

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3684902号  
(P3684902)

(45) 発行日 平成17年8月17日(2005.8.17)

(24) 登録日 平成17年6月10日(2005.6.10)

(51) Int. Cl.<sup>7</sup>

G06F 3/06

F I

G06F 3/06 301G

G06F 3/06 540

請求項の数 12 (全 17 頁)

(21) 出願番号	特願平11-55079	(73) 特許権者	000005108
(22) 出願日	平成11年3月3日(1999.3.3)		株式会社日立製作所
(65) 公開番号	特開2000-250713 (P2000-250713A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成12年9月14日(2000.9.14)	(74) 代理人	100075096
審査請求日	平成15年1月15日(2003.1.15)		弁理士 作田 康夫
		(72) 発明者	藤本 和久
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所中央研究 所内
		審査官	木村 貴俊

最終頁に続く

(54) 【発明の名称】 ディスクアレイ制御装置

(57) 【特許請求の範囲】

【請求項1】

ホストコンピュータとのインターフェース部とを含むチャンネルインターフェース部と、  
ディスク装置とのインターフェース部を含むディスクインターフェース部と、  
前記ディスク装置に格納するデータの一部を一時的に格納するキャッシュメモリ部と、  
1つの前記チャンネルインターフェース部と1つの前記キャッシュメモリ部との間を1対

1つの前記ディスクインターフェース部と1つの前記キャッシュメモリ部との間を1対  
1に接続する第2のアクセスバスと、

前記第1のアクセスバスのアクセス回数と前記第2のアクセスバスのアクセス回数の比  
に基いて、前記第1のアクセスバスの駆動周波数及び前記第2のアクセスバスの駆動周波  
数を制御する制御手段とを有することを特徴とするディスクアレイ制御装置。

10

【請求項2】

前記制御手段は、複数の前記第1のアクセスバスのアクセス回数と複数の前記第2のア  
クセスバスのアクセス回数の比に基いて、各第1のアクセスバスの駆動周波数及び各第2  
のアクセスバスの駆動周波数を制御することを特徴とする請求項1に記載のディスクアレ  
イ制御装置。

【請求項3】

前記制御手段は、前記第1のアクセスバスの駆動周波数及び前記第2のアクセスバスの  
駆動周波数を設定したテーブルを有し、前記第1のアクセスバスのアクセス回数と前記第

20

2のアクセスパスのアクセス回数の比を計算し、前記第1のアクセスパスのアクセス回数と前記第2のアクセスパスのアクセス回数の比に応じた前記第1のアクセスパスの駆動周波数及び前記第2のアクセスパスの駆動周波数を前記テーブルから選択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制御装置。

【請求項4】

前記チャンネルインターフェース部は、前記第1のアクセスパスに対し、前記制御手段で制御された前記第1のアクセスパスの駆動周波数で、データを送出し、

前記ディスクインターフェース部は、前記第2のアクセスパスに対し、前記制御手段で制御された前記第2のアクセスパスの駆動周波数で、データを送出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制御装置。

10

【請求項5】

前記キャッシュメモリ部は、前記第1のアクセスパスに対し、前記制御手段で制御された前記第1のアクセスパスの駆動周波数で、データを送出し、

前記キャッシュメモリ部は、前記第2のアクセスパスに対し、前記制御手段で制御された前記第2のアクセスパスの駆動周波数で、データを送出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制御装置。

【請求項6】

ホストコンピュータとのインターフェース部とを含むチャンネルインタフェース部と、  
ディスク装置とのインターフェース部を含むディスクインタフェース部と、  
前記ディスク装置に格納するデータの一部を一時的に格納するキャッシュメモリ部と、  
1つの前記チャンネルインターフェース部と1つの前記キャッシュメモリ部との間を1対1に接続する第1のアクセスパスと、

20

1つの前記ディスクインターフェース部と1つの前記キャッシュメモリ部との間を1対1に接続する第2のアクセスパスと、

前記第1のアクセスパスのアクセス回数と前記第2のアクセスパスのアクセス回数の比に基いて、前記第1のアクセスパスの駆動周波数及び前記第2のアクセスパスの駆動周波数を制御する制御手段とを有し、

前記第1のアクセスパスの駆動周波数と前記第2のアクセスパスの駆動周波数とが異なることを特徴とするディスクアレイ制御装置。

30

【請求項7】

前記第2のアクセスパスの駆動周波数は、前記第1のアクセスパスの駆動周波数よりも高いことを特徴とする請求項6に記載のディスクアレイ制御装置。

【請求項8】

ホストコンピュータとのインターフェース部とを含むチャンネルインタフェース部と、  
ディスク装置とのインターフェース部を含むディスクインタフェース部と、  
前記ディスク装置に格納するデータの一部を一時的に格納するキャッシュメモリ部と、  
前記チャンネルインタフェース部及び前記ディスクインタフェース部と前記キャッシュメモリ部との間の接続を切り替えるセレクタ部と、

1つの前記チャンネルインターフェース部と1つの前記セレクタ部との間を1対1に接続する第1のアクセスパスと、

40

1つの前記ディスクインターフェース部と1つの前記セレクタ部との間を1対1に接続する第2のアクセスパスと、

1つの前記セレクタ部と1つの前記キャッシュメモリ部とを接続する第3のアクセスパスと、

前記第1のアクセスパスのアクセス回数と前記第2のアクセスパスのアクセス回数の比に基いて、前記第1のアクセスパスの駆動周波数及び前記第2のアクセスパスの駆動周波数を制御する制御手段とを有することを特徴とするディスクアレイ制御装置。

【請求項9】

前記キャッシュメモリ部は、複数の前記第3のアクセスパスのアクセス回数の比に基い

50

て、各第3のアクセスパスの動作周波数を制御する他の制御手段を有することを特徴とする請求項8に記載のディスクアレイ制御装置。

【請求項10】

前記チャンネルインターフェース部は、前記第1のアクセスパスに対し、前記制御手段で制御された前記第1のアクセスパスの駆動周波数で、データを送出し、

前記ディスクインターフェース部は、前記第2のアクセスパスに対し、前記制御手段で制御された前記第2のアクセスパスの駆動周波数で、データを送出し、

前記セクタ部は、各第3のアクセスパスに対し、前記キャッシュメモリ部の前記他の制御手段で制御された各第3のアクセスパスの駆動周波数で、データを送出することを特徴とする請求項9に記載のディスクアレイ制御装置。

10

【請求項11】

前記キャッシュメモリ部は、各第3のアクセスパスに対し、前記キャッシュメモリ部の前記他の制御手段で制御された各第3のアクセスパスの駆動周波数で、データを送出することを特徴とする請求項9又は10に記載のディスクアレイ制御装置。

【請求項12】

前記セクタ部は、前記第1のアクセスパスに対し、前記制御手段で制御された前記第1のアクセスパスの駆動周波数で、データを送出し、

前記セクタ部は、前記第2のアクセスパスに対し、前記制御手段で制御された前記第2のアクセスパスの駆動周波数で、データを送出することを特徴とする請求項8～11の何れかに記載のディスクアレイ制御装置。

20

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、データを分割して複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【0002】

【従来の技術】

コンピュータの主記憶のI/O性能に比べて、2次記憶装置として用いられる磁気ディスク装置を使ったサブシステムのI/O性能は3～4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上する努力が各所でなされている。サブシステムを構成し、データを分割して複数の磁気ディスク装置に格納する手段、いわゆるディスクアレイと呼ばれるシステムが知られている。

30

【0003】

例えば、従来技術では、図2に示すようにホストコンピュータ50とディスクアレイ制御装置2間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、磁気ディスク装置20のデータ及びディスクアレイ制御装置2に関する制御情報を格納する共有メモリ部15を備え、キャッシュメモリ部14および共有メモリ部15は全チャンネルIF部11及びディスクIF部12からアクセス可能な構成となっている。この従来技術では、チャンネルIF部11及びディスクIF部12と共有メモリ部15、またチャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間は1：1に接続されている。

40

【0004】

また、他の従来技術では、図3に示すようにホストコンピュータ50とディスクアレイ制御装置3間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置3間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、磁気ディスク装置20のデータ及びディスクアレイ制御装置3に関する制御情報を格納する共有メモリ部15を備え、各チャンネルIF部11及びディスクIF部12と共有メモリ部15間は

50

共有バス130で接続され、各チャンネルIF11部及びディスクIF部12とキャッシュメモリ部14間は共有バス131により接続されている。

【0005】

【発明が解決しようとする課題】

ディスクアレイシステムに対する高性能化要求に対して、これまでは、ディスクアレイ制御装置の大規模化や構成要素の高速化、例えばチャンネルIF部内のプロセッサ数やキャッシュメモリ容量の増大、高性能プロセッサの適用、内部バス幅の拡大やバス転送能力の向上などで対応してきた。しかしながら、図3に示した従来技術では、内部バスの転送能力がシステムの大規模化および性能向上に追従するのが困難になりつつある。

【0006】

そこで内部バス性能を向上させて高いメモリアクセス性能を得るために、図2に示した従来技術のように、チャンネルIF部及びディスクIF部と共有メモリ及びキャッシュメモリ間を1:1に接続する方法が考えられる。この方法によれば、メモリに接続したアクセスバス数に比例して内部バス性能が増加する。しかしながら、搭載するプロセッサ数の増大に比例して、共有メモリおよびキャッシュメモリへのアクセス要求も増加する。そのため、内部バス性能を最大限に引き出すために、各チャンネルIF部、ディスクIF部-共有メモリ、キャッシュメモリ間のアクセスを効率的に制御する必要がある。

【0007】

本発明の目的は上述の課題を解消し、チャンネルIF部、ディスクIF部-キャッシュメモリ間のアクセスバスを効率的に使用し、キャッシュメモリへのデータ転送スループットの高いディスクアレイ制御装置を提供することにある。

【0008】

【課題を解決するための手段】

上記目的は、ホストコンピュータとのインターフェース部とを含むチャンネルインターフェース部と、ディスク装置とのインターフェース部を含むディスクインターフェース部と、前記ディスク装置に格納するデータの一部を一時的に格納するキャッシュメモリ部と、1つの前記チャンネルインターフェース部と1つの前記キャッシュメモリ部との間を1対1に接続する第1のアクセスバスと、1つの前記ディスクインターフェース部と1つの前記キャッシュメモリ部との間を1対1に接続する第2のアクセスバスと、前記第1のアクセスバスのアクセス回数と前記第2のアクセスバスのアクセス回数の比に基いて、前記第1の

【0020】

また、上記目的は、ホストコンピュータとのインターフェース部とを含むチャンネルインターフェース部と、ディスク装置とのインターフェース部を含むディスクインターフェース部と、前記ディスク装置に格納するデータの一部を一時的に格納するキャッシュメモリ部と、前記チャンネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間の接続を切り替えるセレクト部と、1つの前記チャンネルインターフェース部と1つの前記セレクト部との間を1対1に接続する第1のアクセスバスと、1つの前記ディスクインターフェース部と1つの前記セレクト部との間を1対1に接続する第2の

【0026】

【発明の実施の形態】

以下、本発明の実施例を、図面を用いて説明する。

【0027】

(実施例1)

10

20

30

40

50

図 1 に、本発明の一実施例を示す。

【 0 0 2 8 】

ディスクアレイ制御装置 1 は、ホストコンピュータ 5 0 との 2 つのインターフェース部 ( チャネル I F 部 ) 1 1、磁気ディスク装置 2 0 との 2 つのインターフェース部 ( ディスク I F 部 ) 1 2、2 つのキャッシュメモリ部 1 4 と、8 本のアクセスパス 0 : 1 3 5 とを有する。

【 0 0 2 9 】

チャンネル I F 部 1 1 は、ホストコンピュータ 5 0 との 2 つの I F ( ホスト I F ) 1 0 2、2 つのマイクロプロセッサ 1 0 1、キャッシュメモリ部 1 4 への 1 つのアクセス制御部 ( C M アクセス制御部 ) 1 0 4 を有し、ホストコンピュータ 5 0 とキャッシュメモリ部 1 4 間のデータ転送を実行する。マイクロプロセッサ 1 0 1 及びホスト I F 1 0 2 は内部バス 1 1 0 によって接続され、C M アクセス制御部 1 0 4 は 2 つのホスト I F 1 0 2 へ直接接続されている。

10

【 0 0 3 0 】

ディスク I F 部 1 2 は、磁気ディスク装置 2 0 との 2 つの I F ( ドライブ I F ) 1 0 3、2 つのマイクロプロセッサ 1 0 1、キャッシュメモリ部 1 4 への 1 つのアクセス制御部 ( C M アクセス制御部 ) 1 0 4 を有し、磁気ディスク装置 2 0 とキャッシュメモリ部 1 4 間のデータ転送を実行する。マイクロプロセッサ 1 0 1 及びドライブ I F 1 0 3 は内部バス 1 1 1 によって接続され、C M アクセス制御部 1 0 4 は 2 つのドライブ I F 1 0 2 へ直接接続されている。

20

【 0 0 3 1 】

キャッシュメモリ部 1 4 は、メモリコントローラ 1 0 5 とメモリモジュール 1 0 6 を有し、磁気ディスク装置 2 0 へ記録するデータや、そのデータ及びシステム内の管理情報等を格納する。

【 0 0 3 2 】

1 つのチャンネル I F 部 1 1 と 1 つのキャッシュメモリ部 1 4 間、及び 1 つのディスク I F 部 1 2 と 1 つのキャッシュメモリ部 1 4 間は 1 本のアクセスパス 0 : 1 3 5 で接続する。ここで、各アクセスパス 0 : 1 3 5 の幅は同じである。また、上記個数は一実施例に過ぎず、個数を上記に限定するものではない。

【 0 0 3 3 】

図 4 は、C M アクセス制御部 1 0 4 内の構成を示している。C M アクセス制御部 1 0 4 は、セクタ 3 0 2、アドレス、コマンド、データを一時格納するパケットバッファ 3 0 3、キャッシュメモリ部 1 4 に繋がるアクセスパス 0 : 1 3 5 とのパス I F 3 0 1、データのエラーチェック部 3 0 0、データ転送制御部 3 1 0、データ転送時にデータと並送する同期信号の発生 / 受信部 3 2 0 を有する。セクタ 3 0 2 の 2 つのポートはデータ線 2 1 0 でホスト I F 1 0 2 あるいはドライブ I F 1 0 3 に接続される。また、セクタ 3 0 2 の他の 2 つのポートはパス I F 3 0 1 に接続される。パス I F 3 0 1 は、アクセスパス 0 : 1 3 5 により、キャッシュメモリ部 1 4 と接続される。データ転送制御部 3 1 0 は、制御線 1 : 2 1 1 により、ホスト I F 1 0 2 又はドライブ I F 1 0 3、及びキャッシュメモリ部 1 4 のメモリコントローラ 1 0 5 内のデータ転送制御部 3 1 5 と接続される。データ転送制御部 3 1 0 は、セクタ 3 0 2 の切り替えを行う。同期信号発生 / 受信部 3 2 0 は、同期信号線 2 2 0 により、メモリコントローラ 1 0 5 内の同期信号発生 / 受信部 3 2 0 と接続される。データの送信時には、データ転送制御部 3 1 0 から指示された駆動周波数で、パス I F 3 0 1 と同期してデータと同期信号を並送する。またデータの受信時には、データと並送されてきた同期信号を使用してデータを取り込む。

30

40

【 0 0 3 4 】

図 6 は、C M アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 内の構成を示している。データ転送制御部 3 1 0 は、アービタ 3 1 1、2 つの制御信号発生 / 受信部 3 1 2、駆動周波数選択部 3 1 3 を有する。1 つの制御信号発生 / 受信部 3 1 2 は、制御線 1 : 2 1 1 により、ホスト I F 1 0 2 またはドライブ I F 1 0 3 と接続され、アービトレーションに

50

必要な情報をアービタ311に送る。アービタ311は、アービトレーションを行い、セクタ切り替え信号をセクタ302へ送出する。もう一つの制御信号発生/受信部312は、制御線2:212により、メモリコントローラ105と接続される。また、アービタ311からセクタ切り替え完了の信号を受け、制御線2:212によってアクセス開始信号を送出する。さらに、制御線2:212によって、メモリコントローラ105から送られてきたアクセスパス0:135の駆動周波数変更を指示する信号を駆動周波数選択部313へ送出する。駆動周波数選択部313は、同期信号発生/受信部320が設定可能な駆動周波数のテーブルを持っており、駆動周波数変更を指示する信号によりテーブルから駆動周波数の値を選択し、その値を同期信号発生/受信部320へ送る。

#### 【0035】

図5は、キャッシュメモリ部14内の構成を示している。キャッシュメモリ部14は、メモリコントローラ105とメモリモジュール106を有する。メモリコントローラ105は、チャンネルIF部11及びディスクIF部12に繋がるアクセスパス0:135との4つのパスIF301、セクタ304、アドレス、コマンド、データを一時格納するパケットバッファ303、データのエラーチェック部300、メモリモジュール106へのアクセスを制御するメモリ制御部307、CMアクセス制御部104から送出されたアドレス及びコマンドを解析するアドレス・コマンド(addr, cmd)解析部305、データ転送制御部315、データ転送時にデータと並送する同期信号の発生/受信部320を有する。データ転送制御部315は、制御線2:212でCMアクセス制御部104内のデータ転送制御部310に接続される。また、データ転送制御部315は、セクタ304の切り替えを行う。また、同期信号発生/受信部320は、同期信号線220でCMアクセス制御部104内の同期信号発生/受信部320に接続される。データの送信時には、データ転送制御部315から指示された駆動周波数で、パスIF301と同期してデータと同期信号を並送する。またデータの受信時には、データと並送されてきた同期信号を使用してデータを取り込む。

#### 【0036】

図8は、メモリコントローラ105内のアドレス(addr)、コマンド(cmd)解析部305の構成を示している。addr、cmd解析部305は、バッファ351、アドレス(addr)抽出部352、コマンド(cmd)抽出部353、カウンタ354を有する。addr、cmd解析部305では、メモリコントローラ105に接続される4本のアクセスパス0:135それぞれに1つずつ割り当てられたバッファ351に、アドレス、コマンドを格納する。addr抽出部352及びcmd抽出部353では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、バッファ351にアドレス及びコマンドを格納する毎に、バッファ毎に割り当てたカウンタをカウントアップする。

#### 【0037】

図7は、メモリコントローラ105のデータ転送制御部315内の構成を示している。データ転送制御部315は、アービタ311、2つの制御信号発生/受信部312、駆動周波数演算部314、駆動周波数選択部313を有する。1つの制御信号発生/受信部312は、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310と接続され、CMアクセス制御部104からのアクセス要求をアービタ311に伝える。アービタ311は、アービトレーションを行い、セクタ切り替え信号をセクタ304へ送出する。もう一つの制御信号発生/受信部312は、制御線4:214により、メモリ制御部307と接続され、アービタ311からセクタ切り替え完了の信号を受け、制御線4:214によってメモリ制御部307へメモリアクセス開始信号を送出する。駆動周波数演算部314は、addr、cmd解析部305内のカウンタのカウント数を読み出し、4本のアクセスパス0:135のアクセス回数の比を計算し、駆動周波数選択部313へ送る。駆動周波数選択部313は、同期信号発生/受信部320が設定可能な駆動周波数のテーブルを持っており、そのテーブルを参照して各アクセスパスの駆動周波数の比がアクセス回数の比に近くなるように各アクセスパスの駆動周波数を選択し、その駆動周

10

20

30

40

50

波数の値を同期信号発生/受信部320へ送る。また駆動周波数選択部313は、制御信号発生/受信部312、制御線2:212を介して、アクセスパス0:135の駆動周波数変更を指示する信号をデータ転送制御部310内の駆動周波数選択部313へ送出する。

**【0038】**

チャンネルIF部11からキャッシュメモリ部14へのアクセス回数及びディスクIF部12からキャッシュメモリ部14へのアクセス回数は、ホストコンピュータからのアクセスの種類によって変化する。このため、アクセス回数の多いアクセスパスのスループットは下がり、それがネックとなって装置全体のスループットが制限される。上記のようにアクセス回数の比をもとにアクセスパス0:135毎にデータ転送速度を変えることにより、アクセス回数の多いアクセスパスのスループットが高まる。これにより、アクセス回数の多いアクセスパスのスループットのネックが解消され、装置全体のスループットの向上が可能となる。

10

**【0039】**

また、アクセスパスの物理的な駆動周波数は固定しておき、データ転送制御部315のアービタ311において、アクセス回数の多い順にアクセスパスに優先度を設定してアービトレーションを行うことにより、アクセス回数の多いアクセスパスの実効的なデータ転送速度を上げることが可能となる。これによっても、アクセス回数の多いアクセスパスのスループットを高めることができ、装置全体のスループットの向上が可能となる。

**【0040】**

通常アクセス回数の差は、チャンネルIF部11とキャッシュメモリ部14間のアクセスパスと、ディスクIF部12とキャッシュメモリ部14間のアクセスパスの間で生じる。したがって、チャンネルIF部11とキャッシュメモリ部14間のアクセスパスと、ディスクIF部12とキャッシュメモリ部14間のアクセスパスの間で、データ転送速度を等しくしたり、異ならせたりすることにより、アクセスパス毎にデータ転送速度を制御する場合に比べて、データ転送速度の制御が容易になる。

20

**【0041】**

本実施例では、チャンネルIF部11あるいはディスクIF部12に繋がるアクセスパス0:135とのパスIF301の数を4つとした。上記個数は一実施例に過ぎず、個数を上記に限定するものではない。

30

**【0042】**

次に、キャッシュメモリ部14へのアクセス時の手順について述べる。キャッシュメモリ部14へアクセスする場合、マイクロプロセッサ101は、ホストIF102あるいはドライブIF103へ、キャッシュメモリ部14へのアクセス開始を指示する。

**【0043】**

アクセス開始の指示を受けたホストIF102あるいはドライブIF103は、制御線1:211によりCMアクセス制御部104内のデータ転送制御部310へアクセス開始を示す信号と、アクセス先のメモリコントローラ105を示す信号を送出する。それとともに、データ線210を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

40

**【0044】**

CMアクセス制御部104は、データ線210を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってパスIF301の使用権を決定し、セクタ302を切り替えるとともに、制御線2:212によってメモリコントローラ105内のデータ転送制御部315へアクセス開始を示す信号と、アクセス先のメモリコントローラ105を示す信号を出す。

**【0045】**

メモリコントローラ105内のデータ転送制御部315は、制御線2:212によってアクセス開始の信号と、アクセス先のメモリコントローラ105を示す信号を受けると、C

50

Mアクセス制御部104へ、制御線2:212によってアクセス承認を返す。

【0046】

CMアクセス制御部104はアクセス承認を受けると、パケットバッファ303からアドレス、コマンド、データ(データの書き込み時のみ)を読み出し、セレクタ302、パシフ301を介してアクセスバス0:135へ送出する。

【0047】

メモリコントローラ105は、アクセスバス0:135を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をパケットバッファ303に格納する。また、送られてきたアドレス、コマンドは、アドレス(addr)、コマンド(cmd)解析部305へ取り込み、解析の結果得られたメモリの制御情報をメモリ制御部307へ送出する。また、アクセス先メモリコントローラ105を調べ、自メモリコントローラと異なる場合はエラーとみなし障害処理を行う。

10

【0048】

メモリコントローラ105内のデータ転送制御部315は、アービトレーションを行ってメモリアクセス権を決定し、セレクタ304を切り替える。

【0049】

データの書き込み時は、パケットバッファ303からデータを読み出し、セレクタ304を介してメモリモジュール106へ書き込む。データの読み出し時は、メモリモジュール106からデータを読み出し、セレクタ304、パシフ301を介してアクセスバス0:135へ送出する。

20

【0050】

メモリモジュール106へのアクセスが終了すると、データ転送制御部315においてアクセス状況を示すステータスを生成し、CMアクセス制御部104を経由して、ホストIF102またはドライブIF103へ送出する。

【0051】

また図10に示すように、1つのCMアクセス制御部104からキャッシュメモリ部14へ接続されるアクセスバス0:135の本数が多いディスクアレイ制御装置1においても、本実施例を実施する上で問題はない。ホストIF及びドライブIFとして、今後はファイバチャネル等のスループットが100MB/s以上の高速IFが使用されることが多くなると考えられる。アクセスバス0:135のスループットは、ホストIF及びドライブIFのスループットと同等にする必要がある。このため上記のように高速IFを使用する場合、1つのCMアクセス制御部104からキャッシュメモリ部14へ接続されるアクセスバス0:135の本数が多い図10の構成が有効となる。

30

【0052】

(実施例2)

図11に、本発明の他の実施例を示す。

【0053】

ディスクアレイ制御装置1は、チャンネルIF部11、ディスクIF部12、セレクタ部13、キャッシュメモリ部14と、アクセスバス0:135、アクセスバス1:136を有する。

40

【0054】

チャンネルIF部11、ディスクIF部12、及びキャッシュメモリ部14は、図1に示す実施例1の構成と同様である。

【0055】

セレクタ部13には、2つのチャンネルIF部11、2つのディスクIF部12からそれぞれ2本ずつ、計8本のアクセスバス0:135が接続される。また、セレクタ部13には、2つのキャッシュメモリ部14へのアクセスバス1:136が2本ずつ、計4本接続される。

【0056】

アクセスバス0:135とアクセスバス1:136の間には上記のようなバス数の関係が

50



あるため、セレクト部 1 3 ではチャンネル I F 部 1 1 及びディスク I F 部 1 2 からの 8 本のアクセスパス 0 : 1 3 5 からの要求の内、キャッシュメモリ部 1 4 へのアクセスパス 1 : 1 3 6 の数に相当する 4 個だけを選択して実行する機能を持つ。

【 0 0 5 7 】

ここで、上記個数は一実施例に過ぎず、個数を上記に限定するものではない。1つのセレクト部 1 3 からキャッシュメモリ部 1 4 へ接続されるアクセスパスの数を、チャンネル I F 部 1 1 及びディスク I F 部 1 2 から 1 つのセレクト部 1 3 に接続されるアクセスパスの数より少なくし、チャンネル I F 部 1 1 とディスク I F 部 1 2 の合計数よりもセレクト部 1 3 の数が少なくなるように上記個数を設定すると、キャッシュメモリ部 1 4 へ接続されるアクセスパス数を削減することができる。キャッシュメモリ部 1 4 の L S I ピンネック及びパッケージのコネクタネックという問題が生じた場合、上記のようにすることで、L S I のピンネック及びパッケージのコネクタネックを解消することができる。

10

【 0 0 5 8 】

C M アクセス制御部 1 0 4 は、図 4 に示す構成と同様であるが、制御線 2 : 2 1 2 は、セレクト部 1 3 内のデータ転送制御部 3 1 5 に接続される。

【 0 0 5 9 】

図 9 は、セレクト部 1 3 内の構成を示している。セレクト部 1 3 は、チャンネル I F 部 1 1 及びディスク I F 部 1 2 に繋がるアクセスパス 0 : 1 3 5 との 8 つのパス I F 3 0 1、メモリコントローラ 1 0 5 に繋がるアクセスパス 1 : 1 3 6 との 4 つのパス I F 3 0 1、両者間を互いに接続するセレクト部 3 0 6、パケットバッファ 3 0 3、データのエラーチェック部 3 0 0、C M アクセス制御部 1 0 4 から送出されたアドレス及びコマンドを解析するアドレス ( a d r )、コマンド ( c m d ) 解析部 3 0 5、データ転送制御部 3 1 5、データ転送時にデータと並送する同期信号の発生 / 受信部 3 2 0 を有する。データ転送制御部 3 1 5 は、制御線 2 : 2 1 2 で C M アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 に接続され、制御線 3 : 2 1 3 でメモリコントローラ 1 0 5 内のデータ転送制御部 3 1 5 に接続される。また、データ転送制御部 3 1 5 は、セレクト部 3 0 6 の切り替えを行う。パケットバッファ 3 0 3 は、アクセスパス 0 : 1 3 5 側のパスの駆動周波数とアクセスパス 1 : 1 3 6 側の駆動周波数が異なる場合、周波数の差を吸収するために、転送するデータの一部または全部をバッファリングする。アクセスパス 0 : 1 3 5 側の同期信号発生 / 受信部 3 2 0 は、同期信号線 2 2 0 で C M アクセス制御部 1 0 4 内の同期信号発生 / 受信部 3 2 0 に接続され、アクセスパス 1 : 1 3 6 側の同期信号発生 / 受信部 3 2 0 は、同期信号線 2 2 0 でメモリコントローラ 1 0 5 内の同期信号発生 / 受信部 3 2 0 に接続される。データの送信時には、データ転送制御部 3 1 5 から指示された駆動周波数で、パス I F 3 0 1 と同期してデータと同期信号を並送する。またデータの受信時には、データと並送されてきた同期信号を使用してデータを取り込む。

20

30

【 0 0 6 0 】

a d r、c m d 解析部 3 0 5 は、図 8 に示す構成と同様に、バッファ 3 5 1、アドレス ( a d r ) 抽出部 3 5 2、コマンド ( c m d ) 抽出部 3 5 3、カウンタ 3 5 4 を有する。a d r、c m d 解析部 3 0 5 では、メモリコントローラ 1 0 5 に接続される 8 本のアクセスパス 0 : 1 3 5 それぞれに 1 つずつ割り当てられたバッファ 3 5 1 に、アドレス、コマンドを格納する。a d r 抽出部 3 5 2 及び c m d 抽出部 3 5 3 では、アクセスするメモリコントローラ 1 0 5 とアクセスの種類を割り出し、データ転送制御部 3 1 5 内のアービタ 3 1 1 へ送出する。また、バッファ 3 5 1 にアドレス及びコマンドを格納する毎に、バッファ毎に割り当てたカウンタをカウントアップする。

40

【 0 0 6 1 】

データ転送制御部 3 1 5 は、図 7 に示す構成と同様に、アービタ 3 1 1、2 つの制御信号発生 / 受信部 3 1 2、駆動周波数演算部 3 1 4、駆動周波数選択部 3 1 3 を有する。1 つの制御信号発生 / 受信部 3 1 2 は、制御線 2 : 2 1 2 で C M アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 に接続される。また、C M アクセス制御部 1 0 4 からのアクセス要求をアービタ 3 1 1 に伝える。アービタ 3 1 1 は、a d r、c m d 解析部 3 0 5 からの情

50

報に基づいてアービトレーションを行い、セクタ切り替え信号をセクタ306へ送出する。もう一つの制御信号発生/受信部312は、図7とは異なり、制御線3:213でメモリコントローラ105内のデータ転送制御部315に接続される。また、アービタ311からセクタ切り替え完了の信号を受け、制御線3:213によってメモリコントローラ105内のデータ転送制御部315にアクセス開始信号を送出する。駆動周波数演算部314は、*adr*、*cmd*解析部305内のカウンタのカウンタ数を読み出し、8本のアクセスパス0:135のアクセス回数の比を計算し、駆動周波数選択部313へ送る。駆動周波数選択部313は、同期信号発生/受信部320が設定可能な駆動周波数のテーブルを持っており、そのテーブルを参照して各アクセスパスの駆動周波数の比がアクセス回数の比に近くなるように各アクセスパスの駆動周波数を選択し、その駆動周波数の値をアクセスパス0:135側の同期信号発生/受信部320へ送る。また駆動周波数選択部313は、アクセスパス0:135の制御信号発生/受信部312、制御線2:212を介して、アクセスパス0:135の駆動周波数変更を指示する信号をデータ転送制御部310内の駆動周波数選択部313へ送出する。また、駆動周波数選択部313は、アクセスパス1:136側の制御信号発生/受信部312を介してメモリコントローラ105から駆動周波数変更を示す信号を受け、設定可能な駆動周波数のテーブルから駆動周波数の値を選択し、その値をアクセスパス1:136側の同期信号発生/受信部320へ送る。

10

**【0062】**

チャンネルIF部11からキャッシュメモリ部14へのアクセス回数及びディスクIF部12からキャッシュメモリ部14へのアクセス回数は、ホストコンピュータからのアクセスの種類によって変化する。このため、アクセス回数の多いアクセスパスのスループットは下がり、それがネックとなって装置全体のスループットが制限される。上記のようにアクセス回数の比をもとにアクセスパス0:135毎にデータ転送速度を変えることにより、アクセス回数の多いアクセスパスのスループットが高まる。これにより、アクセス回数の多いアクセスパスのスループットのネックが解消され、装置全体のスループットの向上が可能となる。

20

**【0063】**

メモリコントローラ105は、図5に示す構成と同様で、その中のデータ転送制御部315は、図7に示す構成と同様で、実施例1と同様の処理を行う。

**【0064】**

セクタ部13とキャッシュメモリ部14間の各アクセスパス1:136それぞれのアクセス回数が、各キャッシュメモリ部へのアクセスの偏りによって変化する場合、上記のようにアクセスパス1:136毎にデータ転送速度を可変にし、アクセス回数の多いアクセスパスのスループットを高めることによって、装置全体のスループットの向上が可能となる。図11に示すディスクアレイ制御装置1では、この制御を、上記のチャンネルIF部11及びディスクIF部12とセクタ部13間での制御と合わせて行うことにより、どちらか一方だけ制御を行う場合より、装置全体のスループットの向上が可能となる。

30

**【0065】**

また、アクセスパスの物理的な駆動周波数は固定しておき、データ転送制御部315のアービタ311において、アクセス回数の多い順にアクセスパスに優先度を設定してアービトレーションを行うことにより、アクセス回数の多いアクセスパスの実効的なデータ転送速度を上げることが可能となる。これによっても、アクセス回数の多いアクセスパスのスループットを高めることができ、装置全体のスループットの向上が可能となる。

40

**【0066】**

図12に示すように、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間をスイッチ(SW)16を用いた相互結合網110で接続するディスクアレイ制御装置1では、SW16内に図9に示すセクタ部13内のデータ転送制御部315及び同期信号発生/受信部320の機能を持たせることにより、各アクセスパス毎に駆動周波数を可変にすることが可能となる。それによって、図12に示すディスクアレイ制御装置1においても、装置全体のデータ転送スループットを向上することが可能となる。

50

## 【0067】

(実施例3)

実施例1または2において、ディスクIF部12とキャッシュメモリ部14間のデータ転送速度をチャンネルIF部11とキャッシュメモリ部14間のデータ転送速度より大きくする。また、アクセスパスのデータ転送速度は可変とせず、固定する。

## 【0068】

ディスクアレイ制御装置では、一般的にディスクIF部12からキャッシュメモリ部14へのアクセス回数は、チャンネルIF部11からキャッシュメモリ部14へのアクセス回数以上となる。ディスクアレイではデータに冗長性を持たせるため、データの書き込み時にパリティを生成して、データとともに磁気ディスク装置に記録する。したがって、データの書き込み時には、ディスクIF部12とキャッシュメモリ部14間でパリティ生成のためのアクセスが発生する。これは、一般的にライトペナルティと呼ばれている。このため、ディスクIF部12からキャッシュメモリ部14へのアクセス回数は、チャンネルIF部11からキャッシュメモリ部14へのアクセス回数以上となる。

10

## 【0069】

通常、ディスクIF部12とキャッシュメモリ部14間のアクセス回数とチャンネルIF部11とキャッシュメモリ部14間のアクセス回数の比が最大で約3:1となるため、本実施例では、ディスクIF部12とキャッシュメモリ部14間のデータ転送速度がチャンネルIF部11とキャッシュメモリ部14間のデータ転送速度の3倍になるように設定する。

## 【0070】

上記のように、アクセス回数の多いアクセスパスのデータ転送速度をアクセス回数の少ないアクセスパスのデータ転送速度より大きく設定しておくことにより、アクセス回数の多いアクセスパスのスループットが装置全体のスループットを制限することがなくなり、装置全体のスループットの向上が可能となる。

20

## 【0071】

(実施例4)

図13に、本発明の他の実施例を示す。

## 【0072】

ディスクアレイ制御装置1は、ホストコンピュータ50との2つのインターフェース部(チャンネルIF部)11、磁気ディスク装置20との2つのインターフェース部(ディスクIF部)12、2つのキャッシュメモリ部14と、12本のアクセスパス0:135を有する。

30

## 【0073】

チャンネルIF部11は、図1に示す実施例1の構成と同様である。

## 【0074】

ディスクIF部12は、CMアクセス制御部104からキャッシュメモリ部14へのアクセスパス0:135が4本に増えている以外は、図1に示す実施例1の構成と同様である。

## 【0075】

キャッシュメモリ部14は、メモリコントローラ105に接続されるアクセスパス0:135の本数が6本に増えている以外は、図1に示す実施例の構成と同様である。

40

## 【0076】

ここで、上記個数は一実施例に過ぎず、個数を上記に限定するものではない。

## 【0077】

本実施例の特徴は、ディスクIF部12とキャッシュメモリ部14間のアクセスパス0:135の本数をチャンネルIF部11とキャッシュメモリ部14間のアクセスパス0:135の本数より多くするという点である。

## 【0078】

ディスクアレイ制御装置では、一般的にディスクIF部12からキャッシュメモリ部14へのアクセス回数は、チャンネルIF部11からキャッシュメモリ部14へのアクセス回数

50

以上となる。ディスクアレイではデータに冗長性を持たせるため、データの書き込み時にパリティを生成して、データとともに磁気ディスク装置に記録する。したがって、データの書き込み時には、ディスクIF部12とキャッシュメモリ部14間でパリティ生成のためのアクセスが発生する。これは、一般的にライトペナルティと呼ばれている。このため、ディスクIF部12からキャッシュメモリ部14へのアクセス回数は、チャンネルIF部11からキャッシュメモリ部14へのアクセス回数以上となる。

#### 【0079】

実施例1で説明した方法により、アクセスパスのデータ転送速度を可変にできない場合には、アクセス回数の多いアクセスパスの本数をアクセス回数の少ないアクセスパスの本数より多くすることにより、アクセス回数の多いアクセスパスのスループットを高くすることができる。これにより、アクセス回数の多いアクセスパスのスループットのネックが解消され、装置全体のスループットの向上が可能となる。

10

#### 【0080】

また図14に示すように、1つのCMアクセス制御部104からキャッシュメモリ部14へ接続されるアクセスパス0:135の本数が多いディスクアレイ制御装置1においても、本実施例を実施する上で問題はない。ホストIF及びドライブIFとして、今後はファイバーチャネル等のスループットが100MB/s以上の高速IFが使用されることが多くなると考えられる。アクセスパス0:135のスループットをホストIF及びドライブIFのスループットと同等にする必要がある。このため上記のように高速IFを使用する場合、1つのCMアクセス制御部104からキャッシュメモリ部14へ接続されるアクセスパス0:135の本数が多い図14の構成が有効となる。

20

#### 【0081】

##### 【発明の効果】

本発明によれば、アクセス回数の多いアクセスパスのスループットによって、ディスクアレイ装置全体のスループットが制限されることを防ぐことが可能となる。それによって、ディスクアレイ制御装置のデータ転送スループットが向上する。

##### 【図面の簡単な説明】

【図1】本発明によるディスクアレイ制御装置の構成を示す図。

【図2】従来のディスクアレイ制御装置の構成を示す図。

【図3】従来のディスクアレイ制御装置の構成を示す図。

30

【図4】本発明によるディスクアレイ制御装置内のCMアクセス制御部の構成を示す図。

【図5】本発明によるディスクアレイ制御装置内のキャッシュメモリ部の構成を示す図。

【図6】本発明によるディスクアレイ制御装置内のデータ転送制御部の構成を示す図。

【図7】本発明によるディスクアレイ制御装置内のデータ転送制御部の他の構成を示す図。

【図8】本発明によるディスクアレイ制御装置内のアドレス(adr)、コマンド(cmd)解析部の構成を示す図。

【図9】本発明によるディスクアレイ制御装置内のセクタ部の構成を示す図。

【図10】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図11】本発明によるディスクアレイ制御装置の他の構成を示す図。

40

【図12】本発明によるディスクアレイ制御装置の他の構成を示す図。

【図13】本発明によるディスクアレイ制御装置の他の構成を示す図。

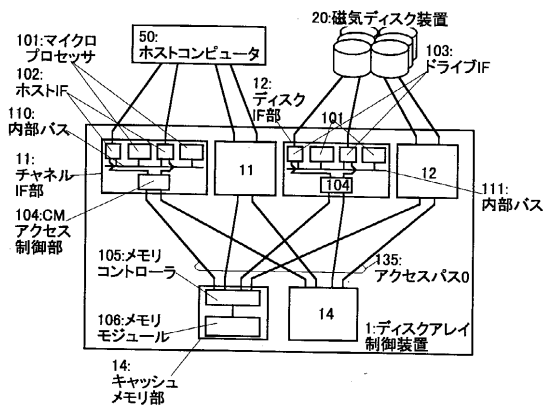
【図14】本発明によるディスクアレイ制御装置の他の構成を示す図。

##### 【符号の説明】

1...ディスクアレイ制御装置、11...チャンネルIF部、12...ディスクIF部、14...キャッシュメモリ部、20...磁気ディスク装置、50...ホストコンピュータ、101...マイクロプロセッサ、102...ホストIF、103...ドライブIF、104...CMアクセス制御部、105...メモリコントローラ、106...メモリモジュール、110、111...内部バス、135...アクセスパス0。

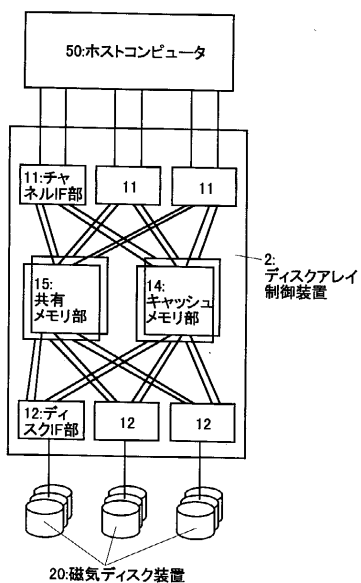
【 図 1 】

図1



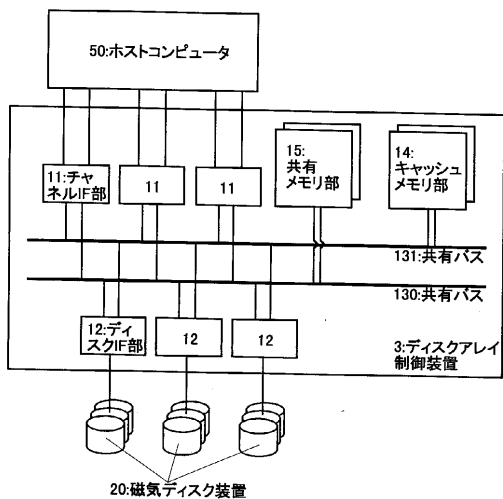
【 図 2 】

図2



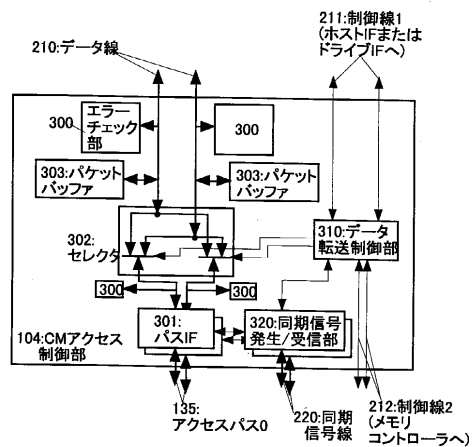
【 図 3 】

図3

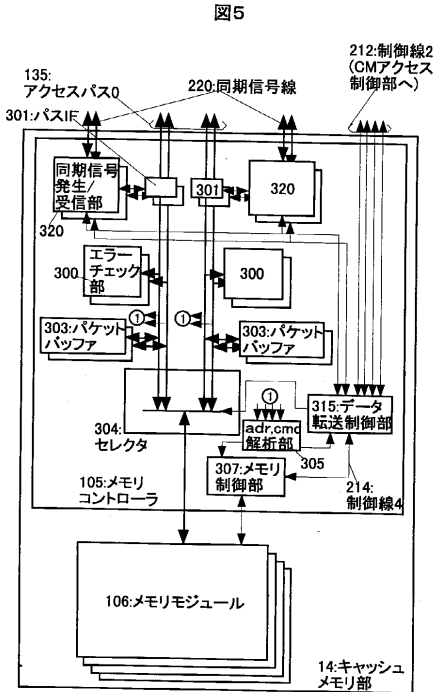


【 図 4 】

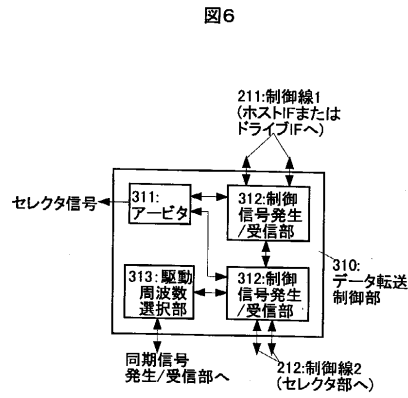
図4



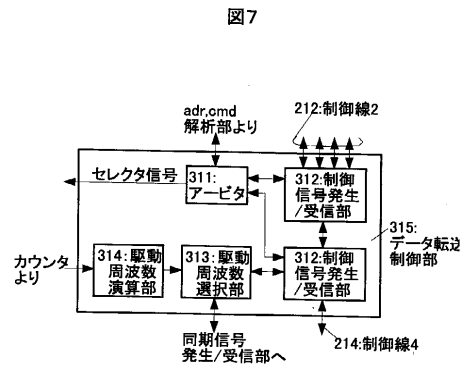
【 図 5 】



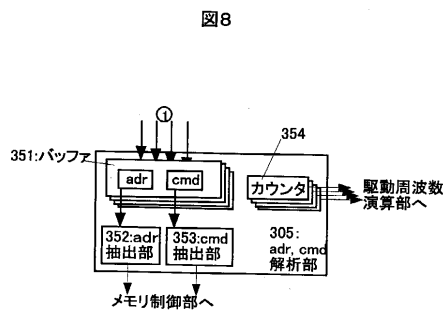
【 図 6 】



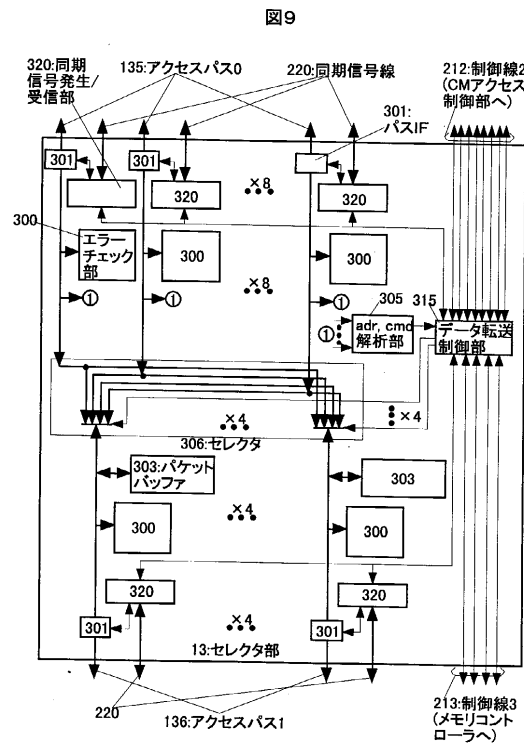
【 図 7 】



【 図 8 】

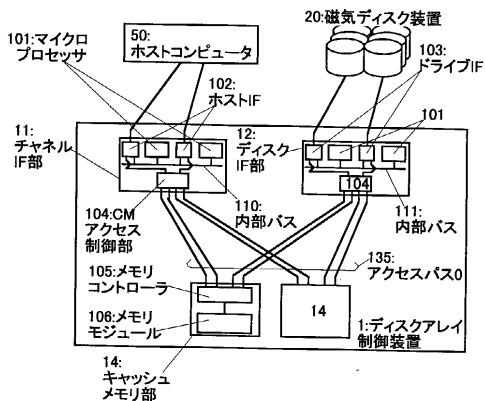


【 図 9 】



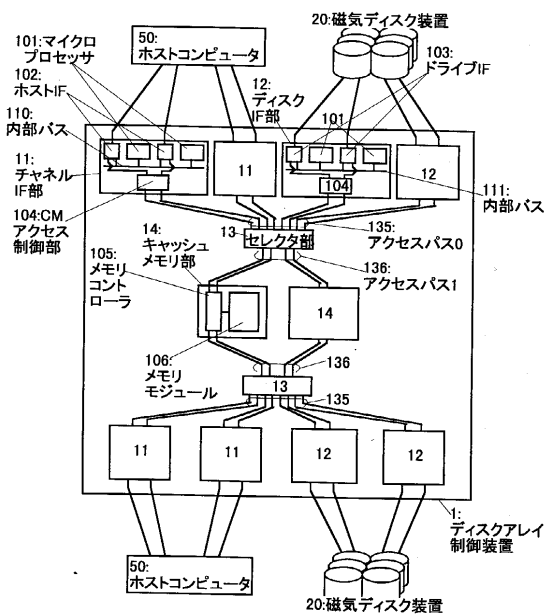
【 図 1 0 】

図10



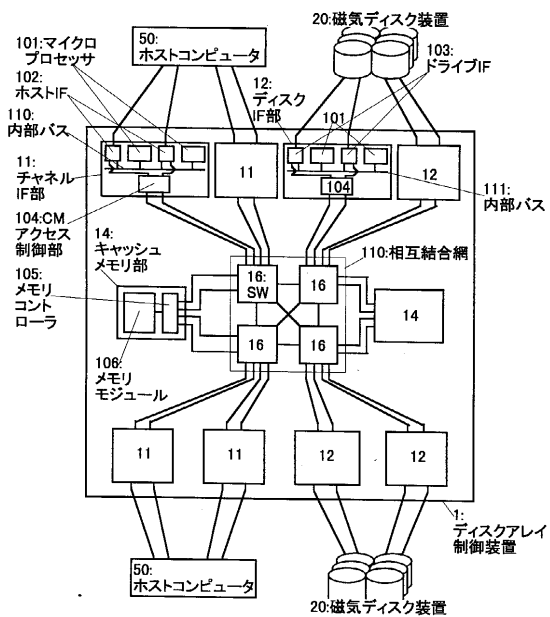
【 図 1 1 】

図11



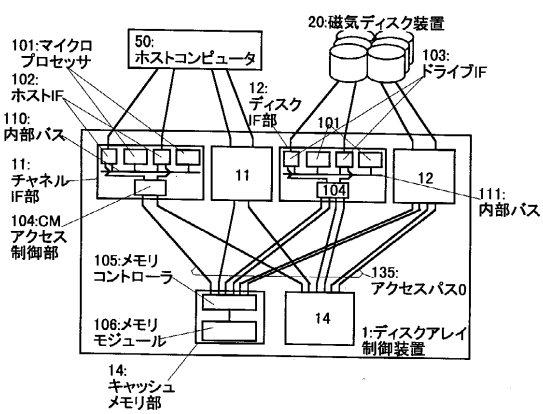
【 図 1 2 】

図12



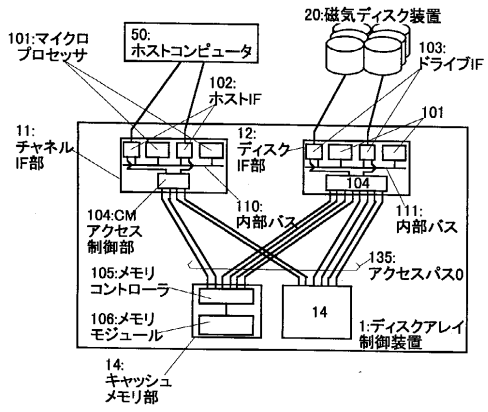
【 図 1 3 】

図13



【 図 1 4 】

図14





---

フロントページの続き

- (56)参考文献 特開平05 - 158797 (JP, A)  
特開平07 - 020994 (JP, A)  
特開平06 - 019627 (JP, A)  
特開平09 - 034646 (JP, A)  
特開平09 - 231158 (JP, A)  
特開平10 - 333836 (JP, A)

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G06F 3/06- 3/08

G06F 12/00-12/16

G06F 13/10-13/38