

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5147586号
(P5147586)

(45) 発行日 平成25年2月20日 (2013. 2. 20)

(24) 登録日 平成24年12月7日 (2012.12.7)

(51) Int. Cl.	F I		
G06F 3/06 (2006.01)	G06F 3/06	3 O 1 Z	
G06F 13/10 (2006.01)	G06F 13/10	3 4 O A	
G06F 12/08 (2006.01)	G06F 12/08	5 5 7	
	G06F 12/08	5 7 9	
	G06F 12/08	5 1 1 Z	

請求項の数 14 (全 20 頁)

(21) 出願番号	特願2008-197052 (P2008-197052)	(73) 特許権者	000005108
(22) 出願日	平成20年7月30日 (2008. 7. 30)		株式会社日立製作所
(65) 公開番号	特開2010-33466 (P2010-33466A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成22年2月12日 (2010. 2. 12)	(74) 代理人	100093861
審査請求日	平成22年9月3日 (2010. 9. 3)		弁理士 大賀 真司
		(72) 発明者	高橋 秀明
			神奈川県小田原市中里322番2号 株式会社日立製作所 R A I D システム事業部内
		審査官	木村 貴俊
		(56) 参考文献	特開2001-256003 (JP, A)
)
			特開2006-107080 (JP, A)
)

最終頁に続く

(54) 【発明の名称】 ストレージ装置及びその制御方法

(57) 【特許請求の範囲】

【請求項1】

計算機からの記憶資源に対するリード及びライト要求を処理する制御資源をモジュール化し、少なくとも1つのモジュールを備え、他のモジュールを順次前記少なくとも一つのモジュールに追加できるように構成されたストレージ装置であって、

前記複数のモジュール同士を高速内部バス規格で接続する接続路と、

前記複数のモジュールのそれぞれの管理部に接続する管理ネットワークと、

前記管理ネットワークに接続する管理装置と

を備え、

前記複数の管理部の少なくとも一つがアクティブとし、他の管理部がスタンバイとして設定され、

前記管理装置は前記複数の管理部のうち前記複数のモジュールを統一管理する当該アクティブの管理部にアクセスするインターフェースを備え、

前記複数のモジュールに新たなモジュールが接続されると、前記新たなモジュールの管理部をスタンバイに設定し、

前記複数のモジュールのうち1のモジュールの接続を解除すると、当該接続を解除するモジュールの管理部がアクティブの場合、当該接続を解除するモジュールの管理部をスタンバイにし、前記複数のモジュールのうち他のモジュールの管理部をアクティブにし、当該接続を解除するモジュールの接続を解除する

ストレージ装置。

【請求項 2】

前記計算機から、前記複数のモジュールのうち、第1のモジュールへのリード又はライトコマンドを、前記接続路を介して、第2のモジュールが受領し、当該コマンドを処理する

請求項1記載のストレージ装置。

【請求項 3】

前記第1のモジュールと前記第2のモジュールとはそれぞれ、

前記計算機を接続する第1のアダプタと、

記憶資源を接続する第2のアダプタと、

前記計算機との間で送受信されるデータを一時記憶するキャッシュメモリと、

前記リード及びライト要求に対するプログラムを実行するマイクロプロセッサと前記接続路で接続された前記複数のモジュールの管理情報を格納するローカルメモリとを有するマイクロプロセッサパッケージと、

前記第1のアダプタと、前記第2のアダプタと、前記キャッシュメモリと、そして、前記マイクロプロセッサと、を互いに接続するスイッチとを備える

請求項2記載のストレージ装置。

【請求項 4】

前記接続路は、前記第1のモジュールのスイッチと、前記第2のスイッチとの間に形成されている

請求項3記載のストレージ装置。

【請求項 5】

前記複数のモジュールのうち、第1のモジュールの前記管理部が前記アクティブのモードにあり、第2のモジュールの前記管理部が前記スタンバイのモードにあり、前記第1のモジュールの前記管理部が前記第2のモジュールの前記管理部から前記第2のモジュールの管理情報を取得し、当該取得した前記第2のモジュールの管理情報を前記第1のモジュールの管理情報とともに保存する

請求項1記載のストレージ装置。

【請求項 6】

前記第1のモジュールの前記管理部は前記保存した管理情報を前記第2のモードの前記管理部にコピーする

請求項5記載のストレージ装置。

【請求項 7】

前記第1のモジュールの前記管理部が前記アクティブから前記スタンバイにモードを変更し、この変更に合わせて、前記第2のモジュールの前記管理部が前記スタンバイから前記アクティブにモードを変更する

請求項5記載のストレージ装置。

【請求項 8】

前記第2のモジュールの前記管理部が前記第1のモジュールの前記管理部をオフし、前記スタンバイからアクティブにモードを変更する

請求項5記載のストレージ装置。

【請求項 9】

前記アクティブのモードにある前記管理部は第1の識別情報を備え、前記スタンバイのモードにある前記管理部は第2の識別情報を備え、前記管理装置は前記第1の識別情報を認識して、前記アクティブのモードにある前記管理部にアクセスする

請求項1記載のストレージ装置。

【請求項 10】

前記第1のアダプタと、前記第2のアダプタと、前記キャッシュメモリと、そして、前記マイクロプロセッサと、の少なくとも一つは、省電力モードをとり得る

請求項3記載のストレージ装置。

【請求項 11】

10

20

30

40

50

前記スイッチは省電力モードをとり得ない
請求項 10 記載のストレージ装置。

【請求項 12】

前記キャッシュメモリを当該キャッシュメモリの使用率または前記マイクロプロセッサの稼働率に基づいて、省電力モードにするか否か判断し、

前記キャッシュメモリを省電力モードにする際、

当該キャッシュメモリに格納されたデータを他のキャッシュメモリにコピーし、

前記管理情報を更新する

請求項 10 に記載のストレージ装置。

【請求項 13】

前記マイクロプロセッサは、当該マイクロプロセッサを備えるマイクロプロセッサパッケージ内の前記ローカルメモリに格納された前記管理情報を参照して、前記リード及びライト要求に対するプログラムを実行する

請求項 3 記載のストレージ装置。

【請求項 14】

計算機からの記憶資源に対するリード及びライト要求を処理する制御資源をモジュール化し、少なくとも 1 つのモジュールを備え、他のモジュールを順次前記少なくとも一つのモジュールに追加できるように構成されたストレージ装置の制御方法であって、

複数のモジュールのそれぞれに存在する管理部のうち、前記複数のモジュールに対して統一的な管理を実行する管理部を設定し、

この管理部に管理ユーザがアクセスし、

管理ユーザがこの管理部を介して前記複数のモジュールを統一的に管理するようにしたストレージ装置の制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はストレージ装置及びその制御方法に係わり、特に、スケーラビリティに優れたグリッドストレージシステムに係る制御技術に関するものである。

【背景技術】

【0002】

サーバなどのホスト計算機に接続され、ホスト計算機からのデータを格納し、格納したデータをホスト計算機に提供するストレージ装置が存在する。ストレージ装置は、主として、複数のハードディスクドライブ（HDD）をアレイ状に備えたディスクアレイ装置として構成される。ホスト計算機は、ストレージ装置に対して上位装置となる。この種のストレージ装置を構成する制御装置として、下記公報記載のように、複数のディスクアレイ制御ユニットからなるディスクアレイ制御装置が存在する。

【特許文献 1】特開 2001 - 256003 号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

この種のストレージ装置において、ホスト計算機等の上位装置と記憶デバイスとの間でデータの転送に伴う制御を行なう制御装置をモジュール化し、ユーザが複数のモジュールを順次増設することによって、データ処理能力を順次拡張できるように設計されたストレージ装置が存在する。

【0004】

このストレージ装置は、グリッドストレージとして知られており、ユーザが扱うデータ量に合わせてストレージ装置の規模を拡張できるため、すなわち、スケーラビリティに優れているため、ユーザにとって、必要以上の初期投資を避けることができるという利点がある。

10

20

30

40

50

【 0 0 0 5 】

モジュール化された制御装置 (Disk Array Controller : D K C) はそれ自身 1 つの制御資源として上位装置に対して動作する。一方、モジュール化された D K C が複数接続されたものも同様に、複数の制御資源としてではなく 1 つの制御資源として上位装置に対して動作する。

【 0 0 0 6 】

モジュール化された制御装置 (Disk Array Controller : D K C) はそれ自身 1 つの制御資源として上位装置に対して動作する。一方、モジュール化された D K C が複数接続されたものも同様に、複数の制御資源としてではなく 1 つの制御資源として上位装置に対して動作する。

10

【 0 0 0 7 】

D K C は、例えば、S V P (Service Processor) と称される管理部を備えている。管理部はデータを直接扱わないが、ストレージ装置の障害情報をユーザに参照させたり、ストレージ装置の構成情報をユーザに設定、変更、及び参照させたり、ストレージ装置の保守に関する情報をユーザに参照させるなどの各種管理のためのユーザインターフェースを備えている。

【 0 0 0 8 】

ところで、複数の D K C モジュールを接続したグリッドストレージにおいて、それぞれの D K C モジュールの S V P が互いに関連なく動作すると、グリッドストレージを統一された制御資源として管理できないという問題がある。

20

【 0 0 0 9 】

そこで、本発明は、ストレージシステムに制御モジュールを順次追加できるように設計されても、上位装置に対して、統一された制御資源として管理できるようにしたストレージ装置及びその制御方法を提供することを目的とするものである。

【 課題を解決するための手段 】

【 0 0 1 0 】

本発明はこの目的を達成するために、複数の制御モジュールのそれぞれに存在する管理部のうち、複数のモジュールに対して統一的な管理を実行する管理部を設定し、この管理部に管理ユーザがアクセスし、管理ユーザがこの管理部を介して複数のモジュールを統一的に管理できるようにしたことを特徴とするものである。

30

【 発明の効果 】

【 0 0 1 1 】

本発明は、ストレージシステムに制御モジュールを順次追加できるように設計されても、上位装置に対して、統一された制御資源として管理できるようにしたストレージ装置及びその制御方法を提供することができる。

【 発明を実施するための最良の形態 】

【 0 0 1 2 】

次に、本発明の実施形態について説明する。図 1 は本発明に係わるストレージ装置全体の概要を示す斜視図である。このストレージ装置は、二つの D K C モジュール 1 0 A と 1 0 B とが互いに接続されて、一つのシステムを構成している。

40

【 0 0 1 3 】

二つの D K C モジュールは 1 9 インチ標準ラックのような汎用的フレーム 1 8 に着脱自在に搭載される。ストレージ装置は、二つの D K C モジュール 1 0 A , 1 0 B のそれぞれに対して、二つの D K U モジュール (1 2 A , 1 2 B) (1 2 C , 1 2 D) を備えている。

【 0 0 1 4 】

D K U モジュールは、D K C モジュールから転送されるユーザデータを格納する部位であり、ハードディスクドライブなどの不揮発性記憶媒体により構成される H D D - B O X 、Serial Attached SCSI (S A S) などのデータ転送プロトコルに対応した SAS Switch (S S W) 1 4 、および電源 (P S) 1 6 により構成される。記憶資源は、H D D の他、フ

50

ラッシュメモリなどの半導体メモリでもよい。SSWは主に、記憶媒体へのデータ転送・記憶媒体からDKCモジュールへのデータ読み出し時のパイパス制御、障害報告などを行う。

【0015】

上位ホストとの間のパスを増やしたい、記憶容量を増やしたい、などのユーザの要求に従って、DKCモジュールおよびDKUモジュールを、モジュール毎に、標準ラックに対して増設でき、反対に減設することが可能である。

【0016】

DKCモジュールは、単独で一つのストレージ装置として上位装置に対して機能することができる。すなわち、DKCモジュール10Aは、システムに、DKCモジュール10Bが組み込まれていない場合には、上位ホストからDKU12A, 12Bに対するデータのライト及びリード要求を単独で制御することができ、一方、DKCモジュール10AとDKCモジュール10Bとが相互に結合された場合には、二つのDKCモジュールが一体となって一つの高性能・高機能なストレージ装置として動作する。すなわち、例えば、DKCモジュール10Aのポートに上位ホストから送られたコマンドを、DKCモジュール10Bが処理できるし、その逆も可能である。

【0017】

図2は、図1に示すストレージ装置を備えるストレージシステムのハードウェアのブロック構成図を示すものである。ストレージシステムは、DKCモジュール10AとDKCモジュール10Bとが、PCIのような、パソコンの内部バスアーキテクスチャ、好ましくは、PCIエクスプレスのような高速シリアルデータ転送プロトコルを実現するパス26(Xパス)によって接続されている。上位ホスト28は、SANなどのネットワーク30を介して、DKCモジュール10A及び12Bに接続している。

【0018】

DKCモジュール10A(12B)は、大きくは、処理系装置(CTL)20A(20B)、電源装置22A(22B)、および管理系装置24A(24B)に分けられる。

【0019】

CTL(20A, 20B)はデータを実際に処理する部位であり、図2に示された後述の各要素を備えている。図3は、これら各要素の構成をより詳しく説明するブロック図である。

【0020】

以下、図2と図3に基づいて、DKCモジュールを構成するハードウェア要素を説明する。上位ホスト28は、上位ホストのデータプロトコルに対応したプロトコルチップ(PRC)を有するチャンネルアダプタパッケージ(CHA-PK)に接続している。LRはローカルルータであり、Mは揮発性のMemory DIMMである。

【0021】

記憶媒体のデータプロトコルに対応したプロトコルチップを有するディスクアダプタパッケージ(DKA-PK)は、図1のDKUモジュールに接続している。符号32A, 32Bは、それぞれDKUの記憶領域を表す内部/外部論理ボリュームである。内部論理ボリュームとは、DKCに接続されたDKUによって形成される論理ボリュームであり、外部論理ボリュームとは、外部のストレージシステムの記憶デバイスによって形成される論理ボリュームである。

【0022】

さらに、CTL(20A, 20B)は、上位ホスト28と記憶媒体との間でのデータのバッファリングを行うキャッシュメモリパッケージ(CM-PK)と、命令・演算処理を行うマイクロプロセッサパッケージ(MP-PK)と、CHA-PK、DKA-PK、CM-PK、そしてMP-PKの相互間でのデータやコマンドのやり取りを橋渡し(Switching)する、スイッチパッケージ(SW-PK)と、を備えている。

【0023】

CM-PKのCMAはCache Memory Adapterであり、MCHはMemory Controller Hub

10

20

30

40

50

であり、そしてLMはLocal Memoryであり、SW-PKのMPAはMicro Processor Pass Adapterである。CTLの各要素間でのデータの転送は、例えば、PCI-EXPRESSのような内部高速データ転送プロトコルにより実現される。

【0024】

各DKCモジュールのCTLは、耐障害性の観点から複数のクラスタによって構成されている。各クラスタは、CHA、DKA、CM、MPがSWによって相互に接続された構成を備えている。CHA、DKA、CM、MP及びSWそれぞれのハードウェア資源は、パッケージ化されており、ユーザの使用環境や要求に従って各パッケージを増設したり、あるいは減設したりできる。“PK”はパッケージを意味する。

【0025】

これらパッケージを今ある第1のDKCモジュールに追加できない場合には、第2のDKCモジュールを第1のDKCモジュールに追加して、その追加されたDKCモジュールにユーザが望むパッケージを増設すればよい。

【0026】

既述の二つのDKCモジュール間のパス26は、各DKCモジュールのSW-PK間に形成されている。パスの耐障害性の観点から、各DKCモジュール間のパスは冗長化され、複数のルートが存在する。そのために、各DKCモジュールには、SWのパッケージが少なくとも1pair(SWパッケージが2枚)実装され、DKCモジュール10Aの1つのSWパッケージと、DKCモジュール10Bの1つのSWパッケージとの間にパスが形成される。DKCモジュールに1pairのSWパッケージがあるとすると、それぞれのSWパッケージについてパスが形成される。

【0027】

二つのDKCモジュールがXパスで結合されることにより、二つのDKCモジュールはそれぞれ他のDKCモジュールのハードウェア資源を共有することができ、二つのDKCモジュールが結合されたストレージシステムは、1つのDKCモジュールよりも高性能な一つの制御資源を実現することができる。

【0028】

例えば、二つのDKCモジュールの一方のDKCモジュールの処理系装置のMPは、SW及びパス26を介して他のDKCモジュールのCM、CHA、DKA、そしてMPにアクセスすることができる。

【0029】

従って、一方のモジュールのMPは、他モジュール配下の対ホストパス(上位ホスト28)、および対ディスクパス(内部/外部論理VOL)32A, 32Bにアクセスすることができる。すなわち、命令・演算処理を行う各々のMPは、自モジュール・他モジュールを問わず、独立して命令・演算処理を行うことができる。

【0030】

例えば、管理サーバはSVPを介して、DKCモジュールの構成情報として、あるCHAの上位ホスト向けポートに接続する論理ボリューム(Logical Unit:LU)を予め管理テーブルに設定できる。

【0031】

この時、例えば、ストレージ装置は、DKCモジュール10AのCHAポートに、他のDKCモジュール10Bの配下の論理ボリュームに対するパスを管理テーブルに定義することも可能である。

【0032】

これは、既述のXパスにより二つのDKCモジュール間でハードウェア資源を共有できるがため、可能になる。上位ホスト28は、DKCモジュール10AとDKCモジュール10Bとを区別することなく、1つのストレージ装置として認識する。

【0033】

管理系装置24A, 24Bはデータを直接は扱わないが、処理系装置20A, 20Bの障害情報をユーザに参照させたり、処理系装置の構成情報をユーザに設定、変更、及び参

10

20

30

40

50

照させたり、処理系装置の保守に関する情報をユーザに参照させるなどの各種管理情報を管理クライアントである管理サーバ42, 44に提供するインターフェースを備えている。

【0034】

主要なハードウェアとして、管理系装置は、管理系情報を管理し、ユーザとのインターフェースとなるSVP(Service Processor)と、第1のDKCモジュール10AのSVP0と第2のDKCモジュール10BのSVP1との間のLAN通信のためのHUBを備えている。SVPの障害が、二つのDKCモジュールの管理系装置同士の結合に影響しないように、一方のDKCモジュールのSVPはHUB経由で他モジュールのSVPと結合している。

10

【0035】

第1のDKCモジュールのHUB0と第2のDKCモジュールのHUB1の間にはモジュール間の直接の通信のためのLAN38が形成されている。また、HUB0とHUB1はLANネットワーク40に接続されており、このLANネットワーク40には第1の管理サーバ42と第2の管理サーバ44が接続されている。

【0036】

複数のDKCモジュール同士が結合する場合、互いに連結する複数の管理系装置はアクティブ/スタンバイの主従関係を持つ。複数のSVPの一つがアクティブな状態となり、残りのSVPがスタンバイの状態となる。

【0037】

アクティブモードのSVPは、自モジュールの管理情報にアクセスして、これを参照したり、変更、新たに管理情報を設定することができるばかりでなく、LAN38を経由して他モジュールのスタンバイモードのSVPにアクセスして、他モジュールの管理情報にもアクセスすることができる。

20

【0038】

例えば、SVP0がアクティブモードであるとする、SVP0の自モジュールであるDKCモジュール10Aの処理系装置20Aから収集した管理情報と、スタンバイ状態のSVP1が所属する他モジュールであるDKCモジュール20Bの処理系装置12Bから収集した情報とを統合管理し、管理ユーザへ提供する。

【0039】

一方、スタンバイ状態のSVP1は、自モジュール及び他モジュールの管理情報には通常アクセスせず、アクティブ状態のSVP0からLAN38を経由して定期的にアクティブSVP0が持つ全ての管理情報の取得を行う。したがって、スタンバイモードのSVPはアクティブモードのSVPに同期化されていることになる。

30

【0040】

アクティブモードのSVPが管理系ユーザとのI/Fを持ち、ユーザはアクティブのSVPにアクセスすることにより、二つのDKCモジュールが結合したストレージ装置を一つのDKCとして管理可能とする。

【0041】

管理系装置のアクティブ/スタンバイの関係は、動的に変更され、アクティブ状態のSVPがスタンバイ状態になり、スタンバイ状態のSVPがアクティブ状態に切り替わることもある。

40

【0042】

この切り替わりの契機は、SVPへの設定の変更や、アクティブモードのSVPに障害が生じた場合などである。障害がアクティブモードのSVPに発生すると、このSVPがスタンバイモードに切り替わり、もともとスタンバイモードのSVPがアクティブモードに切り替わる。これによって、図2のストレージシステムは、管理系ユーザに対して、複数のDKCモジュールが統一された管理系装置を継続して有するように動作可能である。

【0043】

アクティブ状態のSVPを介して、ユーザが複数のDKCモジュールを統一して管理で

50

きるようにするために、管理サーバ42, 44に対して、アクティブ/スタンバイの状態を持つ複数の管理系を一つに見せるに方式について説明する。

【0044】

ストレージシステムの管理者は、アクティブであるSVPのIPアドレスを特定のものになるように予め決めておき、管理サーバはそのアクティブのSVPのIPアドレスに対してアクセスするようにする。

【0045】

複数のSVPの間で、アクティブ/スタンバイのモードの切り替えが行われた際には、複数のSVP間でIPアドレスの切り替えを行い、アクティブのSVPは予め決めたIPアドレスを持つ。

【0046】

例えば、予めアクティブモードとなるSVPのIPアドレスを“***.***.***.15”と定めると、管理サーバが常に***.***.***.15にアクセスするようにする。

【0047】

DKCモジュール10AのSVP0がアクティブであり、DKCモジュール10BのSVP1がスタンバイの状態にあるとき、SVP0はIPアドレス“***.***.***.15”を有し、SVP1は、IPアドレスとして、例えば、“***.***.***.14”を有する。

【0048】

後述するように、SVP0に障害が発生するなど何らかの理由で、複数のSVPに、アクティブ/スタンバイの切り替えが必要である際には、新たにアクティブモードとなるSVP1がSVP0のIPアドレスを“***.***.***.14”に変更し、自身のIPアドレスを“***.***.***.15”に変更する。

【0049】

これにより、複数存在する管理サーバからは、アクティブモードのSVP、即ち、IPアドレスとして特定の値“***.***.***.15”を持つSVPが認識され、スタンバイモードのSVPは管理サーバによって認識されない。

【0050】

なお、SVPは次の各機能を管理ユーザ(管理サーバ)に提供する。ストレージシステムの構成情報・ユーザによって設定された情報の参照・変更、ストレージシステムの障害情報の参照、ストレージシステムの保守操作のユーザI/F(障害部位交換作業、回復作業時の装置オペレーションユーザI/F)、モニタ情報(ポートIOPS(I/O per Second)、CM使用率、MP使用率、LEDV IOPS等ハードウェアの稼動状況)の参照、各種ユーザ管理アプリケーションのハードウェアプラットフォーム(Webエンジン機能)の設定、参照、及び電源状態情報の参照。

【0051】

構成情報とは、DKCモジュールに実装されているハードウェア資源の有無・種類などの情報である。例えば、ユーザがオプションで追加するCM-PK(キャッシュメモリパッケージ)の枚数、CM-PKに搭載するキャッシュメモリの種類・容量、CHAの種類・枚数、ユーザの固有設定値などが構成情報として処理系装置のマイクロプログラムによって取り扱われる。

【0052】

構成情報は、DKCモジュールの停止時、MP-PKのフラッシュメモリなどの不揮発メモリ内、および、それらのバックアップとしてSVPのハードディスクなどの不揮発媒体に格納されている。

【0053】

構成情報は、ストレージ装置の起動を契機として、MP-PKの不揮発メモリからMP-PKの揮発性ローカルメモリ(LM)にロードされる。LMの構成情報は、随時最新の情報に更新され、装置計画停止時は、計画停止処理の中で各MP-PKの不揮発性のフラッシュメモリに格納される。

【0054】

10

20

30

40

50

電源状態情報とは、パッケージ単位での電源状態の情報である。パッケージは省電力モードをとれるようになっている。MP - PKは、自分が属するDKCモジュールや他のDKCモジュールの各種パッケージにアクセスする際に、それらパッケージの電源状態を把握しておく必要がある。

【0055】

電源状態情報は、DKCモジュールの立上げが完了した後、マイクロプログラム中の省電力モード機能のJOBが動作し始めた段階で、MP - PKのLMのある特定のメモリ領域が確保され、その領域に電源状態情報が管理される。

【0056】

電源状態情報は、装置が起動している間、省電力モード機能のJOBが管理し、電源状態情報は常に更新される。装置の計画停止時は、状態情報は保持される必要が無いため、計画停止処理後の電源オフのタイミングで揮発される。

10

【0057】

モニタ情報は、MP - PKのLM内に格納される。モニタ情報は、電源状態情報と同じように、装置の立上げ完了後、制御マイクロプログラム中のモニタ機能のJOBが動作し始めた段階で、MP - PKのLMのある特定のメモリ領域が確保され、その領域でモニタ情報が管理される。

【0058】

モニタ情報は、装置が起動している間、モニタ機能のJOBが管理し、モニタ情報は状態に合わせて常に更新される。装置計画停止時は、状態情報は保持しておく必要が無いため、計画停止処理後の電源オフのタイミングで揮発される。

20

【0059】

図4は、DKCモジュールの代表MP - PKによって実行されるマイクロプログラムが管理する構成情報・電源状態情報・モニタ情報の管理テーブルの一例であり、図4Aはストレージシステムが1つのモジュールから構成されているもののテーブルであり、図4Bはストレージシステムが2つのモジュールから構成されているもののテーブルである。

【0060】

ストレージシステムが複数のモジュールからなるグリッド構成である場合、複数のモジュールに存在する複数のMP - PKの中で全てのMP - PKのMPは独立してマイクロプログラムを実行する。

30

【0061】

これらの管理テーブルは、例えば、全MP - MPのLMに格納されている。全MPのそれぞれは管理テーブルなど共通の情報を把握し、各々独立して動作することができる。各MPはあるハードウェア資源にアクセスする際、アクセス先のハードウェアの有無、省電力モードの状態を予め把握する必要があるため、それぞれが構成情報、電源状態情報等を例えばLMに有している。モジュール増設時は、図4Aの“実装情報”の“モジュール増設”のbitが立ち、図4Bのように、モジュール1のテーブル領域に各種構成情報等が反映・更新されていく。

【0062】

図2において、DKCモジュール10A(12B)は電源装置22A(22B)を有している。電源装置は、PS0 - PS3までの4つの電源と、電源制御を行なうPower Controller(PCTL)、とを備えている。電源から、処理系装置のCHA - PKなどの各要素、及び管理系装置に電力が供給される。

40

【0063】

電源は最大4台、最小2台がDKCモジュールに設置される。電源1台あたりに2台のDC-DC Converterを内蔵し、電源2台でも合計4系統の電源境界をつくることのできる。従って、電源が最小の2台であっても、Converterが1台故障しても4系統の電源境界を維持することができる。符号36A, 36Bはそれぞれ、PCTL-電源-処理系装置間の電源制御に使用される制御信号である。

【0064】

50

次に、ストレージシステムが、複数のDKCモジュールの各管理系装置のモードを切り替える動作について説明する。図5は、その動作の詳細を示すフローチャートである。ストレージシステムの管理者は、DKCモジュール10AのSVP0に“アクティブモード”を設定する(500)。SVPは所定の制御フラグを備え、このフラグに“アクティブモード”又は“スタンバイモード”が設定される。SVPは、この制御フラグをチェックして、自身のモードを判定してそのモードにしたがって行動する。

【0065】

次いで、管理者は、DKCモジュール10Aに他のモジュールが接続されているか否かを判定し(502)、図2のようにDKCモジュール10Bが接続されている場合には、この判定を肯定し、SVP1に“スタンバイ”を設定する(504)。一方、この判定を

10

【0066】

SVP0とSVP1との間ではLAN38を介してハートビートパケットを転送しあっており、SVP1がSVP0からのハートビートを検出できないときには、SVP1は、SVP0になんらかの原因によって異常が生じたと判断する(506)。

【0067】

その後、SVP1は、LAN38を介してSVP0の電源を強制的にオフし(508)、自身のステータスを“スタンバイモード”から“アクティブモード”に切り替える(510)。

【0068】

20

これが正常に終了すると、ストレージシステムに対する管理権が、SVP0からSVP1に切り替える処理を行なう(510)。この処理が正常に終了できない場合には、SVP1は、ストレージシステムの管理者に異常終了を報告する(514)。SVP0の電源がオフされることにより、異常モードから正常モードに復帰して、SVP0とSVP1との間でストレージシステムに対する管理権にコンフリクトが発生しないようにされている。

【0069】

なお、DKCモジュール10AのHUBは、異常となったSVP0とは独立して動作しているため、DKCモジュール10BのSVP1は、DKCモジュール10BばかりでなくDKCモジュール10AのPCTLなど各ハードウェア資源にアクセスすることができる。

30

【0070】

一方、SVP0に異常がなく、ストレージシステムにDKCモジュールの増設がある場合には(516, 518)、DKCモジュールの増設に合わせて、管理者は、増設されたモジュールのSVP1に“スタンバイモード”を設定する(520, 524)。

【0071】

一方、ストレージシステムからDKCモジュールが減設された場合には(518)、減設されたモジュールのSVPが“スタンバイモード”にあるか否かを判定し(522)、それが肯定された場合には、減設処理を行い、減設されたモジュールのSVPが“アクティブモード”にあると判定された場合には、被減設対象のモジュールのSVPは、“スタンバイモード”であるので、これを“アクティブモード”に切り替えた後(526)、減設処理を行う(528)。

40

【0072】

“スタンバイモード”であるSVPのステータスを“アクティブモード”に切り替え、“アクティブモード”のSVPを“スタンバイモード”に切り替える際、ユーザは“アクティブモード”のSVPから“スタンバイモード”のSVPに対して切り替え操作を行うことができる。

【0073】

“アクティブモード”のSVPが“スタンバイモード”のSVPに対して管理系情報の転送をLAN38経由で開始し、転送を正常終了した時点でSVP同士の間で管理情報の

50

同期がとられ、ここで、モジュール間でSVPのステータス・モードの切り替えが行われる。

【0074】

なお、図2のように2つのDKCモジュールが存在するケースで、ユーザが自発的に2つのSVPの間でステータスの切り替えを行ってもよい(530, 532)。例えば、ハードウェアメンテナンス時にこの切り替えが行なわれる。

【0075】

ユーザによる、SVPのステータスの切り替え操作の際、“アクティブモード”のSVPが管理サーバに提供するGUI画面の一例を図6で示す。切り替え操作は“アクティブ”のSVPからのみ可能である。

10

【0076】

図6のGUI画面上では、DKCモジュール別にそれぞれFRONT側、REAR側から見た装置構成・装置状態が一ユーザに目で分かるようになっている。網掛け部はそのコンポーネントが未実装であることを示す。また、異常があるコンポーネントは、コンポーネント単位での点滅表示等で異常部位を容易に特定できる。

【0077】

SVPのステータスの情報は、図のようにGUI画面上で表示され、ユーザはどちらのモジュールのSVPが“アクティブモード”又は“スタンバイモード”が容易に判断できるようになっている。また、切り替えの操作もGUI画面上で操作できるようになっている。

20

【0078】

既述のように、複数のモジュール間で結合が形成され、上位ホストが複数のモジュールのハードウェア資源を共有する場合、上位ホストからモジュールへのI/Oの頻度が少ない軽負荷状態では、モジュールのハードウェア資源を省電力化することが望ましい。

【0079】

省電力化は、ハードウェア資源のパッケージ単位で省電力モードのオン/オフによって行なわれる。省電力モードとは、例えばMP-PKであれば、CLK周波数の低減、MP単位でのスタンバイモードへの移行、また、CM-PKであれば、メモリDIMM単位でのスタンバイモードへの移行をいう。

【0080】

ただし、省電力モードの形態として、PKに給電される電源をPKの入力端でON/OFFさせる方式、または、PKへの給電をAC-DCの出力端でON/OFFさせる方式でもよい。図7は、処理系装置の各種PKの省電力モードの可否とその実施形態を示す。

30

【0081】

PK単位に対する省電力モードのON/OFF切り替え制御は、DKCモジュール10AとDKCモジュール10Bとからなるストレージ装置にある、風数のMP-PKのうち、代表のMP-PKが実行するマイクロプログラムによって達成される。

【0082】

省電力モードのON/OFF切り替え制御のためには、装置構成情報、モニタ情報、装置の障害情報などを必要とし、制御マイクロプログラムはこれらの情報を参照し、切り替え制御を行う。

40

【0083】

PKの中で、SW-PKを省電力モードに切り替えることはできない。例えば、CM-PKでは、負荷状況によってその使用率が変動するため、使用率が低い状態でCM-PKを省電力モードへ移行することができる。しかしながら、SW-PKは、先ず自DKCモジュール内でCHA、DKA、CM-PK、MP-PK間でのデータの橋渡しを行っており、DKCモジュール間でのデータの橋渡しを行うものであるから、CM-PKのように選択的に省電力モードへの切り替えを行うことはできない。

【0084】

次に、2つのDKCモジュールが一つのストレージシステムを構成するケースで、CM

50

- P Kへの負荷が変動することに応じてC M - P Kの一部が省電力モードへ移行 (C M OFF処理)、また、省電力モードから通常モードへ復帰 (C M ON処理)する例の動作を図8のフローチャートで説明する。このケースにおけるC M ON/OFF処理の動作フローについて説明する。

【 0 0 8 5 】

ここでは、図2のように2つのD K Cモジュールが接続されたケースで、D K Cモジュール1 0 Aの管理系のS V P 0が“アクティブモード”にあり、D K Cモジュール1 0 Bの管理系S V P 1が“スタンバイモード”にあるとする。

【 0 0 8 6 】

代表のM P - Pは制御マイクロプログラムを実行し、自身のL M中の構成情報を参照する(8 0 0)。また、M P - P KはL M中のモニタ情報(ポートI O P S、C M使用率、M P使用率、L E D V I O P S等)を取得する(8 0 2)。

【 0 0 8 7 】

制御マイクロプログラムを実行するM Pは、使用率 2 5 %のC M - P Kが2枚以上、かつ動作率 2 5 %のM P - P Kが1つ以上あるかを判定する(8 0 4)。

【 0 0 8 8 】

ここで、D K Cモジュール1 0 AのC M - P K、D K Cモジュール1 0 BのC M - P Kが使用率 2 5 %、D K Cモジュール1 0 Aの1つのM P - P Kが動作率 2 5 %とする。

【 0 0 8 9 】

そこで、制御マイクロプログラムは、条件に該当するC M P KのO F F処理を行う(8 0 6)。この処理の詳細は図9に示されている。制御マイクロプログラムは、O F FされるC M - P KからO F FされないC M - P Kにデータのコピーを開始する(8 0 6 A)。制御マイクロプログラムは、L M内の管理情報としてコピー処理の管理情報をL Mに更新登録する(8 0 6 B)。

【 0 0 9 0 】

コピー処理の管理情報とは、主として、コピーするデータのセグメント毎の管理I D、コピー完了フラグ、コピー元情報(セグメントのD K CモジュールI D、キャッシュアドレス)、及びコピー先情報(セグメントのD K CモジュールI D、キャッシュアドレス)により構成される。

【 0 0 9 1 】

コピー処理が完了し、コピー完了フラグがクリアされたセグメントの管理情報は、随時新しい管理情報が登録されるに従ってクリアされ更新される。上記コピー処理の管理情報のテーブルを図13に示す。

【 0 0 9 2 】

開始にあたって、制御マイクロプログラムは、データのコピー元、コピー先のアドレスの設定、コピー処理を担当する担当M P - P Kを決定する。ここでは、D K Cモジュール1 0 BのC M - P Kがコピー元、D K Cモジュール1 0 AのC M - P Kがコピー先、担当M P - P KがD K Cモジュール1 0 AのM P - P Kとなる。

【 0 0 9 3 】

制御マイクロプログラムはコピーを開始すると、既述の8 0 4の条件に従って、O F F処理開始条件の判定(8 0 6 C)、一定時間でのC MからC Mへのデータのコピー(8 0 6 D)、管理系S V P (アクティブモードにあるS V P)が持つコピーの管理情報の更新(8 0 6 E)、の処理の完了(先に定義した、コピー完了フラグが全て立つ状態)まで繰り返される(8 0 6 F)。

【 0 0 9 4 】

ここで一定時間としたのは、ユーザの使用状態によって、C M使用率、M P使用率等が動的に変動するため、ある一定時間間隔として、“8 0 6 C”のO F F処理開始条件での判定が考慮される必要があるからである。

【 0 0 9 5 】

10

20

30

40

50

制御マイクロプログラムが、コピーが完了したと判定すると、コピー元CM - PK (DKCモジュール10BのCM - PK)をOFFする省電力モードへ移行させる(806I、806J)。

【0096】

さらに、制御マイクロプログラムは、LM中の管理情報にコピー元CM - PKの省電力モードを更新登録する(806K)。

【0097】

制御マイクロプログラムは、CM - PKに係わるコピーを繰返し処理する過程で、OFF処理を開始する条件が成立しないと判定すると、その時点でOFF処理を中止し(806G)、管理情報を更新する(806H)。すなわち、図13に示すコピー処理の管理情報を全てクリアして、コピーのタスクを全てクリアする。

10

【0098】

制御マイクロプログラムが、ステップ804を否定判定し、OFF済みのCM - PKがあり(808)、かつ、使用率75%となるCM - PKが1枚以上存在することを判定すると(810)、オフ済みのCM - PKをオンさせる処理を行う(812)。

【0099】

コピー先CM - PKがオンされると、全てのMP - PKのLMの管理情報が更新され、オフされていたDKCモジュール10BのCM - PKが利用可能となり、75%を超えていたCM使用率は緩和される。なお、図4Aで管理される構成情報、電源状態情報は全てのMP - PKが所有する。全てのMPが上記情報を持つことにより、各MPが独立して動作できる。

20

【0100】

次に、CM - PKの省電力モードについて説明する。CM - PKのメモリDIMMのスタンバイモードを省電力モードとする。図10にCM - PKの内部ブロック図を示す。符号100は、DC/DCコンバータであり、符号102はOscillator(振動子)である。DIMM(Dual Inline Memory Module)106がキャッシュメモリアダプタ(CMA)104に接続している。

【0101】

符号108はPROM(Programmable Read Only Memory)である。PROMは、PK固有情報(バージョン情報および部品情報など)を格納している。符号110はバックボードコネクタであり、符号112は電源を制御するマイコンである。バックボードコネクタとCMAとは一時記憶される処理系データのやり取りを行う。

30

【0102】

メモリDIMMのスタンバイモードでは、電源部(電源制御マイコン含む)およびメモリDIMM以外のLSI等が電源オフの状態となり、メモリDIMM各々がデータを消去させないためだけの必要最低限のリフレッシュ周期で、自分自身でリフレッシュ動作を繰り返している。

【0103】

メモリDIMMの省電力モードはセルフ - リフレッシュモードとも呼ばれ、停電時のキャッシュメモリのバックアップモード時などでも、このセルフ - リフレッシュモードが使用される。

40

【0104】

複数のDKCモジュールが一つのストレージシステムを構成するケースでのリード・ライト処理のフローチャートを以下に示す。複数のDKCモジュールのうちの代表のMP - PKが制御マイクロプログラムを実行し、LM内の構成情報・電源状態情報を参照・管理し、省電力モードにあるパッケージの有無に関わらず、リード/ライト処理を行う。

【0105】

図11は、リード処理を説明するフローチャートである。CHAが上位ホストからリードコマンドを受信する(1100)。CHAのLRは、LR配下のメモリ内に格納されている管理テーブルを参照して、このリードコマンドの処理を担当する担当MP - PKを決

50

定し、LRは担当MP PKに対して、ステップ1102に示すように、リードコマンドを送信する（これをルーティングという）。

【0106】

LRのメモリ内の管理テーブルには、論理ボリュームの一つの管理単位であるLogical Unit (LU) に対して担当するMPを関連付ける情報が格納されており、LRはリードコマンドのアドレスからLUを特定し、担当MPを決定することができる。

【0107】

LRの管理テーブルの情報は、装置起動時にMPのLMからロードされ、装置構成の変更（例えば、CHA/DKA/MP-PKの増減設、対上位ホストパスの変更など）、および電源状態情報の変更のタイミングで情報が随時更新される。従って、例えば、省電力モードに入っているMP-PKに対しては、そのMP-PKは構成上存在していないものと見なされ、コマンドのルーティングは行われない。

【0108】

ここで、LRが担当MPを決定するに当たり、担当MPが基本DKCモジュールにあるか、増設DKCモジュールにあるかは関係ない。全てのモジュール内のハードウェア資源は、2つのモジュール間のXパスにより、モジュール間の跨ぎを意識することなくお互いに共有される。

【0109】

また、省電力モードに移行できるPK種は、既述のとおり、SW-PK以外のPKとなる。SW-PKは、モジュール内およびモジュール間でのハードウェア資源間のデータの橋渡しを行っており、装置起動時は常時アクティブの状態となっている。

【0110】

LRからリードコマンドを受信したMPは、LMを参照する(1104)。MPのLMには、自身が管理するCMのデータのアドレス情報が格納されている。MPはコマンドに含まれるアドレス情報からリードデータがCM上に存在するかどうかの判定を行う(1106)。

【0111】

リードデータがCM上に存在する場合は、MPはコマンド送信元のCHAに対してCM上のデータのアドレスを返す(1108)。当該リードデータがCM上に存在しない場合は、MPはCM上にリードデータがステージング可能な領域が存在するかどうかの判定を行う(1114)。

【0112】

CM上にリードデータがステージング可能な領域が存在する場合は、MPはDKAに対して記憶媒体上に存在する当該リードデータのアドレス情報を渡す(1120)。

【0113】

CM上にリードデータがステージング可能な領域が存在しない場合は、MPは(Least Recently Used (LRU) 則) に従ってDKAに対してデステージする最も最近使われていないデータ(オールドデータ)のアドレスを渡す(1116)。

【0114】

MPからオールドデータのアドレスを受信したDKAのLRは、CM上のオールドデータを記憶媒体上にデステージする(1118)。

【0115】

MPから受信した記憶媒体上のアドレス情報を元に、DKAのLRは記憶媒体上のデータをCM上にステージングする(1122)。CHAのLRはCM上のリードデータを読み出す(1110)。CHAはリードデータを上位ホストに返す(1112)。

【0116】

次に、ライト処理について説明する。CHAが上位ホストからライトコマンドを受信する(1200)。CHAのLRはLR配下のメモリ内に格納されている管理テーブルを参照して、当該ライトコマンドの処理を担当する担当MPを決定し、LRは担当MPに対してライトコマンドを送信する(1202)。LRのメモリ内の管理テーブルには、LUに

10

20

30

40

50

対して担当するMPを関連付ける情報が格納されており、LRはライトコマンドのアドレスからLUを特定し、担当MPを決定することができる。ライトコマンドを受けたCHA PKがDKCモジュール10Aのものであっても、担当MP - PKはDKCモジュール10Bのものでよいし、ライトコマンドを受けたCHA PKがDKCモジュール10Bのものであっても、担当MP - PKはDKCモジュール10Aのものでよいしその逆でもよい。

【0117】

LRからライトコマンドを受信したMPは、LMを参照する(1204)。MPのLMには、自身が管理するCMのデータのアドレス情報が格納されている。MPはCmdに含まれるアドレス情報から当該ライトデータのアドレスがCM上に存在するかどうかの判定を行う(1206)。

10

【0118】

ライトデータのアドレスがLMの管理テーブル上に存在する場合は、MPはコマンド送信元のCHAに対してCM上のライトする領域の同一アドレスを(1208)返す。

【0119】

CHAのLRはCM上の同一アドレスにデータを更新記録する(1210)。

【0120】

当該ライトデータのアドレスがLMの管理テーブル上に存在しない場合(1206)は、MPはCM上にライト可能な領域が存在するかどうかの判定を行う(1214)。

【0121】

CM上にライト可能な領域が存在する場合は、MPはDKAに対してライト可能な領域のアドレス情報を渡す(1216)。CM上にライト可能な領域が存在しない場合は、MPはDKAに対してデステージする最も最近使われていないデータのアドレスを渡す(1220)。

20

【0122】

MPからオールドデータのアドレスを受信したDKAのLRは、CM上のオールドデータを記憶媒体上にデステージする(1222)。CHAのLRはCM上にデータをライトする(1218)。CHAは上位ホストにライト完了を報告する(1212)。

【0123】

以上の実施形態では、二つのDKCモジュールを結合したシステムであったが、3基以上のDKCモジュールを結合するシステムにも本発明は応用可能である。

30

【図面の簡単な説明】

【0124】

【図1】本発明に係わるストレージ装置全体の概要を示す斜視図である。

【図2】図1に示すストレージ装置を備えるストレージシステムのハードウェアのブロック構成図である。

【図3】図2のストレージ装置の各要素の構成をより詳しく説明するブロック図である。

【図4A】DKCモジュールの代表MP - PKによって実行されるマイクロプログラムが管理する構成情報・電源状態情報・モニタ情報の管理テーブルの一例であり、ストレージ装置が1つのモジュールから構成されているもののテーブルである。

40

【図4B】ストレージ装置が2つのモジュールから構成されているもののテーブルである。

【図5】ストレージ装置が、複数のDKCモジュールの各管理系装置のモードを切り替える動作の詳細を示すフローチャートである。

【図6】“アクティブモード”のSVPが管理サーバに提供するGUI画面の一例である。

【図7】処理系装置の各種PKの省電力モードの可否とその実施形態を示すテーブルである。

【図8】2つのDKCモジュールが一つのストレージ装置を構成するケースで、CM - PKへの負荷が変動することに応じてCM - PKの一部が省電力モードへ移行(CM OFF処理

50

)、また、省電力モードから通常モードへ復帰 (CM ON処理) する例の動作を示すフローチャートである。

【図9】制御マイクロプログラムは、条件に該当するキャッシュメモリパッケージのOFF処理の詳細は示すフローチャートである。

【図10】キャッシュメモリパッケージのブロック図である。

【図11】ストレージシステムのリード処理を説明するフローチャートである。

【図12】そのライト処理を説明するフローチャートである。

【図13】コピー処理に関する管理情報のテーブルである。

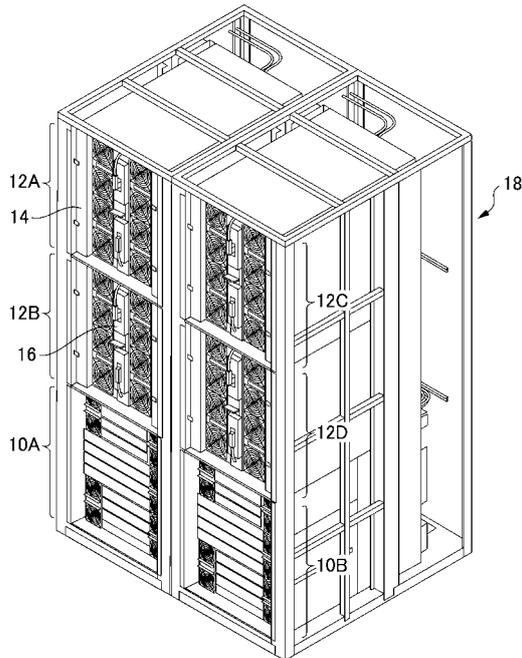
【符号の説明】

【0125】

10A, 10B: DKCモジュール(制御資源)、12A, 12B, 12C, 12D: DKUモジュール(記憶資源)、18: 標準ラック(汎用的フレーム)、26: 高速シリアルデータ転送プロトコルを実現するバス(Xパス)、22A, 22B: 電源装置、24A, 24B: 管理系装置、30: SANなどのネットワーク

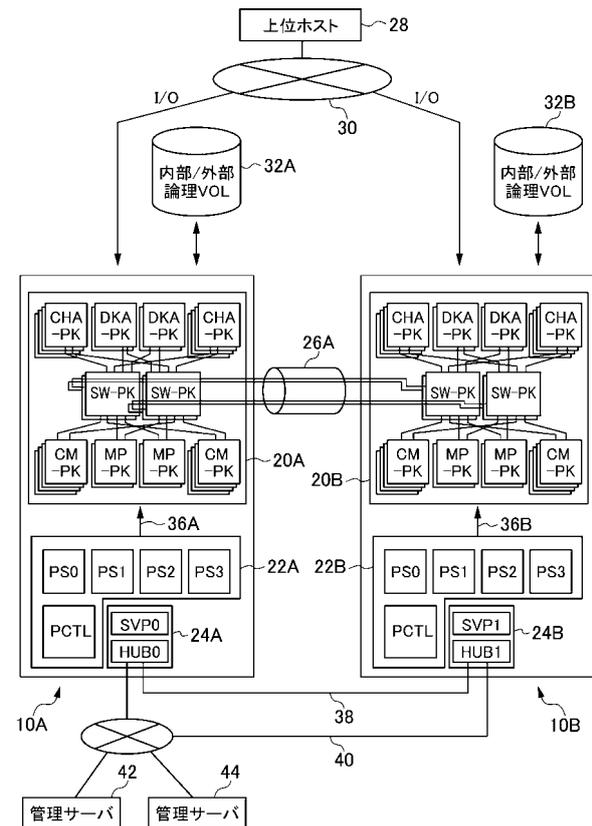
【図1】

図1



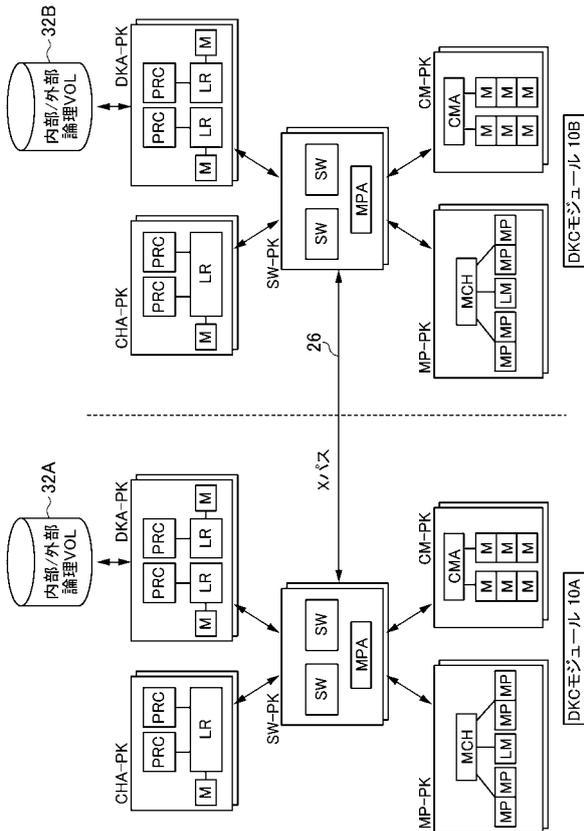
【図2】

図2



【図3】

図3



【図4A】

図4A

構成情報	実装情報	モジュール増設	1							
			モジュール1				モジュール0			
		実装	#n	...	#1	#0	#n	...	#1	#0
		CHA	0	...	1	1	1	...	1	1
		DKA	1	...	1	1	1	...	1	1
		...								
		MP-PK	0	...	0	0	0	...	1	1
		DIMM type	モジュール1				モジュール0			
			#n	...	#1	#0	#n	...	#1	#0
			3	...	3	3	0	...	0	0
		MODE設定 (パフォーマンス tuning etc)	#0n	...	#05	#04	#03	#02	#01	#00
			0	...	0	1	0	...	1	1
			...							
			#mn	...	#m5	#m4	#m3	#m2	#m1	#m0
			0	...	0	0	1	0	0	0
		...								
電源状態情報	省電力モード		モジュール1				モジュール0			
		Loc	#n	...	#1	#0	#n	...	#1	#0
		CHA	0	...	0	0	0	...	0	0
		DKA	0	...	0	0	0	...	0	0
		...								
		MP-PK	0	...	F	F	0	...	F	F
		...								
モニタ情報	ポート IOPS (kIOPS)		モジュール1				モジュール0			
		#nA	...	#1A	#0A	#nA	...	#1A	#0A	
		0	...	0	5	5	...	10	10	
		...								
		#nP	...	#1P	#0P	#nP	...	#1P	#0P	
		0	...	20	20	0	...	0	5	
		...								
		CM 使用率 (%)	モジュール1				モジュール0			
			#n	...	#1	#0	#n	...	#1	#0
			0	...	50	60	10	...	30	40
		...								

【図4B】

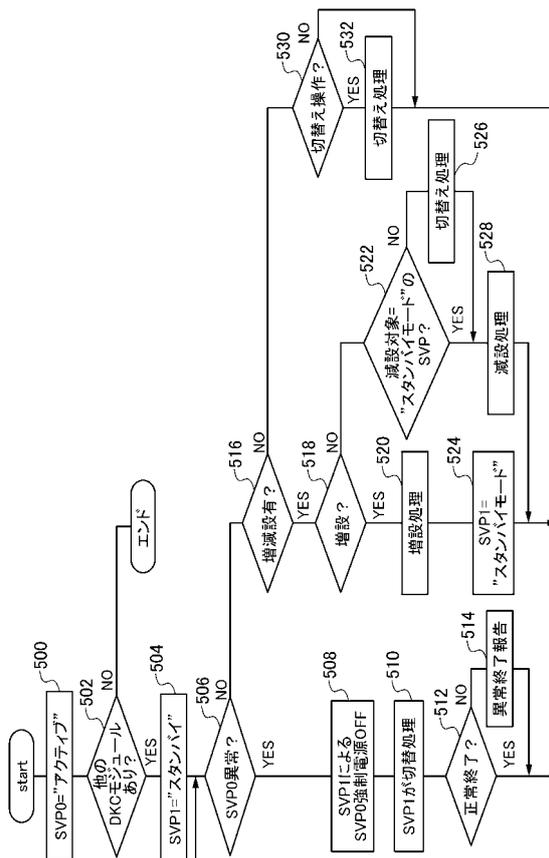
図4B

構成情報	実装情報	モジュール増設	0							
			モジュール1				モジュール0			
		実装	#n	...	#1	#0	#n	...	#1	#0
		CHA	u	...	u	u*	1	...	1	1
		DKA	u	...	u	u	1	...	1	1
		...								
		MP-PK	u	...	u	u	0	...	1	1
		DIMM type	モジュール1				モジュール0			
			#n	...	#1	#0	#n	...	#1	#0
			u	...	u	u	0	...	0	0
		MODE設定 (パフォーマンス tuning etc)	#0n	...	#05	#04	#03	#02	#01	#00
			0	...	0	1	0	...	1	1
			...							
			#mn	...	#m5	#m4	#m3	#m2	#m1	#m0
			0	...	0	0	1	0	0	0
		...								
電源状態情報	省電力モード		モジュール1				モジュール0			
		Loc	#n	...	#1	#0	#n	...	#1	#0
		CHA	u	...	u	u	0	...	0	0
		DKA	u	...	u	u	0	...	0	0
		...								
		MP-PK	u	...	u	u	0	...	F	F
		...								
モニタ情報	ポート IOPS (kIOPS)		モジュール1				モジュール0			
		#nA	...	#1A	#0A	#nA	...	#1A	#0A	
		u	...	u	u	5	...	10	10	
		...								
		#nP	...	#1P	#0P	#nP	...	#1P	#0P	
		u	...	u	u	0	...	0	5	
		...								
		CM 使用率 (%)	モジュール1				モジュール0			
			#n	...	#1	#0	#n	...	#1	#0
			u	...	u	u	10	...	30	40
		...								

*u: undefined

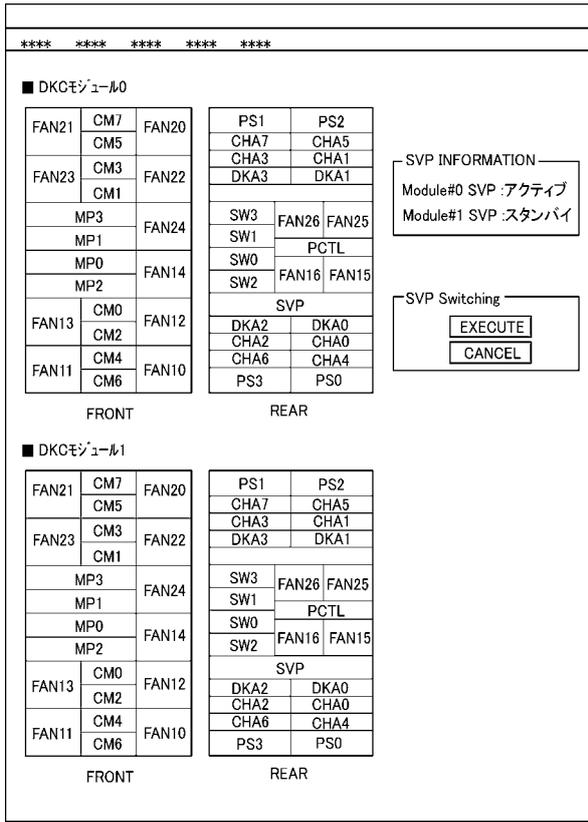
【図5】

図5



【図6】

図6



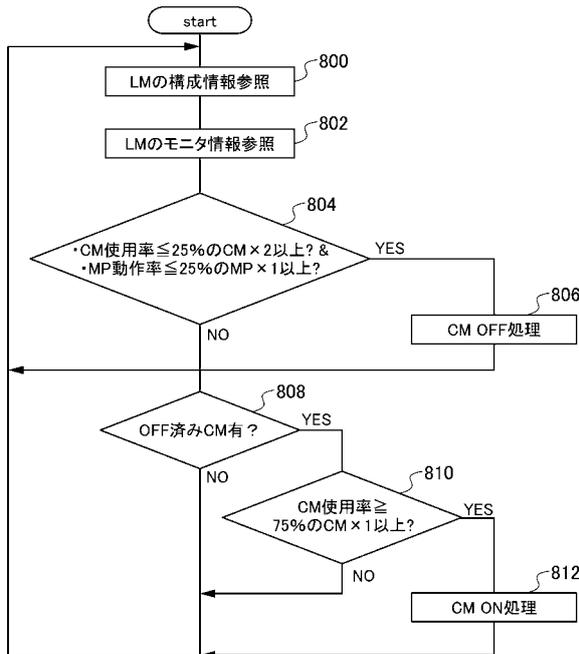
【図7】

図7

PK種	省電力モード	実施形態
CHA	可	PRC/LRのStandbyモード移行、CLK周波数の低減など
DKA	可	PRC/LRのStandbyモード移行、CLK周波数の低減など
CM-PK	可	メモリDIMMのStandbyモード移行など
SW-PK	不可	-
MP-PK	可	MPのStandbyモード移行、CLK周波数の低減など

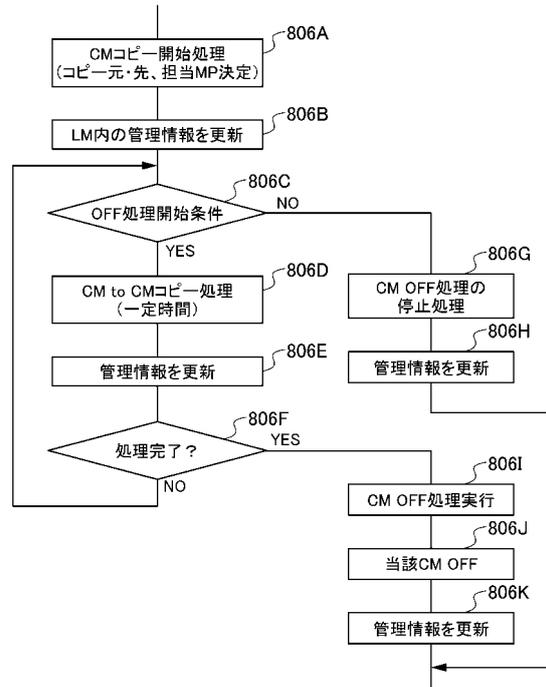
【図8】

図8



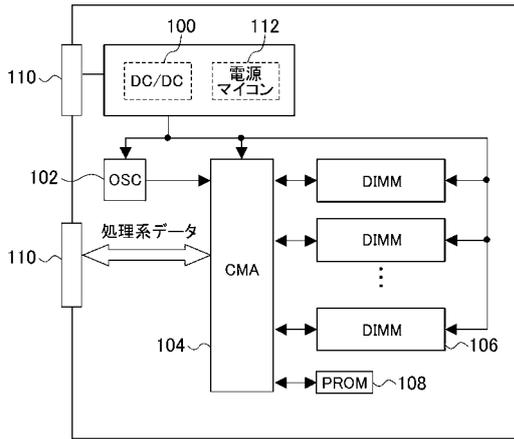
【図9】

図9



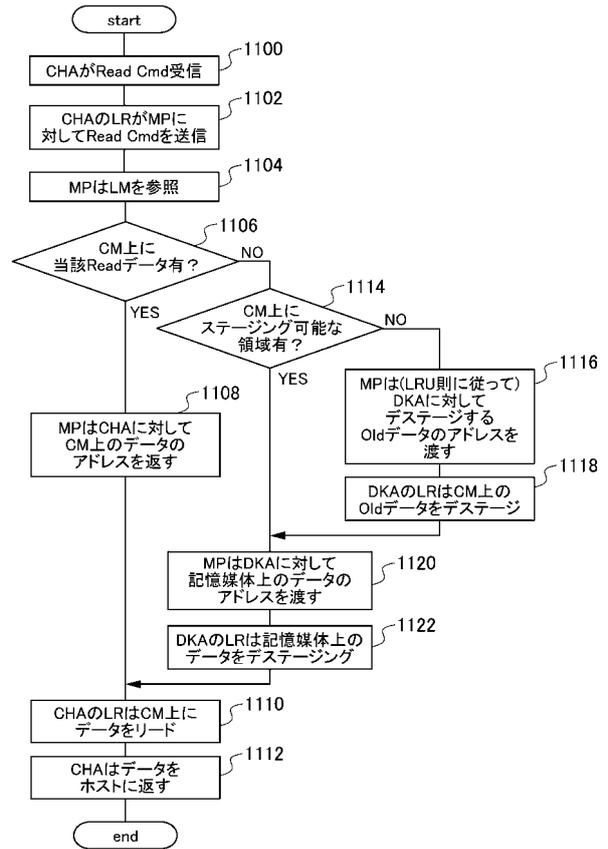
【図10】

図10



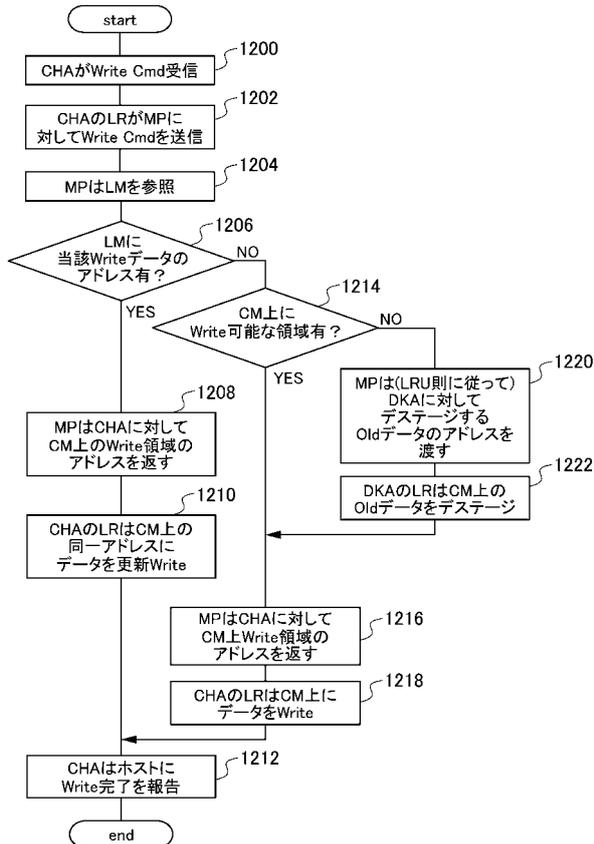
【図11】

図11



【図12】

図12



【図13】

図13

セグメント管理ID	コピー完了フラグ	コピー元		コピー先	
		DKCモジュールID	キャッシュアドレス	DKCモジュールID	キャッシュアドレス
0000	1	0	00EF1234	1	000A1056
0001	1	0	00EF1235	1	000A1057
0002	0	0	00EF1236	1	000A1058
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

フロントページの続き

(58)調査した分野(Int.Cl. , DB名)

G 0 6 F 3 / 0 6 - 3 / 0 8
G 0 6 F 1 2 / 0 0 - 1 2 / 1 6
G 0 6 F 1 3 / 0 0 - 1 3 / 4 2