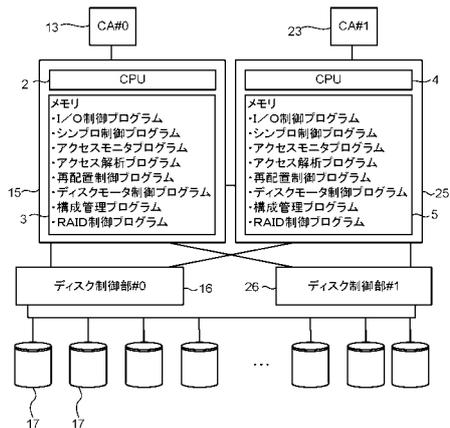
【 図 1 】



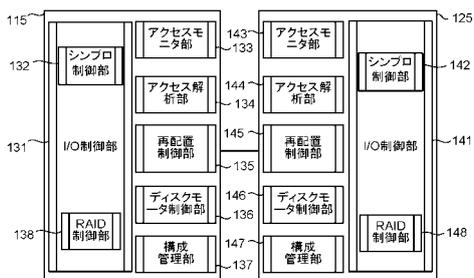
【 図 2 】



【 図 3 】



【 図 4 】



【 図 6 】

アクセス管理テーブル

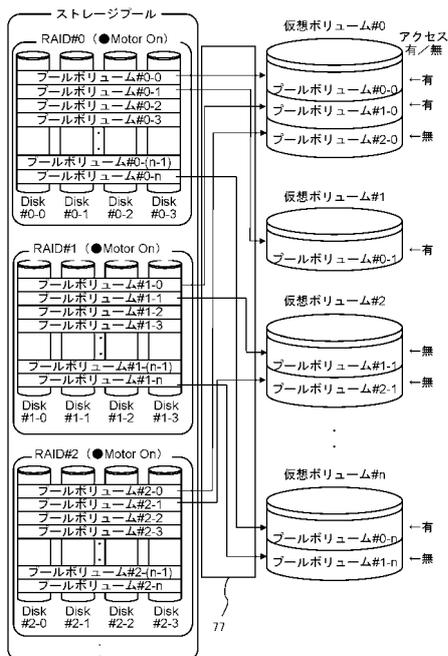
プールボリューム	仮想ボリューム	最新アクセス時刻
Pool Volume#0-0	Virtual Volume#0	AAAAAAA
Pool Volume#0-1	Virtual Volume#1	BBBBBBBB
Pool Volume#0-2	Non(未割当)	00000000
Pool Volume#0-3	Non(未割当)	00000000
:	:	:
Pool Volume#0-n	Virtual Volume#n	CGCGCGCC
Pool Volume#1-0	Virtual Volume#0	AAAAAAA
Pool Volume#1-1	Virtual Volume#2	00000000 ←アクセスなし
Pool Volume#1-2	Non(未割当)	00000000
Pool Volume#1-3	Non(未割当)	00000000
:	:	:
Pool Volume#1-n	Virtual Volume#n	00000000 ←アクセスなし
Pool Volume#2-0	Virtual Volume#0	00000000 ←アクセスなし
Pool Volume#2-1	Virtual Volume#2	00000000 ←アクセスなし
Pool Volume#2-2	Non(未割当)	00000000
Pool Volume#2-3	Non(未割当)	00000000
:	:	:
Pool Volume#2-n	Non(未割当)	00000000
:	:	:

【 図 5 】

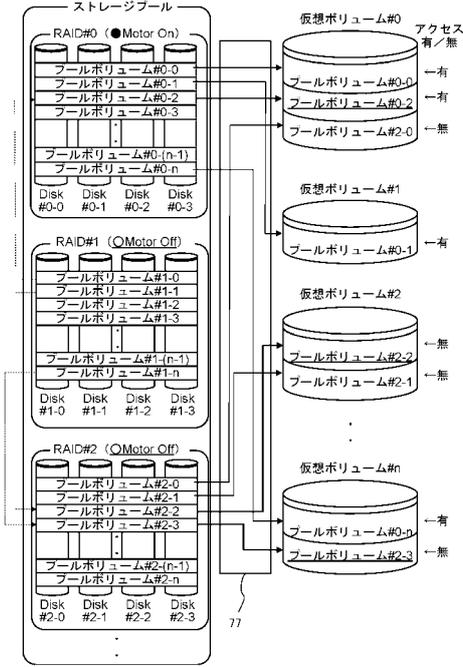
77

プールボリューム	プールボリューム LBA	仮想ボリューム	仮想LBA
Pool Volume#0-0	PV LBA#0-0	Virtual Volume#0	Virtual LBA#0-0
Pool Volume#0-1	PV LBA#0-1	Virtual Volume#1	Virtual LBA#1-0
Pool Volume#0-2	PV LBA#0-2	Non(未割当)	Non(未割当)
Pool Volume#0-3	PV LBA#0-3	Non(未割当)	Non(未割当)
:	:	:	:
Pool Volume#0-n	PV LBA#0-n	Virtual Volume#n	Virtual LBA#n-0
Pool Volume#1-0	PV LBA#1-0	Virtual Volume#0	Virtual LBA#0-1
Pool Volume#1-1	PV LBA#1-1	Virtual Volume#2	Virtual LBA#2-0
Pool Volume#1-2	PV LBA#1-2	Non(未割当)	Non(未割当)
Pool Volume#1-3	PV LBA#1-3	Non(未割当)	Non(未割当)
:	:	:	:
Pool Volume#1-n	PV LBA#1-n	Virtual Volume#n	Virtual LBA#n-1
Pool Volume#2-0	PV LBA#2-0	Virtual Volume#0	Virtual LBA#0-2
Pool Volume#2-1	PV LBA#2-1	Virtual Volume#2	Virtual LBA#2-1
Pool Volume#2-2	PV LBA#2-2	Non(未割当)	Non(未割当)
Pool Volume#2-3	PV LBA#2-3	Non(未割当)	Non(未割当)
:	:	:	:
Pool Volume#2-n	PV LBA#2-n	Non(未割当)	Non(未割当)
:	:	:	:

【 図 7 】



【図 8】

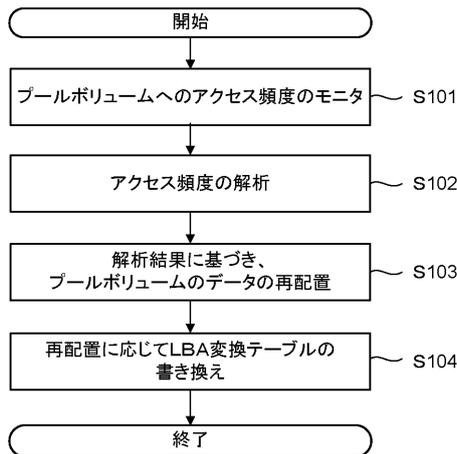


【図 9】

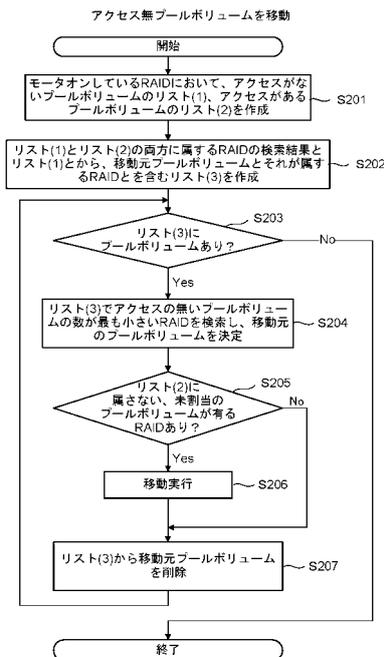
22

プールボリューム	プールボリューム LBA	仮想ボリューム	仮想LBA
Pool Volume#0-0	PV LBA#0-0	Virtual Volume#0	Virtual LBA#0-0
Pool Volume#0-1	PV LBA#0-1	Virtual Volume#1	Virtual LBA#1-0
Pool Volume#0-2	PV LBA#0-2	Virtual Volume#0	Virtual LBA#0-1
Pool Volume#0-3	PV LBA#0-3	Non(未割当)	Non(未割当)
:	:	:	:
Pool Volume#0-n	PV LBA#0-n	Virtual Volume#n	Virtual LBA#n-0
Pool Volume#1-0	PV LBA#1-0	Non(未割当)	Non(未割当)
Pool Volume#1-1	PV LBA#1-1	Non(未割当)	Non(未割当)
Pool Volume#1-2	PV LBA#1-2	Non(未割当)	Non(未割当)
Pool Volume#1-3	PV LBA#1-3	Non(未割当)	Non(未割当)
:	:	:	:
Pool Volume#1-n	PV LBA#1-n	Non(未割当)	Non(未割当)
Pool Volume#2-0	PV LBA#2-0	Virtual Volume#0	Virtual LBA#0-2
Pool Volume#2-1	PV LBA#2-1	Virtual Volume#2	Virtual LBA#2-1
Pool Volume#2-2	PV LBA#2-2	Virtual Volume#2	Virtual LBA#2-0
Pool Volume#2-3	PV LBA#2-3	Virtual Volume#n	Virtual LBA#n-1
:	:	:	:
Pool Volume#2-n	PV LBA#2-n	Non(未割当)	Non(未割当)
:	:	:	:

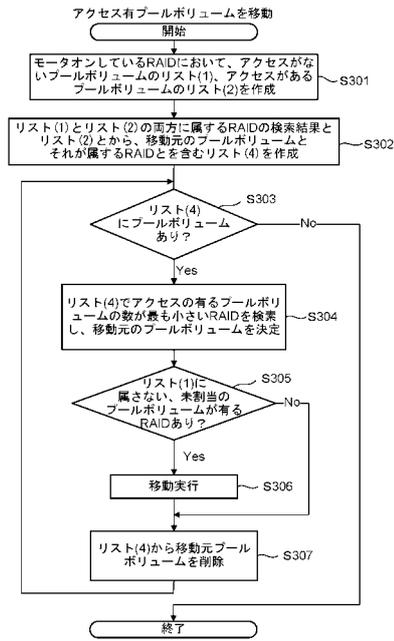
【図 10】



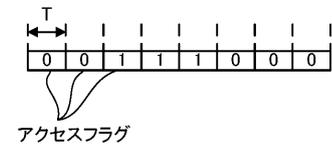
【図 11】



【 図 1 2 】



【 図 1 3 】



【 図 1 4 】

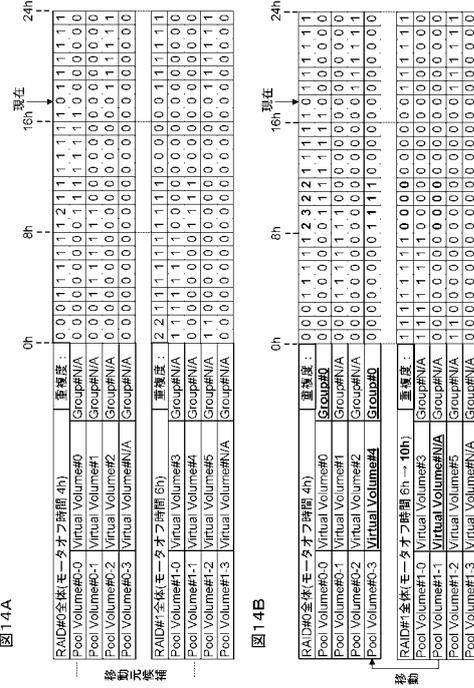


図 14A

図 14B

【 図 1 5 】

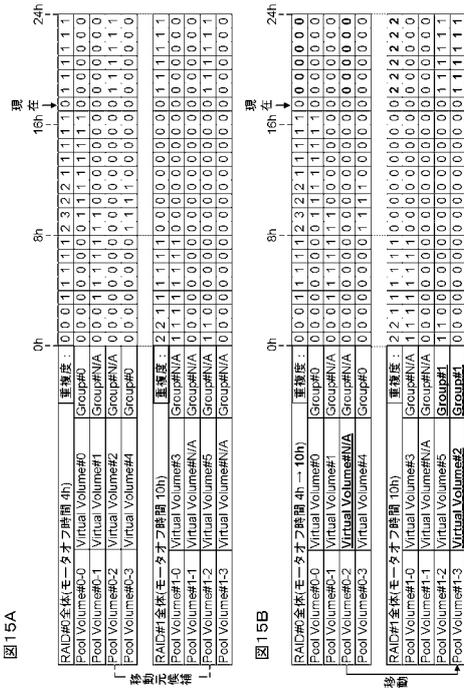
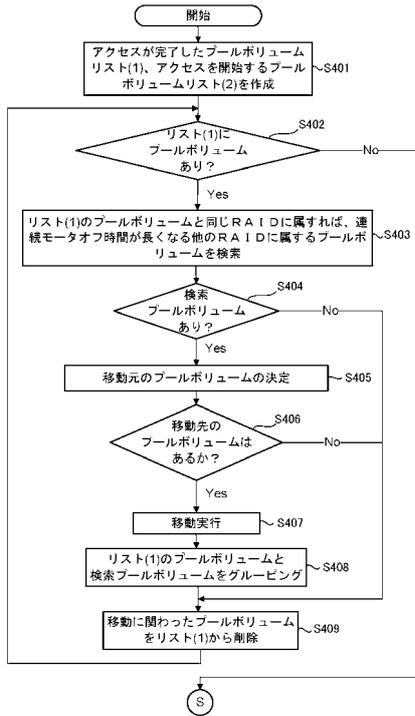


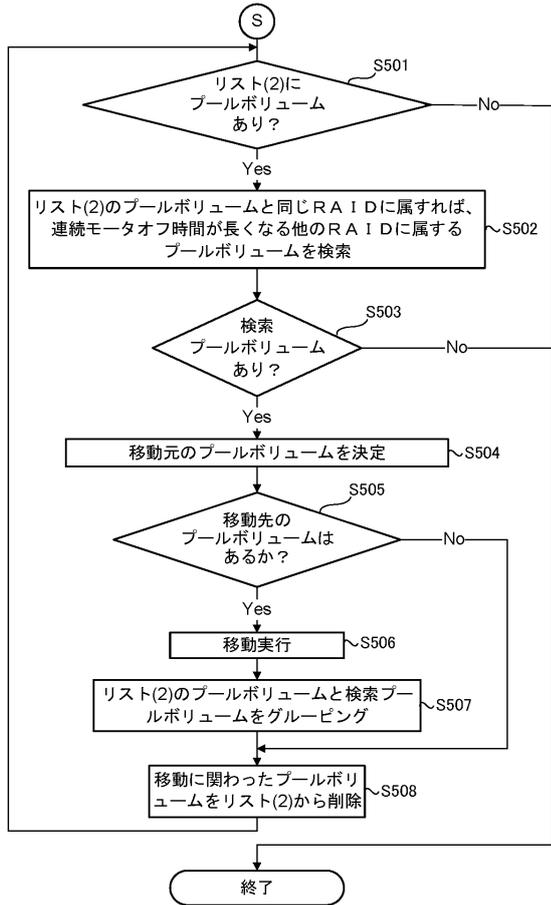
図 15A

図 15B

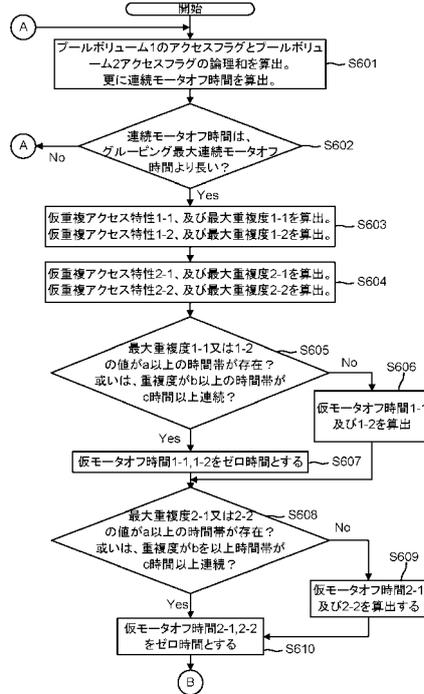
【 図 1 6 】



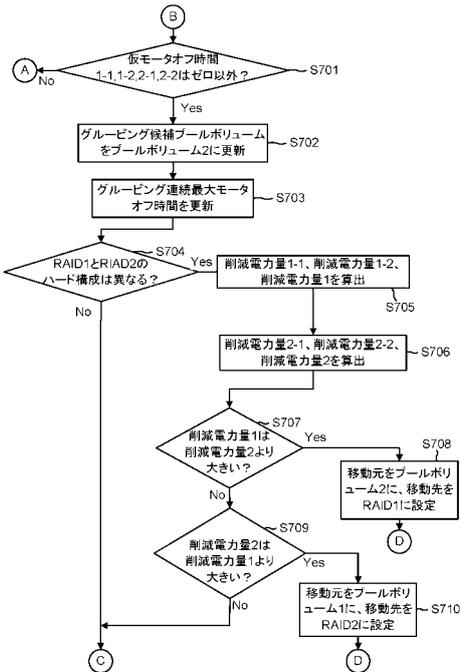
【図17】



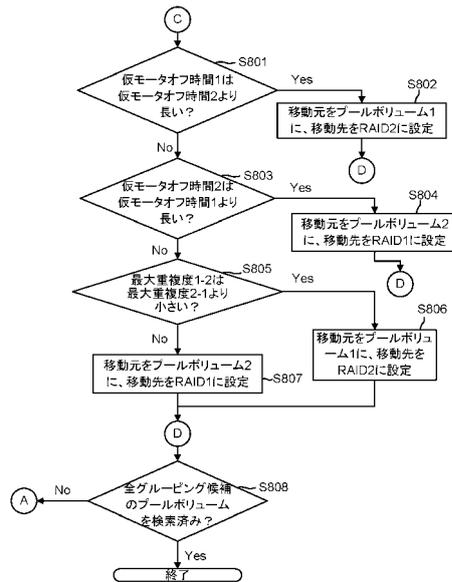
【図18】



【図19】



【図20】



フロントページの続き

Fターム(参考) 5B014 EB04

5B065 BA01 BA05 BA06 CC02 CH19 ZA14

5B082 CA05 CA11 FA06