

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-252358  
(P2006-252358A)

(43) 公開日 平成18年9月21日(2006.9.21)

(51) Int. Cl.	F I	テーマコード (参考)
<b>G06F 3/06 (2006.01)</b>	G06F 3/06 302A	5B005
<b>G06F 12/08 (2006.01)</b>	G06F 3/06 540	5B065
	G06F 12/08 503Z	
	G06F 12/08 541Z	
	G06F 12/08 551Z	
	審査請求 有 請求項の数 31 O L (全 22 頁) 最終頁に続く	

(21) 出願番号 特願2005-70175 (P2005-70175)  
(22) 出願日 平成17年3月11日 (2005.3.11)

(71) 出願人 000004237  
日本電気株式会社  
東京都港区芝五丁目7番1号  
(74) 代理人 100093595  
弁理士 松本 正夫  
(72) 発明者 桑田 篤史  
東京都港区芝五丁目7番1号 日本電気株式会社内  
Fターム(参考) 5B005 JJ01 JJ12 KK15 MM12 MM23  
NN01 VV13  
5B065 BA01 CA30 CH01

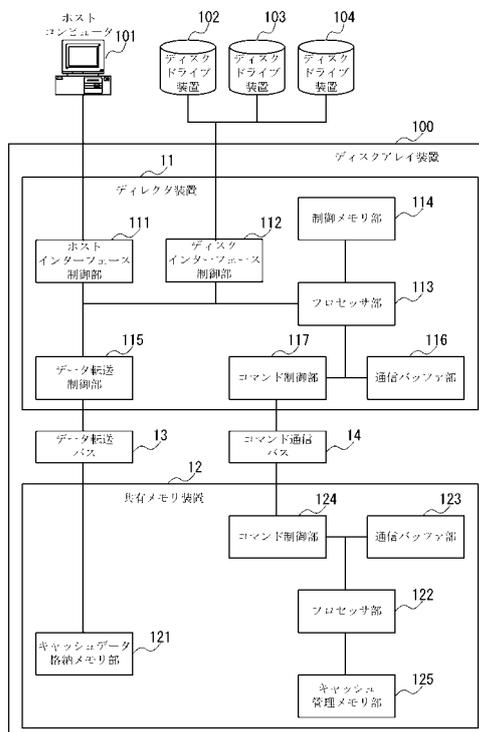
(54) 【発明の名称】 ディスクアレイ装置及びその共有メモリ装置、ディスクアレイ装置の制御プログラム及び制御方法

(57) 【要約】

【課題】 高速スループットバスを使用してキャッシュ制御を高速化したディスクアレイ装置を提供する。

【解決手段】 外部インタフェース制御部111、データ転送制御部112、制御メモリ113、プロセッサ114、コマンド制御部115、通信バッファ116を持つディレクタ装置11と、キャッシュデータ格納メモリ121、コマンド制御部122、通信バッファ123、プロセッサ124、キャッシュ管理メモリ125を持つ共有メモリ装置12とを有するディスクアレイ装置100は、ディレクタ装置11と共有メモリ装置12とが、データ転送制御部112、122間でデータ転送バス13を介して接続され、コマンド制御部115、122間でコマンド通信バス14を介して接続される。データ転送バス13とコマンド通信バス14とは、転送レートの高いシリアルバスである。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

外部装置とディスクドライブ装置との間のデータの入出力を管理するディレクタ装置と、入出力データに対するキャッシュメモリを有する共有メモリ装置とを備えるディスクアレイ装置において、

前記ディレクタ装置が、前記共有メモリ装置に、前記入出力データに対するキャッシュメモリの制御を指示するためのコマンドを送信し、

前記共有メモリ装置が、前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行することを特徴とするディスクアレイ装置。

10

**【請求項 2】**

前記ディレクタ装置が、

前記コマンドを送信し、かつ前記共有メモリ装置から送られる前記コマンドに対する処理結果を受信するコマンド制御部を備え、

前記共有メモリ装置が、

前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行する処理部と、

前記ディレクタ装置からのコマンドを受信し、かつ前記共有メモリ装置からの前記コマンドに対する処理結果を送信するコマンド制御部を備えることを特徴とする請求項 1 に記載のディスクアレイ装置。

20

**【請求項 3】**

前記ディレクタ装置と前記共有メモリ装置のコマンド制御部が、互いに転送速度の大きい通信バスによって接続され、

前記ディレクタ装置と前記共有メモリ装置のコマンド制御部間で、前記キャッシュメモリの状態に関する情報をやり取りすることを特徴とする請求項 2 の記載のディスクアレイ装置。

**【請求項 4】**

前記ディレクタ装置が、通信バッファ部を有し、

前記コマンドを前記通信バッファに格納することで、前記共有メモリ装置に対する制御動作から開放されることを特徴とする請求項 1 から請求項 3 の何れか 1 項に記載のディスクアレイ装置。

30

**【請求項 5】**

前記ディレクタ装置は、前記共有メモリ装置から送られる前記コマンドに対する処理結果を前記通信バッファで受信することを特徴とする請求項 4 に記載のディスクアレイ装置。

**【請求項 6】**

前記共有メモリ装置は、前記ディレクタ装置から送られる前記コマンドを受信して格納し、前記コマンドに対する処理結果を格納する通信バッファ部を有することを特徴とする請求項 1 から請求項 5 の何れか 1 項に記載のディスクアレイ装置。

**【請求項 7】**

複数の前記ディレクタ装置と前記共有メモリ装置を備え、

複数の前記ディレクタ装置と複数の前記共有メモリ装置とが、相互に前記コマンド制御部を介して接続されることを特徴とする請求項 2 に記載のディスクアレイ装置。

40

**【請求項 8】**

前記ディレクタ装置は、通信バッファ部を有し、複数の前記メモリ装置から送られる前記コマンドに対する複数の処理結果を前記通信バッファで一括して受信することを特徴とする請求項 7 に記載のディスクアレイ装置。

**【請求項 9】**

複数の前記共有メモリ装置は、複数の前記ディレクタ装置から送られる前記コマンドを一括して受信し、かつ前記コマンドに対する処理結果を格納する通信バッファ部を有することを特徴とする請求項 7 又は請求項 8 の何れか 1 項に記載のディスクアレイ装置。

50

## 【請求項 10】

複数の前記ディレクタ装置を、前記外部装置からのデータ要求を受け付けるホストディレクタ装置と、前記ディスクドライブ装置が接続される他のディレクタ装置とに分けて構成したことを特徴とする請求項 7 から請求項 9 の何れか 1 項に記載のディスクアレイ装置。

## 【請求項 11】

複数の前記ディレクタ装置が、それぞれ前記外部装置と前記ディスクドライブ装置が接続される構成としたことを特徴とする請求項 7 から請求項 9 の何れか 1 項に記載のディスクアレイ装置。

## 【請求項 12】

複数の前記ディレクタ装置と、単一の前記共有メモリ装置を備え、

10

複数の前記ディレクタ装置が、前記共有メモリ装置の処理部に対してキャッシュメモリの制御を指示するコマンドを送信することを特徴とする請求項 1 に記載のディスクアレイ装置。

## 【請求項 13】

単一の前記ディレクタ装置と、複数の前記共有メモリ装置を備え、

前記ディレクタ装置が、複数の前記共有メモリ装置に対してキャッシュメモリの制御を指示するコマンドを送信することを特徴とする請求項 1 に記載のディスクアレイ装置。

## 【請求項 14】

前記共有メモリ装置に、前記ディスクドライブ装置へのライトバック処理において前記キャッシュメモリのデータに対するパリティ演算処理を行うパリティ演算部を設けることを特徴とする請求項 1 から請求項 13 の何れか 1 項に記載のディスクアレイ装置。

20

## 【請求項 15】

前記パリティ演算部が、前記キャッシュメモリのデータ転送パスと別のパスで前記キャッシュメモリと接続されることを特徴とする請求項 14 に記載のディスクアレイ装置。

## 【請求項 16】

前記ディレクタ装置と前記共有メモリ装置を、互いに切り離して別装置として構成することを特徴とする請求項 1 から請求項 15 の何れか 1 項に記載のディスクアレイ装置。

## 【請求項 17】

外部装置とディスクドライブ装置との間のデータの入出力を管理するディレクタ装置と、入出力データに対するキャッシュメモリを有する共有メモリ装置とを備えるディスクアレイ装置の共有メモリ装置であって、

30

前記ディレクタ装置から送信される前記入出力データに対するキャッシュメモリの制御を指示するためのコマンドに基づいて、前記入出力データに対する前記キャッシュメモリの制御を実行する構成としたことを特徴とするディスクアレイ装置の共有メモリ装置。

## 【請求項 18】

前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行する処理部と、

前記ディレクタ装置のコマンド制御部から送信される前記コマンドを受信し、かつ前記コマンドに対する処理結果を、前記前記ディレクタ装置のコマンド制御部に対して送信するコマンド制御部を備えることを特徴とする請求項 17 に記載のディスクアレイ装置の共有メモリ装置。

40

## 【請求項 19】

前記コマンド制御部を介して、前記ディレクタ装置のコマンド制御部と互いに通信バスによって接続され、

前記ディレクタ装置のコマンド制御部との間で、前記キャッシュメモリの状態に関する情報をやり取りすることを特徴とする請求項 18 の記載のディスクアレイ装置の共有メモリ装置。

## 【請求項 20】

前記ディレクタ装置から送られる前記コマンドを受信して格納し、前記コマンドに対する処理結果を格納する通信バッファ部を有することを特徴とする請求項 17 から請求項 19

50

の何れか 1 項に記載のディスクアレイ装置の共有メモリ装置。

【請求項 2 1】

複数の前記ディレクタ装置と前記共有メモリ装置を備え、

複数の前記ディレクタ装置と複数の前記共有メモリ装置とが、相互に前記コマンド制御部を介して接続されることを特徴とする請求項 1 8 に記載のディスクアレイ装置の共有メモリ装置。

【請求項 2 2】

複数の前記共有メモリ装置は、複数の前記ディレクタ装置から送られる前記コマンドを一括して受信し、かつ前記コマンドに対する処理結果を格納する通信バッファ部を有することを特徴とする請求項 2 1 に記載のディスクアレイ装置の共有メモリ装置。

10

【請求項 2 3】

前記共有メモリ装置に、前記ディスクドライブ装置へのライトバック処理において前記キャッシュメモリのデータに対するパリティ演算処理を行うパリティ演算部を設けることを特徴とする請求項 1 7 から請求項 2 2 の何れか 1 項に記載のディスクアレイ装置の共有メモリ装置。

【請求項 2 4】

前記パリティ演算部が、前記キャッシュメモリのデータ転送パスと別のパスで前記キャッシュメモリと接続されることを特徴とする請求項 2 3 に記載のディスクアレイ装置の共有メモリ装置。

【請求項 2 5】

前記ディレクタ装置から切り離して別装置として構成することを特徴とする請求項 1 7 から請求項 2 4 の何れか 1 項に記載のディスクアレイ装置の共有メモリ装置。

20

【請求項 2 6】

外部装置とディスクドライブ装置との間のデータの入出力を管理するディレクタ装置と、入出力データに対するキャッシュメモリを有する共有メモリ装置とを備えるディスクアレイ装置において、前記データの入出力を制御する制御プログラムであって、

前記ディレクタ装置のプロセッサ及び前記共有メモリ装置に設けたプロセッサ上で実行され、

前記ディレクタ装置のプロセッサに、前記共有メモリ装置に、前記入出力データに対するキャッシュメモリの制御を指示するためのコマンドを送信する機能を持たせ、

30

前記共有メモリ装置のプロセッサに、前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行する機能を持たせる

ことを特徴とするディスクアレイ装置の制御プログラム。

【請求項 2 7】

前記ディレクタ装置のプロセッサに、

前記コマンドを送信し、かつ前記共有メモリ装置から送られる前記コマンドに対する処理結果を受信する機能を持たせ、

前記共有メモリ装置のプロセッサに、

前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行する機能と、

40

前記ディレクタ装置からのコマンドを受信し、かつ前記共有メモリ装置からの前記コマンドに対する処理結果を送信する機能を持たせることを特徴とする請求項 2 6 に記載のディスクアレイ装置の制御プログラム。

【請求項 2 8】

前記ディレクタ装置のプロセッサと前記共有メモリ装置のプロセッサに対し、前記ディレクタ装置と前記共有メモリ装置間で、前記キャッシュメモリの状態に関する情報をやり取りする機能を持たせることを特徴とする請求項 2 7 に記載のディスクアレイ装置の制御プログラム。

【請求項 2 9】

外部装置とディスクドライブ装置との間のデータの入出力を管理するディレクタ装置と、

50

入出力データに対するキャッシュメモリを有する共有メモリ装置とを備えるディスクアレイ装置における、前記データの入出力を制御する制御方法であって、

前記ディレクタ装置のプロセッサから、前記共有メモリ装置に備えたプロセッサに対して、前記入出力データに対するキャッシュメモリの制御を指示するためのコマンドを送信するステップと、

前記共有メモリ装置のプロセッサが、前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行するステップとを有することを特徴とするディスクアレイ装置の制御方法。

【請求項 30】

前記ディレクタ装置のプロセッサが、

前記コマンドを送信し、かつ前記共有メモリ装置から送られる前記コマンドに対する処理結果を受信するステップを有し、

前記共有メモリ装置のプロセッサが、

前記ディレクタ装置からのコマンドに基づいて前記入出力データに対する前記キャッシュメモリの制御を実行するステップと、

前記ディレクタ装置からのコマンドを受信し、かつ前記共有メモリ装置からの前記コマンドに対する処理結果を送信するステップを有することを特徴とする請求項 29 に記載のディスクアレイ装置の制御方法。

【請求項 31】

前記ディレクタ装置のプロセッサと前記共有メモリ装置のプロセッサ間で、前記キャッシュメモリの状態に関する情報をやり取りするステップを有することを特徴とする請求項 30 に記載のディスクアレイ装置の制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ディスクアレイ装置、キャッシュメモリ管理方法、キャッシュメモリ管理プログラム及びキャッシュメモリに関し、特に、高速スループットバスを用いたディスクアレイ装置及びその共有メモリ装置、ディスクアレイ装置の制御プログラム及び制御方法に関する。

【背景技術】

【0002】

従来 of ディスクアレイ装置の 1 例について、図 11 を参照して説明する。

【0003】

図 11 において、従来 of ディスクアレイ装置は、外部インタフェース 1111 及び 1121、データ転送管理部 1112 及び 1122、プロセッサ 1113 及び 1123、管理領域制御部 1114 及び 1124 をそれぞれ有する複数のディレクタ装置 1110 及び 1120 と、キャッシュデータ格納メモリ 1131 及び 1141、キャッシュ管理メモリ 1132 及び 1142 をそれぞれ有する複数の共有メモリ装置 1130 及び 1140 とを備えており、プロセッサ 1113 及び 1123 が、共有メモリ装置 1130 及び 1140 の管理領域を操作して、キャッシュデータ格納メモリ 1131 及び 1141、キャッシュ管理メモリ 1132 及び 1142 の管理や処理を行う。

【0004】

上記のような従来 of ディスクアレイ装置の 1 例が、例えば特開 2004 - 139260 号公報（特許文献 1）で示されている。

【0005】

特許文献 1 は、上位ホストサーバからのコマンド処理を各マイクロプロセッサへコマンドを転送し、複数のマイクロプロセッサにて分散処理する方式により、インタフェース部のマイクロプロセッサのボトルネックを軽減し、ストレージシステムの性能低下を防ぐディスク装置の構成を開示している。

【特許文献 1】特開 2004 - 139260 号公報

10

20

30

40

50

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0006】

しかしながら、上述した従来のディスクアレイ装置においては、以下に述べるような課題が存在する。

## 【0007】

第1の課題は、従来のディスクアレイ装置では、ディレクタ装置上のプロセッサが、共有メモリ装置上のキャッシュメモリの制御を行うことから、ディレクタ装置のローカルバス、ディレクタ装置と共有メモリ装置間の共有バス、共有メモリ装置内のメモリバスと複数階層のバスを介してメモリアクセスを行う必要があり、メモリアクセスに時間を長く要するということである。

10

## 【0008】

第2の課題は、従来の技術として示した複数のディレクタ装置を複数設けたマルチプロセッサシステムによって分散処理を行う構成としても、ディレクタ装置上のプロセッサでのキャッシュ制御処理（メモリアクセス処理）において、プロセッサキャッシュを使用することが難しいので、ディレクタ装置のプロセッサによるキャッシュメモリ制御処理の高速化が難しいということである。

## 【0009】

第3の課題は、クロックアップなどの基盤技術の向上によりデータ転送能力が向上しても、共有キャッシュメモリの制御に関し、高速スループットバスを活用した処理時間の短縮は難しいことである。

20

## 【0010】

本発明の目的は、上記課題を解決し、キャッシュメモリ制御の処理を高速化することができるディスクアレイ装置及びその共有メモリ装置、ディスクアレイ装置の制御プログラム及び制御方法を提供することにある。

## 【課題を解決するための手段】

## 【0011】

上記目的を達成するため本発明は、ディレクタ装置上のプロセッサによって共有メモリ装置上のキャッシュメモリの制御を行う代わりに、ディレクタ装置上のプロセッサからの通信により、共有メモリ装置上のプロセッサが、共有メモリ装置上のキャッシュメモリの制御を行うことを特徴とする。

30

## 【0012】

この構成によって、本発明は、メモリ操作において、共有メモリ装置上のプロセッサがメモリバスを直接制御し、キャッシュ制御に要する処理時間を短縮することができる。また、ディスクアレイ装置がキャッシュ制御中の場合でも、ディレクタ装置上のプロセッサがプロセッサキャッシュを使用することができる。さらに、複数のディレクタ装置から構成されなくても、単一のディレクタ装置によって、キャッシュメモリ制御に要する処理時間を短縮することができる。

## 【発明の効果】

## 【0013】

本発明のディスクアレイ装置及びその共有メモリ装置、ディスクアレイ装置の制御プログラム及び制御方法によれば、以下のような効果が達成される。

40

## 【0014】

第1の効果は、共有メモリ装置のキャッシュメモリ制御に要する処理時間を短縮することが可能となることである。

## 【0015】

その理由は、ディレクタ装置上のプロセッサによって共有メモリ装置上のキャッシュメモリの制御を行う代わりに、ディレクタ装置上のプロセッサからの通信により、共有メモリ装置上のプロセッサが、共有メモリ装置上のキャッシュメモリの制御を行うようにしたためである。

50

## 【0016】

第2の効果は、共有メモリ装置上のプロセッサが、共有メモリ装置上のキャッシュメモリの制御を行うことから、ディレクタ装置のプロセッサ間での処理の競合を防止するためのロック処理が必要なくなるので、ロック処理に要する時間も節約されるということである。

## 【発明を実施するための最良の形態】

## 【0017】

次に、本発明の実施の形態について、図面を参照して詳細に説明する。

## 【0018】

(第1の実施の形態)

図1は、本発明の第1の実施の形態によるディスクアレイ装置100のハードウェア構成を示すブロック図である。

10

## 【0019】

図1において、ディスクアレイ装置100は、そのハードウェア構成として、データ転送バス13及びコマンド通信バス14を介して互いに接続されたディレクタ装置11及び共有メモリ装置12を有する。

## 【0020】

ディレクタ装置11は、ホストコンピュータ101及びディスクドライブ102、103、104との間で共有メモリ装置12を管理するコマンドについて通信し、共有メモリ装置12にその管理コマンドを送信する装置であって、プログラム制御によりホストインタフェース制御部111、ディスクインタフェース部112、プロセッサ部113、制御メモリ部114、データ転送制御部115、通信バッファ部116、コマンド制御部117の機能を実現する。

20

## 【0021】

共有メモリ装置12は、ディレクタ装置11から共有メモリ装置12を管理するコマンドを受信することによって、キャッシュデータ格納メモリ部121、プロセッサ部122、通信バッファ部123、コマンド制御部124、キャッシュ管理メモリ部125の各機能を実現する。

## 【0022】

ディレクタ装置11と共有メモリ装置12とは、データ転送制御部115とキャッシュデータ格納メモリ部121間がデータ転送バス13を介して接続され、コマンド制御部117、124がコマンド通信バス14を介して接続される。

30

## 【0023】

データ転送バス13とコマンド通信バス14とは、転送レートの高いシリアルバスであって、例えば、InfiniBand等のバスである。

## 【0024】

まず、ディレクタ装置11の構成について説明する。

## 【0025】

ホストインタフェース制御部111は、ホストコンピュータ101、データ転送制御部115、プロセッサ部113等と接続しており、プロセッサ部113からの命令に従い、ホストコンピュータ101から受信したキャッシュデータを要求するコマンドをプロセッサ部113に送信し、データ転送制御部115から受信したキャッシュデータをホストコンピュータ101送信する機能を有する装置である。

40

## 【0026】

ディスクインタフェース部112は、ディスクドライブ102～104、プロセッサ部113、データ転送制御部115等と接続しており、プロセッサ部113からの命令に従い、ディスクドライブ102～104にキャッシュデータを要求するコマンドを送信し、ディスクドライブ102～104から受信したキャッシュデータをデータ転送制御部115へ送信する機能を有する。

## 【0027】

50

プロセッサ部 1 1 3 は、ホストインタフェース制御部 1 1 1、ディスクインタフェース部 1 1 2、制御メモリ部 1 1 4、データ転送制御部 1 1 5、通信バッファ部 1 1 6、コマンド制御部 1 1 7 と接続し、ホストインタフェース制御部 1 1 1 から受信したコマンドに従ってディスクインタフェース部 1 1 2、制御メモリ部 1 1 4、データ転送制御部 1 1 5、通信バッファ 1 1 6 等に命令する機能を有する。

【 0 0 2 8 】

より詳細には、プロセッサ部 1 1 3 は、このデータ転送に先立って、共有メモリ装置 1 2 にキャッシュページオープンを示すコマンドをコマンド制御部 1 1 5 から送信させるための命令を通信バッファ部 1 1 6 に格納する。

【 0 0 2 9 】

ここで、キャッシュページとは、キャッシュデータ格納メモリ 1 2 1 に格納されているキャッシュデータに対応した領域を指しており、後述するプロセッサ 1 2 2 が返却するメモリアドレス情報とは、このキャッシュデータに対応した領域（キャッシュページ）のメモリアドレスである。

【 0 0 3 0 】

プロセッサ 1 1 3 は、さらに、プロセッサ 1 2 2 から返送されたこれらの情報をもとにデータ転送を行い、データ転送完了後に今度はキャッシュページクローズを示すコマンドを送信する機能を有する。ここでは、クローズすべきキャッシュページの論理アドレスとキャッシュステート情報を送信する。

【 0 0 3 1 】

また、後述するキャッシュステート情報とは、このキャッシュページに有効なデータが格納されているかどうかを示す情報である。このキャッシュステート情報は、空のキャッシュページにデータを格納した場合には有効化し、未書き込みデータをディスクに新たに書き落とした場合などには変更する。

【 0 0 3 2 】

制御メモリ部 1 1 4 は、プロセッサ 1 1 3 によって処理されるデータを一時的に格納するプロセッサキャッシュとしての機能を有する。

【 0 0 3 3 】

データ転送制御部 1 1 5 は、データ転送バス 1 3、ホストインタフェース制御部 1 1 1、ディスクインタフェース部 1 1 2、プロセッサ部 1 1 3 と接続しており、プロセッサ部 1 1 3 からの命令に従い、共有メモリ装置 1 2 からデータ転送バス 1 3 を介して受信したデータを、ホストインタフェース制御部 1 1 1 に送信し、ディスクインタフェース部 1 1 2 から受信したキャッシュデータをデータ転送バス 1 3 を介して共有メモリ装置 1 2 へ送信する機能を有する。

【 0 0 3 4 】

通信バッファ部 1 1 6 は、プロセッサ部 1 1 3、コマンド制御部 1 1 7 と接続しており、プロセッサ部 1 1 3 からの命令を格納し、この命令をコマンド制御部 1 1 7 に送信する機能を有する。

【 0 0 3 5 】

コマンド制御部 1 1 7 は、コマンド通信バス 1 4、プロセッサ部 1 1 3、通信バッファ部 1 1 6 と接続しており、通信バッファ部 1 1 6 から送信された命令に従ってコマンド通信バス 1 4 を介して共有メモリ装置 1 2 のコマンド制御部 1 2 2 と通信を行う機能を有する。

【 0 0 3 6 】

すなわち、コマンド制御部 1 1 7 は、共有メモリ装置 1 2 のコマンド制御部 1 2 4 に対して、通信バッファ部 1 1 6 からの命令によって送信を指示された上記共有メモリ装置 1 2 にキャッシュページオープンを示すコマンドを送信する。また、このコマンドのレスポンスとしてコマンド制御部 1 2 4 から受信したメモリアドレス情報、キャッシュステート情報、新規キャッシュデータ要求コマンド等を受信し、通信バッファ部 1 1 6 に格納するとともに、プロセッサ部 1 1 3 に通知する。

10

20

30

40

50

## 【0037】

次いで、共有メモリ装置12の構成について説明する。

## 【0038】

キャッシュデータ格納メモリ部121は、データ転送バス13と接続しており、キャッシュメモリとしてデータを格納する機能を有する。

## 【0039】

プロセッサ122は、通信バッファ部123、コマンド制御部124、キャッシュ管理メモリ部125と接続しており、通信バッファ部123から上記コマンドを取り込んでキャッシュ管理メモリ125上でキャッシュページオープン制御等のキャッシュメモリの制御に関する処理を行う。

10

## 【0040】

より詳細には、プロセッサ122は、指示された論理アドレスがキャッシュヒットした場合にはヒットしたキャッシュページについて、メモリアドレス情報とキャッシュステート情報をプロセッサ114に返送する。一方、キャッシュミスした場合には追い出し制御により新規に割り当てたキャッシュページについて、メモリアドレス情報とキャッシュステート情報をプロセッサ114に返送する。

## 【0041】

通信バッファ部123は、コマンド制御部124及びプロセッサ122と接続し、コマンド制御部124及びプロセッサ122とデータの送受信を行い、受信したデータを格納する機能を有する装置である。

20

## 【0042】

コマンド制御部124は、コマンド通信バス14、プロセッサ部122、通信バッファ部123と接続しており、コマンド通信バス14を介してコマンド制御部117から受信したコマンドを通信バッファ部123に格納し、割り込み信号によってプロセッサ122に通知する装置である。

## 【0043】

キャッシュ管理メモリ部125は、キャッシュデータ格納メモリの割り当て状態を管理する。

## 【0044】

本発明の第1の実施の形態によるディスクアレイ装置100の構成の特徴は、共有メモリ装置に、プロセッサ122及びコマンド制御部124を有することである。また、プロセッサ122、コマンド制御部124間の通信を仲介する通信バッファ部123を有することである。

30

## 【0045】

さらに、ディレクタ装置11に、ホストインタフェース部111及びディスクインタフェース部112を有することである。

## 【0046】

また、プロセッサ114、122間で、メモリアドレス情報に加え、キャッシュステート情報を送受信することである。

## 【0047】

図2を参照すると、図1に示された共有メモリ装置のプロセッサ部、通信バッファ部及びコマンド制御部の詳細な構成が示されている。

40

## 【0048】

図2に示すように、通信バッファ部123は、プロセッサ部122及びコマンド制御部124とデータ通信を行う装置である。

## 【0049】

本実施の形態において、通信バッファ部123は、それぞれ複数の送信バッファ部123-1及び受信バッファ部123-2によって構成され、また、コマンド制御部124は、送信制御部124-1及び受信制御部124-2から構成されている。

## 【0050】

50

図2において、通信バッファ部123を構成する送信バッファ部123-1及び受信バッファ部123-2は、それぞれFIFO(First In First Out)構造になっている。

【0051】

送信制御部124-1は、プロセッサ部122が送信バッファ123-1に情報を書き込んで送信制御部124-1への送信指示を出すと、シリアルバスを通じてデータを送信する。

【0052】

受信制御部124-2は、シリアルバスを通じてデータを受信すると、受信したデータを受信バッファ123-2に書き込み、割り込み信号によってプロセッサ部432に通知する。

10

【0053】

以上詳細に本実施の形態の構成を述べたが、図2のシリアルバス及びFIFO構造のバッファについては、当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な構成の説明は省略する。

【0054】

なお、本実施の形態の実施例では、通信バッファ部として、プロセッサ部のローカルメモリの一部を使用してもよい。その場合には、通信バッファ部のアクセスにおいてプロセッサキャッシュを利用することもできる。

【0055】

また、本実施の形態では、共有メモリ装置12の例で説明したが、ディレクタ装置11においても同様である。

20

【0056】

次に、本実施の形態によるディスクアレイ装置のリードライト動作を説明する。

【0057】

図3は、第1の実施の形態によるディスクアレイ装置100のリードライト動作において、ホストディレクタ装置11と共有メモリ装置12の動作を示すフローチャートである。

【0058】

図3に示すように、ディレクタ装置11は、ステップ311において、ホストコンピュータ101から、キャッシュページオープンを示すコマンドを受信すると、ステップ312において、通信バッファ部116上にこのキャッシュページオープンコマンドを格納し、コマンド制御部117が、このコマンドを共有メモリ装置12に送信する。その後プロセッサ部113は、通信のレスポンスを待つが、その間には別のコマンド処理を行うことができる。

30

【0059】

共有メモリ装置12は、ステップ321において、ディレクタ装置11からキャッシュページオープンコマンドを受信すると、ステップ322において、キャッシュ管理メモリ部125上にてキャッシュページの検索処理を行う。

【0060】

次いで、キャッシュページの検索処理の結果キャッシュミスとなった場合には、ステップ323において、追い出し処理により新規に割り付ける処理を行う。

40

【0061】

次いで、キャッシュページの検索処理の結果キャッシュヒットとなった場合には、ステップ324において、そのキャッシュページがオープン中であるときには解放されるまで待つ。この間プロセッサ部122は、別のキャッシュ処理を行うことができる。

【0062】

以上のステップ323又はステップ324の処理により、使用するキャッシュ領域が確定すると、共有メモリ装置12は、ステップ325において、キャッシュページオープンコマンドのレスポンスとして、メモリアドレスとキャッシュステート情報をディレクタ装

50

置 1 1 に送信する。

【 0 0 6 3 】

ディレクタ装置 1 1 のプロセッサ部 1 1 6 は、ステップ 3 1 3 において、コマンド制御部 1 1 7 からの割り込み信号の受信により、キャッシュページオープン処理の完了を確認する。

【 0 0 6 4 】

次いで、プロセッサ部 1 1 3 は、送られてきたキャッシュステート情報を参照し、ステップ 3 1 4 において、必要なデータ転送を行う。必要なデータ転送とは、リード処理の場合、キャッシュヒットならば共有メモリ装置 1 2 からホストコンピュータ 1 0 1 へのデータ転送であり、キャッシュミスならディスクドライブ 1 0 2 ~ 1 0 4 から共有メモリ装置 1 2 へのデータ転送及び、共有メモリ装置 1 2 からホストコンピュータ 1 0 1 へのデータ転送である。一方、ライト処理の場合、ホストコンピュータ 1 0 1 から共有メモリ装置 1 2 へのデータ転送を行う。また、必要ならば共有メモリ装置 1 2 からディスクドライブ 1 0 2 ~ 1 0 4 へのデータ転送を行う。

10

【 0 0 6 5 】

データ転送が完了すると、プロセッサ部 1 1 3 が、ステップ 3 1 5 において、ステップ 3 1 2 と同様にして、キャッシュページクローズコマンドを作成して、コマンド制御部 1 1 7 が共有メモリ装置 1 2 に送信する。

【 0 0 6 6 】

プロセッサ部 1 2 2 は、ステップ 3 2 6 において、ステップ 3 2 1 と同様にして、キャッシュページクローズコマンドを受信すると、ステップ 3 2 7 において、排他制御の解除を行う。ここで同一のキャッシュページの使用待ちをしている処理があればその処理が使用可能となる。

20

【 0 0 6 7 】

次いで、共有メモリ装置 1 2 は、ステップ 3 2 8 において、ステップ 3 2 5 と同様にして、キャッシュページクローズコマンドのレスポンスをディレクタ装置 1 1 に送信する。

【 0 0 6 8 】

ディレクタ装置 1 1 は、ステップ 3 1 6 において、ステップ 3 1 3 と同様にして、プロセッサ 1 2 2 からのレスポンスを受信すると、ステップ 3 1 7 において、ホストコンピュータ 1 0 1 から受信したコマンドの処理を完了する。

30

【 0 0 6 9 】

本実施の形態は、共有メモリ装置 1 2 上でのキャッシュ制御を、ディレクタ装置 1 1 のプロセッサ 1 1 3 により行う代わりに、ディレクタ装置 1 1 のプロセッサ 1 1 3 からコマンドを送信された共有メモリ装置 1 2 上の単一のプロセッサ 1 2 2 により行うので、共有メモリ装置 1 2 のプロセッサ 1 2 2 がメモリ操作においてメモリバスを直接制御し、かつ、ディレクタ装置 1 1 のプロセッサ 1 1 6 がプロセッサキャッシュを使用できることから、キャッシュ制御に要する処理時間を短縮することができる。

【 0 0 7 0 】

なお、ディレクタ装置 1 1 によるライトバック処理は、キャッシュデータ格納メモリ 1 2 1 へのデータの書き込み処理と同期して行ってもよいし、非同期でもよい。

40

【 0 0 7 1 】

図 4 は、本実施の形態におけるディスクアレイ装置の処理について、ディレクタ装置と共有メモリ装置との通信の内容を時系列で示す図である。

【 0 0 7 2 】

図 4 を参照すると、本実施の形態における通信は、まず初めにステップ 4 1 0 において、ディレクタ装置 1 1 が、共有メモリ装置 1 2 にキャッシュページオープン指示を行う。ここでは、ホストコンピュータ 1 0 1 から要求されたコマンドの論理アドレス情報をこの通信内に付随させる。

【 0 0 7 3 】

次に、ステップ 4 2 0 において、ステップ 4 1 0 のレスポンスとして、共有メモリ装置

50

12が、ディレクタ装置11に割り当てたキャッシュページのメモリアドレス情報とキャッシュステート情報をディレクタ装置11に送信する。

【0074】

次に、ステップ430において、ディレクタ装置11は、オープンした共有メモリ装置12のキャッシュページとの間で、ホストコンピュータ101と共有メモリ装置12間のデータ転送や、ディスク102～104と共有メモリ装置12間のデータ転送を行う。(キャッシュ検索によるキャッシュのヒット/ミスの場合分け必要?)

【0075】

データ転送が完了すると、ステップ440において、ディレクタ装置11が、共有メモリ装置12にキャッシュページクローズ指示を行う。ここでは、論理アドレスとキャッシュステート情報をこの通信内に付随させる。

【0076】

最後に、ステップ450において、ステップ440のレスポンスとして、共有メモリ装置12が、ディレクタ装置11に処理終了を通知することによって、ディスクアレイ装置100の処理が終了となる。

【0077】

(第1の実施の形態による効果)

このようにして、本実施の形態は、共有メモリ装置12上でのキャッシュ制御を、ディレクタ装置11上のプロセッサ113により行うかわりに、ディレクタ装置11上のプロセッサ113からの通信に基づき共有メモリ装置12上のプロセッサ122により行うので、メモリ操作において共有メモリ装置12上のプロセッサ122がメモリバスを直接制御し、かつディレクタ装置11上のプロセッサ113がプロセッサキャッシュを使用することができることから、キャッシュメモリ制御に要する処理時間を短縮することができる。

【0078】

さらに、キャッシュ制御のためのディレクタ装置と共有メモリ装置間の通信処理を、プロセッサが直接実施するのではなく制御部に指示して行わせるだけなので、通信を行うことによるオーバーヘッドを少なくすることができ、処理の高速化を実現する。

【0079】

またコマンド通信バス14に転送レートの高いシリアルバスを使用することで、転送情報にメモリアドレスやキャッシュステート情報等を含む複数の情報を載せることができ、これによって転送時間の短縮が図られる。

【0080】

(第2の実施の形態)

図5は、本発明の第2の実施の形態によるディスクアレイ装置のハードウェア構成を示すブロック図である。

【0081】

図5を参照すると、本発明の第2の実施の形態によるディスクアレイ装置500が示されている。以下、第1の実施の形態と重複する説明を適宜省略し、本実施の形態によるディスクアレイ装置500の構成を説明する。

【0082】

図5が示すように、第2の実施の形態のディスクアレイ装置500は、データ転送バス55、56間及び、コマンド通信バス57、58間がそれぞれ接続されているディスクアレイ部50-1、50-2から構成されている。

【0083】

また、本実施の形態のディスクアレイ装置500は、第1の実施の形態によるディスクアレイ装置100と同様に、ディスクアレイ部50-1が、ホストディレクタ装置51及び共有メモリ装置53を有し、ディスクアレイ部50-2が、ディスクディレクタ装置52及び共有メモリ装置54を有する。

【0084】

10

20

30

40

50

本実施の形態によるディスクアレイ装置 500 が、第 1 の実施の形態によるディスクアレイ装置 100 と異なる点は、ディスクアレイ部 50 - 1、50 - 2 と複数のディスクアレイ部から構成されている点、ホストディレクタ装置 51 が、ディスクインタフェース部を有さない点、ディスクディレクタ装置 52 が、ホストインタフェース部を有さない点、データ転送バス 55、56 同士が接続されている点、コマンド通信バス 57、58 同士が接続されている点である。

【0085】

図 5 において、ホストディレクタ装置 51 のプロセッサ 513 は、ホストコンピュータ 501 から受信したコマンドを判別して通信バッファ部 516 上に作成したコマンドを、共有メモリ装置 53、54 に対して送信する。

10

【0086】

ディスクディレクタ装置 52 は、ディスクインタフェース制御部 522 を介してディスクドライブ 502、503、504 と接続されており、ホストディレクタ装置 51 からの命令によって、共有メモリ装置 53、54 と通信を行う。

【0087】

ホストディレクタ装置 51、ディスクディレクタ装置 52 及び共有メモリ装置 53、54 は、それぞれプロセッサ部 (513、523、532、542)、通信バッファ部 (516、526、533、543)、及びコマンド制御部 (517、527、534、544) を備えている。

【0088】

ホストディレクタ装置 51 及びディスクディレクタ装置 52 がそれぞれ備えるデータ転送制御部 515、525 は、シリアルバスなどの高速転送バスによって構成されたデータ転送バス 55、56 でキャッシュデータ格納メモリ 531、541 と接続されている。

20

【0089】

全てのコマンド制御部 (517、527、534、544) は、シリアルバスなどの高速転送バスで構成されたコマンド通信バス 57、58 で相互接続されている。

【0090】

本実施の形態によるディスクアレイ装置におけるリードライトの動作について説明する。

【0091】

本実施の形態によるディスクアレイ装置のリードライト動作は、第 1 の実施の形態によるディスクアレイ装置のリードライト動作と同様のため、図 2 を用い、重複部分を適宜省略して説明する。

30

【0092】

なお、本実施の形態によるリードライト動作が第 1 の実施の形態によるリードライト動作と異なる点は、複数の共有メモリ装置 53、54 が、ホストディレクタ装置 40 - 1 と通信する点、また、必要ならば複数の共有メモリ装置 53、54 からディスクドライブ 502 ~ 504 へのデータ転送を行う点、さらにその際、ホストディレクタ装置 51 とディスクディレクタ装置 52 との間で必要に応じて通信を行う点である。

【0093】

特に本実施の形態では、ホストディレクタ装置 51 のプロセッサ部 513 が、ステップ 213 において、送られてきたキャッシュステート情報を参照し、ステップ 214 において、必要なデータ転送を共有メモリ装置 53、54 との間で行う。必要なデータ転送とは、リード処理の場合、キャッシュヒットならば共有メモリ装置 53、54 からホストコンピュータ 501 へのデータ転送であり、キャッシュミスならディスクドライブ 502 ~ 504 から共有メモリ装置 53、54 へのデータ転送及び、共有メモリ装置 53、54 からホストコンピュータ 501 へのデータ転送である。一方、ライト処理の場合、ホストコンピュータ 501 から共有メモリ装置 33、34 へのデータ転送及び、必要ならば共有メモリ装置 33、34 からディスクドライブ 502 ~ 504 へのデータ転送を行う。

40

【0094】

50

このとき、ホストディレクタ装置 5 1 とディスクディレクタ装置 5 2 との間で必要に応じて通信を行う。

【0095】

(第2の実施の形態による効果)

このようにすることで、本実施の形態は、共有メモリ装置 5 3、5 4 上でのキャッシュ制御を、複数のディレクタ装置 5 1、5 2 上の各プロセッサ部 5 1 3、5 2 3 により行う代わりに、複数のディレクタ装置 5 1、5 2 上の各プロセッサ部からの通信に基づいて、各共有メモリ装置 5 3、5 4 上の単一のプロセッサ部 5 3 2、5 4 2 により行うので、共有メモリ装置 5 3、5 4 のプロセッサ部 5 3 2、5 4 2 がメモリ操作においてメモリバスを直接制御し、かつ、複数のディレクタ装置 5 1、5 2 の各プロセッサ部 5 1 3、5 2 3 がプロセッサキャッシュを使用することができることから、キャッシュ制御に要する処理時間を短縮することができる。

10

【0096】

また、共有メモリ装置 5 3、5 4 上のプロセッサが、共有メモリ装置上のキャッシュメモリの制御を行うことから、ディレクタ装置のプロセッサ間での処理の競合を防止するためのロック処理が必要なくなるので、ロック処理に要する時間も節約され、処理のより高速化が図られる。

【0097】

(第3の実施の形態)

本発明の第3の実施の形態は、その基本的構成は上記の第2の実施の形態の通りであるが、ホストディレクタ装置とディスクディレクタ装置間の通信をしなくてすむようにさらに工夫されている。

20

【0098】

図6は、本発明の第3の実施の形態によるディスクアレイ装置 6 0 0 の構成を示すブロック図である。

【0099】

図6を参照すると、本実施の形態によるディスクアレイ部 6 0 - 1、6 0 - 2 は、それぞれ第1の実施の形態によるディスクアレイ装置 1 0 0 (図1参照)と同様の構成である。

【0100】

したがって、本実施の形態によると、複数のディレクタ装置 6 1、6 2 上のプロセッサ部 6 1 3、6 2 3 からの通信により、共有メモリ装置 6 3、6 4 上のプロセッサ部 6 3 2、6 4 2 がキャッシュ管理制御を行うので、メモリ操作において共有メモリ装置 6 3、6 4 のプロセッサ部 6 3 2、6 4 2 がメモリバスを直接制御し、かつ、ディレクタ装置 6 1、6 2 のプロセッサ部 6 1 3、6 2 3 がプロセッサキャッシュを使用することができることとなり、複数のディレクタ装置からなる構成であっても、キャッシュ制御に要する処理時間を短縮することが可能となる。

30

【0101】

また、本実施の形態によるディレクタ装置 6 1 は、第2の実施の形態のホストディレクタ装置 3 1 (図3参照)と異なり、ホストインタフェース制御部 6 1 1 及びディスクインタフェース制御部 6 1 2 を有しており、また、ディレクタ装置 6 2 も、第2の実施の形態のディスクディレクタ装置 3 2 (図3参照)と異なり、ディレクタ装置 6 1 と同様にして、ホストインタフェース制御部 6 2 1 及びディスクインタフェース制御部 6 2 2 を有している。

40

【0102】

(第3の実施の形態による効果)

このように、本実施の形態によると、ディレクタ装置 6 1、6 2 は、第1の実施の形態によるディレクタ装置 1 1 と同様に、それぞれホストインタフェース制御部 6 1 1、6 2 1 及びディスクインタフェース制御部 6 1 2、6 2 2 を有しているので、第2の実施の形態による効果と比較して、共有メモリ装置 6 3、6 4 からメモリアドレスを受信した後に

50

データ転送を実施する際、ディレクタ装置 6 1、6 2 間の通信を行うことなく、各ディレクタ装置でコマンド処理を全て完了することが可能となる。

【0103】

(第4の実施の形態)

本発明の第4の実施の形態は、その基本的構成は上記の第3の実施の形態の通りであるが、共有メモリ装置からディスクドライブへのデータのライトバック処理においてのパリテイ演算処理についてさらに工夫している。

【0104】

図7は、本発明の第4の実施の形態によるパリテイ演算処理機能を有する共有メモリ装置の構成を示すブロック図である。

10

【0105】

図7を参照すると、共有メモリ装置73は、第3の実施の形態における図6に示される共有メモリ装置63、64と同様の構成であるが、共有メモリ装置63、64の構成と比べ、パリテイ演算部736を備えているので、RAID制御に必要なパリテイ演算を共有メモリ装置73内において閉じた状態で行うことができる。

【0106】

したがって、ディレクタ装置によるパリテイ演算処理の負荷を軽減できることとなった。

【0107】

また、パリテイ演算部736は、キャッシュデータ格納メモリ部731及びプロセッサ部732と接続しており、プロセッサ部732の指示により、キャッシュデータ格納メモリ部731がディレクタ装置71、72とデータの送受信を行うデータ転送バス75とは別経路によって、データをキャッシュデータ格納メモリ部731に送信する構成となっている。

20

【0108】

したがって、データ転送バス75の競合が軽減され、転送率の向上を実現することができる。

【0109】

図8は、第4の実施の形態におけるディスクアレイ装置のライトバック処理を説明するフローチャートである。

30

【0110】

図8を参照すると、第4の実施の形態におけるライトバック処理は、まずステップ810において、書き込み用データページ、旧データ用ページ、旧パリテイ用ページ、新パリテイ用ページをオープンする。

【0111】

次にステップ820において、旧データ用ページ、旧パリテイ用ページに、ディスクドライブからデータを読み出す。

【0112】

次にステップ830において、パリテイ演算を指示するコマンドをディレクタ装置から共有メモリ装置73に通信する。プロセッサ732は、このコマンドを受信すると、パリテイ演算部736にパリテイ演算を指示してパリテイ演算を行わせる。

40

【0113】

次に、ステップ840において、新たなデータ、新たなパリテイをディスクに書き込む。

【0114】

最後に、ステップ850において、書き込み用データページ、旧データ用ページ、旧パリテイ用ページ、新パリテイ用ページをクローズする。

【0115】

(第4の実施の形態による効果)

このように、本実施の形態では、パリテイ演算処理にかかるデータが共有メモリ装置7

50

3 内部でのみ処理されるので、パリティ演算処理にかかるデータの転送時間が短縮されることで、装置全体の性能を向上できるという効果が得られる。

【0116】

また、パリティ演算処理を、ディレクタ装置のプロセッサで行うかわりに、共有メモリ装置73のプロセッサ732で行うので、ディレクタ装置によるパリティ演算処理の負荷を軽減できることとなり、通信を行うことによるオーバーヘッドを減少させるという効果が得られる。

【0117】

なお、本実施の形態において、パリティ演算部736は、共有メモリ装置73内のデータコピー機能などを用いてもよいし、また、プロセッサ部732に同様の機能を持たせてもよい。

10

【0118】

(第5の実施の形態)

本発明の第5の実施の形態は、その基本的構成は上記の第2の実施の形態の通りであるが、ディスクディレクタ装置が追加され、また、共有メモリ装置が1つとなった構成となっている。

【0119】

図9は、本発明の第5の実施の形態によるディスクアレイ装置900の構成を示すブロック図である。

【0120】

図9を参照すると、ディスクアレイ装置900は、1つのホストディレクタ装置91と、複数のディスクディレクタ装置92A、92Bと、1つの共有メモリ装置93とから構成されている。

20

【0121】

(第5の実施の形態による効果)

本実施の形態は、第2の実施の形態と同様に、共有メモリ装置93上でのキャッシュ制御を、複数のディレクタ装置91、92A、92B上の各プロセッサ部913、923A、923Bにより行う代わりに、共有メモリ装置93上の単一のプロセッサ部932により行うので、プロセッサ部932がメモリ操作においてメモリバスを直接制御し、かつ、各プロセッサ部913、923A、923Bがプロセッサキャッシュを使用することができることから、キャッシュ制御に要する処理時間を短縮することができる。

30

【0122】

(第6の実施の形態)

本発明の第6の実施の形態は、その基本的構成は上記の第3の実施の形態の通りであるが、共有メモリ装置が追加され、また、ディレクタ装置が1つとなった構成となっている。

【0123】

図10は、本発明の第6の実施の形態によるディスクアレイ装置1000の構成を示すブロック図である。

【0124】

図10を参照すると、ディスクアレイ装置1000は、1つのディレクタ装置と、複数の共有メモリ装置とから構成されている。

40

【0125】

(第6の実施の形態による効果)

本実施の形態は、第3の実施の形態と同様に、複数の共有メモリ装置1003、1004上でのキャッシュ制御を、ディレクタ装置1001上のプロセッサ部1013により行う代わりに、各共有メモリ装置1003、1004上の単一のプロセッサ部1032、1042により行うので、各プロセッサ部1032、1042がメモリ操作においてメモリバスを直接制御し、かつ、プロセッサ部1013がプロセッサキャッシュを使用することができることから、キャッシュ制御に要する処理時間を短縮することができる。

50

## 【0126】

以上好ましい実施の形態をあげて本発明を説明したが、本発明は必ずしも、上記実施の形態に限定されるものでなく、その技術的思想の範囲内において様々に変形して実施することができる。

## 【産業上の利用可能性】

## 【0127】

情報処理システムにおいて必要とされるデータ容量が年々増加傾向にあり、パソコンから大型コンピュータに至るまで外部ストレージ装置を接続する傾向となっている。特に複数の情報処理システムでストレージを共有化することで別々にストレージを持つ場合の容量の無駄をなくするためのSANを構築する場合がある。この場合スイッチ装置と小型ストレージ装置を多数組み合わせる方式もあるが、バックアップソリューションなど高度なソリューションを実現するために大型ストレージを導入する場合もある。

## 【0128】

本発明は、単体でも多数のホスト接続ポートや多数のディスクドライブ、大容量のキャッシュメモリを搭載した大型ストレージ装置において、より性能を改善した大型ストレージ装置を提供するために適用することができる。

## 【図面の簡単な説明】

## 【0129】

【図1】本発明の第1の実施の形態によるディスクアレイ装置100の構成を示すブロック図である。

【図2】第1の実施の形態による共有メモリ装置のプロセッサ部、通信バッファ部及びコマンド制御部の詳細な構成を示すブロック図である。

【図3】第1の実施の形態によるディスクアレイ装置のリードライト動作を示すフローチャートである。

【図4】第1の実施の形態によるディレクタ装置と共有メモリ装置との通信の内容を時系列で示す図である。

【図5】第2の実施の形態によるディスクアレイ装置の構成を示すブロック図である。

【図6】第3の実施の形態によるディスクアレイ装置の構成を示すブロック図である。

【図7】第4の実施の形態による共有メモリ装置の構成を示すブロック図である。

【図8】第4の実施の形態によるディスクアレイ装置のライトバック処理を説明するフローチャートである。

【図9】第5の実施の形態によるディスクアレイ装置の構成を示すブロック図である。

【図10】第6の実施の形態によるディスクアレイ装置の構成を示すブロック図である。

【図11】従来のディスクアレイ装置の構成の一例を示すブロック図である。

## 【符号の説明】

## 【0130】

11、61、62、1001：ディレクタ装置

12、53、54、63、64、93、1003、1004：共有メモリ装置

51、91：ホストディレクタ装置

52、92A、92B：ディスクディレクタ装置

124-1：送信制御部

124-2：受信制御部

123-1：送信バッファ

123-2：受信バッファ

100、500、600、900、1000：ディスクアレイ装置

111：外部インタフェース制御部

112、312、322、613、623：データ転送制御部

113、313、323、614、624：制御メモリ

45、114、124、314、324、334、344、615、625、634、644、734：プロセッサ

10

20

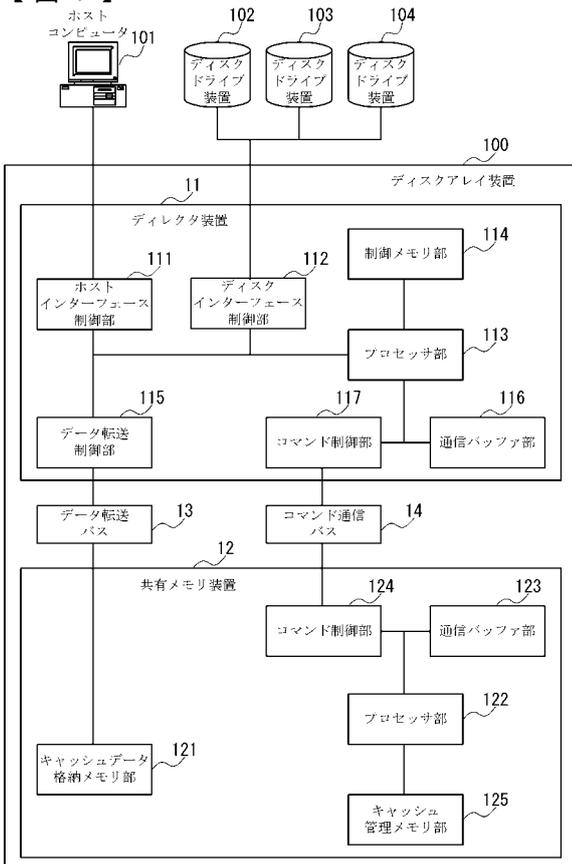
30

40

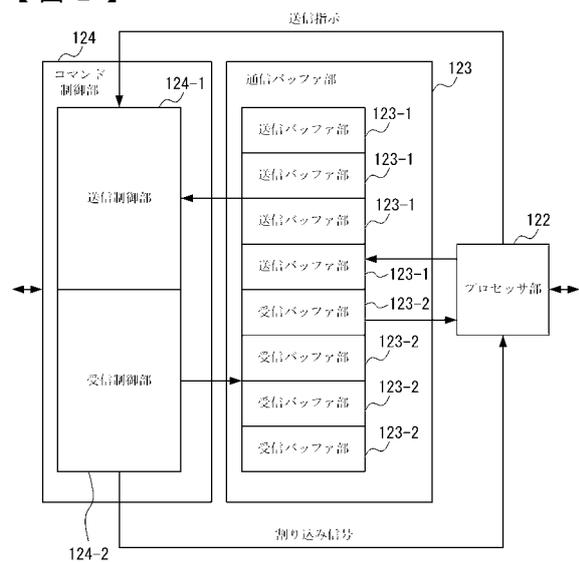
50

- 1 1 5、1 2 2、3 1 5、3 2 5、3 3 2、3 4 2、6 1 6、6 2 6、6 3 2、6 4 2、7 3 2 : コマンド制御部
- 1 1 6、1 2 3、3 1 6、3 2 6、3 3 3、3 4 3、6 1 7、6 2 7、6 3 3、6 4 3、7 3 3 : 通信バッファ
- 1 2 1、3 3 1、3 4 1、6 3 1、6 4 1、7 3 1 : キャッシュデータ格納メモリ
- 1 2 5、3 3 5、3 4 5、6 3 5、6 4 5、7 3 5 : キャッシュ管理メモリ
- 5 1 1、6 1 1、6 2 1、9 1 1 : ホストインタフェース制御部
- 5 2 1、6 1 2、6 2 2、9 2 2 A、9 2 2 B : ディスクインタフェース制御部
- 7 3 6 : パリティ演算部

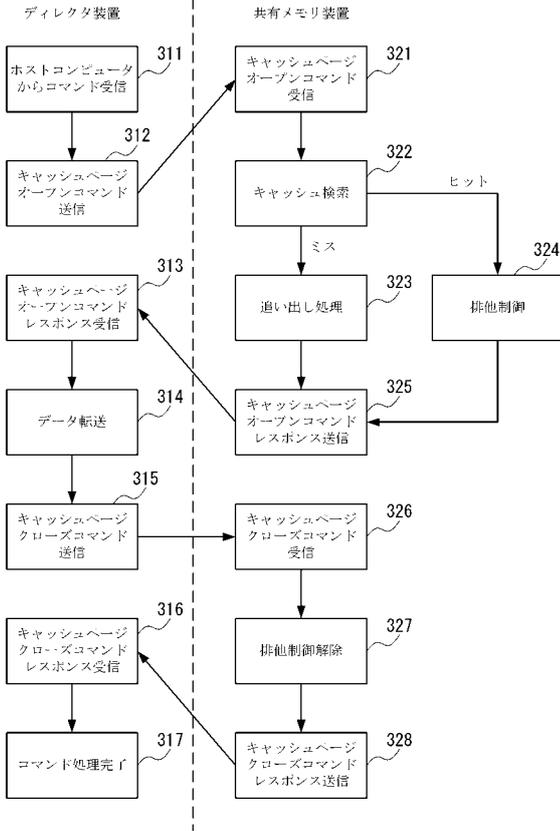
【図 1】



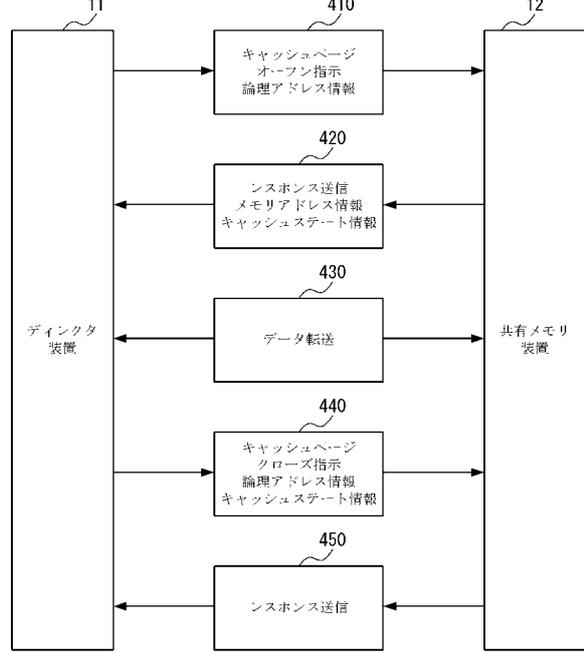
【図 2】



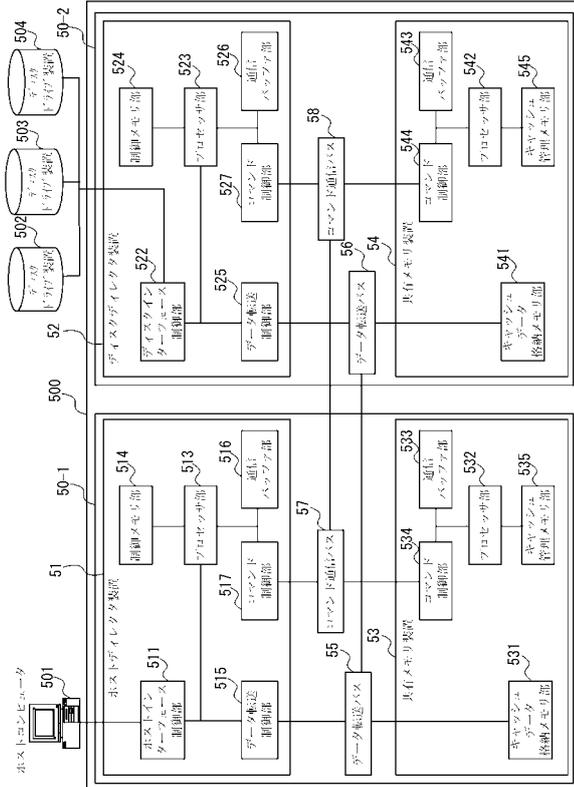
【図3】



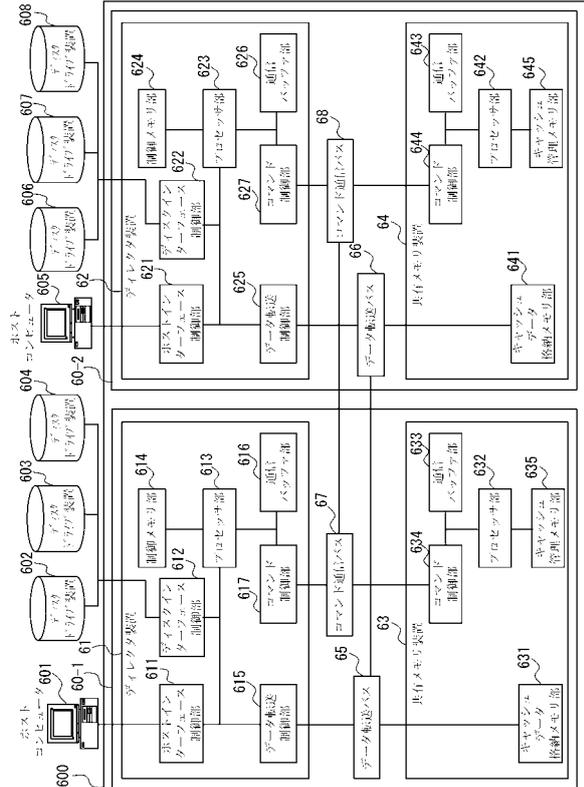
【図4】



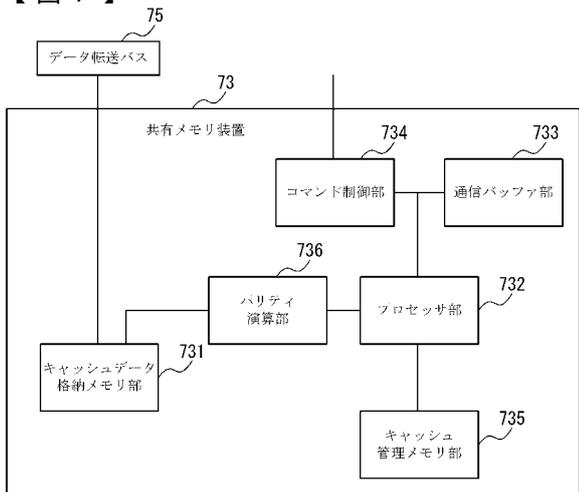
【図5】



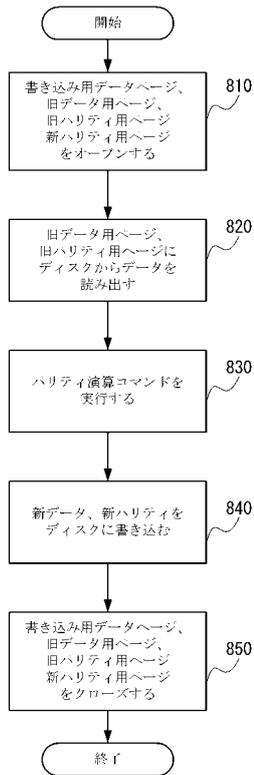
【図6】



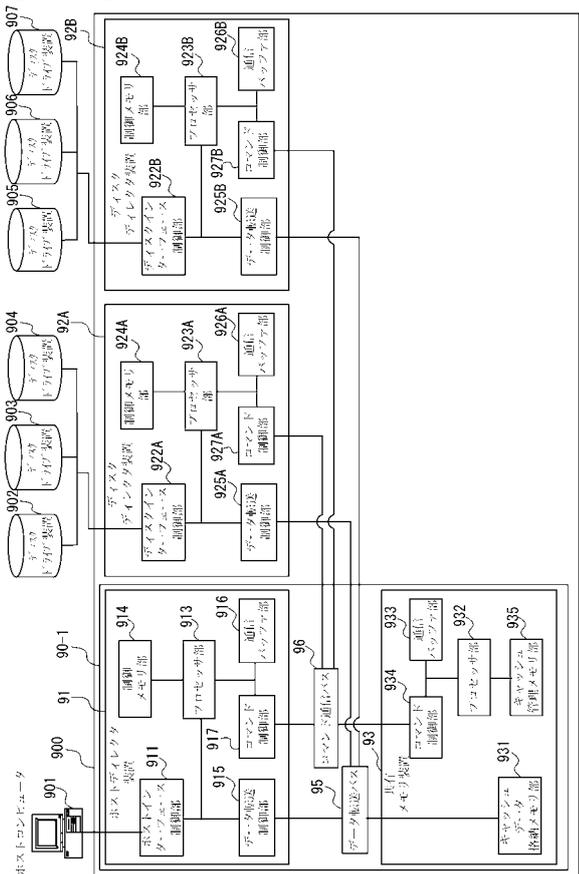
【 図 7 】



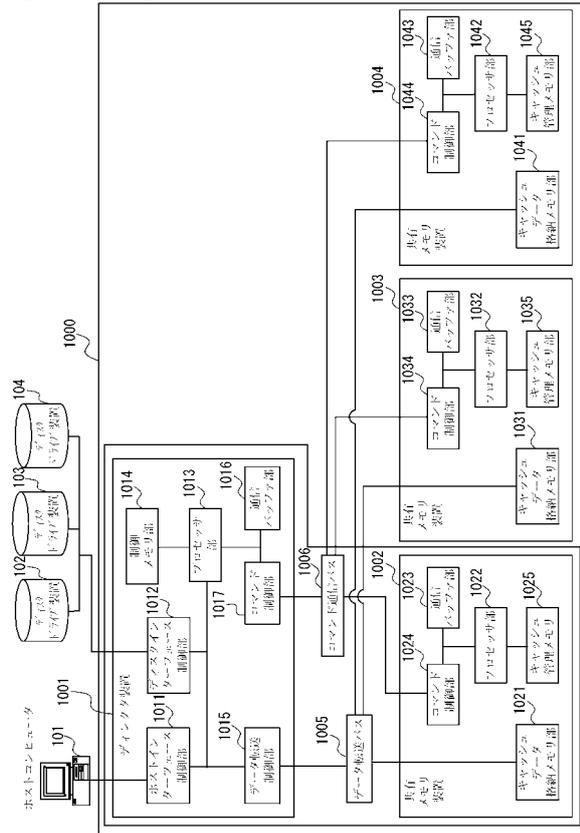
【 図 8 】



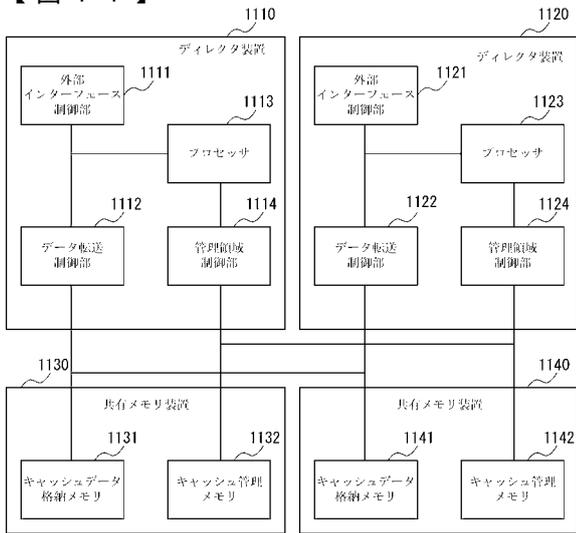
【 図 9 】



【 図 10 】



【図 11】



---

フロントページの続き

(51) Int.Cl.

F I

テーマコード(参考)

G 0 6 F 12/08 5 5 7

G 0 6 F 12/08 5 7 5