

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5060876号
(P5060876)

(45) 発行日 平成24年10月31日(2012.10.31)

(24) 登録日 平成24年8月10日(2012.8.10)

(51) Int. Cl.		F I	
G06F 3/06	(2006.01)	G06F 3/06	
G06F 13/10	(2006.01)	G06F 3/06	540
G06F 12/00	(2006.01)	G06F 13/10	340A
		G06F 12/00	510A

請求項の数 18 (全 26 頁)

(21) 出願番号	特願2007-224743 (P2007-224743)	(73) 特許権者	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成19年8月30日(2007.8.30)	(74) 代理人	100093861 弁理士 大賀 真司
(65) 公開番号	特開2009-59096 (P2009-59096A)	(72) 発明者	森 一 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所 内
(43) 公開日	平成21年3月19日(2009.3.19)	審査官	菅原 浩二
審査請求日	平成21年12月16日(2009.12.16)		

最終頁に続く

(54) 【発明の名称】 ストレージシステム及びストレージシステムの消費電力低減方法

(57) 【特許請求の範囲】

【請求項1】

計算機と、この計算機から受信するデータを格納する複数のディスク装置及びその複数のディスク装置から構成される複数の論理ユニットを有するストレージ装置とを含むストレージシステムであって、

前記複数の論理ユニットを用いて構成されるグループ及びそのグループが起動中か否かを少なくとも含むグループ情報を格納するグループ管理テーブルと、

前記グループに含まれる論理ユニットに対応するディスク装置の電源をON/OFFする電力制御部と、

前記複数のディスク装置に格納されるデータの格納位置と対応する論理ユニットの論理ユニット番号、前記データのビット列を代表する値及び前記論理ユニットと対応する前記計算機が指定する番号を含む管理情報を格納するデータ管理テーブルと、

前記計算機からデータを受信すると、そのデータのビット列を代表する値を算出する算出部と、

この算出部で算出された値と同一の値が前記データ管理テーブルに格納されているか否かを判断する判断部と、

この判断部で同一の値が格納されていないと判断した場合は、前記受信したデータのビット列を代表する値を前記データ管理テーブルに格納するとともに前記グループ管理テーブルに基づいて起動しているグループを構成する論理ユニット番号と対応するディスク装置に前記データを格納する第1の格納部と、

10

20

前記判断部で同一の値が格納されていると判断した場合は、前記データ管理テーブルで管理される管理情報に基づいて、前記受信したデータを重複して格納することを除去する重複除去処理を実行すべきであるか否かを判断し、この判断結果に応じて、前記グループ管理テーブルで管理されているグループ情報に基づいて起動しているグループを構成する論理ユニット番号と対応する前記ディスク装置への格納の要否を制御する第2の格納部と

を備えることを特徴とするストレージシステム。

【請求項2】

データを前記ディスク装置に仮格納する論理ユニット番号及び前記データのビット列を代表する値を少なくとも含む未比較データ情報を格納する未比較データ管理テーブルを備え、

10

前記第2の格納部は、

前記受信したデータのビット列を代表する値と同一の値が前記未比較データ管理テーブルに格納されているか否かを判断する未比較判断部と、

この未比較判断部で同一の値が格納されていると判断した場合は、前記グループ管理テーブルに基づいて起動しているグループを構成する論理ユニット番号と対応するディスク装置に前記受信したデータを格納するとともに前記受信したデータを格納する前記ディスク装置と対応する論理ユニット番号及び前記受信したデータのビット列を代表する値を前記未比較データ管理テーブルに格納する未比較データ処理部と、

前記未比較判断部で同一の値が格納されていないと判断した場合は、前記データを重複して格納することを除去する重複除去処理を行う重複除去処理部とを含む、

20

ことを特徴とする請求項1記載のストレージシステム。

【請求項3】

前記重複除去処理部は、

前記受信したデータのビット列を代表する値と同一の値を持つデータが格納されているディスク装置が起動しているか否かを前記グループ管理テーブルに基づいて判断する状態判断部と、

この状態判断部で前記ディスク装置が起動中であると判断した場合は、前記ディスク装置から前記値と対応するデータを読み出し、その読み出したデータと前記受信したデータとが一致するか否かを判断するデータ判断部と、

30

このデータ判断部でデータが一致すると判断した場合は、前記データ管理テーブルの前記読み出したデータの論理ユニット番号と対応させて前記受信したデータの前記計算機が指定する番号を格納する第3の格納部と、

前記データ判断部でデータが一致しないと判断した場合は、前記読み出したグループと対応するディスク装置に前記受信したデータを格納するとともに前記受信データの管理情報を前記データ管理テーブルに格納する第4の格納部とを含む、

ことを特徴とする請求項2記載のストレージシステム。

【請求項4】

前記重複除去処理部は、

前記状態判断部で前記ディスク装置が起動中でないと判断した場合は、前記グループ管理テーブルで管理されるグループの起動ができるか否かを判断する起動判断部を含み、

40

前記起動判断部で前記グループの起動できると判断した場合に前記省電力制御部は前記他のグループと対応するディスク装置の電源をONし、

前記ディスク装置の電源をONした場合に前記データ判断部は、前記データが一致であるか否かの判断を行なうこと、

を特徴とする請求項3記載のストレージシステム。

【請求項5】

前記起動判断部は、予め設定された起動できるグループのグループ数と、現在起動しているグループ数とを比較して前記他のグループが起動できるか否かの判断を行なうこと、

を特徴とする請求項4記載のストレージシステム。

50

【請求項 6】

前記重複除去処理部は、

前記起動判断部で前記グループの起動ができないと判断した場合は、起動しているグループと対応するディスク装置に前記受信したデータを格納するとともに前記受信データの管理情報を前記データ管理テーブルに格納する第 5 の格納部を含む、

ことを特徴とする請求項 4 記載のストレージシステム。

【請求項 7】

前記グループは、RAIDグループである、

ことを特徴とする請求項 1 記載のストレージシステム。

【請求項 8】

データのビット列を代表する値、起動候補の前記グループ及び前記値のカウント数を含む起動候補情報を格納する起動候補テーブルと、

前記未比較データ管理テーブルに未比較データ情報が格納されている場合に、前記値が同一なデータを同一の前記グループとなるように再配置する最適化処理を行う最適化処理部とを備える、

ことを特徴とする請求項 2 記載のストレージシステム。

【請求項 9】

前記データ最適化処理部は、通常の処理中に前記グループの起動/停止時が発生した場合、停止中の前記グループが起動した場合、前記未比較データ管理テーブルに格納された未比較データ数が予め設定された閾値以上になった場合、ユーザ又は前記計算機によって指定された場合の少なくともいずれかを契機として処理を開始すること、

を特徴とする請求項 8 記載のストレージシステム。

【請求項 10】

計算機と、この計算機から受信するデータを格納する複数のディスク装置及びその複数のディスク装置から構成される複数の論理ユニットを有するストレージ装置とを含むストレージシステムの消費電力低減方法であって、

前記ストレージ装置は、

前記複数の論理ユニットを用いて構成されるグループ及びそのグループが起動中か否かを少なくとも含むグループ情報を格納するグループ管理テーブルと、

前記グループに含まれる論理ユニットに対応するディスク装置の電源をON/OFFする電力制御部と、

前記複数のディスク装置に格納されるデータの格納位置と対応する論理ユニットの論理ユニット番号、前記データのビット列を代表する値及び前記論理ユニットと対応する前記計算機が指定する番号を含む管理情報を格納するデータ管理テーブルとを含み、

前記計算機からデータを受信すると、そのデータのビット列を代表する値を算出するステップと、

この算出された値と同一の値が前記データ管理テーブルに格納されているか否かを判断するステップと、

この判断するステップで同一の値が格納されていないと判断した場合、前記受信したデータのビット列を代表する値を前記データ管理テーブルに格納するとともに前記グループ管理テーブルに基づいて起動しているグループを構成する論理ユニット番号と対応するディスク装置に前記データを格納する第 1 のステップと、

前記判断するステップで同一の値が格納されていると判断した場合は、前記データ管理テーブルで管理される管理情報に基づいて、前記受信したデータを重複して格納することを除去する重複除去処理を実行すべきであるか否かを判断し、この判断結果に応じて、前記グループ管理テーブルで管理されているグループ情報に基づいて起動しているグループを構成する論理ユニット番号と対応する前記ディスク装置への格納の可否を制御する第 2 のステップと、

を有することを特徴とするストレージシステムの消費電力低減方法。

【請求項 11】

10

20

30

40

50

前記ストレージ装置は、データを前記ディスク装置に仮格納する論理ユニット番号及び前記データのビット列を代表する値を少なくとも含む未比較データ情報を格納する未比較データ管理テーブルを備え、

前記第2のステップは、

前記受信したデータのビット列を代表する値と同一の値が前記未比較データ管理テーブルに格納されているか否かを判断する未比較ステップと、

この未比較ステップで同一の値が格納されていると判断した場合は、前記グループ管理テーブルに基づいて起動しているグループを構成する論理ユニット番号と対応するディスク装置に前記受信したデータを格納するとともに前記受信したデータを格納する前記ディスク装置と対応する論理ユニット番号及び前記受信したデータのビット列を代表する値を前記未比較データ管理テーブルに格納する未比較データを処理するステップと、

前記未比較ステップで同一の値が格納されていないと判断した場合は、前記データを重複して格納すること除去するステップとを含む、

ことを特徴とする請求項10記載のストレージシステムの消費電力低減方法。

【請求項12】

前記除去するステップは、

前記受信したデータのビット列を代表する値と同一の値を持つデータが格納されているディスク装置が起動しているか否かを前記グループ管理テーブルに基づいて判断する状態を判断するステップと、

この状態を判断するステップで前記ディスク装置が起動中であると判断した場合は、前記ディスク装置から前記値と対応するデータを読み出し、その読み出したデータと前記受信したデータとが一致するか否かをデータ判断するステップと、

このデータ判断するステップでデータが一致すると判断した場合は、前記データ管理テーブルの前記読み出したデータの論理ユニット番号と対応させて前記受信したデータの前記計算機が指定する番号を格納する第3のステップと、

前記データ判断するステップでデータが一致しないと判断した場合は、前記読み出したグループと対応するディスク装置に前記受信したデータを格納するとともに前記受信データの管理情報を前記データ管理テーブルに格納する第4のステップとを含む、

ことを特徴とする請求項11記載のストレージシステムの消費電力低減方法。

【請求項13】

前記除去するステップは、

前記状態を判断するステップで前記ディスク装置が起動中でないと判断した場合は、前記グループ管理テーブルで管理されるグループの起動ができるか否かを判断する起動を判断するステップを含み、

前記起動を判断するステップで前記グループの起動ができると判断した場合に前記省電力制御部は前記他のグループと対応するディスク装置の電源をONし、

前記ディスク装置の電源をONした場合に、前記データ判断するステップは前記データが一致であるか否かの判断を行なうこと、

を特徴とする請求項12記載のストレージシステムの消費電力低減方法。

【請求項14】

前記起動を判断するステップは、予め設定された起動できるグループのグループ数と、現在起動しているグループ数とを比較して前記他のグループが起動できるか否かの判断を行なうこと、

を特徴とする請求項13記載のストレージシステムの消費電力低減方法。

【請求項15】

前記除去するステップは、

前記起動を判断するステップで前記グループの起動ができないと判断した場合は、起動しているグループと対応するディスク装置に前記受信したデータを格納するとともに前記受信データの管理情報を前記データ管理テーブルに格納する第5のステップを含む、

ことを特徴とする請求項13記載のストレージシステムの消費電力低減方法。

10

20

30

40

50

【請求項 16】

前記グループは、RAIDグループである、
ことを特徴とする請求項10記載のストレージシステムの消費電力低減方法。

【請求項 17】

前記ストレージ装置は、データのビット列を代表する値、起動候補の前記グループ及び前記値のカウント数を含む起動候補情報を格納する起動候補テーブルを備え、

前記未比較データ管理テーブルに未比較データ情報が格納されている場合に、前記値が同一なデータを同一の前記グループとなるように再配置する最適化処理を行うステップを有する、

ことを特徴とする請求項11記載のストレージシステムの消費電力低減方法。

10

【請求項 18】

前記最適化処理を行うステップは、通常の処理中に前記グループの起動/停止時が発生した場合、停止中の前記グループが起動した場合、前記未比較データ管理テーブルに格納された未比較データ数が予め設定された閾値以上になった場合、ユーザ又は前記計算機によって指定された場合の少なくともいずれかを契機として処理を開始すること、

を特徴とする請求項17記載のストレージシステムの消費電力低減方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージシステム及びストレージシステムの消費電力低減方法に関し、特にディスクを用いたストレージシステム及びストレージシステムの消費電力低減方法に適用しても好適なものである。

20

【背景技術】

【0002】

業務データ等のバックアップ、アーカイブにおいて、ディスクアレイに格納するデータ量の増大を抑制しデータ容量効率を高めるため、重複するデータをディスクに格納しないデータ重複除去に関する技術（以下、データ重複除去ともいう）が公開されている。データ重複除去技術は、新たにディスクに書き込むデータ、いわゆるライトデータが、ディスクに格納済みのデータと同一内容の場合、最終的に重複するデータはディスクに書き込まない技術である。ライトデータがディスクに格納済みのデータと同一内容であるか否かの検証には、一般的にハッシュを用いた高速な検索を利用している（例えば、特許文献1参照）。

30

【0003】

一方、ディスクアレイは、多数のディスクを備えた装置である。ディスクアレイに備えられた多数のディスクを同時に稼働させると、ディスクアレイの消費電力は大きくなってしまふ。この問題の解決策として、ディスクアレイで同時に稼働可能なディスクの数を制限することにより、ディスクアレイ全体での消費電力を削減し、省電力効果をあげる技術が公開されている（例えば、特許文献2参照）。

【特許文献1】米国特許第6928526号公報

【特許文献2】米国特許第5423046号公報

40

【発明の開示】

【発明が解決しようとする課題】

【0004】

従来、ディスクを用いたストレージシステムでは、ディスクに格納されたデータの重複除去に際し、消費電力を抑制する制御を同時に実施することは考慮されていない。このためデータ重複除去を実施するディスクアレイにおいて、稼働可能なディスク数の制限による消費電力の削減を同時に実施した場合、データ重複除去によるディスクへのアクセスを考慮していないため、効率的なディスク稼働の制御ができない。その結果、無駄なディスクの稼働が必要となり、最適な省電力効果が発揮できないという問題が発生する。

【0005】

50

本発明は、以上の点を考慮してなされたもので、データ重複除去とディスク稼働の制御を同時に実施した場合、データ重複除去によるデータ容量の効率化と、ディスク稼働の制御による省電力効果を最大限に発揮することのできるストレージシステム及びストレージシステムの消費電力低減方法を提案しようとするものである。

【課題を解決するための手段】

【0006】

本発明は、計算機と、この計算機から受信するデータを格納する複数のディスク装置及びその複数のディスク装置から構成される複数の論理ユニットを有するストレージ装置を含むストレージシステムであって、前記複数の論理ユニットを用いて構成されるグループ及びそのグループが起動中か否かを少なくとも含むグループ情報を格納するグループ管理テーブルと、前記グループに含まれる論理ユニットに対応するディスク装置の電源をON/OFFする電力制御部と、前記複数のディスク装置に格納されるデータの格納位置と対応する論理ユニットの論理ユニット番号、前記データのビット列を代表する値及び前記論理ユニットと対応する前記計算機が指定する番号を含む管理情報を格納するデータ管理テーブルと、前記計算機からデータを受信すると、そのデータのビット列を代表する値を算出する算出部と、この算出部で算出された値と同一の値が前記データ管理テーブルに格納されているか否かを判断する判断部と、この判断部で同一の値が格納されていないと判断した場合は、前記受信したデータのビット列を代表する値を前記データ管理テーブルに格納するとともに前記グループ管理テーブルに基づいて起動しているグループを構成する論理ユニット番号と対応するディスク装置に前記データを格納する第1の格納部と、前記判断部で同一の値が格納されていると判断した場合は、前記データ管理テーブルで管理される管理情報に基づいて、前記受信したデータを重複して格納することを除去する重複除去処理を実行すべきであるか否かを判断し、この判断結果に応じて、前記グループ管理テーブルで管理されているグループ情報に基づいて起動しているグループを構成する論理ユニット番号と対応する前記ディスク装置への格納の要否を制御する第2の格納部とを備えるものである。

【発明の効果】

【0007】

本発明によれば、データ重複除去と起動可能な省電力グループを制限する消費電力の抑制を同時に実施した場合、データ重複除去によるデータ容量の効率化と、ディスク稼働の制御による省電力効果を最大限に発揮することができる。

【発明を実施するための最良の形態】

【0008】

以下、本発明の実施形態について図面を参照して説明する。

【0009】

第1の実施形態：

(1-1) 構成の説明

図1は、ストレージシステム1の構成の一例を示す図である。図1では業務用途等のアプリケーションが稼働する計算機であるホスト100と、ホスト100から送信されたデータを格納するディスクアレイ200は、SAN (Storage Area Network) を介して接続されている。なお、図1では計算機としてホスト100のみ記載しているが、実際には複数台の計算機を有していてもよい。

【0010】

ディスクアレイ200は、SANに接続されるファイバーチャネルインターフェイス(以下、FC/IF)250と、管理用インターフェイス(以下、管理I/F)260を有する。なお管理I/F260は、ディスクアレイ200のデータを管理サーバ300などのほかの装置とやり取りするためのインターフェイスである。FCI/F250は、ディスクアレイ200から読み出されたデータやディスクアレイ200に書き込まれるデータをホスト100などとやり取りするためのインターフェイスである。

【0011】

10

20

30

40

50

また、ディスクアレイ 200 は、ディスクアレイ 200 を制御する CPU 210 と、各種管理テーブルや各種プログラムなどを格納するメモリ 220 と、ユーザデータを格納するキャッシュ 230 を有する。さらに、ディスクアレイ 200 は複数のディスク 291 ~ 298 を制御するディスクコントローラ 240 を有する。

【0012】

ディスクコントローラ 240 には、既述のディスク 291 ~ 298 が接続されている。これらディスク群（ディスク 291 ~ 298）は、各々 RAID（Redundant Arrays of independent（inexpensive）Disc）と呼ばれる冗長構成を持つ RAID グループ 271 ~ 272 が設定されている。また、ディスク 291 ~ 298 内の記憶領域は、SCSI の論理ユニット（LU：Logical Unit）281 ~ 283 としてホスト 100 からアクセスされる。図 1 では LU 281 ~ 283 は、上記ディスク群の中の複数のディスクに跨って存在する記憶領域として示されているが、一つのディスク内の記憶領域から構成される LU でもよい。また、LU は RAID グループを跨って設定されてもよいものとする。

10

【0013】

メモリ 220 には、省電力グループ制御プログラム 221 と、省電力グループ管理テーブル 222 と、データ重複除去プログラム 223 と、重複データ管理テーブル 224 と、重複未比較データ管理テーブル 225 と、データプール管理テーブル 226 と、起動候補省電力グループ管理テーブル 227 が格納されている。これらの詳細な説明については後述する。

【0014】

（1 - 2）ディスクアレイの消費電力の抑制

この第 1 の実施形態では、省電力グループ制御プログラム 221 は、省電力グループ管理テーブル 222 を利用して、ディスクアレイ 200 に含まれるディスク 291 ~ 298 の起動及び停止を制御することで、消費電力の削減を実現する。

20

【0015】

省電力グループ制御プログラム 221 は、ディスクアレイ 200 の仕様、またはネットワーク管理者等により、RAID グループ単位や、LU 単位、またはディスク単位からなる省電力グループを予め設定されている。また、同時に稼働できる省電力グループの上限数である上限値も予め設定されている。

【0016】

図 2 で示す省電力グループ管理テーブル 222 は、省電力グループを管理するためのテーブルである。省電力グループ管理テーブル 222 は省電力識別番号欄 2221、状態欄 2222、RAID グループ番号欄 2223、LU 番号欄 2224 及びボリューム利用率欄 2225 を有している。

30

【0017】

省電力識別番号欄 2221 には省電力グループごとの識別子が格納される。状態欄 2222 には省電力識別番号ごとに省電力グループの稼働状態が格納される。RAID グループ番号欄 2223 には省電力の制御グループに内包される RAID グループ番号が格納される。LU 番号欄 2224 には LU の識別子が格納される。ボリューム利用率欄 2225 には省電力グループごとのボリュームの利用率が格納される。

40

【0018】

省電力管理テーブル 222 には、例えば、図 2 に示すように、省電力識別番号欄 2221 に格納される省電力識別番号“P - 01”に対応して、状態欄 2222 には状態として“稼働中”、RAID グループ番号欄 2223 には RAID グループ番号として“RG01”、LU 番号欄 2224 には LU 番号として“LU01”、“LU02”及びボリューム利用率欄 2225 にはボリューム利用率として“50%”が格納される。

【0019】

省電力グループ制御プログラム 221 は、省電力グループ管理テーブル 222 のうち、省電力識別番号欄 2221 に格納される省電力識別番号と状態欄 2222 に格納される状態から、現在稼働している省電力グループ数を把握し、常にディスクアレイ 200 に含ま

50

れる稼働中の省電力グループ数が前述した上限値を超えないよう制御する。

【 0 0 2 0 】

具体的には、この第1の実施形態では、省電力グループ制御プログラム221の省電力グループはRAIDグループ単位とする。特定のLUに対するアクセスが発生した場合に、省電力グループ管理テーブル222のRAIDグループ番号欄2223に格納されるRAIDグループ番号と、状態欄2222に格納される状態と、RAIDグループに含まれるLU番号欄2224に格納されるLU番号とを参照する。該特定のLUを含む特定のRAIDグループの状態欄2222に格納される状態から省電力グループ制御プログラム221は起動が必要か判断する。さらに起動が必要な場合は、省電力グループ制御プログラム221は上限値による制限を受けるか否かを判断する。そして、特定のRAIDグループが利用制限を受けない場合には、省電力グループ制御プログラム221は、該特定のLUへのデータアクセス処理の実行を許容し、実際に該特定のLUを含む特定のRAIDグループを起動させ、ディスク上のデータにアクセス可能とする。一方、特定のRAIDグループが利用制限を受ける場合には、省電力グループ制御プログラム221は、データアクセス処理を待機させる。上記は公知で開示されている技術であり、本実施例に関連する内容を除いて詳細は省略している。

10

【 0 0 2 1 】

なお、この第1の実施形態では、一例として省電力グループをRAIDグループ単位として説明している。しかし省電力グループは、LU単位、ディスク単位、同一電源から電力を供給されている筐体やシャーシ単位、1台以上のディスクアレイ単位、としてもよいもとする。

20

【 0 0 2 2 】

このように単位が変わると図2で示す省電力グループに内包される要素が増減する。例として、省電力グループの単位をディスクアレイ単位とし、複数のディスクアレイに跨って省電力の制御を実施する場合のストレージシステム2の構成を図10に示す。ディスクアレイ201~203は、図1のディスクアレイ200と同様な構成である。このため、ディスクアレイ201~203の構成については説明を省略する。

【 0 0 2 3 】

図10の例では、省電力制御の対象である各ディスクアレイ200~203と接続された計算機である電力制御サーバ401上で省電力グループ制御プログラム221が動作する。また、電力制御サーバ401は省電力グループ制御プログラム221で用いられる省電力グループ管理テーブル222がメモリに格納される。

30

【 0 0 2 4 】

なお、電力制御サーバ401は独立した専用のサーバとして示してあるが、他の装置、例えばストレージ仮想化装置や、FC(Fiber Channel)スイッチ内に内蔵されていてもかまわないものとする。電力制御サーバ401上の省電力グループ制御プログラム221は、各ディスクアレイ201~203からの稼働・停止等の状態欄2222に格納される“状態”に関する情報を収集し、それを元に各ディスクアレイ201~203の稼働・停止を制御する。すなわち、省電力グループ制御プログラム221は、ホスト100からの各ディスクアレイ201~203へのアクセスを監視し、停止中のディスクアレイへのアクセスがある場合、同時に稼働可能なディスクアレイ数を予め定めた上限値を超えるかどうかを判定し、その結果を元にディスクアレイの稼働・停止を制御する。

40

【 0 0 2 5 】

なお、この制御と同時に各ディスクアレイ201~203自体が個別に省電力グループ制御プログラム221をもち、各ディスクアレイ内部の筐体やシャーシ単位、RAIDグループ単位、LU単位、ディスク単位などで省電力グループを設定し、この省電力グループの稼働状態を制御してもよい。当然、これらの制御を電力制御サーバ401に集約させて制御してもよい。

【 0 0 2 6 】

また、図10に示すストレージシステム2とは異なり、電力制御サーバ401を持たず

50

、各ディスクアレイ 201 ~ 203 が省電力グループ制御プログラム 221 を持ち、そのうち 1 台のディスクアレイの省電力グループ制御プログラム 221 が電力制御サーバ 401 と同様の機能を有し、残りのディスクアレイの稼動・停止の制御を実施してもよい。この場合、電力グループ制御プログラム 221 が稼動しているディスクアレイが停止する時は、他の稼動しているディスクアレイに省電力制御と稼動状況を示す情報を引き継がせるものとする。

【0027】

また、全てのディスクアレイ 201 ~ 203 それぞれが常に他のディスクアレイが保持する省電力グループの情報と同期を取ることで、各ディスクアレイ 201 ~ 203 で個別に省電力制御を実施してもよい。

10

【0028】

なお、第 1 の実施形態では、省電力の制御する内容を RAID グループ単位でのディスクの停止と起動としているが、各省電力グループ単位ごとに個別にもつ省電力状態、例えばディスク単位でのスリープモードへの遷移などとしてもよいものとする。

【0029】

(1-2) データ重複除去

第 1 の実施形態では、データ重複除去プログラム 223 が重複データ管理テーブル 224 を利用することで、ディスクに書き込まれるデータのうちデータの内容が同一のデータの書き込みを制限し、ディスク 291 ~ 298 に格納されるデータ容量を削減する効果を実現する。

20

【0030】

図 3 は、重複データ管理テーブル 224 を示す図である。重複データ管理テーブル 224 は、ディスク 291 ~ 298 に格納される重複データを管理するためのテーブルである。重複データ管理テーブル 224 は、ハッシュ値欄 2241、実データ格納アドレス欄 2242 及びホスト側指定アドレス欄 2243 を有している。

【0031】

ハッシュ値欄 2241 にはデータのハッシュ値が格納される。実データ格納アドレス欄 2242 には、LU 番号：LBA が格納される。ホスト側指定アドレス欄 2243 には、LU 番号：LBA が格納される。

【0032】

重複データ管理テーブル 224 には、例えば図 3 に示すように、ハッシュ値欄 2241 に格納されたハッシュ値 “a” に対応して実データ格納アドレス欄 2242 に “LU01 : 100000”、ホスト側指定アドレス欄 2243 に “001 : 101000” が格納される。

30

【0033】

次に、重複データ管理テーブル 224 を利用したデータ重複処理プログラム 223 の処理について説明する。ホスト 100 からのライトデータと該ライトデータの書き込み先を指定した論理アドレス（以下、ホスト側指定アドレス）をディスクアレイ 200 が受信した際に、データ重複処理プログラム 223 は、ライトデータのデータを代表する数値であるハッシュ値を算出し、重複データ管理テーブル 224 のハッシュ値欄 2241 に格納されたハッシュ値と比較することで、既にディスク 291 ~ 298 に書き込まれているデータとライトデータが同一内容の可能性があるか否かを判断する。

40

【0034】

重複データ管理テーブル 224 に該ライトデータのハッシュ値と同一の値がない場合、データ重複処理プログラム 223 は、該ライトデータを新規のデータとしてディスクに書き込み、重複データ管理テーブル 224 に新規の行を追加し、該ライトデータのハッシュ値とディスクに書き込み先の論理アドレスをハッシュ値欄 2241 と実データ格納アドレス欄 2242 に格納し、ホスト側指定アドレスをホスト側指定アドレス欄 2243 に格納する。

【0035】

50

一方、重複データ管理テーブル 2 2 4 に該ライトデータのハッシュ値と同一の値がある場合、データ重複処理プログラム 2 2 3 は既に同一データを書き込まれている可能性があるとして判断する。この場合、データ重複処理プログラム 2 2 3 は重複データ管理テーブル 2 2 4 を参照し、該ライトデータのハッシュ値と同一のハッシュ値の実データ格納アドレス欄 2 2 4 2 に格納された論理アドレスが示すディスク上のデータを読み出し、該ライトデータとデータ内容を比較する。

【 0 0 3 6 】

比較の結果、データ内容が一致した場合、データ重複処理プログラム 2 2 3 は該ライトデータをディスクに書き込まず、重複データ管理テーブル 2 2 4 の該ライトデータのハッシュ値の行のホスト側指定アドレス欄 2 2 4 3 にホスト側指定アドレスを追加する。

10

【 0 0 3 7 】

一方、比較の結果、データ内容が一致しない場合、ハッシュ値は同じだが元データは異なるシノニムが発生している。この場合、データ重複処理プログラム 2 2 3 は該ライトデータをディスクに書き込み、重複データ管理テーブル 2 2 4 の該ライトデータのハッシュ値と同一ハッシュ値の行の実データ格納アドレス欄 2 2 4 2 に該ライトデータの書き込み先の論理アドレスを追加し、ホスト側指定アドレス欄 2 2 4 3 にホスト側指定アドレスを追加する。

【 0 0 3 8 】

つまり、シノニムの発生したハッシュ値は、同一ハッシュ値でも異なるデータをディスク上に持つ。すなわち、重複データ管理テーブル 2 2 4 は、一つのハッシュ値に対して、実データ格納アドレスとホスト側指定アドレスの組み合わせを複数持つことになる。

20

【 0 0 3 9 】

なお、ホスト 1 0 0 からのリードの際、データ重複除去プログラム 2 2 3 は重複データ管理テーブル 2 2 4 のホスト側指定アドレス欄 2 2 4 3 を参照し、対応する実データ格納アドレス欄 2 2 4 2 に格納された論理アドレスが示すディスクからデータを読み出し、ホスト 1 0 0 に返す。なお、上記は公知で開示されている技術であり、本実施形態に関連する内容を除いて詳細な説明を省略している。

【 0 0 4 0 】

ところで、ディスクアレイ 2 0 0 はデータプール管理テーブル 2 2 6 を用いて、“実ボリューム LU 番号”に示される複数の LU を、ホスト 1 0 0 に“ホスト側ボリューム番号”に示される仮想的な 1 つ以上の LU として認識させるストレージボリューム仮想化の仕組みを有している。これは一般的に開示されているストレージの仮想化技術であり、本実施形態では仮想化の仕組み自体は重要ではないため説明を省略する。

30

【 0 0 4 1 】

なお、データプール管理テーブル 2 2 6 は、図 4 に示すように、ホスト欄 2 2 6 1、ホスト側ボリューム番号欄 2 2 6 2 及び実ボリューム LU 番号欄 2 2 6 3 を有している。ホスト欄 2 2 6 1 には、ホストを示す情報が格納される。ホスト側ボリューム番号欄 2 2 6 2 にはホスト側のボリューム番号が格納される。実ボリューム LU 番号欄 2 2 6 3 にはホスト側のボリューム番号と対応する実ボリューム LU 番号が格納される。データプール管理テーブル 2 2 6 には、ホスト欄 2 2 6 1 に格納されるホスト “ A ” に対応してホスト側ボリューム番号欄 2 2 6 2 にホスト側ボリューム番号として “ 0 0 1 ”、実ボリューム LU 番号欄 2 2 6 3 に実ボリューム LU 番号として “ LU 0 1 , LU 0 2 , LU 0 3 , LU 0 4 , LU 0 5 , LU 0 6 ” が格納される。

40

【 0 0 4 2 】

この第 1 の実施形態では、既述のデータ重複除去を適用する範囲を、データプール管理テーブル 2 2 6 の実ボリューム LU 番号欄 2 2 6 2 に示されるホストが認識している仮想 LU に対応する複数の LU 単位としている。勿論、データ重複除去が適用される範囲は、仮想化されたボリューム単位に限定しなくてもよい。例えばホストに割り当てられた 1 つ以上の LU 単位や、データ重複除去プログラム 2 2 3 により予め設定されたホストと関連のない 1 つ以上の LU 単位や、 RAID グループ単位、ディスクアレイ単位などでもよい

50

ものとする。

【0043】

以上がストレージシステム1, 2の構成である。ただし、この第1の実施形態のストレージシステム1, 2の構成は本発明が適用されるストレージシステムの一例である。すなわち、この第1の実施形態では、データ重複除去に関するデータ重複除去プログラム223と、重複データ管理テーブル224が有する機能はディスクアレイ200のメモリ220に格納されることとしているが、これに限るものではない。例えば、これらデータ重複除去を実現する機能はディスクアレイ200の外部に位置し、SANを介してホスト100とディスクアレイ200のデータの仲介をする独立した計算機にデータ重複除去装置として存在してもよいとする。また、上記機能はデータ重複除去装置ではなく別の装置、例えば仮想化装置等に実装される構成でもよい。

10

【0044】

上記の場合、ディスクアレイ200と外部に存在するデータ重複除去プログラム223と、重複データ管理テーブル224は、ディスクアレイ200の管理I/F260またはFCI/F250を介して接続され、後述するデータ重複除去を考慮した省電力制御で必要となる既述の省電力グループ管理テーブル222と、重複データ管理テーブル224、及びデータプール管理テーブル226と、重複未比較データ管理テーブル225との情報を遣り取りできるものとする。

【0045】

重複未比較データ管理テーブル225は、同時に起動可能なRAIDグループ数を制限する上限値により同一ハッシュ値をもつデータがディスクからリードできず、その結果データ内容の比較を実施しないでディスクに書き込まれたライトデータ(以下、未比較データともいう)が格納されるテーブルである。

20

【0046】

重複未比較データ管理テーブル225は、図5に示すように、データ仮格納アドレス欄2251、ホスト側指定アドレス欄2252及びハッシュ値欄2253を有している。データ仮格納アドレス欄2251には、データ重複除去処理でデータを仮格納するアドレスが格納される。ホスト側指定アドレス欄2252にはホスト側指定アドレスが格納される。ハッシュ値欄2253は、データ仮格納アドレスに格納されるデータのハッシュ値が格納される。重複未比較データ管理テーブル225には、例えば図5に示すように、データ仮格納アドレス欄2251に格納されるデータ“LU03:100000”に対応してホスト側指定アドレス欄2252にホスト側指定アドレスとして“001:109000”、ハッシュ値欄2253にハッシュ値“a”が格納される。

30

【0047】

(1-3)データ重複除去を考慮した省電力制御

データ重複除去において、ホスト100からのライトデータをディスクアレイ200が受信した際、前述の通り、データ重複除去プログラム223は該ライトデータが既にディスク291~298に書かれたデータか否かを判断するためハッシュ等を利用した高速な検索を実施する。

【0048】

しかし、シノニムが発生している場合を考慮し、ライトデータと同一ハッシュ値が存在した場合、ホスト100からライトデータを受領したディスクアレイ200は、データプール管理テーブル226に示されたデータ重複除去の範囲にあるLUのうち、同一ハッシュ値をもつデータを格納しているディスク全てにアクセスする可能性が生じる。

40

【0049】

一方、ディスクアレイ200は既述の通り省電力グループ制御プログラム221が、省電力グループ管理テーブル222を利用し、同時に起動するRAIDグループ数を上限値で制限し、消費電力の抑制を実施している。

【0050】

このため、データ重複除去において同一ハッシュ値発生時にデータ比較を実施するため

50

ディスク 291 ~ 298 へのリードアクセスが発生する際、リード対象の論理アドレスを含む RAID グループが停止状態であった場合、起動までディスク起動の時間がかかり性能の低下が発生する。また上限値の制限が適用されている場合、該当する RAID グループが起動できず、データ比較が実施できない。

【 0051】

更に、同一ハッシュ値のデータが異なる RAID グループに分散して格納されている場合、同一ハッシュ値が発生するたびに、同一ハッシュ値を含む複数の RAID グループを起動させる必要が生じ、省電力効果の低減となる。

【 0052】

そこで、上記課題を解決する本発明の処理を、図 6 から図 9 のフローチャートを参照して説明する。

10

【 0053】

先ず、ホスト 100 からのライト処理について説明する。図 6 はこのライト処理を示すフローチャートである。この処理は、ホスト 100 からライトデータとその書き込み先のホスト側指定アドレスをディスクアレイ 200 が受信したときに開始される。

【 0054】

先ず、ステップ S1001 において、データ重複除去プログラム 223 は該ライトデータのハッシュ値を算出する。次に、ステップ S1002 において、データ重複除去プログラム 223 は同一ハッシュ値が重複データ管理テーブル 224 のハッシュ値欄 2241 にあるかを判断する。

20

【 0055】

同一ハッシュ値がない場合 (S1002: NO)、新規のデータ内容のライトデータであると判断できるため、ステップ S1003 において、データ重複除去プログラム 223 は重複データ管理テーブル 224 に新規に追加し、該ライトデータのハッシュ値をハッシュ値欄 2241 に書き込み先の論理アドレスを実データ格納アドレス欄 2242 に、ホスト側指定アドレスをホスト側指定アドレス欄 2243 に格納する。その後、ステップ S1004 において、データ重複除去プログラム 223 は省電力グループ管理テーブル 222 を参照し、起動している RAID グループの LU に該ライトデータを書き込む

【 0056】

一方、重複データ管理テーブル 224 に該ライトデータのハッシュ値と同一の値がある場合は (S1002: YES)、既に同一データを書き込まれている可能性があるため、ステップ S1005 において、データ重複除去プログラム 223 は、重複除去の判断の処理を実施する。

30

【 0057】

ステップ S1004 の処理又はステップ S1005 の処理が終了すると、ライト処理は終了する。

【 0058】

次に、重複除去の判断の処理について説明する。図 7 はステップ S1005 の重複除去の判断処理を示すフローチャートである。

【 0059】

ステップ S2001 において、データ重複除去プログラム 223 は、重複未比較データ管理テーブル 225 のハッシュ値欄 2253 に格納されるハッシュ値が、該ライトデータのハッシュ値に一致するか比較する。

40

【 0060】

上記ステップ 2001 の結果が一致しない場合 (S2001: NO)、ステップ S2002 において、データ重複除去プログラム 223 は、重複除去の実施処理を行う。この重複除去の実施処理については図 8 を用いて後述する。

【 0061】

上記ステップ 2001 の結果が一致する場合 (S2001: YES)、既にディスク上には同一ハッシュ値をもつ未比較データが存在する。該ライトデータのデータ内容の比較

50

をするには、重複データ管理テーブル 2 2 4 と、重複未比較データ管理テーブル 2 2 5 に登録された同一ハッシュ値のデータを全てディスクからリードする必要がある。本実施形態では、起動 R A I D グループ数の制限のため全てのディスクを起動できない。ディスクからの同一ハッシュ値のデータを全てリードできないため、ステップ S 2 0 0 3 において、データ重複除去プログラム 2 2 3 は省電力グループ管理テーブル 2 2 2 を参照し、起動している R A I D グループの L U にデータ内容の比較を実施していない該ライトデータを書き込む。

【 0 0 6 2 】

その後、ステップ S 2 0 0 4 において、データ重複除去プログラム 2 2 3 は、重複未比較データ管理テーブル 2 2 5 に該ライトデータを未比較データとして格納する。すなわち、ステップ S 2 0 0 3 で該ライトデータを書き込んだディスクの論理アドレスをデータ仮格納アドレス欄 2 2 5 1 に、該ライトデータのホスト側指定アドレスをホスト側指定アドレス欄 2 2 5 2、該ライトデータのハッシュ値をハッシュ値欄 2 2 5 3 に格納する。

10

【 0 0 6 3 】

次に、重複除去の実施処理について説明する。図 8 はこの重複除去の実施処理を示すフローチャートである。

【 0 0 6 4 】

先ず、データ重複除去プログラム 2 2 3 は、ステップ S 3 0 0 1 において、重複データ管理テーブル 2 2 4 の当該ライトデータと同一ハッシュ値の実データ格納アドレス欄 2 2 4 2 に格納される実データ格納アドレスの L U 番号と、省電力グループ管理テーブル 2 2 2 の L U 番号欄 2 2 2 4 に格納される L U 番号とを比較し、当該 L U を含む R A I D グループの状態欄 2 2 2 2 に格納される“状態”が“起動中”であるか否かを判断する。

20

【 0 0 6 5 】

当該 R A I D グループの状態欄 2 2 2 2 が“停止中”の場合 (S 3 0 0 1 : N O)、ステップ S 3 0 0 2 において、データ重複除去プログラム 2 2 3 は、当該 L U が含まれる R A I D グループを起動させることが可能か判断する。起動正否の判断は、省電力グループ制御プログラム 2 2 1 が、省電力グループ管理テーブル 2 2 2 を参照し、上限値の制限内かで判断する。

【 0 0 6 6 】

このステップ S 3 0 0 2 の R A I D グループの起動判断の結果、当該 R A I D グループの起動できない場合は (S 3 0 0 2 : N O)、データ重複除去プログラム 2 2 3 はディスクから同一ハッシュ値のデータをリードできない。この結果、データ重複除去プログラム 2 2 3 は当該ライトデータのデータ内容の比較ができない。このためデータ重複除去プログラム 2 2 3 はステップ S 3 0 0 3 において、省電力グループ管理テーブル 2 2 2 を参照し、ステップ S 3 0 0 1 で確認した R A I D グループとは異なる、起動中の R A I D グループの L U に該ライトデータを書き込む

30

【 0 0 6 7 】

その後、ステップ S 3 0 0 4 において、データ重複除去プログラム 2 2 3 は重複未比較データ管理テーブル 2 2 5 に該ライトデータを未比較データとして格納する。ステップ S 3 0 0 3 で該ライトデータを書き込んだディスクの論理アドレスをデータ仮格納アドレス欄 2 2 5 1、ホスト側指定アドレスをホスト側指定アドレス欄 2 2 5 2 及びハッシュ値をハッシュ値欄 2 2 5 3 に格納する。

40

【 0 0 6 8 】

ステップ S 3 0 0 1 で状態欄 2 2 2 2 が“起動中”の場合 (S 3 0 0 1 : Y E S)、又はステップ S 3 0 0 2 において、上記 R A I D グループの起動は可能であると判断された場合は (S 3 0 0 2 : Y E S)、ステップ S 3 0 0 5 において、データ重複除去プログラム 2 2 3 は、当該 R A I D グループを起動状態にし、重複データ管理テーブル 2 2 4 の該ライトデータのハッシュ値の実データ格納アドレス欄 2 2 4 2 に格納された全ての論理アドレスのデータをディスクからリードし、該ライトデータとデータ内容の比較を実施する。

【 0 0 6 9 】

50

上記ステップS3005の結果、データが一致した場合(S3005: YES)、既に同一データがディスク291~298上に存在するため、データ重複除去プログラム223は、当該ライトデータはディスク291~298に書き込まない。またこの場合、データ重複除去プログラム223は、ステップS3006において、ステップS3005の比較で結果が一致したデータの論理アドレスが格納された重複データ管理テーブル224の実データ格納アドレス欄2242と同じ行のホスト側指定アドレス欄2243に当該ライトデータのホスト側指定アドレスを格納する。

【0070】

上記ステップS3005の結果、データが一致しなかった場合(S3005: NO)、ステップS3007において、データ重複除去プログラム223は、該ライトデータは新規のデータとし、省電力グループ管理テーブル222を参照し、ステップS3005で比較データをリードし起動中のRAIDグループのLUに該ライトデータを書き込む。

【0071】

その後、ステップS3008において、データ重複除去プログラム223は重複データ管理テーブル224の該ライトデータのハッシュ値の行の実データ格納アドレス欄2242に該ライトデータの書き込み先の論理アドレスを追加し、ホスト側指定アドレス欄2243にホスト100から指定されたアドレスを追加する。

【0072】

以上がホスト100からのライトデータのディスクアレイ200における処理手順である。本処理では、同一ハッシュ値を持つデータがないライトデータは、新規のデータとしてディスク291~298上に格納され、算出されたハッシュ値がハッシュ値欄2241に、ディスク上の論理アドレスが実データ格納アドレス欄2242に、ホスト側指定アドレスがホスト側指定アドレス欄2243(重複データ管理テーブル224)に、新規に登録される。

【0073】

また、同一ハッシュ値をもつデータが存在し、比較の結果データ内容が異なることが確定しているライトデータは、データ自体はディスク291~298上に格納され、既にある同一ハッシュ値の行の、ディスク上の論理アドレスが実データ格納アドレス欄2242に、ホスト側指定アドレスがホスト側指定アドレス欄2243(重複データ管理テーブル224)に追加で登録される。

【0074】

さらに、また、省電力グループの上限値によりディスクが起動できず同一ハッシュ値をもつデータとのデータ内容の比較が実施できなかったライトデータは、ライトデータを一時的に書き込んだディスク上の論理アドレスがデータ仮格納アドレス欄2251に、ハッシュ値がハッシュ値欄2253に、ホスト側指定アドレスがホスト側指定アドレス欄2252(重複未比較データ管理テーブル225)に新規に登録される。

【0075】

次に、本発明において、データ重複除去プログラム223が実施するデータ最適化処理について説明する。データ最適化処理とは、重複未比較データ管理テーブル225に登録されている未比較データに対して、データ内容の比較を実施し、データの内容が同一な場合は、通常の方法でデータ重複除去を実施することである。更に、同一ハッシュ値を持ちデータ内容が異なることが確定したデータは、ディスク291~298上のデータ位置を移動させ、同じハッシュ値をもつデータが格納された省電力グループに集められる。

【0076】

データ最適化処理の結果、データ最適化処理後にライトデータを受信し、同一ハッシュ値が発生しディスク上のデータとデータ内容の比較が発生した際、同一ハッシュ値をもつデータが集約された省電力グループのみを起動させることでディスクからの同一キャッシュ値をもつデータのリードが可能となる。

【0077】

続いて、本実施形態におけるデータ最適化処理について説明する。図9はこのデータ最

10

20

30

40

50

適化処理の手順を示すフローチャートである。

【 0 0 7 8 】

先ず、ステップ S 4 0 0 1 において、データ重複除去プログラム 2 2 3 は以下の (1) から (6) の場合をデータ最適化処理の実施契機としてデータ最適化実施を行なうこととする。具体的には、(1) 通常の処理中に R A I D グループの起動 / 停止時が発生する場合、(2) ヘルスチェックなどで停止中の R A I D グループが起動する場合、(3) 上記上限値に余裕がある場合、(4) 省電力グループ管理テーブル 2 2 2 で示される省電力グループ単位である R A I D グループのボリューム利用率欄 2 2 2 5 に格納されるボリューム利用率がユーザの指定した閾値を越えた場合、(5) 重複未比較データ管理テーブル 2 2 5 に格納された未比較データ数が予め設定された閾値以上になった場合、(6) ユーザまたはホスト 1 0 0 のアプリケーションによって実施を指定された場合である。本実施形態では、これら (1) から (6) のいずれかの場合をデータ最適化処理の実施契機とすることとするが、実施契機は、これら (1) から (6) いずれかの場合に限られるものではない。

10

【 0 0 7 9 】

次に、ステップ S 4 0 0 2 において、データ重複除去プログラム 2 2 3 は、重複未比較データ管理テーブル 2 2 5 を参照し、未比較データが存在するか否かを判断する。未比較データが存在しない場合 (S 4 0 0 2 : N O)、データ重複除去プログラム 2 2 3 は、再度ステップ S 4 0 0 1 へ戻り、データ最適化処理開始の契機を待つ。

【 0 0 8 0 】

一方、未比較データが存在する場合 (S 4 0 0 2 : Y E S)、データ重複除去プログラム 2 2 3 は未比較データのデータ内容を比較するために稼動が必要な R A I D グループの候補を選択し、結果を起動候補省電力グループ管理テーブル 2 2 7 に格納する。

20

【 0 0 8 1 】

具体的な一例としては、データ重複除去プログラム 2 2 3 は、未比較データそれぞれに対して、各々のハッシュ値と同一値をもつ重複未比較データ管理テーブル 2 2 5 のデータ仮格納アドレス欄 2 2 5 1 に格納された L U 番号、及び重複データ管理テーブル 2 2 4 の実データ格納アドレス欄に格納された L U 番号を取り出す。次に、データ重複除去プログラム 2 2 3 は、省電力グループ管理テーブル 2 2 2 から上記 L U 番号を含む R A I D グループを選択する。この結果、全ての未比較データのハッシュ値ごとにデータ内容の比較のため起動する必要がある R A I D グループの候補 (以下、起動候補 R A I D グループともいう) が選定される。そして、ステップ S 4 0 0 3 において、データ重複除去プログラム 2 2 3 は、その結果を起動候補省電力グループ管理テーブル 2 2 7 に格納する。

30

【 0 0 8 2 】

ここで起動候補省電力グループ管理テーブル 2 2 7 について説明する。なお、図 1 1 は、起動候補省電力グループ管理テーブル 2 2 7 を示す図である。起動候補省電力グループ 2 2 7 は、データを格納するために起動する R A I D グループの候補を格納するテーブルである。起動候補省電力グループ管理テーブル 2 2 7 は、ハッシュ値欄 2 2 7 1、起動候補省電力グループ欄 2 2 7 2 及び同一ハッシュ値データ数欄 2 2 7 3 を有している。ハッシュ値欄 2 2 7 1 にはハッシュ値が格納される。起動候補省電力グループ欄 2 2 7 2 には起動候補省電力グループが格納される。同一ハッシュ値データ数欄 2 2 7 3 には同一ハッシュ値データ数が格納される。起動候補省電力グループ管理テーブル 2 2 7 には、例えば図 1 1 に示すように、ハッシュ値欄 2 2 7 1 のハッシュ値 “ a ” に対応して起動候補省電力グループ欄 2 2 7 2 には起動候補省電力グループとして “ R G 0 1 , R G 0 2 , R G 0 3 ” という 3 つの R A I D グループが格納され、同一ハッシュ値データ数欄 2 2 7 3 には、R A I D グループ “ R G 0 1 ” と対応して同一ハッシュ値データ数 “ 2 ” が、R A I D グループ “ R G 0 2 ” と対応して同一ハッシュ値データ数 “ 1 ” が、R A I D グループ “ R G 0 3 ” と対応して同一ハッシュ値データ数 “ 1 ” が格納される。

40

【 0 0 8 3 】

なお本実施形態では、省電力グループは R A I D グループ単位としているため、起動候

50

補省電力グループには R A I D グループの識別子を記載する場合について説明することとするが、省電力グループの単位が異なる場合は、省電力識別番号を用いてもよい。

【 0 0 8 4 】

次に、データ重複除去プログラム 2 2 3 は、ステップ S 4 0 0 4 において、起動候補省電力グループ管理テーブル 2 2 7 及び電力グループ管理テーブル 2 2 2 を参照し、省電力グループ制御プログラム 2 2 1 が保持する上限値の範囲で、省電力グループ制御プログラム 2 2 1 に指示をだし起動候補 R A I D グループを稼働させ、未比較データのデータ内容をハッシュ値ごとに比較する。

【 0 0 8 5 】

なお、起動候補 R A I D グループの総数が上限値を超えている場合、データ重複除去プログラム 2 2 3 は、起動候補 R A I D グループを上限値内で順番に立ち上げデータをリードし、ディスクアレイ 2 0 0 のキャッシュ 2 3 0 にリードしたデータを格納し、このデータと比較する。また起動中のディスクの一部を一時領域をとして利用し、上記のキャッシュ 2 3 0 の代わりに利用してもよいものとする。また、ユーザに警告をだして上限値の変更を促してもよいものとする。

【 0 0 8 6 】

ステップ S 4 0 0 4 の比較の結果、データ内容が一致した未比較データは (S 4 0 0 4 : Y E S)、データ重複除去プログラム 2 2 3 が、1 つ以上の重複したデータがディスク上にあると判断する。このため、ステップ S 4 0 0 5 において、データ重複除去プログラム 2 2 3 は、この重複したデータを 1 つ残して他は全て削除する。重複データ削除の優先度は、重複したデータが全て同一 R A I D グループにある時は、未比較データの方から削除する。重複データが異なる R A I D グループにある場合は、起動候補省電力グループ管理テーブル 2 2 7 に格納された同一ハッシュ値の起動候補省電力グループの中で同一ハッシュ値データ数が小さい方の起動候補省電力グループに格納されたデータの方から削除する。

【 0 0 8 7 】

次に、ステップ S 4 0 0 6 において、データ重複除去プログラム 2 2 3 は、削除したデータの重複データ管理テーブル 2 2 4、重複未比較データ管理テーブル 2 2 5、起動候補省電力グループ管理テーブル 2 2 7 の情報を、ディスク 2 9 1 ~ 2 9 8 上のデータ削除に合わせて整理する。

【 0 0 8 8 】

具体的な一例としては、ステップ S 4 0 0 5 で削除したデータが重複未比較データ管理テーブル 2 2 5 に格納されたデータの場合、当該データを重複未比較データ管理テーブル 2 2 5 から削除し、重複データ管理テーブル 2 2 4 に格納しなおす。この格納内容は未比較データの“ホスト側指定アドレス”の情報を、ハッシュ値、データ内容とも一致したデータに該当する重複データ管理テーブル 2 2 4 のホスト側指定アドレス欄 2 2 4 3 に追加するものである。この結果、重複データ管理テーブル 2 2 4 は、1 つのハッシュ値と実データ格納アドレスの組み合わせに対して複数のホスト側指定アドレス持つ場合が存在することとなる。

【 0 0 8 9 】

ステップ S 4 0 0 5 で削除したデータが、重複データ管理テーブル 2 2 4 に格納されたデータだった場合は、ステップ S 4 0 0 6 において、データ重複除去プログラム 2 2 3 は、以下に説明する重複未比較データ管理テーブル 2 2 5 及び、重複データ管理テーブル 2 2 4 の整理を行う。

【 0 0 9 0 】

この整理は、削除したデータと比較した未比較データを示す重複未比較データ管理テーブル 2 2 5 の格納情報から削除し、重複データ管理テーブル 2 2 4 に格納しなおすものである。具体的にはデータ重複除去プログラム 2 2 3 は、重複未比較データ管理テーブル 2 2 5 に格納された“データ仮格納アドレス”、“ハッシュ値”及び“ホスト側指定アドレス”を、重複データ管理テーブル 2 2 4 の実データ格納アドレス欄 2 2 4 2、ハッシュ値

10

20

30

40

50

欄 2 2 4 1 及びホスト側指定アドレス欄 2 2 4 3 に格納する。さらに、データ重複除去プログラム 2 2 3 は、削除したデータを示す重複データ管理テーブル 2 2 4 に格納されていた情報を削除し、当該データのホスト側指定アドレスの情報を、先ほど重複未比較データ管理テーブル 2 2 5 から重複データ管理テーブル 2 2 4 に移動したデータと対応するホスト側指定アドレス欄 2 2 4 3 に追加する。この整理は順番が前後してもかまわないが、結果的に、データ内容が一致した重複データは、ディスクアレイ 2 0 0 上のディスク 2 9 1 ~ 2 9 8 には 1 つしか存在せず、かつ当該データ内容に対するホスト側指定アドレスを示す情報は全て引き継がれるものとする。最後に、データ重複除去プログラム 2 2 3 は、削除したデータの数に合わせて、起動候補省電力グループ管理テーブル 2 2 7 の同一ハッシュ値データ数欄 2 2 7 3 に格納される同一ハッシュ値データ数を削除したデータの数だけ減少させる。

10

【 0 0 9 1 】

一方、ステップ 4 0 0 4 の比較の結果、データ内容が一致しない場合、データ重複除去プログラム 2 2 3 は、当該データは、重複したデータがディスク 2 9 1 ~ 2 9 8 上にないと判断する。この結果、ハッシュ値は同じだがデータ内容は異なるため削除対象とはならない。そして、ステップ S 4 0 0 7 において、データ重複除去プログラム 2 2 3 は、以下に説明する重複未比較データ管理テーブル 2 2 5 及び、重複データ管理テーブル 2 2 4 の整理を行う。

【 0 0 9 2 】

データ重複除去プログラム 2 2 3 は、データ内容が一致しないデータを示す重複未比較データ管理テーブル 2 2 5 の格納情報を削除し、重複データ管理テーブル 2 2 4 に格納しなおす。

20

【 0 0 9 3 】

具体的には、ステップ S 4 0 0 7 において、データ重複除去プログラム 2 2 3 は、当該データを示す重複未比較データ管理テーブル 2 2 5 の格納情報のうち“データ仮格納アドレス”及び“ホスト側指定アドレス”を、重複データ管理テーブル 2 2 4 の同一ハッシュ値の行の“実データ格納アドレス”及び“ホスト側指定アドレス”に移動する。この結果、重複データ管理テーブル 2 2 4 は、1 つのハッシュ値に対して、複数の実データ格納アドレスとホスト側指定アドレスの組み合わせを持つ場合が存在するようになる。

【 0 0 9 4 】

上記ステップ S 4 0 0 4 からステップ S 4 0 0 7 までの重複データの削除は重複未比較データ管理テーブル 2 2 5 に格納された情報がなくなるまで繰り返し実施される。

30

【 0 0 9 5 】

最後に、ステップ S 4 0 0 8 において、データ重複除去プログラム 2 2 3 は、ハッシュ値が同じでデータ内容が異なるデータを、同じ省電力グループに集約させるためデータを移動させ、それに伴う重複データ管理テーブル 2 2 4 及び起動候補省電力グループ管理テーブルを整理する。

【 0 0 9 6 】

具体的な一例としては、データ重複除去プログラム 2 2 3 は、重複データ管理テーブル 2 2 4 の実データ格納アドレス欄 2 2 4 2 省電力グループ管理テーブル 2 2 2 を参照する。そして同一ハッシュ値で異なる省電力グループにデータが格納されている場合、データ重複除去プログラム 2 2 3 は、このデータを移動し、同じ省電力グループにデータを集める。移動元と移動先の決定は、起動候補省電力グループ管理テーブル 2 2 7 を参照し、同一ハッシュ値データ数欄 2 2 7 3 から“同一ハッシュ値データ数”が少ない省電力グループから“同一ハッシュ値データ数”が多い省電力グループにデータを移動させる。このデータの移動に伴い、重複データ管理テーブル 2 2 4 の移動したデータを示す格納情報のうち実データ格納アドレス欄 2 2 4 2 に格納される実データ格納アドレスを移動先のアドレスに変更する。さらに、データ重複除去プログラム 2 2 3 は、起動候補省電力グループ管理テーブル 2 2 7 の同一ハッシュ値データ数欄 2 2 7 3 の“同一ハッシュ値データ数”をデータが異なる省電力グループに移動したのに合わせて変更する。

40

50

【 0 0 9 7 】

上記のステップ S 4 0 0 8 は、起動候補省電力グループ管理テーブル 2 2 7 の同一ハッシュ値データ数欄 2 2 7 3 に格納される“同一ハッシュ値データ数”がすべて一つの“起動候補省電力グループ”に集約されるまで繰り返し実施する。

【 0 0 9 8 】

上記ステップ S 4 0 0 8 の処理が終了するとステップ S 4 0 0 1 の処理へ戻り、上記で説明したデータ最適化実施の契機が訪れると再び上述した処理が実行される。

【 0 0 9 9 】

この第 1 の実施形態のストレージシステム 1 によると、ディスクアレイ 2 0 0 においてデータ重複除去とディスク稼働の制御を同時に実施した場合に、データ重複除去によるデータ容量の効率化と、ディスク稼働の制御による省電力効果を最大限に発揮することができる。

10

【 0 1 0 0 】

なお、上記実施形態は一例であり、上記実施形態と異なる手順で処理を実施してもよいものとする。すなわち、結果的に、データ内容が同じ重複データは削除することでディスク上に必ずデータを一つとし、複数の省電力グループに分散して格納されているハッシュ値が同じでデータ内容が異なるデータは同一省電力グループに集約して格納されればよいものとする。例えば、データ最適化処理に際して省電力グループの起動回数を減らすために、ステップ S 4 0 0 2 から S 4 0 0 8 までを同時に実施してもよいものとする。

【 0 1 0 1 】

第 2 の実施形態：

(2 - 1) 構成の説明

第 2 の実施形態の構成は、上記第 1 の実施形態のストレージシステム 1 の構成と同様であるため省略する。

20

【 0 1 0 2 】

(2 - 2) ディスクアレイの消費電力の抑制

【 0 1 0 3 】

上記第 1 の実施形態と同様のため省略する。ただし、省電力グループの中に 1 つ以上、ディスクアレイ 2 0 0 が稼働中は、常にアクセスが可能な状態で起動している省電力グループ(以下、常時稼働グループと呼ぶ)が定義されている。

30

【 0 1 0 4 】

(2 - 3) データ重複除去

上記第 1 の実施形態と同様のため省略する。

【 0 1 0 5 】

(2 - 4) データ重複除去を考慮した省電力制御

上記第 1 の実施形態と処理の流れは以下の点を除いて同様であるため同様な点については説明を省略し、異なる点について以下に説明する。

【 0 1 0 6 】

上記第 1 の実施形態と異なりこの第 2 の実施形態では、常時稼働グループを未比較データの一時的な格納場所に利用する。この場合、図 7 ステップ S 2 0 0 2 における未比較データは全て常時稼働グループに一時的に保持される。また、データ最適化処理に際しては、ステップ S 4 0 0 5 の重複データの削除処理、ステップ S 4 0 0 8 の同一ハッシュ値のデータ移動処理は、必ず常時稼働グループを削除元、移動元とする。

40

【 0 1 0 7 】

さらに、常時稼働グループが常にアクセス可能である特性を活かし、同一ハッシュ値のデータ数が予め決められた閾値以上になった場合、該当するハッシュ値のデータを全て常時稼働グループに移動させてもよいものとする。

【 0 1 0 8 】

これにより、図 8 のステップ S 3 0 0 1 での省電力グループが起動しているかの判断において、常時稼働グループは常に起動しているため、余分な省電力グループの起動を必要

50

としないデータ重複除去の実施が可能となる。

【0109】

なお、この第2の実施形態では、常時稼動グループはディスク291～298が一つ以上含まれる省電力グループの一つの制御形態としている。しかしながら、このディスク291～298が含まれる制御形態とは異なり、揮発/不揮発性のメモリなどのディスクアレイ200の稼動時にアクセスが可能なデータ記憶領域が含まれる制御形態であってもよい。

【0110】

第3の実施形態：

(3-1) 構成の説明

第3の実施形態の構成は、上記第1の実施形態のストレージシステム1の構成と同様であるため省略する。

【0111】

(3-2) ディスクアレイの消費電力の抑制

上記第1の実施形態と同様のため省略する。

【0112】

(3-3) データ重複除去

上記第1の実施形態と同様のため詳細な説明は省略する。ただし、ホスト100からディスクアレイ200へのライトデータは、データ重複除去プログラム223により、複数のセグメントに分割されてからハッシュ値を計算されることとする。また、セグメントの大きさはデータ重複除去の仕様によって異なってもよいものとする。

【0113】

(3-4) データ重複除去を考慮した省電力制御

上記第1の実施形態と同様な点については説明を省略し、異なる点について以下に説明する。

【0114】

上記第1の実施形態と異なり第3の実施形態では、ライトデータが複数のセグメントに分割されている。これら一つのライトデータから分割された複数のセグメントはホスト100からは連続したデータとして認識される。その結果、ホスト100からのデータリードで同時にディスクから読み出される確立が高くなる。上記第1の実施形態に適用した場合、連続性のあるセグメント群に所属するデータ(以下、関連データと呼ぶ)の一部が、異なる省電力グループ(上記第1の実施形態の場合はRAIDグループ)に格納される場合も発生する。この場合、連続したセグメントを呼び出す際に複数の省電力グループを起動させる必要が生じる。この結果、省電力グループが停止状態などの省電力状態であった場合、複数の省電力グループを起動させる必要が生じ、省電力効果の低減となる。そこで、この第3の実施形態では、連続性のあるデータの関連性を示す情報を保持し、この情報を元にデータ最適化処理の際に、同じ省電力グループとなるようにデータを配置する。

【0115】

具体的には、図3に示す重複データ管理テーブル224と、図5に示す重複未比較データ管理テーブル225に、関連アドレスを格納する関連アドレス欄を追加する。図12は関連アドレス欄を追加した重複データ管理テーブル228、図13は関連アドレス欄を追加した重複未比較データ管理テーブル229を示す図である。

【0116】

図12に示すように重複データ管理テーブル228は、ハッシュ値欄2281、実データ格納アドレス欄2282及びホスト側指定アドレス欄2283に加え、関連アドレス欄2284が設けられている。また、図13に示すように重複未比較データ管理テーブル229は、データ仮格納アドレス欄2291、ホスト側指定アドレス欄2292及びハッシュ値欄2293に加え、関連アドレス欄2294が設けられている。

10

20

30

40

50

【 0 1 1 7 】

図 1 2 及び図 1 3 に示すように、関連アドレス欄 2 2 8 4 は実データ格納アドレス欄 2 2 8 2 と、関連アドレス欄 2 2 9 4 はデータ仮格納アドレス欄 2 2 9 1 とそれぞれ 1 対 1 の関係となっている。また、実データ格納アドレス欄 2 2 8 2 又は、データ仮格納アドレス欄 2 2 9 1 で示されるディスク上のアドレスに格納されたデータの関連データがある場合、当該データのディスク上のアドレスが複数格納される。

【 0 1 1 8 】

このアドレスが複数格納される処理を上記第 1 の実施形態を基に具体的に説明すると、第 1 の実施形態でのライトデータをディスクに書き込むタイミング、つまり、実データ格納アドレスを重複データ管理テーブル 2 2 8、又はデータ仮格納アドレスを重複未比較データ管理テーブル 2 2 9 に格納するタイミングで、関連アドレス欄 2 2 8 4、又は関連アドレス欄 2 2 9 4 に、関連データとしてディスク 2 9 1 ~ 2 9 8 に書き込まれた論理アドレスを登録する。なお、関連データが複数の連続性をもつ場合は、関連アドレス欄 2 2 8 4、2 2 9 4 には複数の論理アドレスが格納される。

【 0 1 1 9 】

さらに、データ最適化処理の際、同一ハッシュ値をもつデータの移動と重複データ管理テーブル 2 2 8、重複未比較データ管理テーブル 2 2 9 などの管理テーブルの整理を実施後に、“関連アドレス”を利用したデータの再配置とテーブルの整理を実施する。なお、これらの処理は同時に実施してもよいものとする。

【 0 1 2 0 】

次に、第 3 の実施形態におけるデータ最適化処理について説明する。図 1 4 はこのデータ最適化処理を示すフローチャートである。なお、図 1 4 におけるステップ S 4 0 0 1 から S 4 0 0 8 は、図 9 で説明した処理と同様であるため説明を省略し、新たに追加されたステップ S 4 0 0 9 について詳細に説明する。

【 0 1 2 1 】

ステップ S 4 0 0 9 において、データ重複除去プログラム 2 2 3 は、重複データ管理テーブル 2 2 8 の実データ格納アドレス欄 2 2 8 2 と関連アドレス欄 2 2 8 4 を参照する。そして、データ重複除去プログラム 2 2 3 は、LU 番号が異なっていると判断した場合、さらに、当該 LU 番号が異なる RAID グループに所属しているか否かを、省電力グループ管理テーブル 2 2 2 を参照し判断する。この結果、データ重複除去プログラム 2 2 3 は異なる RAID グループに所属していると判断した場合、同一の RAID グループにデータを移動し、データ移動に伴い変更したアドレスに合わせて重複データ管理テーブル 2 2 8 の実データ格納アドレス欄 2 2 8 2 と関連アドレス欄 2 2 8 4 を書き換える。データ重複除去プログラム 2 2 3 は、この処理を“実データ格納アドレス”と“関連アドレス”が全て同じ RAID グループに属するまで繰り返す。この結果、“関連データ”は同じ RAID グループ(省電力グループ)上に格納される。

【 0 1 2 2 】

ただし、この際、重複データ管理テーブル 2 2 8 の“ハッシュ値”に対して“実データ格納アドレス”が複数格納されているデータに関しては、同一ハッシュ値のデータを集約するため同一省電力グループに集められているデータのため除外してもよいものとする。

【 0 1 2 3 】

なお、同一省電力グループに、同一ハッシュ値で異なるデータを同一省電力グループへ集約するかを優先するか、または関連データを同一省電力グループに集約するかを優先するかは、ユーザが指定してもよいし、例えば同一ハッシュ値のデータ数と関連データの数を比較し、多い方を優先するなどシステムの的に決定してもよいものとする。

【 0 1 2 4 】

さらに、上記第 3 の実施形態では、一例としてライトデータがデータ重複除去プログラム 2 2 3 により分割された場合で説明したが、ホスト 1 0 0 のアプリケーション等でのデータの連続性が示される情報、例えば差分ファイルなどの特性を利用して関連データとしてもよい。

10

20

30

40

50

【 0 1 2 5 】

また、上記第 1 から第 3 の実施形態では、説明の都合上、処理の主体を省電力グループ管理プログラム 2 2 1 又はデータ重複除去プログラム 2 2 3 とした場合で説明しているが、当該処理は、省電力グループ管理プログラム 2 2 1 又はデータ重複除去プログラム 2 2 3 が CPU 2 1 0 により実行されて実現されるものである。

【 0 1 2 6 】

さらに、上記第 1 から第 3 の実施形態において、取り扱うライトデータの単位はブロック単位でもファイル単位のいずれの単位を用いてもよい。

【 0 1 2 7 】

さらに、上記各実施形態では本発明のストレージシステムをディスクアレイ 2 0 0 に適用した場合で説明したが、これに限られるものではなく、広くストレージシステムに適用できるものである。

【 0 1 2 8 】

また、上記実施の形態では本発明を、ホスト 1 0 0 と、この 1 0 0 から受信するデータを格納する複数のディスク 2 9 1 ~ 2 9 8 及びその複数のディスク 2 9 1 ~ 2 9 8 から構成される複数の論理ユニット 2 8 1 ~ 2 8 3 を有するディスクアレイ 2 0 0 とを含むストレージシステム 1 であって、ディスクアレイ 2 0 0 は、複数の論理ユニット 2 8 1 ~ 2 8 3 を用いて構成される RAID グループ 2 7 1 , 2 7 2 及びその RAID グループ 2 7 1 , 2 7 2 が起動中か否かを少なくとも含むグループ情報を格納する省電力グループ管理テーブル 2 2 2 と、RAID グループに含まれる論理ユニット 2 8 1 ~ 2 8 3 に対応するディスク 2 9 1 ~ 2 9 8 の電源を ON / OFF するディスクアレイコントローラ 2 4 0 と、複数のディスク 2 9 1 ~ 2 9 8 に格納されるデータの格納位置と対応する論理ユニット 2 8 1 ~ 2 8 3 の論理ユニット番号 (LU)、データのビット列を代表するハッシュ値及び論理ユニット 2 8 1 ~ 2 8 3 と対応するホスト 1 0 0 が指定するホスト側指定番号を含む管理情報を格納する重複データ管理テーブル 2 2 4 とを含み、ホスト 1 0 0 からデータを受信すると、そのデータのビット列を代表する値を算出するステップ (S 1 0 0 1) と、この算出されたハッシュ値と同一のハッシュ値が重複データ管理テーブル 2 2 4 に格納されているか否かを判断するステップと (S 1 0 0 2)、この判断するステップで同一のハッシュ値が格納されていないと判断した場合は受信したデータを重複データ管理テーブル 2 2 4 に格納するとともに省電力グループ管理テーブル 2 2 2 に基づいて起動しているグループを構成する論理ユニット番号と対応するディスク 2 9 1 ~ 2 9 8 にデータを格納する第 1 のステップと (S 1 0 0 3 , S 1 0 0 4)、判断するステップで同一のハッシュ値が格納されていると判断した場合は受信したデータを省電力グループ管理テーブル 2 2 2 で管理されるグループ情報及び重複データ管理テーブル 2 2 4 で管理される管理情報に基づいて、ディスク 2 9 1 ~ 2 9 8 に格納する第 2 のステップと (S 1 0 0 5) を有するものに適用した場合について説明したが、これに限るものでない。

【 0 1 2 9 】

また、ディスクアレイ 2 0 0 は、データをディスク 2 9 1 ~ 2 9 8 に仮格納する論理ユニット番号及びデータのビット列を代表するハッシュ値を少なくとも含む未比較データ情報を格納する重複未比較データ管理テーブル 2 2 5 を備え、第 2 のステップは、受信したデータのビット列を代表するハッシュ値と同一のハッシュ値が前記未比較データ管理テーブルに格納されているか否かを判断する未比較ステップと (S 2 0 0 1)、この未比較ステップで同一のハッシュ値が格納されていると判断した場合は、省電力グループ管理テーブル 2 2 2 に基づいて起動しているグループを構成する論理ユニット番号と対応するディスク 2 9 1 ~ 2 9 8 に受信データを格納するとともに受信データを格納するディスク 2 9 1 ~ 2 9 8 と対応する論理ユニット番号及び受信したデータのビット列を代表するハッシュ値を重複未比較データ管理テーブルに格納する未比較データを処理するステップと (S 2 0 0 3 , S 2 0 0 4)、未比較ステップで同一のハッシュ値が格納されていないと判断した場合は、データを重複して格納すること除去するステップ (S 2 0 0 2) を含む場合について説明したいが、これに限るものではない。

10

20

30

40

50

【0130】

さらに、除去するステップは、受信したデータのビット列を代表するハッシュ値と同一のハッシュ値を持つデータが格納されているディスク291～298が起動しているか否かを省電力グループ管理テーブル222に基づいて判断する状態を判断するステップと（S3001）、この状態を判断するステップでディスク291～298が起動中であると判断した場合は、ディスク291～298からハッシュ値と対応するデータを読み出し、その読み出したデータと受信したデータとが一致するか否かをデータ判断するステップと（S3005）、このデータ判断するステップでデータが一致すると判断した場合は、重複データ管理テーブル224の読み出したデータの論理ユニット番号と対応させて受信したデータのホスト100が指定するホスト側指定番号を格納する第3のステップと（S3006）、データ判断するステップでデータが一致しないと判断した場合は、読み出したRAIDグループと対応するディスク291～298に受信したデータを格納するとともに受信データの管理情報を重複データ管理テーブル224に格納する第4のステップ（S3007, S3008）を含む場合について説明したが、これに限るものではない。

10

【0131】

さらに、除去するステップは、状態を判断するステップでディスク291～298が起動中でないと判断した場合は、省電力グループ管理テーブル222で管理されるRAIDグループの起動ができるか否かを判断する起動を判断するステップ（S3002）を含み、起動を判断するステップでRAIDグループの起動ができると判断した場合にディスクアレイコントローラ240他のRAIDグループと対応するディスク291～298の電源をONし、ディスク291～298の電源をONした場合に、データ判断するステップはデータ一致か否かの判断を行なう場合について説明したが、これに限るものではない。

20

【0132】

さらに、除去するステップは、起動を判断するステップでRAIDグループの起動ができないと判断した場合は、起動しているRAIDグループと対応するディスク291～298に受信したデータを格納するとともに受信データの管理情報を重複データ管理テーブルに格納する第5のステップ（S3003, S3004）を含む場合について説明したが、これに限るものではない。

【0133】

ディスクアレイ200は、データのビット列を代表するハッシュ値、起動候補のRAIDグループ及びハッシュ値のカウント数を含む起動候補情報を格納する起動候補省電力グループテーブル227を備え、重複未比較データ管理テーブル225に未比較データ情報が格納されている場合に、ハッシュ値が同一なデータを同一のRAIDグループとなるように再配置する最適化処理を行うステップ（S4001からS4008）を有する場合について説明したが、これに限るものではない。

30

【産業上の利用可能性】

【0134】

本発明は、ストレージシステム及びストレージシステムの消費電力低減方法に広く適用することができる。

【図面の簡単な説明】

40

【0135】

【図1】本発明の第1の実施形態に係わるストレージシステムの構成を示す図である。

【図2】同実施形態に係わる省電力管理テーブルを示す図である。

【図3】同実施形態に係わる重複データ管理テーブルを示す図である。

【図4】同実施形態に係わるデータプール管理テーブルを示す図である。

【図5】同実施形態に係わる重複未比較データ管理テーブルを示す図である。

【図6】同実施形態に係わるライト処理を示すフローチャートである。

【図7】同実施形態に係わる重複除去の判断処理を示すフローチャートである。

【図8】同実施形態に係わる重複除去の実施処理を示すフローチャートである。

【図9】同実施形態に係わるデータ最適化処理の一例を示すフローチャートである。

50

【図10】同実施形態に係わるストレージシステムの構成の変形例を示す図である。

【図11】本発明の第2の実施形態に係わる起動候補省電力グループ管理テーブルを示す図である。

【図12】本発明の第3の実施形態に係わる重複データ管理テーブルを示す図である。

【図13】同実施形態に係わる重複未比較データ管理テーブルを示す図である。

【図14】同実施形態に係わるデータ最適化処理の一例を示すフローチャートである。

【符号の説明】

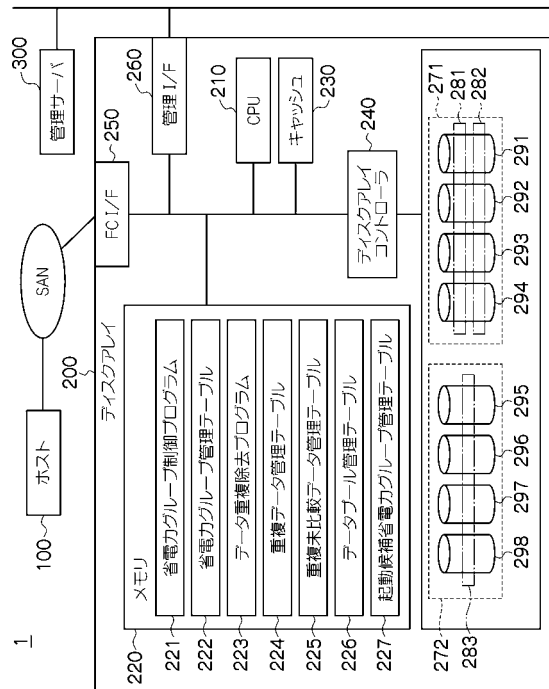
【0136】

1, 2...ストレージシステム、100...ホスト、200...ディスクアレイ、300...管理サーバ、210...CPU、220...メモリ、221...省電力グループ制御プログラム、222...省電力グループ管理テーブル、223...データ重複除去プログラム、224, 228...重複データ管理テーブル、225, 229...重複未比較データ管理テーブル、226...データプール管理テーブル、227...起動候補省電力グループ管理テーブル、240...ディスクアレイコントローラ、271...RAIDグループ、281~283...論理ユニット、291~298...ディスク、401...電力制御サーバ

10

【図1】

図1



【図2】

図2

222

省電力識別番号	状態	RAIDグループ番号	LU番号	ボリューム利用率
P-01	稼働中	RG01	LU01, LU02	50%
P-02	停止	RG02	LU03	10%
P-03	停止	RG03	LU04, LU05	10%
P-04	停止	RG04	LU06	0%

2221 2222 2223 2224 2225

【図3】

図3

224

ハッシュ値	実データ格納アドレス (LU番号:LBA)	ホスト側指定アドレス (LU番号:LBA)
a	LU01:100000	001:101000
	LU01:100001	001:103000
b	LU02:100100	001:114000
c	LU02:102000	001:134000

2241 2242 2243

【図4】

図4

2261	2262	2263
ホスト	ホスト側ボリューム番号	実ボリュームLU番号
A	001	LU01, LU02, LU03, LU04 LU05, LU06

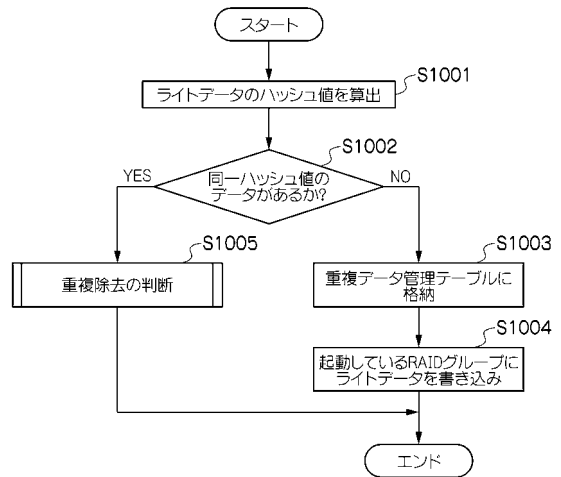
【図5】

図5

2251	2252	2253
データ格納アドレス (LU番号:IBA)	ホスト側指定アドレス	ハッシュ値
LU03:100000	001:109000	a
LU03:100100	001:109500	b
LU04:102000	001:107000	a

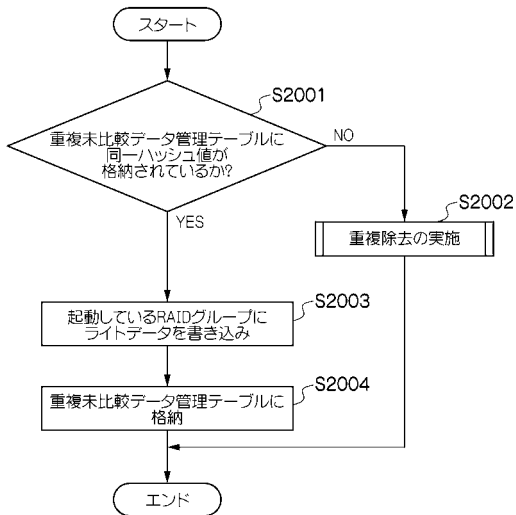
【図6】

図6



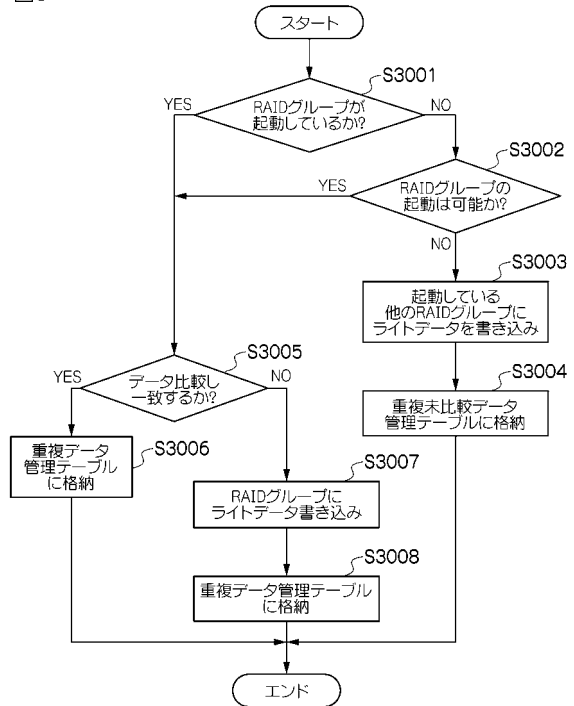
【図7】

図7



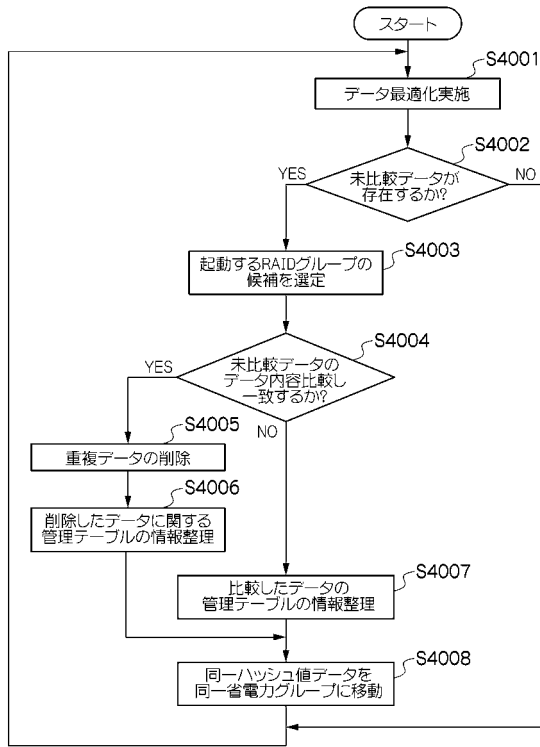
【図8】

図8



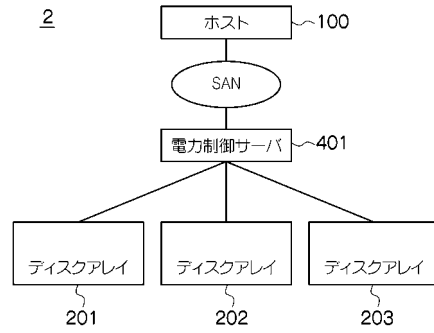
【図9】

図9



【図10】

図10



【図11】

図11

227

2271	2272	2273
ハッシュ値	起動候補省電力グループ	同一ハッシュ値データ数
a	RG01	2
	RG02	1
	RG03	1
b	RG01	1
	RG02	1

【図12】

図12

228

2281	2282	2283	2284
ハッシュ値	実データ格納アドレス (LU番号:LBA)	ホスト側指定アドレス (LU番号:LBA)	関連アドレス (LU番号:LBA)
a	LU01:100000	001:101000	LU02:100100
		001:103000	--
	LU01:100001	001:104000	LU03:100100
b	LU02:100100	001:114000	LU01:100000
			LU02:102000
c	LU02:102000	001:134000	LU01:100000
			LU02:100100

【図13】

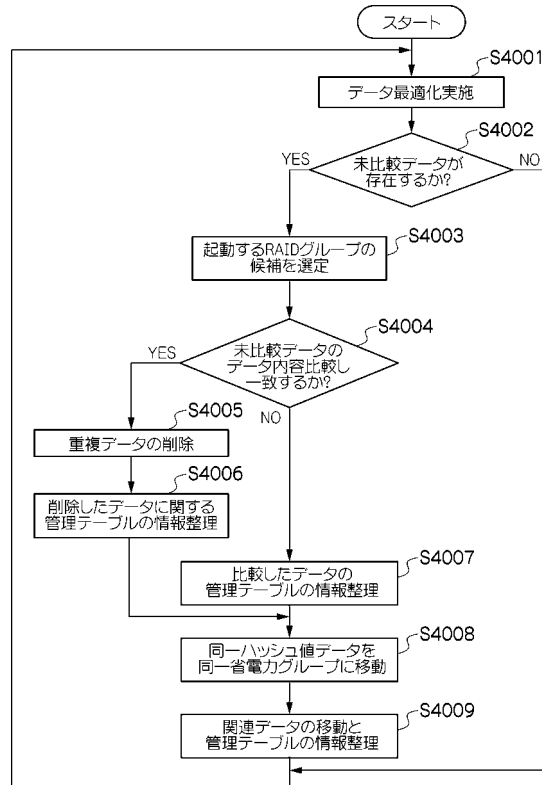
図13

229

2291	2292	2293	2294
データ仮格納アドレス (LU番号:LBA)	ホスト側指定アドレス	ハッシュ値	関連アドレス (LU番号:LBA)
LU03:100000	001:109000	a	--
LU03:100100	001:109500	b	LU01:100001
LU04:102000	001:107000	a	--

【図14】

図14



フロントページの続き

- (56)参考文献 国際公開第2007/089502(WO, A1)
特開2007-133478(JP, A)
特開2007-156597(JP, A)
特開2009-020858(JP, A)
特開2009-080788(JP, A)
米国特許出願公開第2005/0216669(US, A1)

- (58)調査した分野(Int.Cl., DB名)
G06F 3/06
G06F 12/00
G06F 13/10