

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4971895号
(P4971895)

(45) 発行日 平成24年7月11日(2012.7.11)

(24) 登録日 平成24年4月13日(2012.4.13)

(51) Int.Cl. F I
G 0 6 F 3 / 0 6 (2006.01) G 0 6 F 3 / 0 6 3 0 1 A

請求項の数 17 (全 19 頁)

(21) 出願番号	特願2007-178830 (P2007-178830)	(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7月6日(2007.7.6)	(74) 代理人	100093861 弁理士 大賀 真司
(65) 公開番号	特開2009-15724 (P2009-15724A)	(72) 発明者	水野 潤 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1月22日(2009.1.22)	(72) 発明者	那須 弘志 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
審査請求日	平成22年1月22日(2010.1.22)		

最終頁に続く

(54) 【発明の名称】 記憶システム及び消費電力削減方法並びに情報処理装置

(57) 【特許請求の範囲】

【請求項1】

ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する複数の情報処理装置と、前記アーカイブデータを格納する物理ディスクを有するストレージ装置とを含む記憶システムにおいて、

前記ストレージ装置は、

前記物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を管理する管理部

を備え、

前記情報処理装置は、

前記管理部により管理されている前記電源状態情報を取得する取得部と、

前記取得部により取得された前記電源状態情報に基づいて、前記複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する決定部とを備えることを特徴とする記憶システム。

【請求項2】

前記決定部は、

前記電源状態がオンの物理ディスクを使用する情報処理装置があるか否かを判断し、当該情報処理装置がある場合には、当該情報処理装置を選択する

ことを特徴とする請求項1に記載の記憶システム。

【請求項3】

前記決定部は、

前記電源状態がオンの物理ディスクを使用する情報処理装置があるか否かを判断し、当該情報処理装置がない場合には、前記アーカイブデータを読み出し又は書き込む情報処理装置が使用する前記物理ディスクの前記電源状態をオンにする

ことを特徴とする請求項 1 に記載の記憶システム。

【請求項 4】

前記決定部は、

前記アーカイブデータを読み出し又は書き込む情報処理装置が複数存在する場合には、複数の当該情報処理装置の中から、所定の情報処理装置を選択する

ことを特徴とする請求項 2 又は請求項 3 に記載の記憶システム。

10

【請求項 5】

前記管理部は、

前記物理ディスクの電源状態をオフに変更した時刻である電源状態変更時刻を管理し、

前記取得部は、

前記管理部により管理されている前記電源状態変更時刻を取得し、

前記決定部は、

前記電源状態変更時刻が最も早い情報処理装置を選択する

ことを特徴とする請求項 4 に記載の記憶システム。

【請求項 6】

前記決定部は、

使用済み容量の少ない物理ディスクを使用する情報処理装置を選択する

ことを特徴とする請求項 4 に記載の記憶システム。

20

【請求項 7】

前記決定部は、

残容量の多い物理ディスクを使用する情報処理装置を選択する

ことを特徴とする請求項 4 に記載の記憶システム。

【請求項 8】

前記決定部は、

前記アーカイブデータを読み出し又は書き込む前記情報処理装置として、前記電源状態がオンの物理ディスクを使用する前記情報処理装置があるか否かを判断し、該情報処理装置の数が前記ホスト計算機から送信された冗長度を満たしていない場合に、該冗長度に対応する前記情報処理装置を選択して、該情報処理装置が使用する前記物理ディスクの電源状態をオンにする

30

ことを特徴とする請求項 1 に記載の記憶システム。

【請求項 9】

ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する複数の情報処理装置と、前記アーカイブデータを格納する物理ディスクを有するストレージ装置とを含む記憶システムの消費電力削減方法において、

前記物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を取得する第 1 のステップと、

40

前記第 1 のステップにおいて取得した前記電源状態情報に基づいて、前記複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する第 2 のステップと

を備えることを特徴とする消費電力削減方法。

【請求項 10】

前記第 2 のステップでは、

前記電源状態がオンの物理ディスクを使用する情報処理装置があるか否かを判断し、当該情報処理装置がある場合には、当該情報処理装置を選択する

ことを特徴とする請求項 9 に記載の消費電力削減方法。

50

【請求項 1 1】

前記第 2 のステップでは、

前記電源状態がオンの物理ディスクを使用する情報処理装置があるか否かを判断し、当該情報処理装置がない場合には、前記アーカイブデータを読み出し又は書き込む情報処理装置が使用する前記物理ディスクの前記電源状態をオンにする

ことを特徴とする請求項 9 に記載の消費電力削減方法。

【請求項 1 2】

前記第 2 のステップでは、

前記アーカイブデータを読み出し又は書き込む情報処理装置が複数存在する場合には、複数の当該情報処理装置の中から、所定の情報処理装置を選択する

ことを特徴とする請求項 1 0 又は請求項 1 1 に記載の消費電力削減方法。

10

【請求項 1 3】

前記第 1 のステップでは、

前記物理ディスクの電源状態をオフに変更した時刻である電源状態変更時刻を取得し、

前記第 2 のステップでは、

前記電源状態変更時刻が最も早い情報処理装置を選択する

ことを特徴とする請求項 1 2 に記載の消費電力削減方法。

【請求項 1 4】

前記第 2 のステップでは、

使用済み容量の少ない物理ディスクを使用する情報処理装置を選択する

ことを特徴とする請求項 1 2 に記載の消費電力削減方法。

20

【請求項 1 5】

前記第 2 のステップでは、

残容量の多い物理ディスクを使用する情報処理装置を選択する

ことを特徴とする請求項 1 2 に記載の消費電力削減方法。

【請求項 1 6】

前記第 2 のステップでは、

前記アーカイブデータを読み出し又は書き込む前記情報処理装置として、前記電源状態がオンの物理ディスクを使用する前記情報処理装置があるか否かを判断し、該情報処理装置の数が前記ホスト計算機から送信された冗長度を満たしていない場合に、該冗長度に対応する前記情報処理装置を選択して、該情報処理装置が使用する前記物理ディスクの電源状態をオンにする

ことを特徴とする請求項 9 に記載の消費電力削減方法。

30

【請求項 1 7】

ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する情報処理装置において、

前記アーカイブデータを格納するストレージ装置の物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を取得する取得部と、

前記取得部により取得された前記電源状態情報に基づいて、複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する決定部と

を備えることを特徴とする情報処理装置。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶システム及び消費電力削減方法並びに情報処理装置に関し、例えば、ホスト計算機から送信されるデータを格納するアーカイブシステムに適用して好適なものである。

【背景技術】

【0002】

一般に、アーカイブシステムは、それぞれの業務を行うホスト計算機と、ホスト計算機

50

の指示によりデータを読み書きするアーカイブ装置により構成される。アーカイブ装置は、アーカイブノードと呼ばれる計算機（情報処理装置）と、アーカイブノードの指示によりデータを読み書きするストレージ装置から構成される。アーカイブ装置は、ホスト計算機に対して、記憶領域を提供する。アーカイブノードは、ホスト計算機のデータ読み書きの指示を受けると、ストレージ装置にデータの読み書きを指示する。ストレージ装置は、物理ディスクを複数の記憶領域に分割して管理する。ストレージ装置は、アーカイブノードに対して記憶領域を論理ボリュームとして提供する。論理ボリュームは、複数のセグメントから構成され、それぞれのセグメントに対して、物理ディスク上の記憶領域を割り当てることでホスト計算機の指示によるデータの読み書きを可能とする。

【0003】

10

例えば、アーカイブ装置で複数のアーカイブノードを用いてクラスタを構成し、ホスト計算機の指定する冗長度に応じて、アーカイブデータを複数のアーカイブノードに書き込むことにより、一部のアーカイブノードに障害が発生した場合にも、ホスト計算機がアーカイブデータにアクセス可能とする分散アーカイブ技術が特許文献1で公開されている。

【0004】

また、ストレージ装置の特定のハードディスクで構成される全ての論理ボリュームに対して一定期間アクセスが無い場合に、論理ボリュームを構成する前記ハードディスクの電源を遮断もしくは節電モードにする省電力技術が特許文献2で公開されている。

【特許文献1】国際公開第2005/043323号パンフレット

【特許文献2】特開2000-293314号公報

20

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、従来技術の分散アーカイブ技術と、省電力技術を組み合わせて、分散アーカイブ技術を適用したアーカイブシステムにおけるストレージ装置の消費電力を、省電力技術を用いて削減する場合、以下の問題が存在する。

【0006】

分散アーカイブ技術では、アーカイブシステムに書き込んだアーカイブデータがクラスタを構成するアーカイブノードに分散するため、アーカイブノードがストレージ装置上の論理ボリュームに平均的にアクセスし、論理ボリュームにアクセスの無い期間は短くなる。

30

【0007】

そのため、論理ボリュームを構成するハードディスクの電源の遮断もしくは節電モードへの移行ができず、ストレージ装置の消費電力を削減できないという問題が存在する。

【0008】

本発明は以上の点を考慮してなされたもので、ストレージ装置の消費電力を削減させ得る記憶システム及び消費電力削減方法並びに情報処理装置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】

【0009】

40

かかる課題を解決するために本発明の一形態においては、ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する複数の情報処理装置と、前記アーカイブデータを格納する物理ディスクを有するストレージ装置とを含む記憶システムにおいて、前記ストレージ装置は、前記物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を管理する管理部を備え、前記情報処理装置は、前記管理部により管理されている前記電源状態情報を取得する取得部と、前記取得部により取得された前記電源状態情報に基づいて、前記複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する決定部とを備える。

【0010】

50

また、本発明の一形態においては、ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する複数の情報処理装置と、前記アーカイブデータを格納する物理ディスクを有するストレージ装置とを含む記憶システムの消費電力削減方法において、前記物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を取得する第1のステップと、前記第1のステップにおいて取得した前記電源状態情報に基づいて、前記複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する第2のステップとを備える。

【0011】

さらに、本発明の一形態においては、ホスト計算機から送信される冗長度に応じて、前記ホスト計算機から送信されるデータを複製してアーカイブデータを作成する情報処理装置において、前記アーカイブデータを格納するストレージ装置の物理ディスクの電源状態がオンであるか否かの情報である電源状態情報を取得する取得部と、前記取得部により取得された前記電源状態情報に基づいて、複数の前記情報処理装置から、前記アーカイブデータを読み出し又は書き込む情報処理装置を決定する決定部とを備える。

10

【0012】

従って、アーカイブデータの読み出し時に、電源がオフである物理ディスクに対するアクセスを減少することができるため、物理ディスクへのアクセスが無い期間を長くすることができ、その結果、物理ディスクの電源をオフにする時間を長くすることができる。

【発明の効果】

20

【0013】

本発明によれば、ストレージ装置の消費電力を削減させ得る記憶システム及び消費電力削減方法並びに情報処理装置を実現できる。

【発明を実施するための最良の形態】

【0014】

以下図面について、本発明の一実施の形態を詳述する。

【0015】

(1) 第1の実施の形態によるアーカイブシステムの構成

図1は、本発明の第1の実施の形態が適用されたアーカイブシステム(記憶システム)1の概略図である。図示するように、本実施の形態のアーカイブシステム1は、アーカイブクラスタ2と、1又は複数のホスト計算機3と、1又は複数のストレージ装置4と、から構成される。

30

【0016】

アーカイブクラスタ2は、1又は複数のアーカイブノード(情報処理装置)5から構成される。本実施の形態では、ホスト計算機3は、ローカルエリアネットワーク6を介してアーカイブノード5と接続する。アーカイブノード5は、ストレージエリアネットワーク7を介してストレージ装置4と接続する。また、アーカイブノード5は、管理用ネットワーク8を介して、ストレージ装置4と接続するものとしているが、ローカルエリアネットワーク6と、ストレージエリアネットワーク7と、管理用ネットワーク8とは同一のものであってもよい。

40

【0017】

図2は、ホスト計算機3の構成を示した図である。ホスト計算機3は、CPU(Central Processing Unit)11と、メモリ12と、ハードディスク13と、入力装置14と、出力装置15と、アーカイブノード5と通信を行う通信ポート16とから構成される。

【0018】

ホスト計算機3は、アーカイブノード5に対して、冗長度を指定してアーカイブデータの保存を指示する。アーカイブノード5は、指定された冗長度の個数になるようにアーカイブデータを複製し、それぞれのアーカイブデータを異なるアーカイブノード5に送信する。そして、各アーカイブノード5は、ストレージ装置4にアーカイブデータを保存する。

50

【 0 0 1 9 】

また、ホスト計算機 3 は、アーカイブノード 5 に対して、アーカイブデータの読み出しを指示する。アーカイブノード 5 はアーカイブデータの読み出しの要求を受けると、アーカイブクラスタ 2 内に保持するアーカイブデータのうち、要求を受けたアーカイブデータをホスト計算機 3 に送信する。

【 0 0 2 0 】

図 3 は、ストレージ装置 4 の構成を示した図である。ストレージ装置 4 は、当該ストレージ装置 4 の制御を行うコントローラ 2 1 と、メモリ 2 2 と、アーカイブノード 5 との通信に利用する I/O ポート 2 3 や管理ポート 2 4 と、物理ディスク 2 5 とから構成される。

【 0 0 2 1 】

ストレージ装置 4 は、1 又は複数の物理ディスク 2 5 を用いて、論理的なボリュームである論理ボリューム 2 6 を保持し、アーカイブノード 5 からの論理ボリューム 2 6 に対する I/O の要求を受け付け、アーカイブデータを保存する。なお、ストレージ装置 4 は、例えば、論理ボリューム 2 6 を保持せず、個々の物理ディスク 2 5 にアーカイブデータを保存するようにしても良い。

【 0 0 2 2 】

メモリ 2 2 内には、物理ディスク 2 5 及び論理ボリューム 2 6 の関係を管理する論理ボリューム管理表 2 7 と、物理ディスク 2 5 の電源状態を管理する物理ディスク電源状態管理表 2 8 と、電源オフ (OFF) や節電モード等に移行するための期間を管理する省電力待ち時間管理表 2 9 とが含まれる。

【 0 0 2 3 】

図 5 は、論理ボリューム管理表 2 7 の構成を示した図である。論理ボリューム管理表 2 7 は、論理ボリューム 2 6 を識別する論理ボリューム ID 2 7 A の欄と、論理ボリューム 2 6 を構成する物理ディスク 2 5 を識別する物理ディスク ID 2 7 B の欄から構成される。

【 0 0 2 4 】

図 6 は、物理ディスク電源状態管理表 2 8 の構成を示した図である。物理ディスク電源状態管理表 2 8 は、物理ディスク 2 5 を識別する物理ディスク ID 2 8 A の欄と、物理ディスク 2 5 の電源状態を保持する電源状態 2 8 B の欄と、電源状態が変更した時刻を保持する電源状態変更時刻 2 8 C の欄と、から構成される。電源状態 2 8 B の欄には、電源オン (ON) 状態を示す「オン」と、電源オフ状態を示す「オフ」と、節電モード状態を示す「節電」のいずれかの値を保持する。

【 0 0 2 5 】

図 7 は、省電力待ち時間管理表 2 9 の構成を示した図である。省電力待ち時間管理表 2 9 は、電源オフや節電モードへの移行の契機となる物理ディスク 2 5 に対するアクセスが無い期間を定める省電力待ち時間 2 9 A の欄から構成される。

【 0 0 2 6 】

ストレージ装置 4 は、物理ディスク 2 5 に対するアクセスを監視し、物理ディスク 2 5 に対するアクセスが無い期間が、省電力待ち時間管理表 2 9 の省電力待ち時間 2 9 A の欄に保持した期間を超過したら、該当する物理ディスク 2 5 の電源をオフにし、物理ディスク電源状態管理表 2 8 に含まれるレコードのうち該当する物理ディスク 2 5 のレコードの電源状態 2 8 B の欄の値を「オフ」にする。

【 0 0 2 7 】

本実施の形態では、物理ディスク 2 5 に対するアクセスの無い期間が、省電力待ち時間 2 9 A を超過した場合に、物理ディスク 2 5 の電源をオフにするとしているが、物理ディスク 2 5 を節電モードにして、物理ディスク電源状態管理表 2 8 に含まれるレコードのうち該当する物理ディスクのレコードの電源状態 2 8 B の欄の値を「節電」にしてもよい。

【 0 0 2 8 】

また、ストレージ装置 4 は、論理ボリューム ID を指定した電源オンの要求を受け付けると、論理ボリューム管理表 2 7 を参照し、前記要求で指定された論理ボリューム ID に

10

20

30

40

50

関係する物理ディスク 25 を検索し、前記検索した物理ディスク 25 の電源をオンにする。

【0029】

さらに、ストレージ装置 4 は、アーカイブノード 5 から論理ボリューム情報の要求を受け付けると、論理ボリューム管理表 28 に含まれる情報を要求元へ送信する。また、ストレージ装置 4 は、アーカイブノード 5 から物理ディスク 25 の電源状態についての情報の要求を受け付けると、物理ディスク電源状態管理表 28 に含まれる情報を要求元へ送信する。

【0030】

さらに、ストレージ装置 4 は、物理ディスク 25 の電源状態が変化した場合に、あらかじめ登録されたホスト計算機 3 に対して、電源状態が変化したことを通知する電源状態変更通知を送信する。

10

【0031】

図 4 は、アーカイブノード 5 の構成を示した図である。アーカイブノード 5 は、CPU 31 と、メモリ 32 と、ハードディスク 33 と、入力装置 34 と、出力装置 35 と、管理ポート 36 と、ホスト計算機 3 とアーカイブデータを通信する通信ポート 37 と、ストレージ装置 4 とアーカイブデータを通信する I/O ポート 38 と、から構成される。

【0032】

ハードディスク 33 内には、アーカイブノード毎にアーカイブノード 5 が使用する論理ボリューム 26 の使用済み容量や電源状態を管理するボリューム管理表 39 と、アーカイブデータ毎にデータ保存先となるアーカイブノード 5 を管理するマッピング管理表 40 と、が含まれる。

20

【0033】

図 8 は、ボリューム管理表 39 の構成を示した図である。ボリューム管理表 39 は、アーカイブノード 5 を識別するアーカイブノード ID 39A の欄と、アーカイブノード 5 が使用するストレージ装置 4 を識別するストレージ装置 ID 39B の欄と、論理ボリューム 26 を識別する論理ボリューム ID 39C の欄と、アーカイブノード 5 が使用済みの容量を保持する使用済み容量 39D の欄と、論理ボリューム 26 を構成する物理ディスク 25 の電源状態を保持する電源状態 39E の欄と、電源状態が変更した時刻を保持する電源状態変更時刻 39F の欄と、から構成される。電源状態 39D の欄には、電源オン状態を示す「オン」と、電源オフ状態を示す「オフ」と、節電モード状態を示す「節電」のいずれかの値を保持する。

30

【0034】

なお、ボリューム管理表 39 で管理する情報のうち、アーカイブノード ID 39A の欄と、ストレージ装置 ID 39B の欄と、論理ボリューム ID 39C の欄は、あらかじめ管理者が設定ファイルや、GUI (Graphical User Interface) や CLI (Command Line Interface) などの設定用インタフェースを用いて設定するものとする。

【0035】

図 9 は、マッピング管理表 40 の構成を示した図である。マッピング管理表 40 は、アーカイブデータを識別するアーカイブデータ ID 40A の欄と、ホスト計算機 3 が指定したアーカイブデータの冗長度を保持する冗長度 40B の欄と、アーカイブデータの保存先であるアーカイブノード 5 の情報を保持するデータ保存先ノード 40C の欄と、から構成される。同一のアーカイブデータの保存先となるアーカイブノード 5 が複数存在する場合には、データ保存先ノード 40C の欄に、アーカイブノード 5 のアーカイブノード ID を例えば、カンマ区切りで記入する。

40

【0036】

また、ハードディスク 33 内には、ストレージ装置 4 の論理ボリューム管理表 27 や物理ディスク電源状態管理表 28 の情報を取得する電源状態情報生成プログラム 41 と、読み出しを要求されたアーカイブデータを読み取るアーカイブノード 5 を決定するデータ提供プログラム 42 と、が設けられている。

50

【 0 0 3 7 】

以下で、本実施の形態における、アーカイブノード5の動作について説明する。

【 0 0 3 8 】

図10は、アーカイブノード5における電源状態情報生成プログラム41の処理の流れを表わしたフローチャートである。

【 0 0 3 9 】

なお、以下においては、各プログラムに基づくアーカイブノード5のCPU31の処理内容を明確化するため、各種処理の処理主体をプログラムとして説明するが、実際上は、そのプログラムに基づいて対応するCPU31がその処理を行うことは言うまでもない。

【 0 0 4 0 】

まず、電源状態情報生成プログラム41は、ストレージ装置4から電源状態変更通知を受信する(S P 1)。なお、この場合は、各アーカイブノード5の電源状態情報生成プログラム41がストレージ装置4から電源状態変更通知を受信する。

【 0 0 4 1 】

次に、電源状態情報生成プログラム41は、電源状態変更通知の送信元であるストレージ装置4に対して、論理ボリューム情報の要求を送信し、その応答として、ストレージ装置4の論理ボリューム管理表27の情報を取得する(S P 2)。

【 0 0 4 2 】

次に、電源状態情報生成プログラム41は、電源状態変更通知の送信元であるストレージ装置4に対して、物理ディスク電源状態情報の要求を送信し、その応答として、ストレージ装置4の物理ディスク電源状態管理表28の情報を取得する(S P 3)。

【 0 0 4 3 】

次に、電源状態情報生成プログラム41は、アーカイブノード5のボリューム管理表39の全てのレコードに対して、ストレージ装置ID39Bに記載されたストレージ装置4から取得した論理ボリューム情報を用いて、アーカイブノード5と物理ディスク25の関係を結びつける。(S P 4)。

【 0 0 4 4 】

次に、電源状態情報生成プログラム41は、前記結び付けられた物理ディスク25の物理ディスクIDをキーにして、ストレージ装置ID39Bに記載されたストレージ装置4から取得した物理ディスク電源状態情報を検索し、物理ディスクIDが一致したレコードの電源状態39Eの値をボリューム管理表39の電源状態39Eの欄に記入する(S P 5)。

【 0 0 4 5 】

以上が、アーカイブノード5における電源状態情報生成プログラム41の処理の説明である。

【 0 0 4 6 】

本実施の形態では、ストレージ装置4から電源状態変更通知を受信することを電源情報生成プログラムの処理を開始する契機としているが、あらかじめ登録済みのストレージ装置4に対して定期的に電源状態情報生成プログラム41の処理を実行する形態をとっても良い。

【 0 0 4 7 】

また、アーカイブノード5がホスト計算機3からI/O要求を受信したことを、電源状態情報生成プログラム41の開始の契機としても良い。

【 0 0 4 8 】

図11は、アーカイブノード5におけるデータ提供プログラム42の処理の流れを表わしたフローチャートである。

【 0 0 4 9 】

まず、データ提供プログラム42は、ホスト計算機3からアーカイブデータの読み出し要求として、アーカイブデータIDを受信する(S P 11)。なお、この場合は、所定のアーカイブノード5のデータ提供プログラム42がホスト計算機3からアーカイブデータ

10

20

30

40

50

の読み出し要求として、アーカイブデータIDを受信する。

【0050】

次に、データ提供プログラム42は、マッピング管理表40に含まれるレコードから読み出し要求のアーカイブデータIDとアーカイブデータID40Aの欄の値が等しいレコードを検索し、検索結果のレコードのデータ保存先ノード40Cの欄に記載されたアーカイブノードIDを全て取得する(S P 1 2)。

【0051】

次に、データ提供プログラム42は、ボリューム管理表39に含まれるレコードから、前記取得したアーカイブノードIDとアーカイブノードID39Aの欄の値が等しい、かつ、電源状態39Eの欄に記載された値が「オン」であるレコードを検索し、該当するレコードの数を確認する(S P 1 3)。

10

【0052】

次に、データ提供プログラム42は、該当するレコード数の確認(S P 1 3)の結果、電源オンのアーカイブノード5があるか否かをチェックする(S P 1 4)。

【0053】

そして、データ提供プログラム42は、前記検索した電源オンのレコードの数が「0」である、すなわち、電源オンのアーカイブノード5がない場合(S P 1 4 : N O)には、アーカイブデータを保存しているアーカイブノード5から、アーカイブデータの読み出し対象とするアーカイブノード5を1つ選択する(S P 1 5)。

【0054】

20

なお、データ提供プログラム42は、候補が複数存在する場合には、例えば、電源状態変更時刻39Fの欄の値を参照し、電源状態変更時刻が最も早いアーカイブノード5を選択する。また、データ提供プログラム42は、例えば、使用済み容量の少ないアーカイブノード5を選択しても良いし、例えば、残容量の多いアーカイブノード5を選択しても良い。

【0055】

次に、データ提供プログラム42は、ボリューム管理表39から、前記選択したアーカイブノード5のレコードを検索し、ストレージ装置IDの値と、論理ボリュームIDの値を取得し、取得したストレージ装置IDのストレージ装置4に対して、取得した論理ボリュームIDを指定して電源オンの要求を送信する(S P 1 6)。

30

【0056】

これに対して、データ提供プログラム42は、前記検索した電源オンのレコードの数が「0」でない、すなわち、電源オンのアーカイブノード5がある場合(S P 1 4 : Y E S)には、前記検索した電源オンのレコードのアーカイブノード5から、アーカイブデータの読み出し対象とするアーカイブノード5を1つ選択する(S P 1 7)。

【0057】

なお、データ提供プログラム42は、候補が複数存在する場合には、例えば、電源状態変更時刻39Fの欄の値を参照し、電源状態変更時刻が最も早いアーカイブノード5を選択する。また、データ提供プログラム42は、例えば、使用済み容量の少ないアーカイブノード5を選択しても良いし、例えば、残容量の多いアーカイブノード5を選択しても良い。

40

【0058】

次に、データ提供プログラム42は、前記アーカイブデータの読み出し対象として選択したアーカイブノード5に対して、アーカイブデータの読み出しを指示し、得られたアーカイブデータを、読み出し要求を行ったホスト計算機3に送信する(S P 1 8)。

【0059】

以上が、アーカイブノード5におけるデータ提供プログラム42の処理の説明である。

【0060】

また、アーカイブノード5がアーカイブデータの書き込み要求を受けた場合の動作については、従来技術で開示された方法を用いても良いし、他の方法でアーカイブデータの書

50

き込みを行っても良い。

【 0 0 6 1 】

以上が、本実施例におけるアーカイブノード 5 の動作の説明である。

【 0 0 6 2 】

本実施の形態では、アーカイブノード 5 で、電源状態情報生成プログラム 4 1 が動作するものとしているが、管理サーバ等の別装置上で、電源状態情報生成プログラム 4 1 が動作し、アーカイブノード 5 のボリューム管理表 3 9 を生成もしくは更新する構成であってもよい。

【 0 0 6 3 】

本実施の形態によれば、アーカイブシステム 1 では、アーカイブデータの読み出し時に、電源がオフ又は節電モードである物理ディスク 2 5 により構成される論理ボリューム 2 6 に対するアクセスを減少することができるため、物理ディスク 2 5 へのアクセスが無い期間を長くすることができる。その結果、アーカイブシステム 1 では、物理ディスク 2 5 の電源をオフ又は節電モードにする時間を長くすることができ、かくして、ストレージ装置 4 の消費電力を削減するという効果が得られる。

10

【 0 0 6 4 】

(2) 第 2 の実施の形態によるアーカイブシステムの構成

以下、本発明の第 2 の実施の形態について図表を用いて説明する。第 2 の実施の形態におけるアーカイブシステム 1 の構成は、図 1 2 に示すように、第 1 の実施の形態のアーカイブシステム 1 の構成とほぼ同様であるが、アーカイブノード 5 のハードディスク 3 3 内から、データ提供プログラム 4 2 を削除し、ホスト計算機 3 から書き込みを指示されたアーカイブデータを保存するアーカイブノード 5 を決定する配置プログラム 5 1 を追加した部分が異なる。

20

【 0 0 6 5 】

以下、本実施の形態における、アーカイブノード 5 のハードディスク 3 3 内に設置された配置プログラム 5 1 の処理を説明する。

【 0 0 6 6 】

図 1 3 は、アーカイブノード 5 のハードディスク 3 3 内に設置された配置プログラムの処理の流れを表わしたフローチャートである。

【 0 0 6 7 】

まず、配置プログラム 5 1 は、ホスト計算機 3 からアーカイブデータの書き込み要求としてアーカイブデータ及び冗長度を受信する (S P 2 1) 。

30

【 0 0 6 8 】

次に、配置プログラム 5 1 は、ボリューム管理表 3 9 を参照し、アーカイブクラスタ 1 を構成するアーカイブノード 5 が使用する論理ボリューム 2 6 の電源状態を電源状態 3 9 E の欄の値により確認する (S P 2 2) 。

【 0 0 6 9 】

次に、配置プログラム 5 1 は、電源状態の確認 (S P 2 2) の結果、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 < 受信した冗長度の関係を満たすか否かをチェックする (S P 2 3) 。

40

【 0 0 7 0 】

そして、配置プログラム 5 1 は、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 < 受信した冗長度の関係を満たす場合 (S P 2 3 : Y E S) には、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 2 4) 。

【 0 0 7 1 】

次に、配置プログラム 5 1 は、冗長度から電源状態 3 9 E の欄の値が「オン」であるアーカイブノード 5 の個数を引いた数だけ、電源状態 3 9 E の欄の値が「オン」でない、すなわち、電源状態 3 9 E の欄の値が「オフ」であるアーカイブノード 5 を選択する (S P 2 5) 。

50

【 0 0 7 2 】

なお、配置プログラム 5 1 は、候補が複数存在する場合には、例えば、使用済み容量 3 9 D の欄の値が小さいアーカイブノード 5 から選択する。

【 0 0 7 3 】

本実施の形態では、使用済み容量の少ないアーカイブノード 5 から選択するとしているが、論理ボリューム 2 6 の残容量が多いアーカイブノード 5 から選択するようにしても良く、省電力状態の長い論理ボリューム 2 6 を使うアーカイブノード 5 から選択するとしても良い。

【 0 0 7 4 】

次に、配置プログラム 5 1 は、選択したアーカイブノード 5 が使用する論理ボリューム 2 6 を構成する物理ディスク 2 5 の電源をオンにするように、ストレージ装置 4 に指示する (S P 2 6) 。

10

【 0 0 7 5 】

次に、配置プログラム 5 1 は、選択したアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 2 7) 。

【 0 0 7 6 】

これに対して、配置プログラム 5 1 は、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 受信した冗長度の関係を満たす、すなわち、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 < 受信した冗長度の関係を満たさない場合 (S P 2 3 : N O) には、電源状態が「オン」であるアーカイブノード 5 から、冗長度の数だけアーカイブノード 5 を選択し、選択したアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 2 8) 。

20

【 0 0 7 7 】

以上が、本実施の形態におけるアーカイブノード 5 のハードディスク 3 3 内に設置された配置プログラム 5 1 の処理の説明である。

【 0 0 7 8 】

また、アーカイブノード 5 がアーカイブデータの読み出し要求を受けた場合の動作については、従来技術で開示された方法を用いても良いし、第 1 の実施の形態で開示する方法を用いても良いし、他の方法を用いてアーカイブデータの読み出しを行っても良い。

【 0 0 7 9 】

30

本実施の形態によれば、アーカイブシステム 1 では、アーカイブデータの書き込み時に、電源がオフ又は節電モードである物理ディスク 2 5 により構成される論理ボリューム 2 6 に対するアクセスを減少することができるため、物理ディスク 2 5 へのアクセスが無い期間を長くすることができる。その結果、アーカイブシステム 1 では、物理ディスク 2 5 の電源をオフ又は節電モードにできる時間を長くすることができ、かくして、ストレージ装置 4 の消費電力を削減するという効果が得られる。

【 0 0 8 0 】

(3) 第 3 の実施の形態によるアーカイブシステムの構成

以下、本発明の第 3 の実施の形態について図表を用いて説明する。第 3 の実施の形態のアーカイブシステム 1 の構成は、第 1 の実施の形態及び第 2 の実施の形態のアーカイブシステム 1 の構成とほぼ同様であるが、アーカイブノード 5 のハードディスク 3 3 内に配置プログラム 6 1 及びデータ提供プログラム 6 2 を保持する。また、配置プログラム 6 1 の動作は、第 2 の実施の形態で示した配置プログラム 5 1 の動作と異なり、データ提供プログラム 6 2 の動作も、第 1 の実施の形態で示したデータ提供プログラム 4 2 の動作と異なる。

40

【 0 0 8 1 】

さらに、本実施の形態では、アーカイブノード 5 のハードディスク 3 3 内に全データ保持ノード管理表 6 3 が存在する。

【 0 0 8 2 】

図 1 5 は、全データ保持ノード管理表 6 3 の構成を示した図である。全データ保持ノード

50

ド管理表 6 3 は、全てのアーカイブデータを保持するアーカイブノード 5 のアーカイブノード ID を管理する全データ保持ノード 6 3 A の欄から構成される。

【 0 0 8 3 】

以下、本実施の形態におけるアーカイブノード 5 のハードディスク 3 3 内に設定された配置プログラム 6 1 の動作と、データ提供プログラム 6 2 の動作を説明する。

【 0 0 8 4 】

図 1 6 は、アーカイブノード 5 のハードディスク 3 3 内に設定された配置プログラム 6 1 の処理の流れを示したフローチャートである。

【 0 0 8 5 】

まず、配置プログラム 6 1 は、ホスト計算機 3 からアーカイブデータの書き込み要求として、アーカイブデータ及び冗長度を受信する (S P 3 1) 。

10

【 0 0 8 6 】

次に、配置プログラム 6 1 は、ボリューム管理表 3 9 を参照し、全データ保持ノード管理表 6 3 の全データ保持ノード 6 3 A の欄に記載されたアーカイブノード ID と、アーカイブノード ID 3 9 A が等しいレコードの電源状態 3 9 E の値を確認する (S P 3 2) 。

【 0 0 8 7 】

次に、配置プログラム 6 1 は、電源状態の確認 (S P 3 2) の結果、確認した電源状態 3 9 E の値が「オン」であるか否かをチェックする (S P 3 3) 。

【 0 0 8 8 】

そして、配置プログラム 6 1 は、確認した電源状態 3 9 E の値が「オン」以外、すなわち、「オン」でない場合 (S P 3 3 : N O) には、前記検索したレコードのストレージ装置 ID 3 9 B に記載されたストレージ装置 4 に対して、前記検索したレコードの論理ボリューム ID 3 9 C に記載された論理ボリューム ID を指定して物理ディスク 2 5 の電源オンを指示する (S P 3 4) 。

20

【 0 0 8 9 】

これに対して、配置プログラム 6 1 は、確認した電源状態 3 9 E の値が「オン」である場合 (S P 3 3 : Y E S) には、特に処理を行わない。

【 0 0 9 0 】

次に、配置プログラム 6 1 は、全データ保持ノード管理表 6 3 の全データ保持ノード 6 3 A の欄に記載されたアーカイブノード 5 にアーカイブデータの書き込みを指示する (S P 3 5) 。

30

【 0 0 9 1 】

次に、配置プログラム 6 1 は、ボリューム管理表 3 9 を参照し、アーカイブクラスタ 1 を構成するアーカイブノード 5 のうち全データ保持ノード管理表 6 3 の全データ保持ノード 6 3 A の欄に記載されたアーカイブノード 5 以外のアーカイブノード 5 が使用する論理ボリューム 2 6 の電源状態を電源状態 3 9 E の欄の値により確認する (S P 3 6) 。

【 0 0 9 2 】

次に、配置プログラム 6 1 は、電源状態の確認 (S P 3 6) の結果、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 < 受信した冗長度 - 1 の関係を満たすか否かをチェックする (S P 3 7) 。

40

【 0 0 9 3 】

そして、配置プログラム 6 1 は、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 < 受信した冗長度 - 1 の関係を満たす場合 (S P 3 7 : N O) には、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 3 8) 。

【 0 0 9 4 】

次に、配置プログラム 6 1 は、冗長度から電源状態 3 9 E の欄の値が「オン」であるアーカイブノード 5 の個数 + 1 を引いた数だけ、電源状態 3 9 E の欄の値が「オン」でない、すなわち、電源状態 3 9 E の欄の値が「オフ」であるアーカイブノード 5 を選択する (S P 3 9) 。

50

【 0 0 9 5 】

なお、配置プログラム 6 1 は、候補が複数存在する場合には、例えば、使用済み容量 3 9 D の欄の値が小さいアーカイブノード 5 から選択する。

【 0 0 9 6 】

本実施の形態では、使用済み容量の少ないアーカイブノード 5 から選択するとしているが、論理ボリューム 2 6 の残容量が多いアーカイブノード 5 から選択するようにしても良く、省電力状態の長い論理ボリューム 2 6 を使うアーカイブノード 5 から選択するとしても良い。

【 0 0 9 7 】

次に、配置プログラム 6 1 は、選択したアーカイブノード 5 が使用する論理ボリューム 2 6 を構成する物理ディスク 2 5 の電源をオンにするように、ストレージ装置 4 に指示する (S P 4 0) 。

10

【 0 0 9 8 】

次に、配置プログラム 6 1 は、選択したアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 4 1) 。

【 0 0 9 9 】

これに対して、配置プログラム 6 1 は、電源状態 3 9 E の欄の値が「オン」であるアーカイブノード数 受信した冗長度 - 1 の関係を満たす、すなわち、アーカイブノード数 < 受信した冗長度 - 1 の関係を満たさない場合 (S P 3 7 : N O) には、電源状態が「オン」であるアーカイブノード 5 から、冗長度の数 - 1 の個数だけアーカイブノード 5 を選択し、選択したアーカイブノード 5 に対して、アーカイブデータの書き込みを指示する (S P 4 2) 。

20

【 0 1 0 0 】

以上が、本実施の形態におけるアーカイブノード 5 のハードディスク 3 3 内に設置された配置プログラム 6 1 の処理の説明である。

【 0 1 0 1 】

図 1 7 は、アーカイブノード 5 のハードディスク 3 3 内に設定されたデータ提供プログラム 6 2 の処理の流れを示したフローチャートである。

【 0 1 0 2 】

まず、データ提供プログラム 6 2 は、ホスト計算機 3 からアーカイブデータの読み出し要求として、アーカイブデータ I D を受信する (S P 5 1) 。

30

【 0 1 0 3 】

次に、データ提供プログラム 6 2 は、ボリューム管理表 3 9 を参照し、全データ保持ノード管理表 6 3 の全データ保持ノード 6 3 A の欄に記載されたアーカイブノード I D と、アーカイブノード I D 3 9 A が等しいレコードの電源状態 3 9 E の値を確認する (S P 5 2) 。

【 0 1 0 4 】

次に、データ提供プログラム 6 2 は、電源状態の確認 (S P 5 2) の結果、確認した電源状態 3 9 E の値が「オン」であるか否かをチェックする (S P 5 3) 。

【 0 1 0 5 】

40

そして、データ提供プログラム 6 2 は、確認した電源状態 3 9 E の値が「オン」以外、すなわち、「オン」でない場合 (S P 5 3 : N O) には、前記検索したレコードのストレージ装置 I D 3 9 B に記載されたストレージ装置 4 に対して、前記検索したレコードの論理ボリューム I D 3 9 C に記載された論理ボリューム I D を指定して物理ディスク 2 5 の電源オンを指示する (S P 5 4) 。

【 0 1 0 6 】

これに対して、データ提供プログラム 6 2 は、確認した電源状態 3 9 E の値が「オン」である場合 (S P 5 3 : Y E S) には、特に処理を行わない。

【 0 1 0 7 】

次に、データ提供プログラム 6 2 は、全データ保持ノード管理表 6 3 の全データ保持ノ

50

ード63Aの欄に記載されたアーカイブノード5にアーカイブデータの読み出しを指示し、得られたデータを、読み出し要求を行ったホスト計算機3に送信する(S P 55)。

【0108】

以上が、本実施の形態におけるアーカイブノード5のハードディスク33内に設置されたデータ提供プログラム62の処理の説明である。

【0109】

本実施の形態によれば、アーカイブシステム1では、アーカイブデータの読み出し時に、特定の論理ボリューム26を構成する物理ディスク25以外の物理ディスク25の電源をオンにせずに、アーカイブデータを読み出すことができ、かくして、ストレージ装置4の消費電力を削減するという効果が得られる。

10

【産業上の利用可能性】

【0110】

本発明は、ホスト計算機から送信されるデータを格納するアーカイブシステムのほか、データを記憶するこの他種々の記憶システムに広く適用することができる。

【図面の簡単な説明】

【0111】

【図1】アーカイブシステムの構成を示した図である。

【図2】ホスト計算機の構成を示した図である。

【図3】ストレージ装置の構成を示した図である。

【図4】アーカイブノードの構成を示した図である。

20

【図5】論理ボリューム管理表の構成を示した図である。

【図6】物理ディスク電源状態管理表の構成を示した図である。

【図7】省電力待ち時間管理表の構成を示した図である。

【図8】ボリューム管理表の構成を示した図である。

【図9】マッピング管理表の構成を示した図である。

【図10】アーカイブノードにおける電源状態情報生成プログラムの処理の流れを表わしたフローチャートである。

【図11】アーカイブノードにおけるデータ提供プログラムの処理の流れを表わしたフローチャートである(第1の実施の形態)。

【図12】アーカイブノードの構成を示した図である(第2の実施の形態)。

30

【図13】アーカイブノードにおける配置プログラムの処理の流れを表わしたフローチャートである(第2の実施の形態)。

【図14】アーカイブノードの構成を示した図である(第3の実施の形態)。

【図15】全データ保持ノード管理表の構成を示した図である。

【図16】アーカイブノードにおける配置プログラムの処理の流れを表わしたフローチャートである(第3の実施の形態)。

【図17】アーカイブノードにおけるデータ提供プログラムの処理の流れを表わしたフローチャートである(第3の実施の形態)。

【符号の説明】

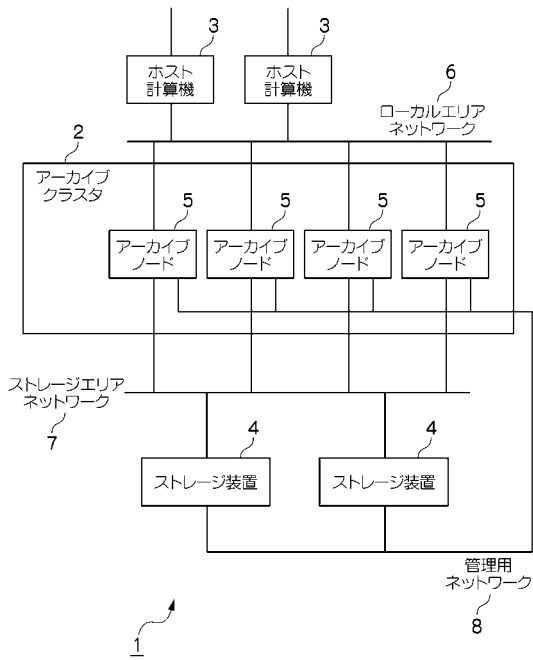
【0112】

40

1 アーカイブシステム、 2 アーカイブクラスタ、 3 ホスト計算機、 4 ストレージ装置、 5 アーカイブノード、 25 物理ディスク、 26 論理ボリューム、 27 論理ボリューム管理表、 28 物理ディスク電源状態管理表、 29 省電力待ち時間管理表、 31 アーカイブノードのCPU、 33 ハードディスク、 39 ボリューム管理表、 39E 電源状態、 40 マッピング管理表、 41 電源状態情報生成プログラム、 42、 62 データ提供プログラム、 51、 61 配置プログラム、 63 全データ保持ノード管理表

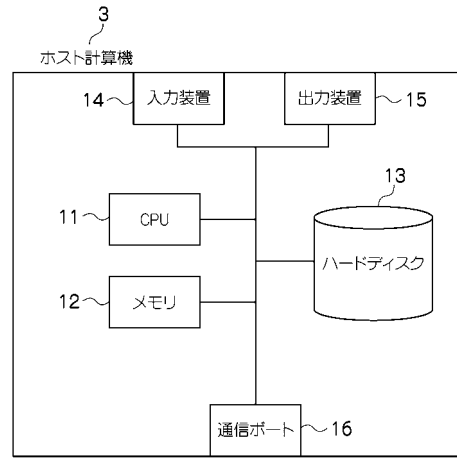
【図1】

図1



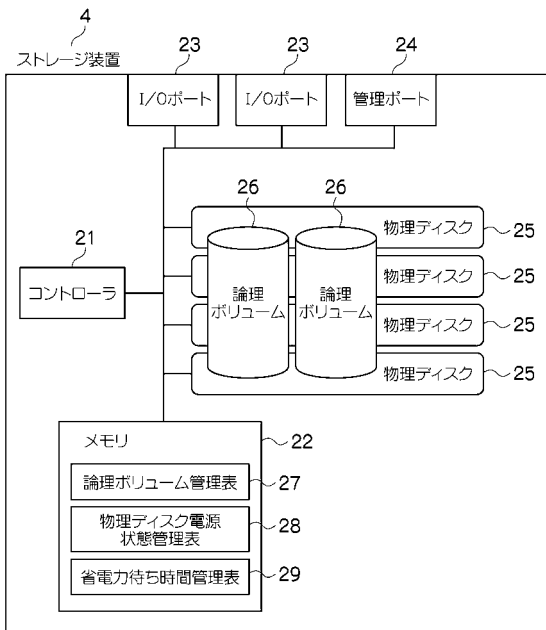
【図2】

図2



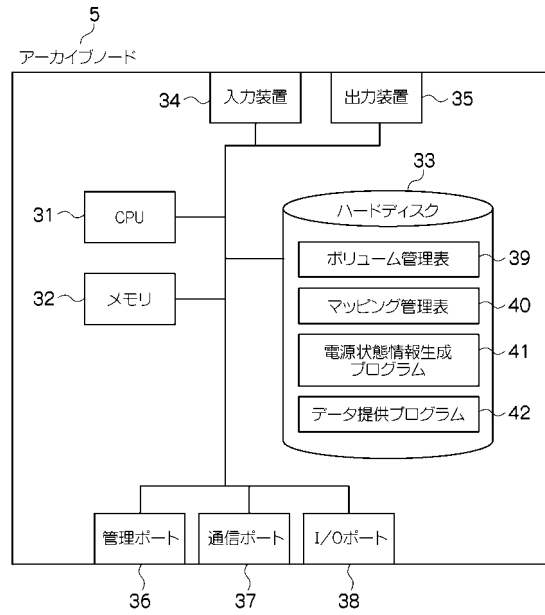
【図3】

図3



【図4】

図4



【図5】

図5

27
論理ボリューム管理表

論理ボリュームID	物理ディスクID
VOL1	PG1
VOL2	PG2
VOL3	PG3
VOL4	PG4

27A 27B

【図8】

図8

39
ボリューム管理表

アーカイブノードID	ストレージ装置ID	論理ボリュームID	使用済み容量	電源状態	電源状態変更時刻
ノード1	ストレージA	VOL1	10GB	OFF	2007/04/01 00:00
ノード2	ストレージA	VOL2	20GB	OFF	2007/03/31 09:00
ノード3	ストレージA	VOL3	10GB	ON	2007/04/15 10:00
ノード4	ストレージA	VOL4	5GB	OFF	2007/04/01 00:00

39A 39B 39C 39D 39E

【図6】

図6

28
物理ディスク電源状態管理表

物理ディスクID	電源状態	電源状態変更時刻
PG1	OFF	2007/04/01 00:00
PG2	OFF	2007/03/31 09:00
PG3	ON	2007/04/15 10:00
PG4	OFF	2007/04/01 00:00

28A 28B 28C

【図7】

図7

29
省電力待ち時間管理表

省電力待ち時間
30 min

29A

【図9】

図9

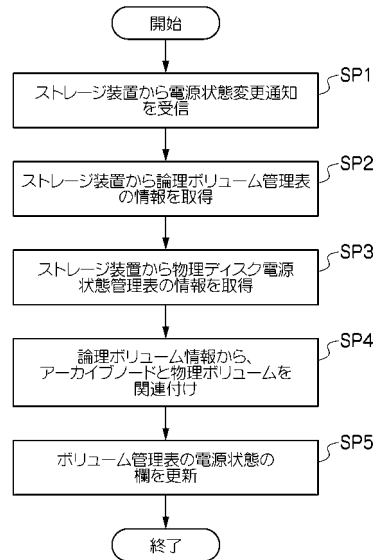
40
マッピング管理表

アーカイブデータID	冗長度	データ保存先ノード
/data/file_a	3	ノード1、ノード2、ノード3
/data/file_b	3	ノード4、ノード1、ノード2
/data/file_c	3	ノード3、ノード4、ノード1
/data/file_d	3	ノード2、ノード3、ノード4

40A 40B 40C

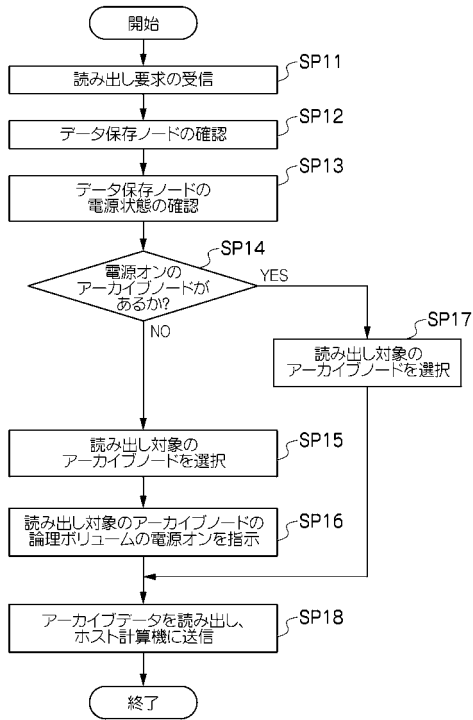
【図10】

図10



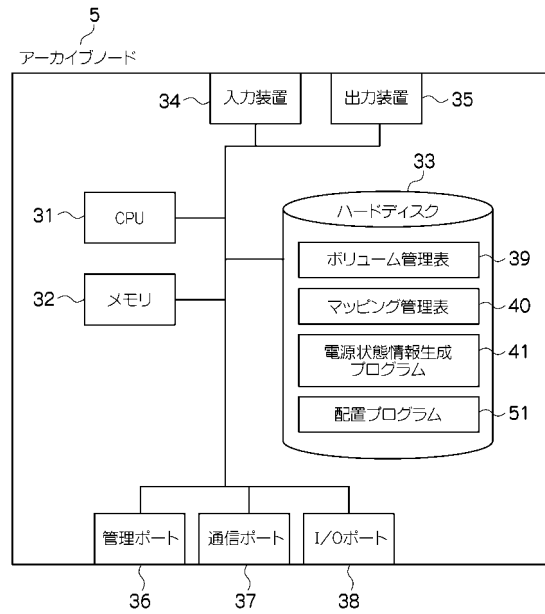
【図11】

図11



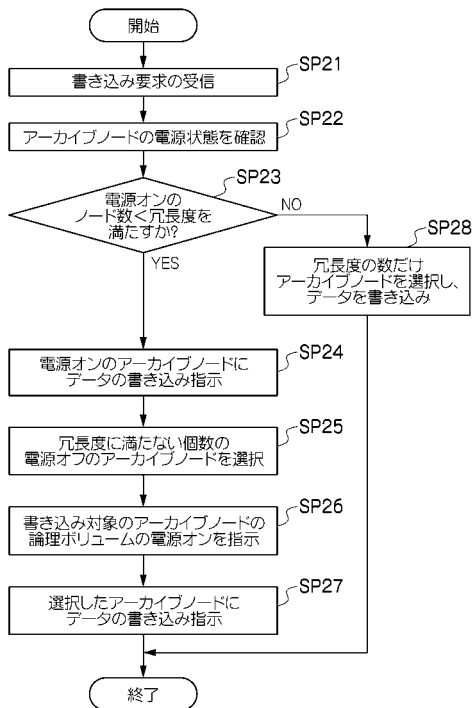
【図12】

図12



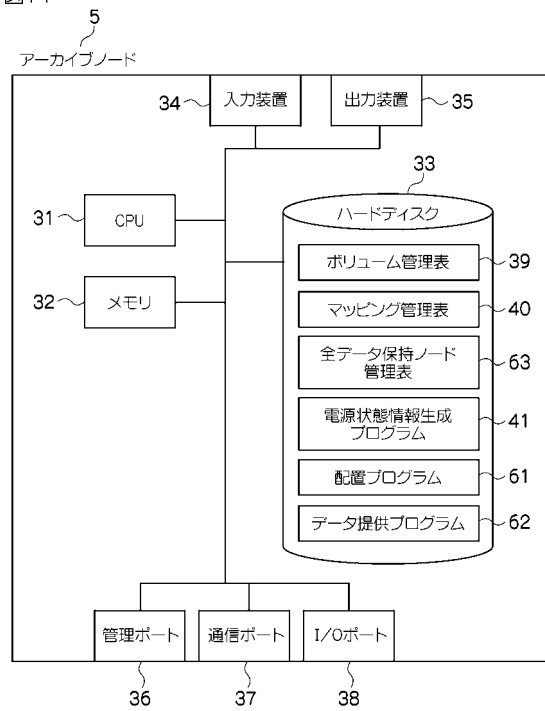
【図13】

図13



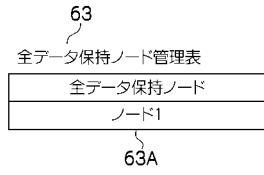
【図14】

図14



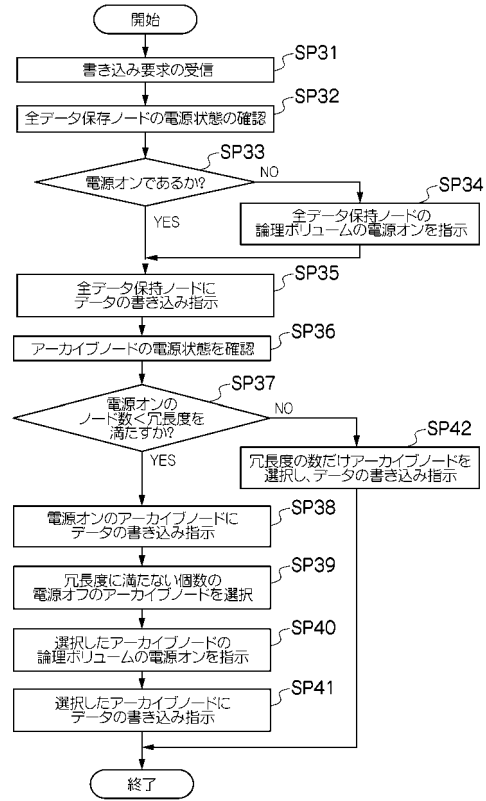
【図15】

図15



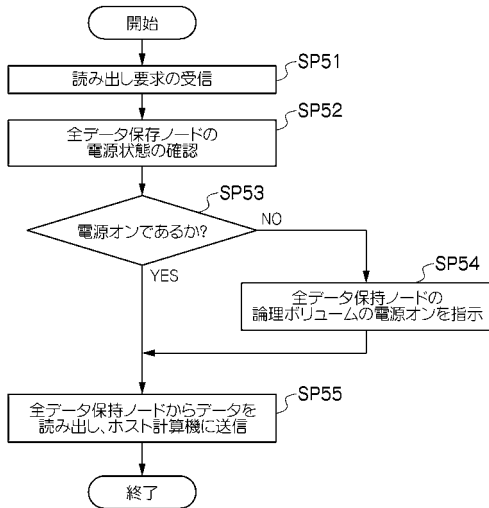
【図16】

図16



【図17】

図17



フロントページの続き

(72)発明者 山本 政行

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

審査官 古河 雅輝

(56)参考文献 特開2007-018407(JP,A)
特開2002-268979(JP,A)
国際公開第2006/098036(WO,A1)
特開平10-283123(JP,A)
特開2002-182862(JP,A)
特開2008-250945(JP,A)
特表2009-505310(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08

G06F 12/00

G06F 13/10 - 13/14