

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5068086号
(P5068086)

(45) 発行日 平成24年11月7日(2012.11.7)

(24) 登録日 平成24年8月24日(2012.8.24)

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

請求項の数 5 (全 33 頁)

(21) 出願番号	特願2007-36301 (P2007-36301)	(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2月16日(2007.2.16)	(74) 代理人	110000279 特許業務法人ウィルフォート国際特許事務所
(65) 公開番号	特開2008-203951 (P2008-203951A)	(72) 発明者	堀 雅則 神奈川県小田原市中里322番2号 株式会社日立製作所 RAIDシステム事業部内
(43) 公開日	平成20年9月4日(2008.9.4)	(72) 発明者	本田 聖志 神奈川県小田原市中里322番2号 株式会社日立製作所 RAIDシステム事業部内
審査請求日	平成21年8月13日(2009.8.13)		

最終頁に続く

(54) 【発明の名称】 記憶制御装置

(57) 【特許請求の範囲】

【請求項1】

第1コントローラと、複数の第1記憶装置とを含む基本筐体と、
第2コントローラと、各々、所定時間アクセスされない場合にスピンドウン状態となり、スピンドウン状態にてアクセスを受領した場合にスピンドアップ状態となる複数の第2記憶装置とを含む第1増設筐体と、

第3コントローラと、各々、前記所定時間アクセスされない場合にスピンドウン状態となり、スピンドウン状態にてアクセスを受領した場合にスピンドアップ状態となる複数の第3記憶装置とを含む第2増設筐体と、

前記基本筐体と、前記第1増設筐体及び前記第2増設筐体の各々とを、接続するスイッチと、

前記基本筐体、前記第1増設筐体及び前記第2増設筐体に接続され、外部電源と前記第1増設筐体とを接続する第1スイッチ及び前記外部電源と前記第2増設筐体とを接続する第2スイッチを含む電源供給回路部と、
を含み、

前記第1コントローラは、前記複数の第2記憶装置の全てがスピンドウン状態の場合に、前記複数の第2記憶装置への電源供給を停止する制御をし、前記第1増設筐体への電源供給を停止する制御をし、

前記第1コントローラは、前記複数の第2記憶装置の少なくとも1つがスピンドウン状態でない場合に、前記第1増設筐体への電源供給を停止する制御をせず、

10

20

前記第 1 コントローラは、前記複数の第 2 記憶装置への電源供給を停止する制御をした後に、前記電源供給回路部に前記外部電源と前記第 1 増設筐体とを接続する前記第 1 スイッチをオフさせるようになっており、

前記電源供給回路部として、前記第 1 増設筐体にそれぞれ異なる電力供給経路を介して電源を供給する第 1 電源供給回路部及び第 2 電源供給回路部を備え、

前記基本筐体は、前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給を停止させる場合、

(1) 先に前記第 1 電源供給回路部から前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給を停止させ、

(2) 前記第 1 電源供給回路部から前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給が正常に停止されたことを確認し、

(3) 次に前記第 2 電源供給回路部から前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給を停止させ、

(4) 前記第 2 電源供給回路部から前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給が正常に停止されたことを確認するように、前記第 1 電源供給回路部及び前記第 2 電源供給回路部をそれぞれ制御し、

前記基本筐体は、前記第 1 電源供給回路部または前記第 2 電源供給回路部のいずれから前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体への電源供給が正常に停止されなかった場合、前記第 1 電源供給回路部または前記第 2 電源供給回路部のうち前記電源供給が正常に停止されなかった方の所定の電源供給回路部と通信が可能か否かを判定し、通信が可能な場合は前記所定の電源供給回路部の内部で異常が生じた旨を出力し、通信が不能な場合は前記所定の電源供給回路部と前記基本筐体との間の通信経路に異常が生じた旨を出力する記憶制御装置。

【請求項 2】

前記第 1 コントローラは、前記電源供給回路部より電源供給を停止させる信号を受信しなかった場合に、前記第 1 スイッチをオンさせ、前記第 1 増設筐体と前記電源供給回路部との接続に異常がある旨通知する、ことを特徴とする請求項 1 に記載の記憶制御装置。

【請求項 3】

前記第 1 コントローラは、前記複数の第 3 記憶装置の全てがスピンドウン状態の場合に、前記複数の第 3 記憶装置への電源供給を停止する制御をし、前記第 2 増設筐体への電源供給を停止する制御をし、

前記第 1 コントローラは、前記複数の第 3 記憶装置の少なくとも 1 つがスピンドウン状態でない場合に、前記第 2 増設筐体への電源供給を停止する制御をしない、

ことを特徴とする請求項 1 に記載の記憶制御装置。

【請求項 4】

前記電源供給回路部は、前記基本筐体に障害が発生したことを検知した場合に、前記第 1 増設筐体に電源を供給する状態に移行する請求項 1 に記載の記憶制御装置。

【請求項 5】

前記基本筐体は、前記複数の第 2 記憶装置の全てがスピンドウン状態である前記第 1 増設筐体に再び電源を供給する場合、当該第 1 増設筐体内の前記複数の第 2 記憶装置を所定数ずつ分けて起動させるように当該第 1 増設筐体を制御する請求項 1 に記載の記憶制御装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶制御装置に関する。

【背景技術】

【0002】

記憶制御装置は、多数のハードディスクドライブをアレイ状に接続した記憶部を備えており、サーバ等のホストコンピュータ（以下「ホスト」）に、論理的な記憶領域（論理ボリューム）を提供する。信頼性等を高めるために、記憶制御装置では、RAID（Redundant Array of Independent Disks）に基づく冗長化された記憶領域をホストに提供する。

【0003】

ところで、企業等で管理すべきデータは日々増大する。従って、データ量の増大に応じて、記憶容量を増加させることのできる記憶制御装置が好ましい。そこで、制御機能を司る基本筐体に、記憶領域を提供する増設筐体を複数個接続させることにより、ユーザの希望に応じた記憶領域を提供可能な記憶制御装置が提案されている（特許文献1）。

10

【0004】

この第1文献に記載の従来技術では、基本筐体への電源供給の開始または停止と、各増設筐体への電源供給の開始または停止とを連動させる。これにより、記憶制御装置を完全に停止させる場合には、基本筐体への電源供給を停止させるだけで、各増設筐体への電源供給を全て自動的に停止させることができる。

【0005】

第2文献に記載の他の従来技術では、RAIDグループを構成する各磁気ディスク装置を節電モードに移行させるようになっている（特許文献2）。

【特許文献1】特開2001-339853号公報

20

【特許文献2】特開2000-293314号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

前記第1文献に記載の従来技術では、基本筐体への電源供給動作と各増設筐体の電源供給動作とを連動させるだけのものであり、記憶制御装置の電力消費量を低減させるという観点からは効果が乏しい。

【0007】

これに対し、前記第2文献に記載の従来技術では、一定期間アクセスされない磁気ディスク装置を節電モードに移行させるため、記憶制御装置の電力消費量を低減させることができる。

30

【0008】

しかし、第2文献に記載の従来技術では、RAIDグループを構成する各磁気ディスク装置の電力消費量を節減するに留まり、さらなる改善の余地がある。即ち、各磁気ディスク装置が節電モードに移行した場合でも、筐体内の制御回路部や冷却ファン、電源回路には電源が供給されており、稼働を続けている。従って、磁気ディスク装置以外の各部における電力消費量について改善の余地がある。

【0009】

本発明は、上記の問題点に鑑みてなされたもので、その目的は、筐体単位で電源の供給及び停止を行うことにより、より一層電力消費量を低減することができるようにした記憶制御装置を提供することにある。本発明の他の目的は、一つの筐体への電源供給の停止が他の筐体に影響を与えないように、筐体単位での省電力を実現することができるようにした記憶制御装置を提供することにある。本発明のさらに別の目的は、各筐体毎の電源供給制御を診断することにより、信頼性及び電力消費量を改善することができる記憶制御装置を提供することにある。本発明のさらなる目的は、後述する実施形態の記載から明らかになるであろう。

40

【課題を解決するための手段】

【0010】

上記課題を解決すべく、本発明の一つの観点に従う、複数の筐体を備える記憶制御装置では、各筐体のうちいずれか1つの筐体は、上位装置との間のデータ通信及び他の筐体の

50

動作を制御するための第1筐体となっており、第1筐体以外の他の筐体は、それぞれ複数の記憶デバイスを備える第2筐体となっており、第1筐体と各第2筐体とは、第1筐体と各第2筐体とをスイッチ接続するための筐体間接続部を介して接続されている。さらに、第1筐体は、記憶制御装置の構成及び電源供給状態を管理するためのシステム構成情報を保持しており、各第2筐体にそれぞれ電源を供給するための電源供給回路部をシステム構成情報に基づいて制御することにより、第2筐体毎にそれぞれ個別に電源の供給または停止を制御する。

【0011】

本発明の実施形態では、第1筐体は、各記憶デバイスのうち、予め設定された所定時間以上のアクセスが無い記憶デバイスについては、電力消費を抑制させる節電モードに移行させる。

10

【0012】

本発明の実施形態では、第1筐体は、各第2筐体のうち、筐体内の全ての記憶デバイスが節電モードに移行している所定の第2筐体を検出された場合には、当該所定の第2筐体内の全ての記憶デバイスへの電源供給を停止させる。

【0013】

本発明の実施形態では、第1筐体は、各第2筐体のうち、筐体内の全ての記憶デバイスが節電モードに移行している所定の第2筐体を検出された場合には、当該所定の第2筐体への電源供給を停止させる。

【0014】

本発明の実施形態では、第1筐体は、所定の第2筐体内の全ての記憶デバイスへの電源供給を停止させた後で、所定の第2筐体への電源供給を停止させる。

20

【0015】

本発明の実施形態では、電源供給回路部は、第1筐体に障害が発生したことを検知した場合に、各第2筐体に電源をそれぞれ供給する状態に移行する。

【0016】

本発明の実施形態では、第1筐体は、所定の第2筐体に再び電源を供給する場合、所定の第2筐体内の記憶デバイスを所定数ずつ分けて起動させるように所定の第2筐体を制御する。

【0017】

本発明の実施形態では、電源供給回路部として、各第2筐体にそれぞれ異なる電力供給経路を介して電源を供給する第1電源供給回路部及び第2電源供給回路部を備え、第1筐体は、所定の第2筐体への電源供給を停止させる場合、(1)先に第1電源供給回路部から所定の第2筐体への電源供給を停止させ、(2)第1電源供給回路部から所定の第2筐体への電源供給が正常に停止されたことを確認し、(3)次に第2電源供給回路部から所定の第2筐体への電源供給を停止させ、(4)第2電源供給回路部から所定の第2筐体への電源供給が正常に停止されたことを確認するように、第1電源供給回路部及び第2電源供給回路部をそれぞれ制御する。

30

【0018】

本発明の実施形態では、第1筐体は、第1電源供給回路部または第2電源供給回路部のいずれから所定の第2筐体への電源供給が正常に停止されなかった場合、第1電源供給回路部または第2電源供給回路部のうち電源供給が正常に停止されなかった方の所定の電源供給回路部と通信が可能か否かを判定し、通信が可能な場合は所定の電源供給回路部の内部で異常が生じた旨を出力し、通信が不能な場合は所定の電源供給回路部と第1筐体との間の通信経路に異常が生じた旨を出力する。

40

【0019】

本発明の実施形態では、第1筐体と電源供給回路部とは、通信ネットワークを介して接続されており、第1筐体は、電源供給の開始または電源供給の停止を指示するためのコマンドを通信ネットワークを介して電源供給回路部に送信することにより、電源供給回路部を制御する。

50

【 0 0 2 0 】

本発明の実施形態では、筐体間接続部と電源供給回路部とは接続されており、第1筐体は、筐体間接続部を介して、電源供給の開始または電源供給の停止を電源供給回路部に指示する。

【 0 0 2 1 】

本発明の実施形態では、第1筐体は、電源供給が停止されている第2筐体に、定期的または不定期に電源を供給することにより、正常に動作するか否かを診断する。

【 0 0 2 2 】

本発明の他の観点に従う、複数の筐体を備えた記憶制御装置は、第1筐体と、それぞれ複数の記憶デバイスを有する第2筐体と、第1筐体及び各第2筐体にそれぞれ外部電源を供給するための電源供給回路部と、第1筐体と前記各第2筐体とをスイッチ接続するための筐体間接続部とを備える。そして、第1筐体は、記憶制御装置の構成及び電源供給状態に関するシステム構成情報を管理するシステム構成情報管理部と、システム構成情報管理部により管理されるシステム構成情報に基づいて、各第2筐体への電源供給及び上位装置と各記憶デバイスとの間のデータ入出力を処理する第1制御部と、電源供給回路部を介して供給される外部電源を所定の電圧に変換することにより、第1制御部に所定の電源を供給する第1筐体内電源部と、を少なくとも備える。各第2筐体は、上位装置により利用されるデータを記憶する複数の記憶デバイスと、第1制御部からの指示に応じて、各記憶デバイスへのデータ入出力及び第2筐体内電源部の動作を制御する第2制御部と、電源供給回路部を介して供給される外部電源を所定の電圧に変換することにより、各記憶デバイス及び第2制御部に電源をそれぞれ供給する第2筐体内電源部と、を少なくとも備える。電源供給回路部は、各第2筐体内電源部と外部電源との間の電力供給経路の途中にそれぞれ設けられるスイッチ部を備え、該各スイッチ部は、第1制御部からの指示に応じて各電力供給経路をそれぞれ連通または遮断させるようになっており、かつ、第1制御部は、システム構成情報に基づいて、各第2筐体のうち当該第2筐体内に存在する各記憶デバイスへのアクセスが不要であると判定された所定の第2筐体については、この所定の第2筐体に接続されているスイッチ部を動作させることにより、所定の第2筐体の第2筐体内電源部と外部電源との間の電力供給経路を遮断させる。

【 0 0 2 3 】

本発明の実施形態では、第1制御部は、(1)各記憶デバイスのうち、予め設定された所定時間以上のアクセスが無い記憶デバイスについては、電力消費を抑制させる節電モードに移行させる第1モードと、(2)各第2筐体のうち、筐体内の全ての記憶デバイスが前記節電モードに移行している所定の第2筐体が発出された場合には、当該所定の第2筐体内の全ての記憶デバイスへの電源供給を停止させる第2モードと、(3)所定の第2筐体内の全ての記憶デバイスへの電源供給を停止させた後で、所定の第2筐体への電源供給を停止させる第3モードとを、それぞれ実行することができる。

【 0 0 2 4 】

本発明の実施形態では、電源供給回路部は、第1制御部との定期的な通信を行うことによって第1制御部に障害が発生したか否かを検出し、第1制御部に障害が発生したことを検出した場合には、各スイッチ部をそれぞれオン状態にして各電力供給回路をそれぞれ連通させる。

【 発明を実施するための最良の形態 】

【 0 0 2 5 】

以下、図面に基づき、本発明の実施の形態を説明する。図1は、本実施形態に係る記憶制御装置の全体概要を示す説明図である。

【 0 0 2 6 】

記憶制御装置は、例えば、複数種類の筐体1, 2を重ねて配置することにより、構成されている。1つの種類の筐体は、基本筐体1であり、他の1つの種類の筐体は、増設筐体2である。

【 0 0 2 7 】

10

20

30

40

50

基本筐体 1 は第 1 筐体に該当し、増設筐体 2 は第 2 筐体に該当する。基本筐体 1 と各増設筐体 2 とは、筐体間接続スイッチ 4 を介して、スイッチ接続される。基本筐体 1 と各増設筐体 2 との間の信号やデータの伝達は、筐体間接続スイッチ 4 を介して行われる。換言すれば、基本筐体 1 を中心として、各増設筐体 2 がスター型接続されている。基本筐体 1 は、記憶制御装置全体の制御を司る筐体であり、上位装置としてのホスト 5 及び各増設筐体 2 と接続される。基本筐体 1 は、ホスト 5 から発行されたコマンドに応じて、増設筐体 2 内のディスクドライブ 2 B にアクセスし、データの読み書きを行う。

【 0 0 2 8 】

基本筐体 1 は、例えば、I/O処理部 1 A と、電源制御部 1 B 及びシステム構成情報管理部 1 C とを備えて構成される。なお、後述の実施例からも明らかのように、基本筐体 1 は、ディスクドライブを備えることもできる。

10

【 0 0 2 9 】

I/O処理部 1 A は、ホスト 5 から発行されたコマンドに基づいて、各増設筐体 2 内のディスクドライブ 2 B にアクセスすることにより、データを書き込んだり、または、データを読み出したりする。

【 0 0 3 0 】

電源制御部 1 B は、基本筐体 1 が有する制御機能の一つであり、各増設筐体 2 に関する電源の供給及び停止を制御する。電源制御部 1 B は、複数種類の省電力モード F 1 ~ F 3 をそれぞれのタイミングで実行させることができる。

【 0 0 3 1 】

20

ドライブ単体の省電力を実現させるモード F 1 では、例えば、一定時間アクセスされないディスクドライブ 2 B のモータへの通電を停止させたり、電子回路の一部を停止させることにより、ドライブ毎の省電力を実現する。

【 0 0 3 2 】

全ドライブへの通電を停止させるモード F 2 では、同一筐体に設けられている全てのディスクドライブ 2 B への通電を停止させることにより、筐体内のディスクドライブ全体について省電力を実現する。なお、モード F 2 では、例えば、冷却ファン 2 D のような、全ディスクドライブ 2 B の停止に伴って停止させることのできる装置への通電を停止させることもできる。

【 0 0 3 3 】

30

筐体全体への通電を停止させるモード F 3 では、電源供給回路部 3 のスイッチ 3 A を開くことにより、外部電源 6 から増設筐体 2 への電源供給を完全に停止させる。このモード F 3 では、増設筐体 2 の各機能は停止され、最も省電力効果が高くなる。

【 0 0 3 4 】

電源制御部 1 B は、自己診断モード F 4 を備えている。自己診断モード F 4 では、筐体全体への通電が停止された増設筐体 2 に対して、定期的にまたは不定期に通電することにより、その増設筐体 2 が正常に動作するか否かを診断する。

【 0 0 3 5 】

システム構成情報管理部 1 C は、記憶制御装置のシステム構成に関する情報を記憶し、管理している。システム構成に関する情報としては、例えば、ディスクドライブ 2 B の配置に関する情報、RAIDグループや論理ボリュームの構成に関する情報、増設筐体 2 と電源供給回路部 3 との接続構成に関する情報等を挙げることができる。本実施形態では、記憶制御装置に関する全ての構成情報を基本筐体 1 が一元的に管理する。電源制御部 1 B は、システム構成情報を参照することにより、目的の増設筐体 2 への通電停止等を行う。

40

【 0 0 3 6 】

電源供給回路部 3 は、例えば、外部電源 6 からの交流電源を各筐体 1, 2 にそれぞれ分配するための回路である。ここでは、基本筐体 1 への電源分配については省略しているが、後述の実施例から明らかのように、電源供給回路部 3 を介して、外部電源 6 と基本筐体 1 とは接続されている。

【 0 0 3 7 】

50

電源供給回路部 3 は、各増設筐体 2 と外部電源 6 とをそれぞれ接続するスイッチ 3 A を備えている。従って、目的の増設筐体 2 への電源供給を停止させる場合は、その増設筐体 2 に接続されたスイッチ 3 A を開く。これにより、外部電源 6 から目的の増設筐体 2 への電力供給経路が遮断される。

【 0 0 3 8 】

電源供給回路部 3 は、各筐体 1 , 2 と独立した回路基板として構成可能である。これに限らず、電源供給回路部 3 を各筐体 1 , 2 内に設けることもできる。即ち、例えば、スイッチ 3 A を各増設筐体 2 内に設ける構成でもよい。

【 0 0 3 9 】

各増設筐体 2 は、例えば、制御基板 2 A と、複数のディスクドライブ 2 B と、電源部 2 C と、冷却ファン 2 D とを備えて構成される。制御基板 2 A は、基本筐体 1 からの指示に基づいて、ディスクドライブ 2 B へのデータ入出力を行う。

【 0 0 4 0 】

ディスクドライブ 2 B は、記憶デバイスに該当する。ここでは、便宜上、ディスクドライブと呼ぶが、ディスク以外の媒体を用いる記憶デバイスでもよい。ディスクドライブ 2 B としては、例えば、ハードディスクドライブ、半導体メモリデバイス、フラッシュメモリデバイス、光ディスクドライブ、光磁気ディスクドライブ、磁気テープデバイス、ホログラフィックスメモリデバイス等を用いることができる。

【 0 0 4 1 】

電源部 2 C は、筐体内電源部に該当する。筐体内電源部 2 C は、例えば、外部電源 6 から電源供給回路部 3 を介して入力された交流電源または直流電源を、所定の電圧を有する直流電源に変換し、制御基板 2 A , 各ディスクドライブ 2 B 及び冷却ファン 2 D にそれぞれ供給するものである。電源部 2 C は、複数種類の電圧を出力可能である。また、外部電源 6 は、交流電源に限らず、直流電源であってもよい。後述の実施例から明らかとなるように、電源部 2 C は、制御基板 2 A からの指示に応じて、各ディスクドライブ 2 B への通電を停止させることができるように構成することができる。

【 0 0 4 2 】

冷却ファン 2 D は、増設筐体 2 内に冷却風を導いて冷却させるものである。冷却ファン 2 D は、制御基板 2 A からの信号によって動作する。即ち、制御基板 2 A は、例えば、冷却ファン 2 D の動作開始や動作停止、回転数を制御することができる。なお、空冷機構として冷却ファン 2 D を挙げたが、増設筐体 2 は水冷機構を備える構成でもよい。水冷機構の場合は、例えば、冷却水循環ポンプや熱交換器の動作を制御することにより、増設筐体 2 の電力消費量を節減する。

【 0 0 4 3 】

上述のように、本実施形態では、基本筐体 1 を中心として各増設筐体 2 がスター型に配置されている。従って、基本筐体 1 との間で通信を行う上では、各増設筐体 2 はそれぞれ独立しており、互いに他の増設筐体 2 に影響を与えない。即ち、ある増設筐体 2 への電源供給が遮断されて完全に停止している場合でも、他の増設筐体 2 は、基本筐体 1 と通信を行うことができる。従って、基本筐体 1 は、目的の増設筐体 2 への電源供給を自由に停止させることができる。

【 0 0 4 4 】

ここで、ディスクドライブ 2 B は、その性能や信頼性によって複数のグループに分類することができる。例えば、FC (Fibre Channel) ディスク等のような高性能のディスクドライブ、例えば、SATA (Serial AT Attachment) ディスク等のような低性能のディスクドライブ等である。各増設筐体 2 には、それぞれ同一種類のディスクドライブ 2 B を内蔵させることができる。即ち、ある増設筐体 2 は、高性能のディスクドライブ 2 B のみを内蔵し、他の増設筐体 2 は、低性能のディスクドライブ 2 B のみを内蔵する。

【 0 0 4 5 】

そして、例えば、高性能のディスクドライブ 2 B には、ホスト 5 によって頻繁に使用されるデータを記憶させ、低性能のディスクドライブ 2 B にはホスト 5 による利用頻度の少

10

20

30

40

50

ないデータ（例えば、バックアップデータやアーカイブデータ）を記憶させる。これにより、低性能のディスクドライブ 2 B を内蔵する増設筐体 2 に電源を供給する時間を短くすることができ、電力消費量を低減できる。

【 0 0 4 6 】

さて、図 1 に示すように、増設筐体 2 (# 1) では、各部に電源が供給されている。増設筐体 2 (# 2) では、所定時間以上アクセスされていないディスクドライブ 2 B について、モード F 1 が適用されている。所定時間以上アクセスされていないディスクドライブ 2 B では、例えば、スピンドルモータの回転を停止または低下させたり、一部の電子回路を停止させる。

【 0 0 4 7 】

増設筐体 2 (# 3) では、モード F 2 が適用されている。即ち、各ディスクドライブ 2 B 及び冷却ファン 2 D への通電がそれぞれ停止されている。増設筐体 2 (# 4) では、モード F 3 が適用されている。このモードでは、スイッチ 3 A (4) が開いており、外部電源 6 と増設筐体 2 (# 4) との間の電力供給経路が遮断されるため、増設筐体 2 (# 4) への電源供給が停止される。

【 0 0 4 8 】

本実施形態の記憶制御装置では、例えば、モード F 1 モード F 2 モード F 3 のように、順番に遷移する。しかし、常に必ずしもこのように遷移するわけではなく、ホスト 5 からの利用状況に応じて、例えば、モード F 3 からモード F 1 へ、モード F 2 からモード F 1 に遷移する場合もある。

【 0 0 4 9 】

後述の実施例で明らかとなるように、電源供給が停止された増設筐体 2 に定期的または不定期に通電し、正常に動作するか否かを診断することもできる。従って、長期間停止されている増設筐体 2 の信頼性を確認しながら、記憶制御装置全体の電力消費量を低減することができる。

【 0 0 5 0 】

後述の実施例で明らかとなるように、電源供給回路部 3 は、定期的に基本筐体 1 と通信を行って、基本筐体 1 に異常が生じたか否かを検出することができる。基本筐体 1 に異常が検出された場合、電源供給回路部 3 は、各スイッチ 3 A をそれぞれオン状態に設定し、外部電源 6 と各内蔵電源 2 C とを接続させる。以下、本実施形態の記憶制御装置を詳細に説明する。

【実施例 1】

【 0 0 5 1 】

図 2 は、記憶制御装置の外観構成を示す斜視図である。記憶制御装置は、例えば、基本ラック 1 0 と増設ラック 1 1 とをケーブル等で接続することにより、構成される。基本ラック 1 0 は、例えば、基本筐体 1 0 0 と、複数の増設筐体 2 0 0 と、複数のバックエンドスイッチ 4 0 0 とを備える。基本ラック 1 0 は、記憶制御装置の最小構成であり、基本ラック 1 0 のみで記憶制御装置の機能を実現することができる。

【 0 0 5 2 】

増設ラック 1 1 は、例えば、複数の増設筐体 2 0 0 を備える。基本ラック 1 0 の記憶容量が足りなくなった場合に、増設ラック 1 1 を基本ラック 1 0 に接続することにより、記憶容量を増加させることができる。なお、バックエンドスイッチ 4 0 0 の全部または一部を増設ラック 1 1 内に設ける構成としてもよい。

【 0 0 5 3 】

図 3 は、基本筐体 1 0 0 を前面側から見た斜視図である。基本筐体 1 0 0 のケーシング 1 0 1 には、その前面側 1 0 1 F に、複数のバッテリー 1 4 0 及び複数のディスクドライブ 2 3 0 がそれぞれ設けられる。基本筐体 1 0 0 は、ディスクドライブ 2 3 0 を備える構成でもよいし、ディスクドライブ 2 3 0 を備えない構成でもよい。ディスクドライブ 2 3 0 を備えない場合、前面 1 0 1 F の所定箇所は、バッテリー等の他の構成部品が搭載されるか、又は、化粧板等で施蓋される。

10

20

30

40

50

【 0 0 5 4 】

図4は、基本筐体100を後面側から見た斜視図である。ケーシング101の後面側101Rには、例えば、複数のコントローラ110と、複数のAC/DC電源部120と、複数の冷却ファン130とがそれぞれ取り付けられる。冷却ファン130は、複数のファンが一体化されたファンアッセンブリとして構成されており、ケーシング101の左右両側にそれぞれ配置される。コントローラ110は、ケーシング101の中央に上下に並んで配置される。AC/DC電源部120は、ケーシング101の下側に、左右に並んで配置される。なお、以下の増設筐体200の説明でも同様であるが、基本筐体100及び増設筐体200の構成は、図示するものに限定されない。

【 0 0 5 5 】

図5は、増設筐体200を前面側から見た斜視図である。増設筐体200のケーシング201には、その前面側201Fに、複数のディスクドライブ230を取り付けることができる。本実施例では、#0～#14の合計15台のディスクドライブ230が取り付けられる場合を例に挙げて説明する。1つの筐体当たりのディスクドライブ搭載数は、15に限定されない。ディスクドライブ230が取り付けられていない空間は、化粧板等で施蓋することにより、外部の塵埃等がケーシング201内に侵入するのを防止する。

【 0 0 5 6 】

図6は、増設筐体200を後面側から見た斜視図である。ケーシング201の後面側201Rには、複数の制御基板210及び複数のAC/DC電源部220がそれぞれ取り付けられている。なお、図6の構成においては、冷却ファンはAC/DC電源部220内に搭載されているが、図10に示すように、増設筐体200には、冷却ファン202を設けることができる。即ち、冷却ファンとAC/DC電源部とを別々の構成部品としてもよい。

【 0 0 5 7 】

図7は、記憶制御装置の構成を模式的に示す説明図である。各ラック10, 11内には、各筐体100, 200に外部からの電源を供給するための電源分配回路部300がそれぞれ設けられている。以下の説明では、電源分配回路部300をPDB (Power Distribution Box) と呼ぶ。

【 0 0 5 8 】

記憶制御装置は、複数の電源供給システムを備えることができる。1つのシステムは、図7中の左側にPDB#0で示すシステムであり、他の1つのシステムは、図7中の右側にPDB#1で示すシステムである。即ち、記憶制御装置の電源供給構造は、冗長化されている。

【 0 0 5 9 】

各PDB300は、例えば、4個, 6個等のような所定数の筐体に外部からの電源を分配することができる。従って、1つのラック内に多数の筐体を取り付けられる場合、複数の電源供給システムのそれぞれにおいて、複数のPDB300が必要となる。なお、これに代えて、ラックに装着可能な最大数の筐体に外部からの電源を分配可能な単一のPDBを、各電源供給システム毎に使用する構成でもよい。

【 0 0 6 0 】

各PDB300は、所定数のスイッチ310を備えている。各スイッチ310は、それぞれ対応する筐体のAC/DC電源部に接続される。図中では、AC/DC電源部を「PS」と略記して示している。

【 0 0 6 1 】

基本筐体100の有する各電源部120は、それぞれ別システムのPDB300のスイッチ310に接続されている。同様に、各増設筐体200の有する各電源部220も、それぞれ別システムのPDB300のスイッチ310にそれぞれ接続されている。従って、いずれか一方のPDB300に障害が生じた場合でも、他方のPDB300から各筐体100, 200に電源を供給することができる。

【 0 0 6 2 】

基本筐体100のコントローラ110は、図中に一点鎖線で示すように、信号線L1を介して、各PDB300にそれぞれ接続されている。詳細は後述するが、コントローラ11

10

20

30

40

50

0 は、各スイッチ 3 1 0 の開閉を個別に制御することができる。コントローラ 1 1 0 は、目的の増設筐体 2 0 0 への電源供給を完全に停止させる場合、その増設筐体 2 0 0 に接続された各系統のスイッチ 3 1 0 をそれぞれ開くことにより、外部電源との電源部 2 2 0 との間の電源供給経路を遮断させる。

【 0 0 6 3 】

なお、各コントローラ 1 1 0 と各 PDB 3 0 0 とは、有線接続でもよいし、無線接続でもよい。無線接続の場合は、PDB 3 0 0 内に、コントローラ 1 1 0 からの無線信号に応じて各スイッチ 3 1 0 を開閉させるための回路を設ければよい。

【 0 0 6 4 】

バックエンドスイッチ 4 0 0 は、基本筐体 1 0 0 と各増設筐体 2 0 0 とをスイッチ接続するための装置である。各バックエンドスイッチ 4 0 0 は、図中点線で示すように、信号線 L 3 を介して、それぞれ所定の増設筐体 2 0 0 と接続されている。なお、図中では、便宜上省略しているが、各信号線 L 3 は冗長化されており、一方の信号線 L 3 が断線した場合でも、他方の信号線 L 3 によって通信可能である。

10

【 0 0 6 5 】

各バックエンドスイッチ 4 0 0 と基本筐体 1 0 0 のコントローラ 1 1 0 とは、信号線 L 2 を介して接続されている。各コントローラ 1 1 0 は、所望の増設筐体 2 0 0 に接続されたバックエンドスイッチ 4 0 0 を介して、所望の増設筐体 2 0 0 と通信する。

【 0 0 6 6 】

図 8 は、図 7 中の構成の一部を抜き出して詳細に示す回路図である。PDB 3 0 0 は、例えば、複数のスイッチ 3 1 0 と、制御回路 3 2 0 と、AC/DC電源部 3 3 0 と、複数のブレーカ 3 4 0 と、を備えて構成される。外部から入力される交流電源は、AC/DC電源部 3 3 0 によって直流電源に変換される。この変換された直流電源は、ブレーカ 3 4 0 及びスイッチ 3 1 0 を介して、基本筐体 1 0 0 内の電源部 1 2 0 や増設筐体 2 0 0 内の電源部 2 2 0 に供給される。

20

【 0 0 6 7 】

制御回路 3 2 0 は、例えば、LAN (Local Area Network) のような信号線 L 1 を介して、コントローラ 1 1 0 に接続されている。制御回路 3 2 0 は、コントローラ 1 1 0 から発行されるコマンドに従って、指定されたスイッチ 3 1 0 をオンまたはオフさせる。ここで、スイッチ 3 1 0 をオンにするとは、スイッチ 3 1 0 の接点を閉じて導通状態にすることを意味し、スイッチ 3 1 0 をオフにするとは、スイッチ 3 1 0 の接点を開いて遮断状態にすることを意味する。

30

【 0 0 6 8 】

図 8 に示すように、バックエンドスイッチ 4 0 0 は、切替回路 4 1 0 を備えている。切替回路 4 1 0 は、信号線 L 2 を介してコントローラ 1 1 0 に接続されている。切替回路 4 1 0 は、各信号線 L 3 を介して、各増設筐体 2 0 0 にそれぞれ接続されている。切替回路 4 1 0 は、コントローラ 1 1 0 からの指示に応じて、指定された増設筐体 2 0 0 と基本筐体 1 0 0 とを接続させる。

【 0 0 6 9 】

図 9 は、記憶制御装置の電源供給構造を模式的に示す説明図である。図 9 中では、紙面の都合上、一部の構成を省略して示す。

40

【 0 0 7 0 】

図 1 0 は、記憶制御装置の内部構成を模式的に示す説明図である。基本筐体 1 0 0 の各コントローラ 1 1 0 は、通信ネットワーク CN 1 を介して、ホスト 1 3 にそれぞれ接続されている。また、各コントローラ 1 1 0 は、図 1 1 に示すように、通信ネットワーク CN 2 を介して管理端末 1 4 にもそれぞれ接続されている。

【 0 0 7 1 】

ホスト 1 3 は、例えば、図外のクライアント端末からの要求に応じて、記憶制御装置にアクセスし、データの読み書きを行う。ホスト 1 3 としては、例えば、メインフレームコンピュータやサーバコンピュータ等を挙げることができる。通信ネットワーク CN 1 として

50

は、例えば、LAN、SAN (Storage Area Network)、インターネットあるいは専用回線等を挙げることができる。

【0072】

LANを用いる場合、ホストコンピュータ13と記憶制御装置とは、TCP/IP (Transmission Control Protocol/Internet Protocol) に従って通信を行う。SANを用いる場合、ホストコンピュータ13と記憶制御装置とは、ファイバチャネルプロトコルに従って通信を行う。また、ホストコンピュータ13がメインフレームコンピュータである場合、例えば、FICON (Fibre Connection: 登録商標)、ESCON (Enterprise System Connection: 登録商標)、ACONARC (Advanced Connection Architecture: 登録商標)、FIBARC (Fibre Connection Architecture: 登録商標) 等の通信プロトコルに従ってデータ転送が行われる。

10

【0073】

さらに、図11に示すように、記憶制御装置には、管理用の通信ネットワークCN2を介して、管理端末14を接続することもできる。管理端末14は、記憶制御装置の各種設定等を行うためのコンピュータ端末である。管理端末14は、図外の管理サーバに接続することもできる。管理サーバは、複数の記憶制御装置を一括して管理可能なコンピュータ装置として構成することができる。

【0074】

図10に戻る。記憶制御装置の制御系統も、電源供給系統と同様に冗長化することができる。#0のコントローラ110と#1のコントローラ110とは、それぞれ配下にある全てのディスクドライブ230にアクセス可能である。即ち、いずれか一方の系統に障害が発生した場合でも、他方の系統からデータアクセスが可能となっている。記憶制御装置の全体動作は、基本筐体100内のコントローラ110によって制御される。従って、各コントローラ110は、基本筐体100内の各ディスクドライブ230のみならず、増設筐体200内の各ディスクドライブ230にもそれぞれアクセス可能である。

20

【0075】

基本筐体100の各電源部120は、各コントローラ110及び基本筐体100内の各ディスクドライブ230に、それぞれ所定電圧の直流電源を供給する。電源部120は、直流5ボルトや直流12ボルト等のような複数種類の電圧を出力可能である。なお、図中では、説明の便宜上、下段側のディスクドライブ230にのみ電源が供給されるかに示されているが、各電源部120は、筐体内の全てのディスクドライブ230に電源を供給する。

30

【0076】

図10中の下側には増設筐体200が示されている。紙面の都合上、1つの増設筐体200のみを示している。上述のように、基本筐体100は、複数の増設筐体200とバックエンドスイッチ400を介してスイッチ接続されている。

【0077】

増設筐体200の電源部220は、電源部120と同様に、外部から入力された電源を所定電圧の直流電源に変換して出力する。筐体内に存在する全てのディスクドライブ230、制御基板210及び冷却ファン202には、各電源部220から直流電源がそれぞれ供給される。なお、説明の便宜のために、各電源部220から直流電源がそれぞれ供給されると述べたが、例えば、通常時においては、いずれか一方の電源部220 (例えば、#0の電源部) のみから各ディスクドライブ230や制御基板210等に直流電源が供給される構成でもよい。ディスクドライブ230や制御基板210等の負荷から各電源部220を見たときのインピーダンスに差をつけることにより、いずれか一方の電源部220から電源を供給させることができる。例えば、いずれかの電源部220からの出力に、他方の電源部220よりも多い段数のダイオードを設ければよい。これにより、障害発生時には、スイッチ切替等を行うことなく、他方の電源部220に自動的に切り替わる。

40

【0078】

ディスクドライブ230は、例えば、ATA (AT Attachment) ディスク、SCSI (Small Computer System Interface) ディスク、FC (Fibre Channel) ディスク等のようなハードデ

50

ィスクドライブとして構成される。これに限らず、例えば、半導体メモリドライブ（フラッシュメモリデバイスを含む）、光ディスクドライブ、光磁気ディスクドライブ等の他の記憶デバイスを用いてもよい。

【 0 0 7 9 】

例えば、4個1組、8個1組等のような所定数のディスクドライブ230によって、RAIDグループ240が構成される。RAIDグループ240は、別々の筐体に存在する複数のディスクドライブ230から構成することもできるし、同一筐体に存在する1つまたは複数のディスクドライブ230から構成することもできる。

【 0 0 8 0 】

図11に示すように、RAIDグループ240には、1つまたは複数の論理ボリューム250を設けることができる。論理ボリューム250は、ホスト13からアクセスされる対象であり、ホスト13により使用されるデータを記憶している。

10

【 0 0 8 1 】

図11は、1つのコントローラ110に着目した場合のブロック図である。コントローラ110は、例えば、ホストインターフェース回路111と、ドライブインターフェース回路112と、データ転送回路113と、キャッシュメモリ114と、プログラムメモリ115と、プロセッサ（図中「MPU」と略記）116と、ブリッジ117と、LANインターフェース118とを備えて構成される。

【 0 0 8 2 】

ホストインターフェース回路111は、ホスト13との間の通信を制御する。ホスト13から発行された各種コマンドやデータは、ホストインターフェース回路111によって受信される。ディスクドライブ230から読み出されたデータやコマンドの処理完了を告げる通知は、ホストインターフェース回路111からホスト13に送信される。

20

【 0 0 8 3 】

ドライブインターフェース回路112は、各ディスクドライブ230との間の通信を制御する。ドライブインターフェース回路112は、論理ブロックアドレス（LBA）とディスクドライブ230の物理的なアドレスとの変換操作等を行う。

【 0 0 8 4 】

データ転送回路113は、コントローラ110内のデータ転送を制御するための回路である。データ転送回路113は、ホストインターフェース回路111とキャッシュメモリ114との間のデータ転送や、ドライブインターフェース回路112とキャッシュメモリ114との間のデータ転送を制御する。

30

【 0 0 8 5 】

プロセッサ116は、一つまたは複数のプロセッサコアを備えている。プロセッサ116は、プログラムメモリ115に記憶された電源管理プログラム等の種々のプログラムを読み込んで実行することにより、後述の電源制御等を実現する。

【 0 0 8 6 】

キャッシュメモリ114は、ホスト13から受信したデータやホスト13により読み出されたデータを記憶する。キャッシュメモリ114には、記憶制御装置のシステム構成に関する情報も記憶されている。キャッシュメモリ114の記憶内容は、バッテリー140によってバックアップされる。

40

【 0 0 8 7 】

ブリッジ117は、プロセッサ116とプログラムメモリ115とを接続する。また、プロセッサ116は、ブリッジ117を介してデータ転送回路113に接続される。

【 0 0 8 8 】

LANインターフェース118は、管理端末14とコントローラ110とを接続するための回路である。ユーザは、管理端末14を介して、コントローラ110に各種の指示を与えたり、記憶制御装置の各種状態に関する情報を読み出して画面に表示させることができる。

【 0 0 8 9 】

50

図12は、増設筐体200のブロック図である。先に、AC/DC電源部220の構成を説明する。電源部220は、例えば、電圧変換回路221と、複数の出力ダイオード222A～222Cと、出力制御スイッチ223とを備えている。

【0090】

電圧変換回路221は、PDB300を介して入力された交流電源を複数種類の直流電圧V1、V2に変換して出力する。例えば、V1は直流5ボルト程度、V2は直流12ボルト程度にそれぞれ設定される。出力ダイオード222A～222Cは、逆流防止のために設けられている。スイッチ223は、直流電圧を出力させるか否かを制御するためのもので、制御基板210によって切り替えられる。

【0091】

各ディスクドライブ230には、V1、V2がそれぞれ供給される。各制御基板210及び各冷却ファン202には、V2がそれぞれ供給される。図12に示すように、2つの電源部220の出力は、それぞれOR接続されている。従って、いずれか一方の電源部220に障害が生じた場合でも、他方の電源部220から制御基板210やディスクドライブ230等に所定電圧の電源が供給される。各電源部220からの出力V1、V2が停止した場合、筐体内の全てのディスクドライブ230への電源供給が同時に停止する。

【0092】

各制御基板210は、信号線P2を介して、各電源部220内の各スイッチ223にそれぞれ接続されている。各制御基板210は、各スイッチ223をそれぞれ個別に開閉させることができる。また、各制御基板210は、信号線P3を介して、各冷却ファン202にもそれぞれ接続されている。各制御基板210は、各冷却ファン202の作動を開始または停止、あるいは回転数を調節することができる。

【0093】

図13は、1つの制御基板210に着目した構成説明図である。制御基板210は、例えば、マイクロプロセッサ部211と、通信インターフェース部212と、制御信号出力回路213とを備えて構成される。

【0094】

マイクロプロセッサ部211は、制御基板210の動作を制御するものであり、例えば、資源管理機能211A、コマンド処理機能211B、電源制御機能211C、冷却ファン制御機能211Dをそれぞれ実現する。

【0095】

資源管理機能211Aは、例えば、増設筐体200内の各資源（ディスクドライブ230、電源部220等）の配置や状態等を管理する機能である。コマンド処理機能211Bは、コントローラ110から入力された指示を実行するための機能である。即ち、コマンド処理機能211Bは、コントローラ110からデータの書込みを要求された場合は、指定されたディスクドライブ230の指定されたアドレスにデータを書き込む。また、コマンド処理機能211Bは、コントローラ110からデータの読出しを要求された場合は、指定されたディスクドライブ230の指定されたアドレスから、指定された長さのデータを読み出してコントローラ110に転送する。

【0096】

電源制御機能211Cは、コントローラ110からの指示に基づいて、スイッチ223を開閉させることにより、電源部220の出力を制御する機能である。冷却ファン制御機能211Dは、コントローラ110からの指示に基づいて、冷却ファン202の作動を制御する機能である。

【0097】

通信インターフェース部212は、バックエンドスイッチ400を介して、基本筐体100との間で通信を行うものである。基本筐体100のコントローラ110から出力されたコマンドやデータは、バックエンドスイッチ400を介して、通信インターフェース部212に入力される。

【0098】

10

20

30

40

50

図14は、論理ボリューム250を管理するためのテーブルT1の構成を示す説明図である。このテーブルT1は、図13に示す電源供給状態管理テーブルT2と共に、システム構成情報を構成する。これら各テーブルT1, T2は、例えば、キャッシュメモリ114内に記憶される。

【0099】

ボリューム管理テーブルT1は、例えば、ボリューム番号(LU#)I11と、容量I12と、RAIDグループ番号I13と、RAIDレベルI14と、ドライブ番号リストI15と、最新アクセス時刻I16とのような項目を対応付けて管理する。なお、図に示す以外の項目を管理する構成でもよい。また、テーブルT1を複数のテーブルに分割して管理する構成でもよい。

10

【0100】

論理ボリューム番号I11は、記憶制御装置内に設けられている全ての論理ボリューム250を一意に特定するための識別情報である。容量I12は、論理ボリューム250の記憶容量を示す情報である。RAIDグループ番号I13は、論理ボリューム250が設けられているRAIDグループ240を一意に特定するための識別情報である。RAIDレベルI14とは、RAIDグループ240のRAIDレベルを示す情報である。RAIDレベルとしては、例えば、RAID1, RAID5, RAID6等を挙げることができる。ドライブ番号リストI15とは、RAIDグループ240を構成しているディスクドライブ230を特定するための情報である。最新アクセス時刻I16とは、ホスト13が論理ボリューム250に最後にアクセスした時刻を示す情報である。なお、紙面の都合上、図示していないが、例えば、所定の期間毎のアクセス頻度や、各論理ボリューム250の空き容量等をテーブルT1で管理することもできる。

20

【0101】

図15は、各筐体100, 200における電源の供給状態を管理するためのテーブルT2を示す説明図である。この電源供給状態管理テーブルT2は、例えば、筐体番号I21と、筐体への電源供給状態I22と、ドライブ番号I23と、スピン状態I24と、第1系統PDBのスイッチ(図中「SW」と略記)番号I25と、そのスイッチ状態I26と、第2系統PDBのスイッチ番号I27と、そのスイッチ状態I28とをそれぞれ対応付けて管理している。テーブルT2は、各ディスクドライブ230の状態等をそれぞれ管理するための管理情報であり、例えば、状態管理マップと呼ぶこともできる。

30

【0102】

筐体番号I21とは、記憶制御装置を構成する各筐体100, 200を一意に特定するための識別情報である。筐体への電源供給状態I22とは、その筐体に外部からの電源が供給されているか否かを示す情報である。ドライブ番号I23とは、その筐体に設けられている各ディスクドライブ230を特定するための識別情報である。

【0103】

スピン状態I24とは、そのディスクドライブ230がスピニングアップしているか、スピンドアウンしているかを示す情報である。スピニングアップ状態の場合は「ON」で示され、スピンドアウン状態の場合は「OFF」で示される。スピニングアップ状態とは、ディスクが所定速度で回転している状態であり、データの読み書きが可能である。スピンドアウン状態とは、ディスクが所定速度以下で回転または停止している状態であり、この状態ではデータの読み書きを行うことはできない。

40

【0104】

第1系統PDBのスイッチ番号I25及び第2系統PDBのスイッチ番号I27とは、その筐体に接続されている第1電源供給系統(#0の系統)及び第2電源供給系統(#1の系統)のPDB内スイッチ310をそれぞれ特定するための情報である。スイッチ状態I26, I28とは、I25, I27で特定されたスイッチ310の開閉状態をそれぞれ示す情報である。スイッチ310が閉じている場合は「ON」で示され、スイッチ310が開いている場合は「OFF」で示される。本実施例では、信頼性を高めるために、スイッチ310は、常時開いているa接点として構成されている。

50

【 0 1 0 5 】

図 1 6 は、各増設筐体 2 0 0 の性質に応じて、データを配置した場合の様子を示す説明図である。各増設筐体 2 0 0 は、FCドライブやSATAドライブのような同一種類のディスクドライブ 2 3 0 を内蔵している。各増設筐体 2 0 0 の有するディスクドライブ 2 3 0 の性能に応じて、図 1 6 に示すように、データを配置することができる。例えば、高性能ディスクドライブ 2 3 0 を有する基本筐体 1 0 0 には、ホスト 1 3 から頻繁にアクセスされるデータ（図中アクセスデータ）を記憶させ、低性能のディスクドライブ 2 3 0 を有する増設筐体 2 0 0 には、ホスト 1 3 からのアクセス頻度が少ないアーカイブデータやバックアップデータを記憶させることができる。アクセス頻度の少ないアーカイブデータやバックアップデータを特定の筐体 2 0 0 に集めることにより、その筐体 2 0 0 に電源を供給する時間を短くすることができ、この結果、記憶制御装置の電力消費量を低減できる。

10

【 0 1 0 6 】

図 1 7 は、電源供給に関する接続構成の初期設定を確認するための処理を示すフローチャートである。どの筐体にどのPDB 3 0 0 のスイッチ 3 1 0 を接続するかは、ユーザによって決定することができる。ユーザは、自らが行った接続に関する構成を、管理端末 1 4 からコントローラ 1 1 0 のテーブル T 2 に登録することができる。しかし、ヒューマンエラーを完全に排除するのは難しい。そこで、図 1 7 に示す処理を実行することにより、ユーザの登録した電源供給の接続構成が正しいか否かを検査する。

【 0 1 0 7 】

本処理は、第 1 電源供給システム及び第 2 電源供給システムのそれぞれについて、別々に実施される。コントローラ 1 1 0 は、検査対象のスイッチ 3 1 0 に制御信号を出力し、このスイッチ 3 1 0 をオフさせる（S 1 0）。コントローラ 1 1 0 は、S 1 0 でオフに設定されたスイッチ 3 1 0 に接続されているはずの電源部 2 2 0 から、停電を示す警報信号が出力されたか否かを確認する（S 1 1）。各電源部 1 2 0、2 2 0 は、外部からの電源供給が途絶えた場合に、警報信号を外部に出力できるように構成されている。

20

【 0 1 0 8 】

S 1 1 でオフに設定されたスイッチ 3 1 0 に接続されているはずの電源部 2 2 0 から警報信号が出力された場合（S11:YES）、コントローラ 1 1 0 は、ユーザにより登録された情報が正しいものと判断する（S 1 2）。これに対し、S 1 1 でオフに設定されたスイッチ 3 1 0 に接続されているはずの電源部 2 2 0 から警報信号が出力されなかった場合（S1 1:NO）、コントローラ 1 1 0 は、ユーザにより登録された情報が誤っていると判断する（S 1 3）。ユーザから登録された情報に誤りが検出された場合、PDB 3 0 0 によって電源供給を停止させないようにする（S 1 4）。

30

【 0 1 0 9 】

コントローラ 1 1 0 は、全てのスイッチ 3 1 0 について接続状態の判定を行ったか否かを判断する（S 1 5）。未判定のスイッチ 3 1 0 が存在する場合（S15:NO）、コントローラ 1 1 0 は、検査対象のスイッチを次のスイッチ 3 1 0 に変更し（S 1 6）、S 1 0 ~ S 1 5 を繰り返す。全てのスイッチ 3 1 0 について検査を完了した場合（S15:YES）、コントローラ 1 1 0 は本処理を終了する。なお、S 1 4 でも示した通り、接続状態に誤りが検出された場合、少なくとも接続の誤っている箇所に関するスイッチ操作は行わない。電源供給に関する接続に誤りが検出された場合、その誤りが正されるまでの間、PDB 3 0 0 による電源供給の停止を禁止することができる。

40

【 0 1 1 0 】

図 1 8 は、1つのモードによる電源供給停止処理を示すフローチャートである。この処理では、各ディスクドライブ 2 3 0 毎に省電力を行う。本処理は、図 1 中で述べたモード F 1 の一例である。

【 0 1 1 1 】

コントローラ 1 1 0 は、テーブル T 1 を参照することにより、所定時間以上アクセスされていないディスクドライブ 2 3 0 が存在するか否かを判定する（S 2 0）。ユーザは、例えば、1時間等のように、任意の値を所定時間として設定できる。

50

【 0 1 1 2 】

コントローラ 1 1 0 は、所定時間以上アクセスされていないディスクドライブ 2 3 0 を検出した場合 (S20: YES)、そのディスクドライブ 2 3 0 をスピンドウン状態に変更させるための指示を、そのディスクドライブ 2 3 0 の存在する増設筐体 2 0 0 に向けて発行する (S 2 1)。コントローラ 1 1 0 は、 S 2 0 で検出されたディスクドライブ 2 3 0 のスピン状態をスピノフとしてテーブル T 2 に記録する (S 2 2)。コントローラ 1 1 0 から発行された指示 (コマンド) は、バックエンドスイッチ 4 0 0 を介して、目的の増設筐体 2 0 0 に伝達される。

【 0 1 1 3 】

コントローラ 1 1 0 からの指示を受領した増設筐体 2 0 0 の制御基板 2 1 0 は、指定されたディスクドライブ 2 3 0 をスピンドウンさせる (S 2 3)。また、制御基板 2 1 0 は、それが可能であれば、冷却ファン 2 0 2 の回転数を低下させる (S 2 4)。

10

【 0 1 1 4 】

例えば、筐体内で他のディスクドライブ 2 3 0 が稼働している場合は、筐体内の温度を保つために、冷却ファン 2 0 2 の回転数は一定値に保持される。これに対し、 S 2 3 で、1 つまたは複数のディスクドライブ 2 3 0 をスピンドウンさせたことにより、冷却能力を低下させることが可能な場合、制御基板 2 1 0 は、冷却ファン 2 0 2 の回転数を低下させることもできる。なお、冷却ファン 2 0 2 の回転数を低下させるか否かは、コントローラ 1 1 0 によって決定してもよいし、あるいは、増設筐体 2 0 0 の制御基板 2 1 0 によって決定してもよい。

20

【 0 1 1 5 】

図 1 9 は、他のモードによる電源供給停止処理を示すフローチャートである。本処理では、筐体内の全ディスクドライブ 2 3 0 への通電停止 (図 1 中のモード F 2) と、筐体全体への電源供給停止 (図 1 中のモード F 3) とがそれぞれ示されている。

【 0 1 1 6 】

コントローラ 1 1 0 は、テーブル T 2 を用いることにより、対象筐体内の全ディスクドライブ 2 3 0 が、スピンドウン状態になっているか否かを判定する (S 3 0)。図 1 8 で述べた通り、各ディスクドライブ 2 3 0 へのアクセスが所定時間以上無い場合は、各ディスクドライブ 2 3 0 は個別にスピンドウン状態に置かれる。スピンドウン状態に置かれたディスクドライブ 2 3 0 についてデータアクセスが発生した場合、そのディスクドライブ 2 3 0 は、スピニアップ状態に変更される。

30

【 0 1 1 7 】

対象筐体内の全ディスクドライブ 2 3 0 がスピンドウン状態であると判定された場合 (S30: YES)、コントローラ 1 1 0 は、対象筐体の制御基板 2 1 0 に向けて、全ディスクドライブ 2 3 0 への通電停止を指示する (S 3 1)。対象筐体の制御基板 2 1 0 は、電源部 2 2 0 内のスイッチ 2 2 3 を開くことにより、出力 V 1 , V 2 をそれぞれ停止させ、筐体内の全ディスクドライブ 2 3 0 への通電を停止させる (S 3 2)。なお、この場合、コントローラ 1 1 0 は、全ディスクドライブ 2 3 0 への通電停止のみならず、冷却ファン 2 0 2 の停止を要求することができる。

【 0 1 1 8 】

対象筐体内の各ディスクドライブ 2 3 0 のうちいずれか 1 つ以上のディスクドライブ 2 3 0 がスピニアップ状態の場合 (S30: NO)、 S 3 7 に移る。

40

【 0 1 1 9 】

コントローラ 1 1 0 は、対象筐体内の全ディスクドライブ 2 3 0 への通電停止を指示した後で、対象筐体に接続された第 1 系統 (#0) の PDB 3 0 0 のスイッチ 3 1 0 をオフさせる (S 3 3)。コントローラ 1 1 0 は、電源供給の停止を知らせる警報信号が第 1 系統側の電源部 (PS#0) 2 2 0 から出力されるのを確認する (S 3 4)。

【 0 1 2 0 】

第 1 系統側の電源部 2 2 0 から警報を受信した場合 (S34: YES)、コントローラ 1 1 0 は、対象筐体に接続された第 2 系統 (#1) の PDB 3 0 0 のスイッチ 3 1 0 をオフさせる (

50

S 3 5)。これにより、第 1 系統及び第 2 系統から対象筐体への電源供給が遮断され、対象筐体は機能を停止する。この際、第 2 系統側の電源部 2 2 0 からコントローラ 1 1 0 に向けて警報信号が出力される。

【 0 1 2 1 】

そこで、コントローラ 1 1 0 は、第 2 系統側の電源部 2 2 0 からの警報信号が受信されたことを確認し、さらに、対象筐体にアクセス不能であることを確認する (S 3 6)。コントローラ 1 1 0 は、対象筐体にアクセス不能であることを確認すると (S36:YES)、記憶制御装置内の全筐体について判断したか否かを判定し (S 3 7)、未判定の筐体が存在する場合 (S37:NO)、対象筐体を次の筐体に変更する (S 3 8)。そして、コントローラ 1 1 0 は、 S 3 0 ~ S 3 7 を繰り返す。

10

【 0 1 2 2 】

S 3 4 で警報信号を受信できなかった場合 (S34:NO)、または、 S 3 6 で対象筐体にアクセスできた場合 (S36:NO) は、何らかの異常が生じている場合である。そこで、コントローラ 1 1 0 は、異常診断処理 (S 3 9 ~ S 4 3) を実行する。

【 0 1 2 3 】

コントローラ 1 1 0 は、警報信号が予定された電源部 2 2 0 とは異なる別の電源部 2 2 0 から出力されたか否かを判定する (S 3 9)。警報信号が出力された場合 (S39:YES)、コントローラ 1 1 0 は、 S 3 3 または S 3 5 でオフに設定したスイッチ 3 1 0 をオンに戻し、PDB 3 0 0 と対象筐体との接続に異常がある旨を警告する (S 4 0)。この警告は、管理端末 1 4 の画面に表示させることができる。あるいは、予め登録されているユーザの電子メールアドレスに電子メールを送信することによって、ユーザに警告を与えることもできる。

20

【 0 1 2 4 】

警報が出力されない場合 (S39:NO)、コントローラ 1 1 0 は、 S 3 3 または S 3 5 で指示を与えた PDB 3 0 0 と正常な通信が可能か否かを判断する (S 4 1)。通信を正常に行える場合 (S41:YES)、コントローラ 1 1 0 は、目的のスイッチ 3 1 0 をオフにできなかった PDB 3 0 0 の備える警報用 LED ランプを点灯させる (S 4 2)。コントローラ 1 1 0 から正常にコマンドを与えることができるにもかかわらず、その PDB 3 0 0 が正常に動作しない場合であるから、その PDB 3 0 0 の内部に障害が発生したと判断できる。

【 0 1 2 5 】

これに対し、PDB 3 0 0 と正常に通信できない場合 (S41:NO)、コントローラ 1 1 0 は、PDB 3 0 0 とコントローラ 1 1 0 との間の通信経路 (L 1) に障害が生じたものと判定し、警告を出力する (S 4 3)。この警告は、 S 4 0 で述べたと同様に、管理端末 1 4 8 に表示させることができる。目的のスイッチ 3 1 0 をオフにできなかった原因は、PDB 3 0 0 の内部異常ではなく、その PDB 3 0 0 とコントローラ 1 1 0 との間の通信経路上に存在すると考えられるためである。

30

【 0 1 2 6 】

図 2 0 は、ホスト 1 3 から発行されたライトコマンドを処理するためのフローチャートを示す。

【 0 1 2 7 】

コントローラ 1 1 0 は、ホスト 1 3 からライトコマンド及びライトデータを受領すると (S 5 0)、ライトデータをキャッシュメモリ 1 1 4 に記憶させる (S 5 1)。コントローラ 1 1 0 は、ライトデータをキャッシュメモリ 1 1 4 に記憶させた時点で、ホスト 1 3 にライトコマンドの処理が完了した旨を通知する (S 5 2)。コントローラ 1 1 0 は、ライトコマンドに明示された論理アドレスをディスクドライブ 2 3 0 の物理的なアドレスに変換し (S 5 3)、アクセス対象のディスクドライブ 2 3 0 を特定する (S 5 4)。以下の説明では、便宜上、アクセス対象のディスクドライブ 2 3 0 をアクセス先ドライブと表現する場合がある。

40

【 0 1 2 8 】

コントローラ 1 1 0 は、テーブル T 2 を用いることにより、 S 5 4 で特定されたアクセ

50

ス先ドライブにアクセス可能か否かを判定する (S55)。アクセス先ドライブがスピ
ンアップ状態である場合、直ちにアクセス先ドライブにアクセス可能である。アクセス先
ドライブにアクセス可能であると判定された場合 (S55:YES)、コントローラ 110 は、ア
クセス先ドライブにライトデータを転送する (S56)。ライトデータは、バックエンド
スイッチ 400 を介して、アクセス先ドライブを有する増設筐体 200 に伝達される。制
御基板 210 は、ライトデータを受領すると、指定されたディスクドライブ 230 (ア
クセス先ドライブである) にライトデータを書き込む。

【0129】

これに対し、アクセス先ドライブへのアクセスが不能であると判定された場合 (S55:NO)
、後述する電源供給開始処理が実行される (S100)。このように、本実施例では、
ライトデータをアクセス先ドライブに転送する前に、アクセス先ドライブにアクセス可能
であるか否かを判断し、アクセス不能な場合は電源供給開始処理を実行する。これに代
えて、アクセス先ドライブへのライトデータの書込みができなかった場合に、電源供給開始
処理を実行させることもできる。

【0130】

図 21 は、ホスト 13 から発行されたリードコマンドを処理するためのフローチャ
ートを示す。コントローラ 110 は、ホスト 13 からリードコマンドを受領すると (S60)
、アクセス先ドライブを特定する (S61)。ここでのアクセス先ドライブとは、読出し
を要求されたデータを記憶しているディスクドライブ 230 である。

【0131】

コントローラ 110 は、ホスト 13 から要求されたデータがキャッシュメモリ 114 に
記憶されているか否かを判定する (S62)。以下の説明では、ホスト 13 により読出し
が要求されているデータをリードデータと呼ぶ場合がある。

【0132】

リードデータがキャッシュメモリ 114 に記憶されている場合 (S62:YES)、後述の S
63、S64 をスキップして、キャッシュメモリ 114 に記憶されているリードデータを
ホスト 13 に送信する (S65)。

【0133】

リードデータがキャッシュメモリ 114 に記憶されていない場合 (S62:NO)、コント
ローラ 110 は、テーブル T2 を用いることにより、アクセス先ドライブにアクセス可能
か否かを判断する (S63)。アクセス先ドライブにアクセス可能であると判定された場合
(S63:YES)、コントローラ 110 は、アクセス先ドライブからリードデータを読み出し
てキャッシュメモリ 114 に記憶させる (S64)。コントローラ 110 は、キャッシュ
メモリ 114 に記憶されたリードデータをホスト 13 に送信する (S65)。

【0134】

これに対し、アクセス先ドライブにアクセス不能であると判定された場合 (S63:NO)、
コントローラ 110 は、電源供給開始処理を実行する (S100)。

【0135】

図 22 は、電源供給開始処理 (S100) の詳細を示すフローチャートである。コント
ローラ 110 は、対象筐体の番号として、アクセス先ドライブを有する増設筐体 200 の
番号をセットする (S101)。

【0136】

コントローラ 110 は、対象筐体に接続された各系統の PDB 300 内のスイッチ 310
をそれぞれオンに設定し、対象筐体に電源を供給させる (S102)。続いて、コント
ローラ 110 は、アクセス先ドライブをスピニングアップさせるためのコマンドを、対象
筐体に送信する。対象筐体の制御基板 210 は、コントローラ 110 からのコマンドに従
って、アクセス先ドライブを起動し、スピニングアップさせる (S104)。

【0137】

ここで、アクセス先ドライブの数が、予め設定される所定数 N 以上である場合、コ
ントローラ 110 は、合計 N 個のアクセス先ドライブを N1 個、N2 個、N3 の 3 段階に分
け

10

20

30

40

50

て順番に起動させることもできる。

【 0 1 3 8 】

図 2 3 は、N 個のアクセス先ドライブを分割して順番に起動させる場合の、電源部 2 2 0 から出力される電流の値の変化を示す説明図である。図 2 3 は、理解の容易ならしめるために、電源部 2 2 0 の出力電流の容量とアクセス先ドライブで消費される電流との関係を模式的に示すものである。

【 0 1 3 9 】

時刻 t_1 において、N 1 個のアクセス先ドライブを起動させると、瞬間的に突入電流が発生する。N 1 個のアクセス先ドライブを起動させたときの突入電流が消滅する時刻 t_2 において、N 2 個のアクセス先ドライブを起動させると、突入電流が発生する。N 2 個のアクセス先ドライブを起動させたときの突入電流が消滅する時刻 t_3 において、N 3 個のアクセス先ドライブを起動させると、突入電流が発生し、短時間後に消滅する。

【 0 1 4 0 】

つまり、アクセス先ドライブを起動させた場合の突入電流が、電源部 2 2 0 の容量を超えないように、ディスクドライブ 2 3 0 を複数回に分割して起動させる。M 1 ~ M 4 は、電源容量とのマージンを示す。スピニングしたディスクドライブ 2 3 0 にアクセスすると、そのディスクドライブ 2 3 0 の消費電流は増大する。アクセス時の消費電流が電源容量を超えないように、マージン M 1 ~ M 4 を確保する。

【 0 1 4 1 】

一例を挙げると、 $N = N_1 = 12$ に設定することができる。つまり、12 個のディスクドライブ 2 3 0 を起動させる場合は、12 個のディスクドライブ 2 3 0 を一度に起動させても、十分なマージン M 1 を得ることができる。N = 15 の場合、 $N_1 = 12$, $N_2 = 2$, $N_3 = 1$ に設定することができる。最初に 12 個のディスクドライブ 2 3 0 を起動させ、次に 2 個のディスクドライブ 2 3 0 を起動させ、最後に 1 個のディスクドライブ 2 3 0 を起動させる。

【 0 1 4 2 】

このように、予め設定された所定数 N 以上のディスクドライブ 2 3 0 を起動させる場合、予め設定された複数の個数 N_1 , N_2 , N_3 の単位で ($N_1 > N_2 > N_3$)、N 個のディスクドライブ 2 3 0 を分割して起動させることもできる。これにより、電源部 2 2 0 の電源容量を超えないように消費電流を制御することができ、かつ、ディスクドライブ 2 3 0 を速やかにスピニングさせることができる。同時に起動可能な最大数 N 1 個のディスクドライブ 2 3 0 を先に起動させるため、これら N 1 個のディスクドライブ 2 3 0 の作動を早期に安定させることができ、直ちにアクセスすることができる。

【 0 1 4 3 】

図 2 4 は、自己診断処理を示すフローチャートである。例えば、バックアップデータ等のようなアクセス頻度の低いデータを記憶している増設筐体 2 0 0 への通電を長時間停止した場合、通電停止中に増設筐体 2 0 0 に障害が発生しても検出することができず、必要な場合にデータを取り出せない可能性がある。

【 0 1 4 4 】

そこで、本実施例では、定期的にはまたは不定期に、休止中の増設筐体 2 0 0 に通電することにより、正常に動作するか否かを確認する。この自己診断処理は、予め設定された所定周期で定期的に行うこともできるし、ユーザが任意に指定した時刻で不定期に行うこともできる。

【 0 1 4 5 】

コントローラ 1 1 0 は、テーブル T 2 を用いて、所定時間以上、PDB 3 0 0 から外部電源が供給されていない増設筐体 2 0 0 が存在するか否かを判定する (S 7 0)。コントローラ 1 1 0 は、所定時間以上通電されていない増設筐体 2 0 0 を発見した場合 (S70:YES)、その増設筐体 2 0 0 に接続されている PDB 3 0 0 内のスイッチ 3 1 0 をオンさせて (S 7 1)、その増設筐体 2 0 0 に外部電源を供給する。

【 0 1 4 6 】

10

20

30

40

50

コントローラ 110 は、診断用コマンドとして予め設定されているコマンド（スピンアップコマンドやスピンドアウンコマンド等）を増設筐体 200 に送信し、その増設筐体 200 内の各ディスクドライブ 230 の動作を確認する（S72）。コントローラ 110 は、その増設筐体 200 内の全ディスクドライブ 230 が正常に動作するか否かを判定する（S73）。コントローラ 110 は、全ディスクドライブ 230 が正常に動作すると判定された場合（S73:YES）、S70 に戻る。いずれか 1 つ以上のディスクドライブ 230 が正常に動作しない場合（S73:NO）、コントローラ 110 は、その増設筐体 200 に異常が生じている旨の警報を出力する（S74）。この警報は、管理端末 14 を介して、ユーザに通知される。また、管理端末 14 が管理サーバに接続されている場合、この警報は、管理端末 14 から管理サーバに送信される。

10

【0147】

図 25 は、各増設筐体 200 における節電モードの変化を模式的に示す。図 25 (a) に示すように、最初は、各ディスクドライブ 230、制御基板 210、冷却ファン 202 のそれぞれに電源部 220 から直流電源が供給されている。図 25 (b) に示すように、所定時間アクセスされなかったディスクドライブ 230 (#1) は、スピンドアウン状態に変更される。

【0148】

図 25 (c) に示すように、スピンドアウン状態となるディスクドライブ 230 が増加し、図 25 (d) に示すように、全てのディスクドライブ 230 がスピンドアウン状態に変化すると、図 25 (e) に示すように、PDB 300 から増設筐体 200 への外部電源の供給そのものが停止する。このように、本実施例では、ディスクドライブ 230 の利用状況に応じて、電力消費量を段階的に節減させていくことができ、最終的には、ほぼ電力消費量を零とすることができる。

20

【0149】

図 26 は、筐体単位で電源供給が停止される様子を示す説明図である。図中斜線で示す筐体 200 (#1, #3) は、電源供給が停止されている。

【0150】

このように構成される本実施例の効果では、ディスクドライブ 230 の利用状況に応じて、各増設筐体 200 への外部電源の供給そのものを停止させるため、従来よりも電力消費量を低減することができる。

30

【0151】

本実施例では、基本筐体 100 と各増設筐体 200 とをバックエンドスイッチ 400 を介してスイッチ接続するため、基本筐体 100 は、内蔵するディスクドライブ 230 が全てスピンドアウンしている増設筐体 200 を発見次第に直ちに電源供給を停止させることができ、電力消費量を低減させることができる。本実施例の構成から得られる他の効果は、既に述べた通りである。

【実施例 2】**【0152】**

図 27、図 28 に基づいて、本発明の第 2 実施例を説明する。本実施例を含む以下の各実施例は、第 1 実施例の変形例に該当する。本実施例では、コントローラ 110 の異常を PDB 300 が検出した場合に、PDB 300 内のスイッチ 310 を全てオンに設定するようにしている。

40

【0153】

図 27 は、記憶制御装置の一部を抜き出して示す回路図である。PDB 300 の制御回路 320A には、診断部 321 が設けられている。診断部 321 は、後述の診断処理を行い、診断結果に基づいて各スイッチ 310 を制御する。

【0154】

図 28 は、診断部 321 により実行される処理を示すフローチャートである。診断部 321 は、所定時間が経過したか否かを判定し（S80）、所定時間が経過した場合（S80:YES）、信号線 L1 を介してコントローラ 110 と通信を試みる（S81）。

50

【 0 1 5 5 】

PDB 3 0 0 内の診断部 3 2 1 は、コントローラ 1 1 0 と正常に通信できたか否かを判定する (S 8 2)。正常に通信できた場合 (S82:YES)、S 8 0 に戻って、所定時間が経過するのを再び待つ。

【 0 1 5 6 】

これに対し、正常に通信できなかった場合 (S82:NO)、診断部 3 2 1 は、PDB 3 0 0 内の各スイッチ 3 1 0 を全てオンに設定する (S 8 3)。S 8 2 で「NO」と判定される場合、コントローラ 1 1 0 に異常が生じており、コントローラ 1 1 0 がスイッチ 3 1 0 を適切にオンオフさせることができないと考えられるためである。

【 0 1 5 7 】

このように構成される本実施例も前記第 1 実施例と同様の効果を奏する。これに加えて、本実施例では、PDB 3 0 0 がコントローラ 1 1 0 の異常を検出した場合に、各スイッチ 3 1 0 を全てオンに設定するため、安全性や信頼性が向上する。全スイッチ 3 1 0 をオンにすることにより、電力消費量を低減することはできなくなる。しかし、各増設筐体 2 0 0 への通電は保証されるため、フェイルセーフの観点から好ましい。

【 実施例 3 】

【 0 1 5 8 】

図 2 9 は、第 3 実施例に係る記憶制御装置の接続構成を模式的に示す説明図である。図 2 9 では、便宜上、バックエンドスイッチ 4 0 0 と各増設筐体 2 0 0 とを接続するための信号線 L 3 を取り除いてある。本実施例では、基本筐体 1 0 0 は、バックエンドスイッチ 4 0 0 を介して、各 PDB 3 0 0 にそれぞれ接続される。従って、本実施例では、信号線 L 1 は廃止される。

【 実施例 4 】

【 0 1 5 9 】

図 3 0 ~ 図 3 2 に基づいて、第 4 実施例を説明する。本実施例では、複数の記憶制御装置を備えるストレージシステムにおいて、1つの記憶制御装置 1 0 0 0 が他の記憶制御装置 1 1 0 0 , 1 2 0 0 の有する記憶領域を仮想化してホスト 1 3 に提供する。

【 0 1 6 0 】

図 3 0 に示すストレージシステムは、複数の記憶制御装置 1 0 0 0 , 1 1 0 0 , 1 2 0 0 を備えている。記憶制御装置 1 0 0 0 は、他の各記憶制御装置 1 1 0 0 , 1 2 0 0 に接続されている接続元の記憶制御装置であり、他の各記憶制御装置 1 1 0 0 , 1 2 0 0 を管理下においている。

【 0 1 6 1 】

接続元の記憶制御装置 1 0 0 0 は、接続先の記憶制御装置 1 1 0 0 , 1 2 0 0 がそれぞれ有する記憶領域を、記憶制御装置 1 0 0 0 の内部で仮想化し、それがあたかも記憶制御装置 1 0 0 0 の内部ボリュームであるかのように、ホスト 1 3 に提供する。従って、接続元の記憶制御装置 1 0 0 0 は、各記憶制御装置 1 0 0 0 , 1 1 0 0 , 1 2 0 0 のシステム構成に関する情報を一元的に管理している。

【 0 1 6 2 】

各記憶制御装置 1 0 0 0 , 1 1 0 0 , 1 2 0 0 は、それぞれ同一の基本構造を備えて構成することもできるし、それぞれ異なる構造を備えてもよい。本実施例では、記憶制御装置 1 0 0 0 , 1 1 0 0 , 1 2 0 0 は、前記第 1 実施例で述べたような構成をそれぞれ備えているものとする。

【 0 1 6 3 】

接続元の記憶制御装置 1 0 0 0 の基本筐体 1 0 0 と、接続先の各記憶制御装置 1 1 0 0 , 1 2 0 0 の各基本筐体 1 0 0 とは通信経路 L 5 を介して接続されている。この通信経路 L 5 は、例えば、SAN等の通信ネットワークとして構成することができる。また、接続元の記憶制御装置 1 0 0 0 の基本筐体 1 0 0 と、接続先の各記憶制御装置 1 1 0 0 , 1 2 0 0 の各 PDB 3 0 0 とは、信号線 L 6 を介して接続されている。

【 0 1 6 4 】

図31は、本実施例による電源供給停止処理を示すフローチャートである。この処理は、接続元記憶制御装置1000の基本筐体100内のコントローラ110によって実行される。

【0165】

接続元のコントローラ110は、処理対象の記憶制御装置として、装置番号の最も少ない記憶制御装置1100を選択する(S300)。コントローラ110は、対象記憶制御装置1100の有する全てのボリューム250について、アクセス不要であるか否かを判定する(S301)。

【0166】

記憶制御装置1100の有する全ボリューム250についてアクセス不要であると判断された場合(S301:YES)、接続元のコントローラ110は、記憶制御装置1100の基本筐体100に向けて、シャットダウンコマンドを発行する(S302)。

【0167】

シャットダウンコマンドを受領した記憶制御装置1100のコントローラ110は、シャットダウン処理を実行する(S303)。シャットダウン処理では、例えば、下位の増設筐体200から順番に電源の供給を停止させていき、最後に基本筐体100の動作も停止させる。

【0168】

接続元のコントローラ110は、記憶制御装置1100に接続されている各PDB300内の各スイッチ310を下位側から順番にオフさせる(S304)。これにより、接続先の記憶制御装置1100の基本筐体100及び各増設筐体200への外部電源の供給が停止される。

【0169】

接続元のコントローラ110は、全ての接続先記憶制御装置について判断したか否かを判定する(S305)。未判定の記憶制御装置が存在する場合(S305:NO)、コントローラ110は、対象の記憶制御装置の装置番号を次の接続先記憶制御装置1200の装置番号に変更し(S306)、S301～S305を繰り返す。

【0170】

図32は、電源供給を開始する場合の処理を示すフローチャートである。接続元のコントローラ110は、対象の記憶制御装置に接続された各PDB300内の各スイッチ310を上位側から順番にオンさせる(S310)。

【0171】

接続先の記憶制御装置では、各系統のPDB300から外部電源が供給されると、まず最初に、基本筐体100を起動させ(S311)、次に、外部電源が供給された順番で、各増設筐体200を起動させる(S312)。

【0172】

このように構成される本実施例も前記第1実施例と同様の効果を奏する。これに加えて、本実施例では、複数の記憶制御装置1000, 1100, 1200を備えるストレージシステムにおいて、より一層、電力消費量を節減することができる。

【実施例5】

【0173】

図33は、第5実施例に係るストレージシステムを示す説明図である。本実施例では、接続先の各記憶制御装置1100, 1200の各PDB300に電源を供給する電源装置15の動作を、接続元記憶制御装置1000の基本筐体100が制御する。

【0174】

このように構成される本実施例も前記第4実施例と同様の効果を奏する。

【0175】

なお、本発明は、上述した実施の形態に限定されない。当業者であれば、本発明の範囲内で、種々の追加や変更等を行うことができる。例えば、各実施例を適宜組み合わせることができる。

10

20

30

40

50

【図面の簡単な説明】

【0176】

【図1】本発明の実施形態の概念を示す説明図である。

【図2】記憶制御装置の外観を示す斜視図である。

【図3】基本筐体を前側から見た斜視図である。

【図4】基本筐体を後側から見た斜視図である。

【図5】増設筐体を前側から見た斜視図である。

【図6】増設筐体を後側から見た斜視図である。

【図7】記憶制御装置の接続構成を模式的に示す説明図である。

【図8】記憶制御装置の一部を抜き出して示す回路図である。

10

【図9】記憶制御装置の電源供給構造を示す説明図である。

【図10】記憶制御装置の構成を簡略化して示す説明図である。

【図11】コントローラの構成を示す説明図である。

【図12】増設筐体の構成を示す説明図である。

【図13】制御基板の構成を示す説明図である。

【図14】ボリュームを管理するためのテーブルを示す説明図である。

【図15】電源供給状態を管理するためのテーブルを示す説明図である。

【図16】ディスクドライブの性質に応じてデータを再配置した様子を示す説明図である。

【図17】電源供給の接続を確認するための処理を示すフローチャートである。

20

【図18】各ドライブ毎に節電するための処理を示すフローチャートである。

【図19】増設筐体毎に電源供給を停止等するための処理を示すフローチャートである。

【図20】ライト処理を示すフローチャートである。

【図21】リード処理を示すフローチャートである。

【図22】電源供給を開始する場合の処理を示すフローチャートである。

【図23】複数のディスクドライブを起動させる場合に、同時に起動可能な最大数のディスクドライブを先に起動させ、続いて残りのディスクドライブを起動させる場合の電流値の変化を示す説明図である。

【図24】通電停止中の増設筐体を診断するための処理を示すフローチャートである。

【図25】ディスクドライブの利用状況に応じて節電モードが段階的に変化していく様子
を示す説明図である。

30

【図26】通電が完全に停止された筐体と通電されている筐体とが同一の記憶制御装置内
で混在している様子を示す説明図である。

【図27】第2実施例に係る記憶制御装置の一部を抜き出して示す回路図である。

【図28】PDBにより実行される診断処理を示すフローチャートである。

【図29】第3実施例に係る記憶制御装置の接続構成を示す説明図である。

【図30】第4実施例に係る記憶制御装置が用いられるストレージシステムの全体構成を
示す説明図である。

【図31】記憶制御装置単位で電源供給を停止させる処理を示すフローチャートである。

【図32】記憶制御装置単位で電源供給を再開させる処理を示すフローチャートである。

40

【図33】第5実施例に係る記憶制御装置が用いられるストレージシステムの全体構成を
示す説明図である。

【符号の説明】

【0177】

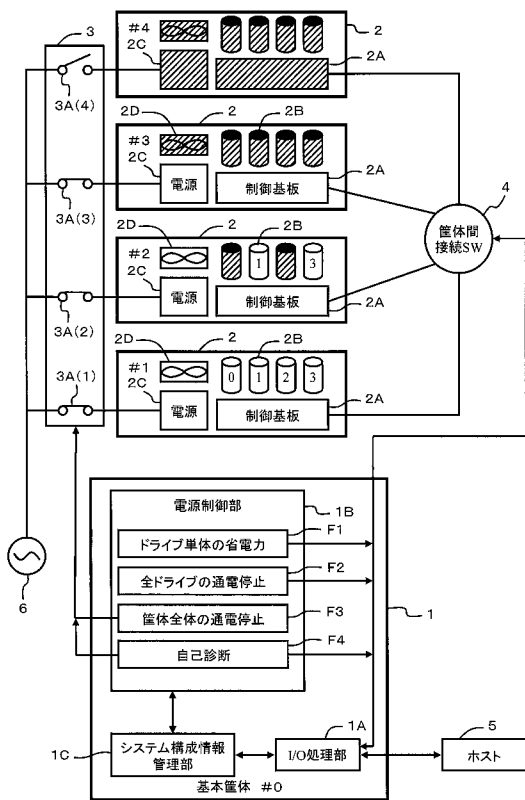
1 ... 基本筐体、1 A ... I/O処理部、1 B ... 電源制御部、1 C ... システム構成情報管理部
、2 ... 増設筐体、2 A ... 制御基板、2 B ... ディスクドライブ、2 C ... 筐体内電源部、2 D
... 冷却ファン、3 ... 電源供給回路部、3 A ... スイッチ、4 ... 筐体間接続スイッチ、5 ... ホ
スト、6 ... 外部電源、10 ... 基本ラック、11 ... 増設ラック、13 ... ホスト、14 ... 管理
端末、15 ... 電源装置、CN1, CN2 ... 通信ネットワーク、100 ... 基本筐体、101 ...
ケーシング、101 F ... 前面側、101 R ... 後面側、110 ... コントローラ、111 ... ホ

50

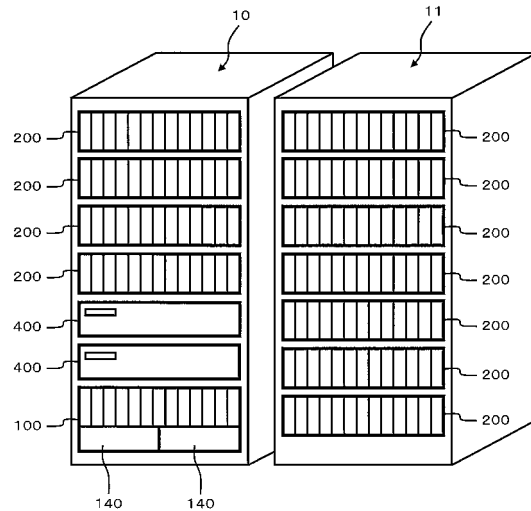
ストインターフェース回路、112...ドライブインターフェース回路、113...データ転送回路、114...キャッシュメモリ、115...プログラムメモリ、116...プロセッサ、117...ブリッジ、118...インターフェース、120...電源部、130...冷却ファン、140...バッテリー、200...増設筐体、201...ケーシング、201F...前面側、201R...後面側、202...冷却ファン、210...制御基板、211...マイクロプロセッサ部、211A...資源管理機能、211B...コマンド処理機能、211C...電源制御機能、211D...冷却ファン制御機能、212...通信インターフェース部、213...制御信号出力回路、220...電源部、221...電圧変換回路、222A~222C...出力ダイオード、223...出力制御スイッチ、230...ディスクドライブ、240...RAIDグループ、250...論理ボリュームボリューム、300...電源分配回路部(PDB)、310...スイッチ、320, 320A...制御回路、321...診断部、330...AC/DC電源部、340...ブレーカ、400...バックエンドスイッチ、410...切替回路、1000, 1100, 1200...記憶制御装置

10

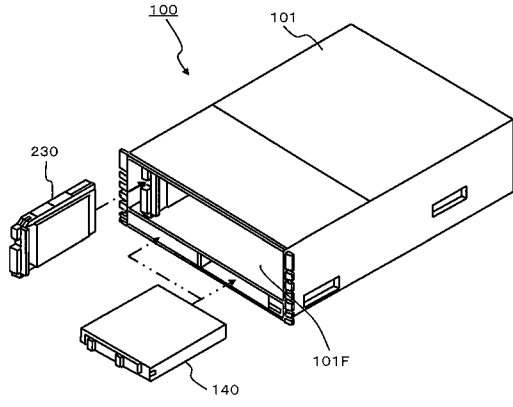
【図1】



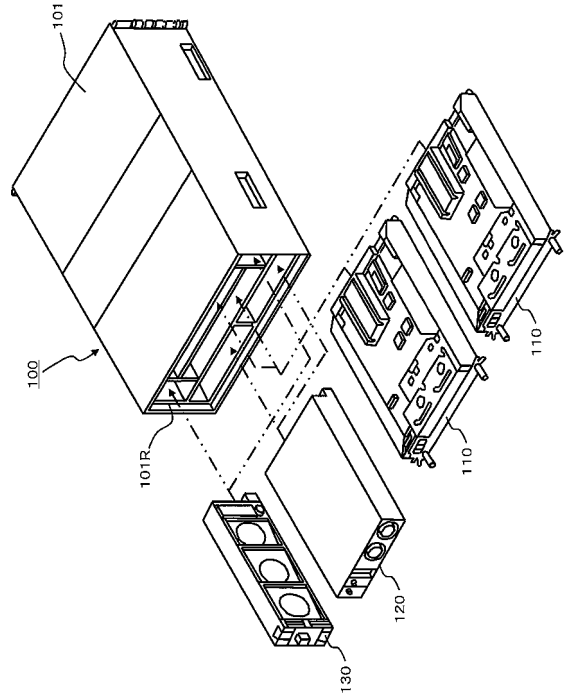
【図2】



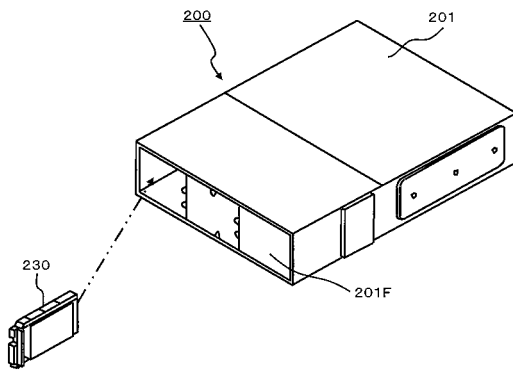
【図3】



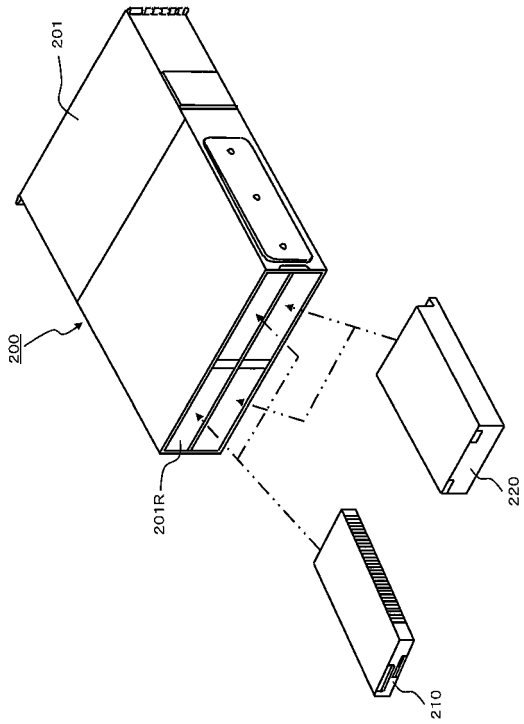
【図4】



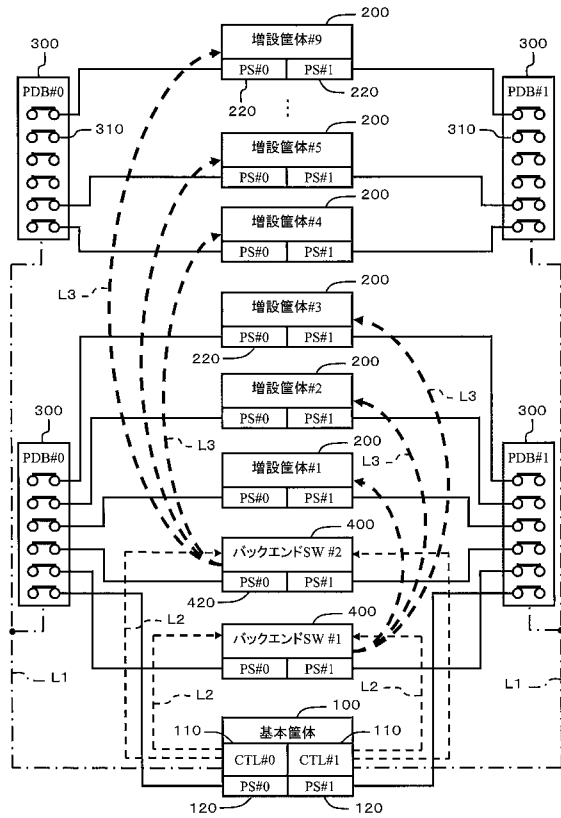
【図5】



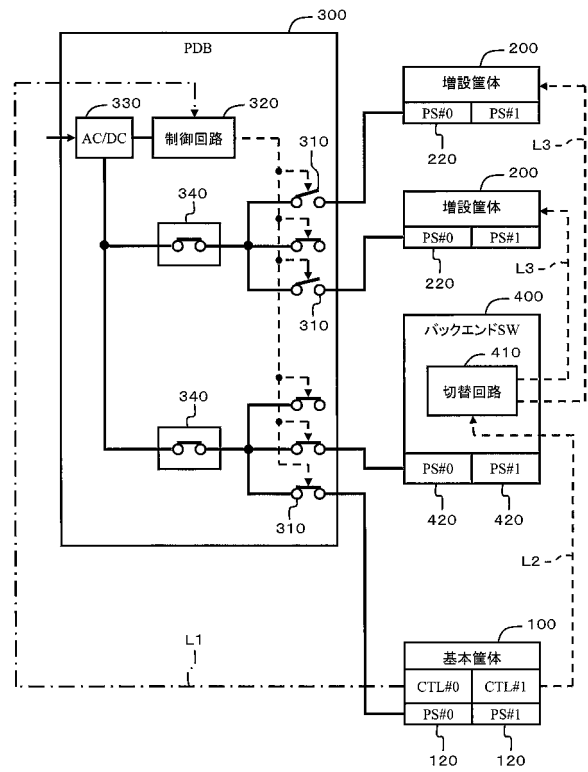
【図6】



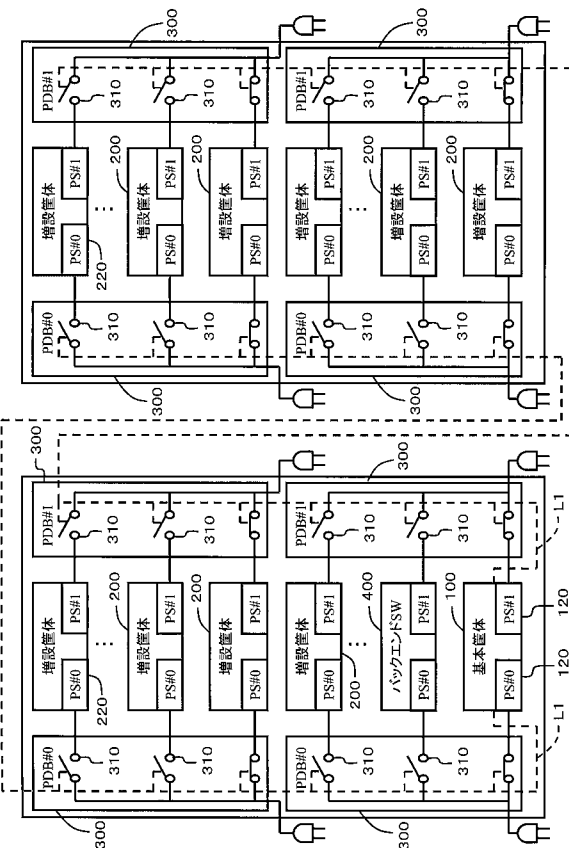
【図7】



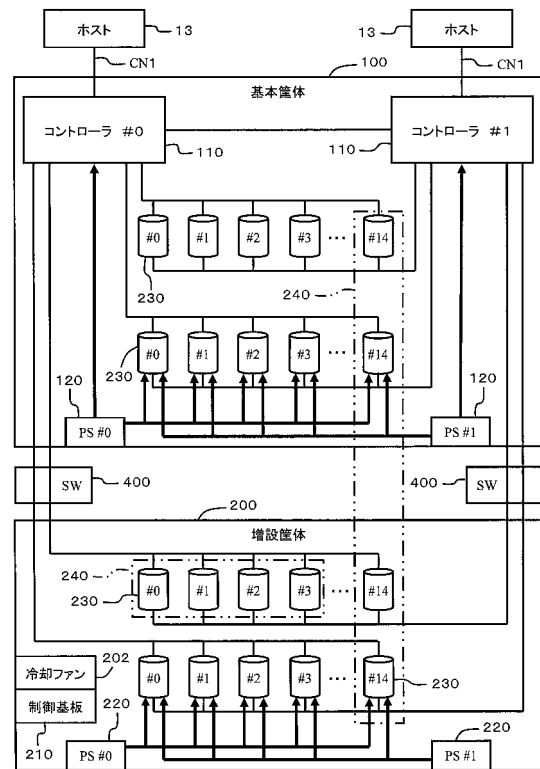
【図8】



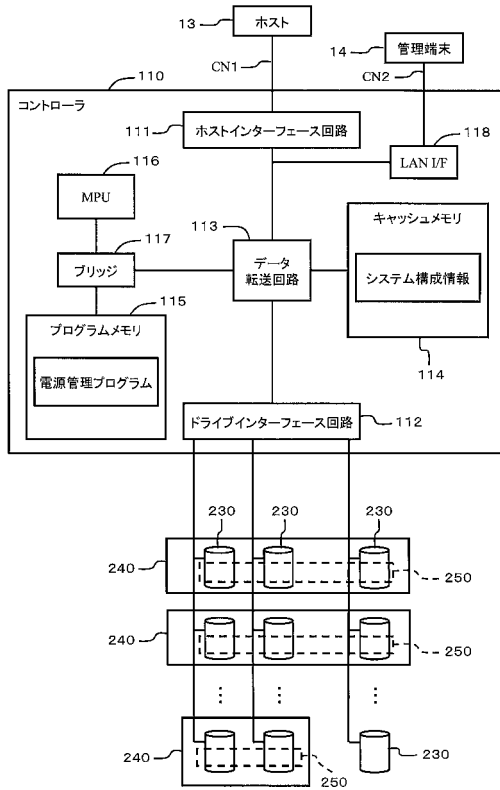
【図9】



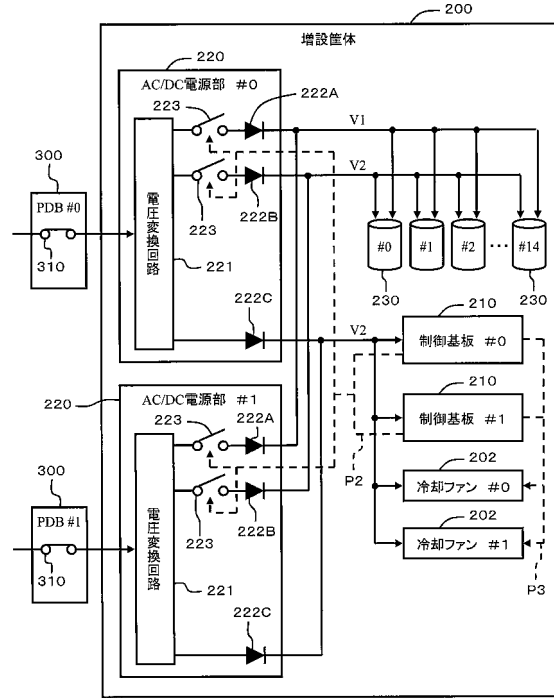
【図10】



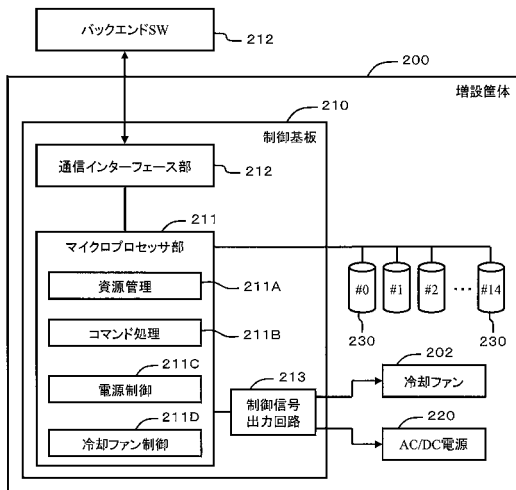
【図 1 1】



【図 1 2】



【図 1 3】



【図 1 4】

T1

ボリューム管理テーブル					
I11	I12	I13	I14	I15	I16
LU#	容量	RAIDグループ#	RAIDレベル	ドライブ#リスト	最新アクセス時刻
LU#0	200MB	RG01	RAID5	HDD01,02,03,04	2006/09/25 12:00:00
...

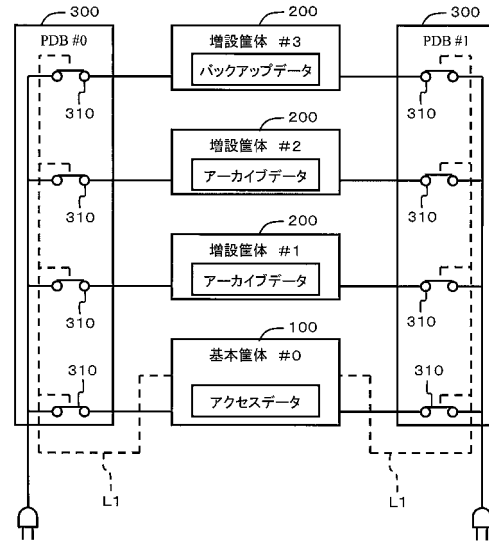
【図15】

T2

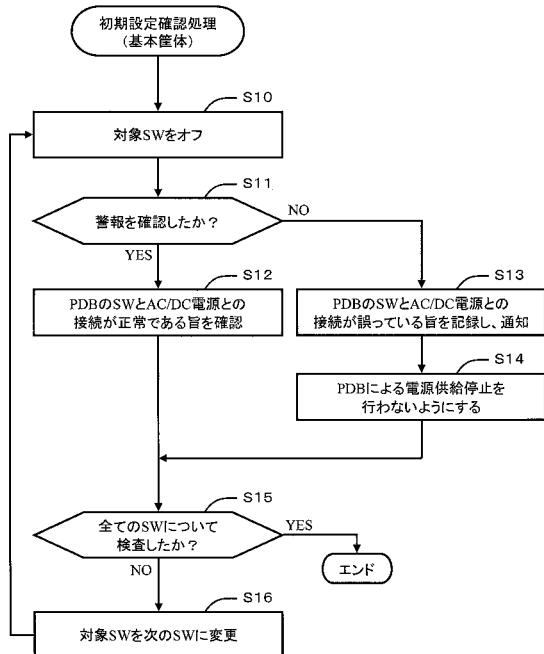
電源供給状態管理テーブル

I21	I22	I23	I24	I25	I26	I27	I28
筐体#	筐体への電源供給状態	ドライブ#	スピン状態	第1PDB SW#	SW状態	第2PDB SW#	SW状態
#0	ON	HDD00	ON	01-0	ON	11-0	ON
		HDD01	ON				
		HDD02	ON				
					
		HDD14	ON				
#1	ON	HDD00	ON	01-1	ON	11-1	ON
		HDD01	OFF				
		HDD02	ON				
					
		HDD14	ON				
#2	ON	HDD00	ON	01-2	ON	11-2	ON
		HDD01	OFF				
		HDD02	ON				
					
		HDD14	ON				
#3	ON	HDD00	ON	01-3	ON	11-4	ON
		HDD01	OFF				
		HDD02	ON				
					
		HDD14	ON				
...							

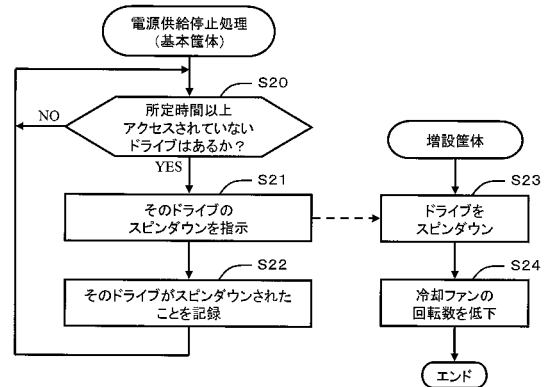
【図16】



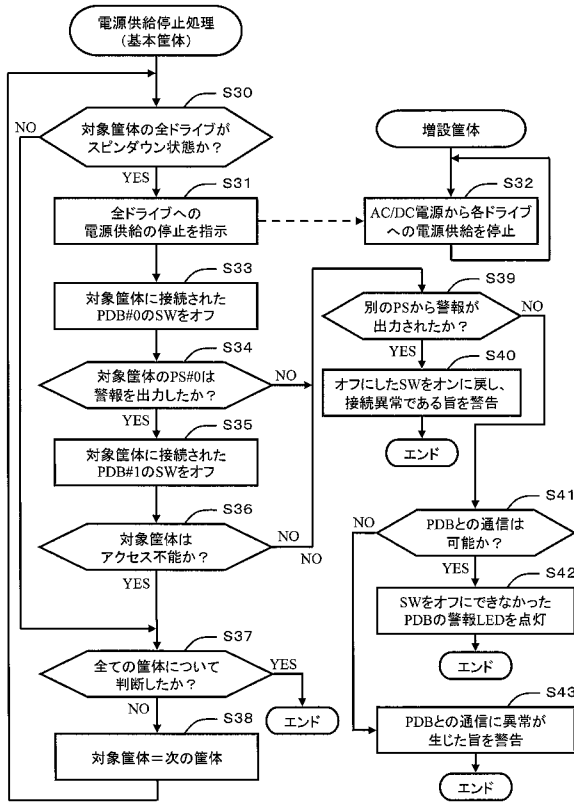
【図17】



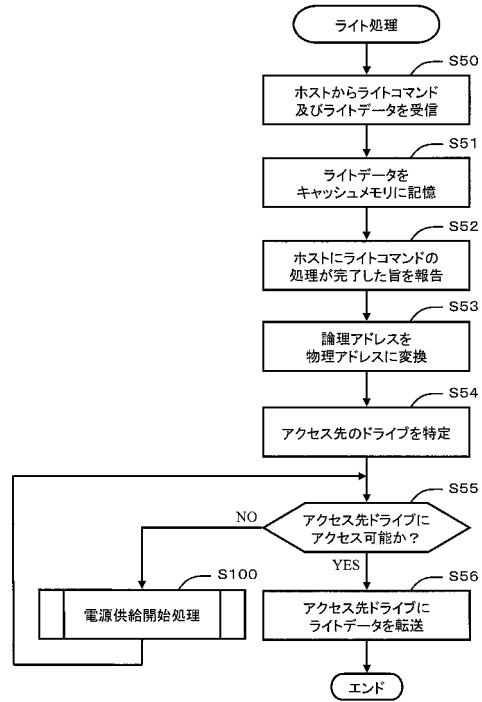
【図18】



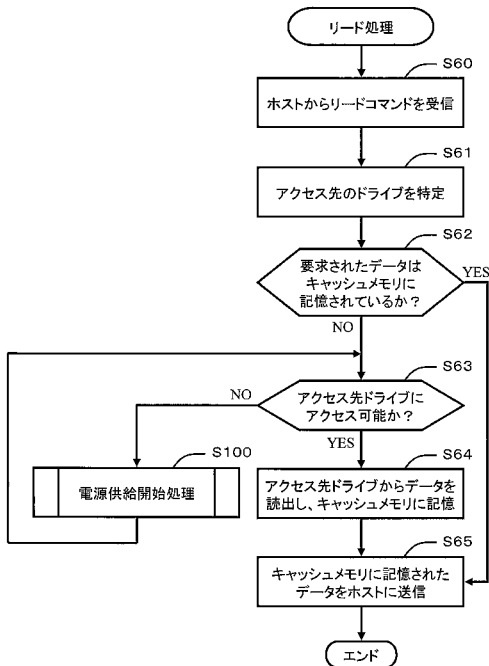
【図19】



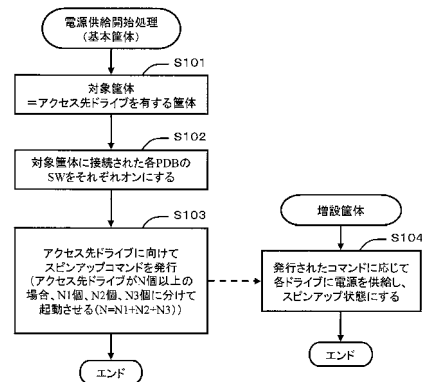
【図20】



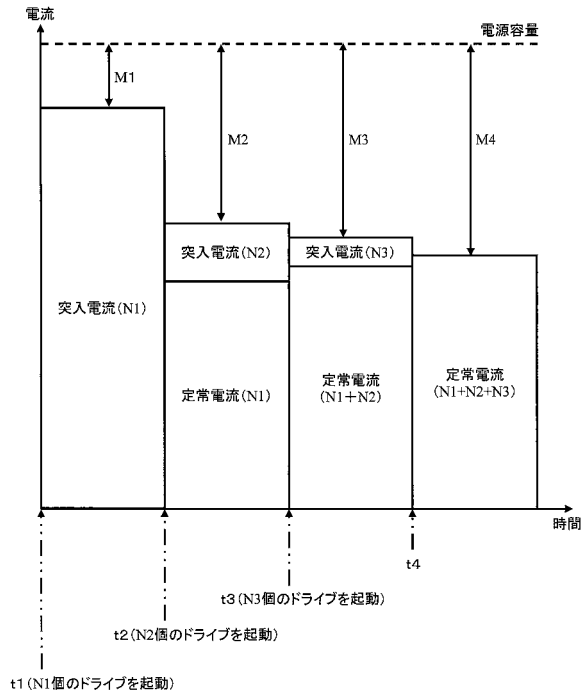
【図21】



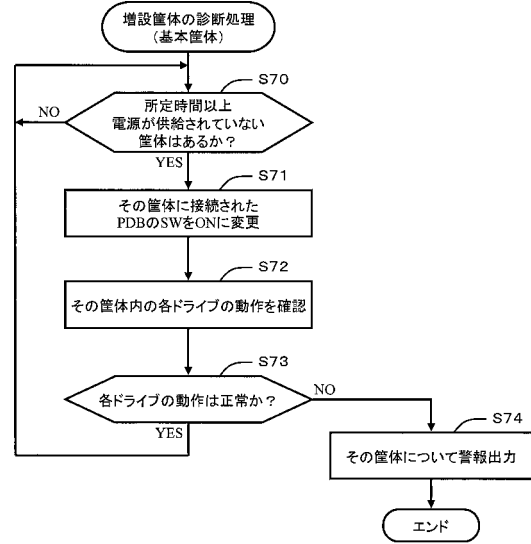
【図22】



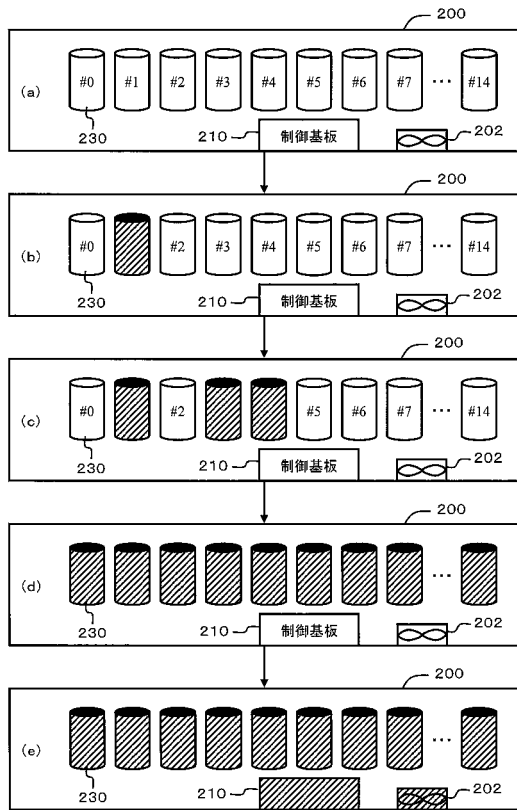
【図 2 3】



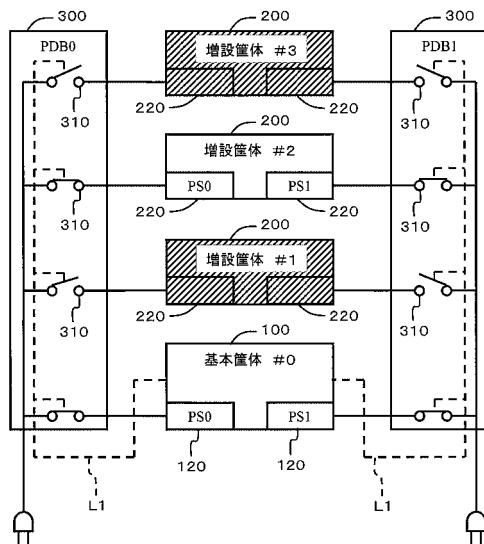
【図 2 4】



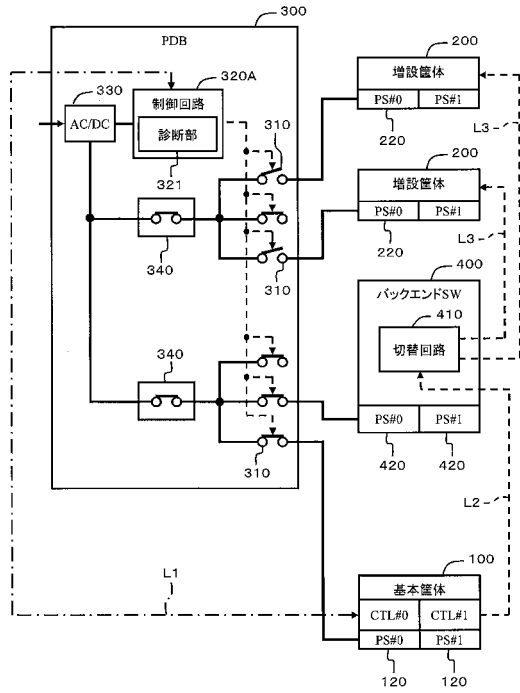
【図 2 5】



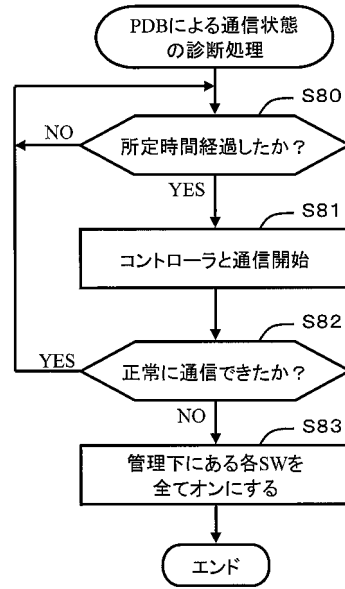
【図 2 6】



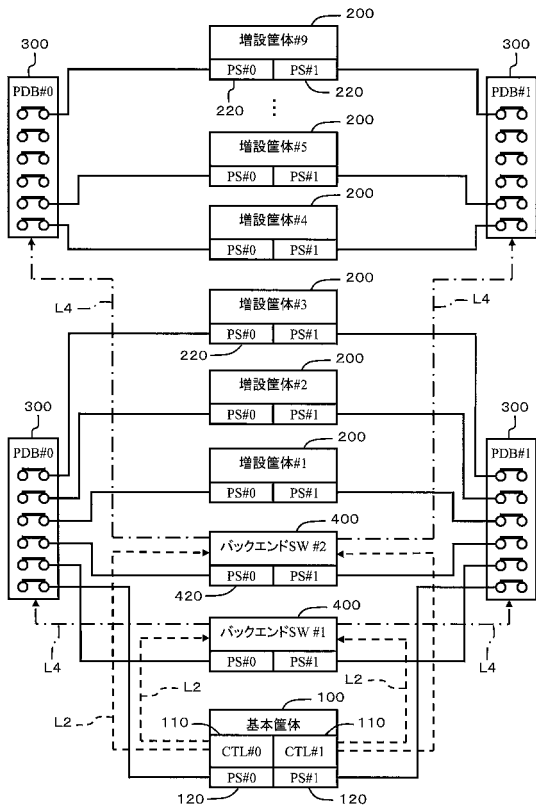
【図 27】



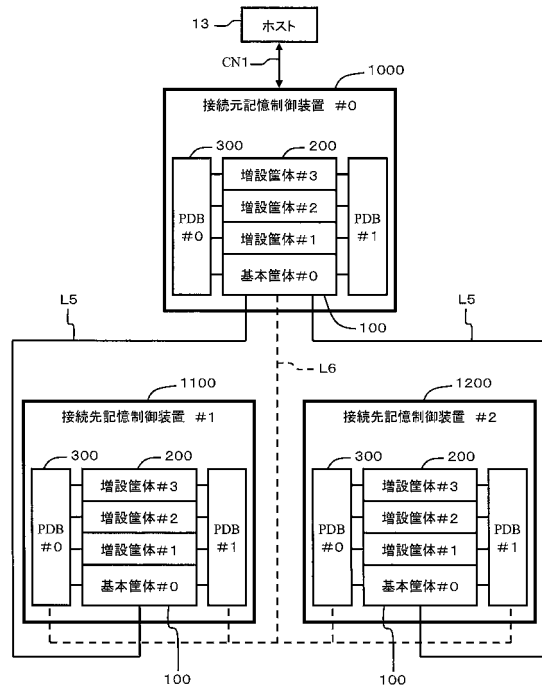
【図 28】



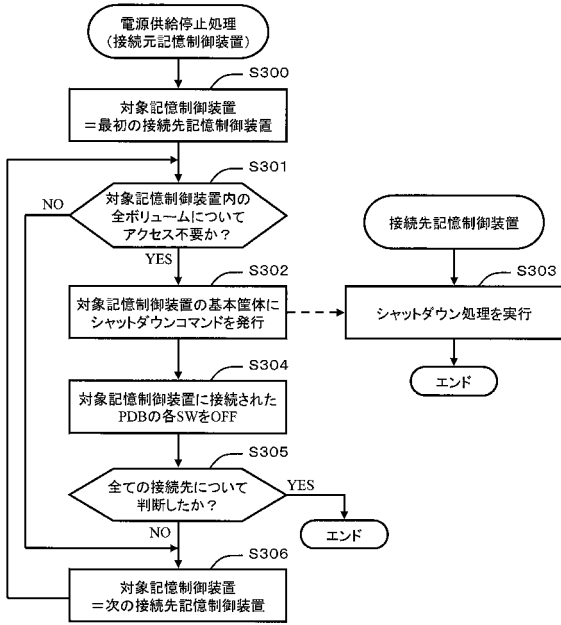
【図 29】



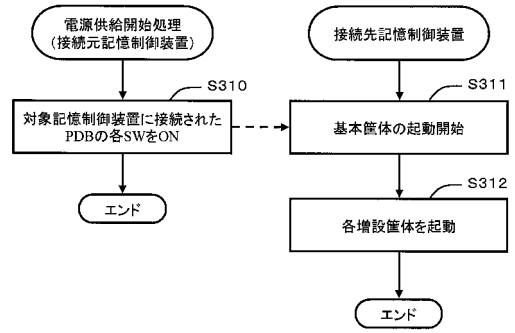
【図 30】



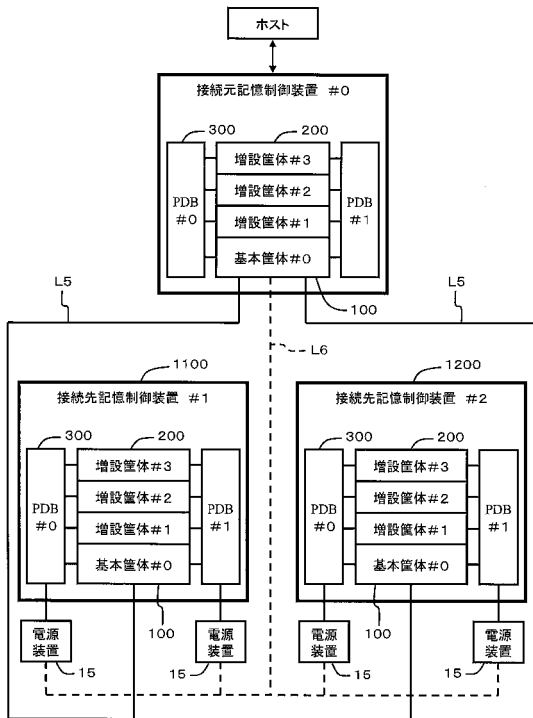
【図31】



【図32】



【図33】



フロントページの続き

審査官 坂東 博司

- (56)参考文献 特開昭62-063325(JP,A)
特開2001-339853(JP,A)
特開2005-234825(JP,A)
特開2000-293314(JP,A)
特開2006-285464(JP,A)
特開2001-337789(JP,A)
特開2008-090352(JP,A)

(58)調査した分野(Int.Cl., DB名)
G06F 3/06