

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2012-504792

(P2012-504792A)

(43) 公表日 平成24年2月23日 (2012. 2. 23)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 3/06 (2006.01)	G06F 3/06 302A	5B005
G06F 13/10 (2006.01)	G06F 13/10 340A	5B014
G06F 12/08 (2006.01)	G06F 12/08 531B	5B065

審査請求 有 予備審査請求 未請求 (全 43 頁)

(21) 出願番号	特願2011-514949 (P2011-514949)	(71) 出願人	000005108
(86) (22) 出願日	平成21年2月17日 (2009. 2. 17)		株式会社日立製作所
(85) 翻訳文提出日	平成23年4月6日 (2011. 4. 6)		東京都千代田区丸の内一丁目6番6号
(86) 国際出願番号	PCT/JP2009/000645	(74) 代理人	110000279
(87) 国際公開番号	W02010/095166		特許業務法人ウィルフォート国際特許事務所
(87) 国際公開日	平成22年8月26日 (2010. 8. 26)	(72) 発明者	佐藤 恵
			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
		(72) 発明者	藤本 健雄
			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

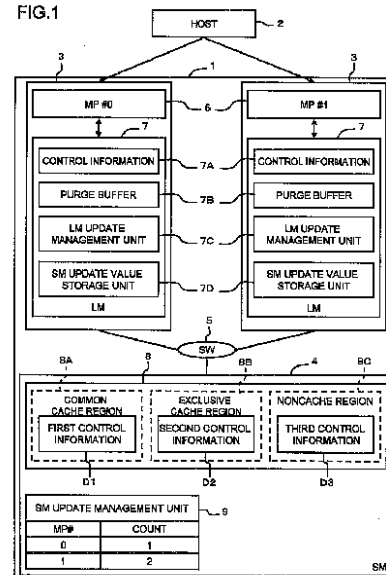
最終頁に続く

(54) 【発明の名称】 記憶制御装置及び記憶制御装置の制御方法

(57) 【要約】

本発明の記憶制御装置は、ページメッセージの通信量を低減することができ、記憶制御装置の処理性能を高めることができる。各マイクロプロセッサは、共有メモリ内の制御情報を更新するたびにページメッセージを作成して保存しておく。一連の更新処理が完了した後で、保存されたページメッセージを各マイクロプロセッサに送信する。制御情報には、その性質に応じた属性が設定される。その属性に応じてキャッシュ制御及びページ制御が実行される。

FIG.1



【特許請求の範囲】

【請求項 1】

ホストコンピュータ（２，３０）と記憶装置（１５１）との間のデータ入出力を制御する記憶制御装置（１，１０）であって、

共有制御情報を記憶する共有メモリ（４，１３１）と、

複数のマイクロプロセッサ（６，１２１）と、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリ（７，１２２）であって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部（７Ｂ，Ｔ１０）と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、

を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させる、

記憶制御装置。

【請求項 2】

複数のマイクロプロセッサパッケージ（３，１２０）を備え、前記各マイクロプロセッサパッケージは、少なくとも一つの前記マイクロプロセッサと、少なくとも一つの前記ローカルメモリと、少なくとも一つの前記パージメッセージ記憶部とを備えており、

前記各マイクロプロセッサパッケージと前記共有メモリとは、スイッチ（５，１６０）を介して接続されており、

前記共有メモリには、前記共有制御情報に加えて、前記共有制御情報の更新状態を管理するための共有制御情報更新管理部（９，Ｔ１３）が設けられており、

前記各ローカルメモリには、前記ローカル制御情報及び前記パージメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部（７Ｃ，Ｔ１１）と、前記共有制御情報更新管理部で管理されている前記更新状態を示す値のコピーを記憶する更新値記憶部（７Ｄ，Ｔ１２）とが記憶されており、

前記共有制御情報更新管理部は、前記各マイクロプロセッサパッケージ内の前記各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、前記各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、前記マイクロプロセッサにより前記共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応する前記カウンタ値が更新されるようになり、

前記各ローカル更新管理部も、前記各マイクロプロセッサ識別情報毎に前記カウンタ値を対応付けて管理しており、さらに、前記各ローカル更新管理部に対応する前記各マイクロプロセッサが前記ローカル制御情報を更新する場合に、その対応する前記マイクロプロセッサに関連するカウンタ値が更新されるようになり、さらに、前記対応する前記マイクロプロセッサ以外の他の前記マイクロプロセッサに関連する前記パージメッセージに基づいて、前記他の前記マイクロプロセッサに対応するカウンタ値が更新されるようになり、

前記各更新値記憶部に対応する前記各マイクロプロセッサが前記共有メモリ内の前記共有制御情報にアクセスした場合に、前記共有制御情報更新管理部で管理されている前記各

10

20

30

40

50

カウンタ値が取得されて、それら取得された前記各カウンタ値が前記更新値記憶部に上書きで記憶されるようになっており、

前記共有メモリは、前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域（８Ａ，１３１１）と、前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域（８Ｂ，１３１２）と、前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域（８Ｃ，１３１３）と、を備えており、

10

前記共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第１情報（Ｄ１，Ｄ１１）が記憶されるようになっており、

前記排他キャッシュ領域には、前記オーナーマイクロプロセッサが比較的高頻度に参照または更新する第２情報（Ｄ２，Ｄ１２）が記憶されるようになっており、

前記非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度で参照または更新される第３情報（Ｄ３，Ｄ１３）が記憶されるようになっており、

前記各マイクロプロセッサは、

前記共有キャッシュ領域に記憶されている情報を更新した場合には、前記更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを他の各マイクロプロセッサにそれぞれ送信し、

20

前記排他キャッシュ領域に記憶されている情報を更新する場合において、自分が前記オーナーマイクロプロセッサである場合は、前記パージメッセージを作成せず、自分が前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを前記オーナーマイクロプロセッサのみに送信し、

前記非キャッシュ領域に記憶されている情報を更新した場合には、前記パージメッセージを作成しないようになっており、

30

さらに、前記各マイクロプロセッサは、前記共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合には、複数の更新を統合するようにして前記パージメッセージを送信させるようになっている、
請求項１に記載の記憶制御装置。

【請求項３】

前記各マイクロプロセッサは、前記制御情報同期管理部に基づいて前記各ローカル制御情報が前記共有制御情報に同期していると判定される場合は、対応する前記各ローカルメモリから前記ローカル制御情報を使用し、前記制御情報同期管理部に基づいて前記各ローカル制御情報が前記共有制御情報に同期していないと判定される場合は、前記共有メモリ内の前記共有制御情報を使用する、
請求項１に記載の記憶制御装置。

40

【請求項４】

前記共有メモリには、前記共有制御情報に加えて、前記共有制御情報の更新状態を管理するための共有制御情報更新管理部（９，Ｔ１３）が設けられており、

前記各ローカルメモリには、前記ローカル制御情報及び前記パージメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部（７Ｃ，Ｔ１１）と、前記共有制御情報更新管理部で管理されている前記更新状態を示す値のコピーを記憶する更新値記憶部（７Ｄ，Ｔ１２）とが記憶されており、

50

前記制御情報同期管理部は、前記共有制御情報更新管理部と、前記各ローカル更新管理部と、前記各更新値管理部とを備えて構成される、
請求項 1 に記載の記憶制御装置。

【請求項 5】

前記共有制御情報更新管理部は、前記各マイクロプロセッサパッケージ内の前記各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、前記各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、前記マイクロプロセッサにより前記共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応する前記カウンタ値が更新されるようになっている、

10

前記各ローカル更新管理部も、前記各マイクロプロセッサ識別情報毎に前記カウンタ値を対応付けて管理しており、さらに、前記各ローカル更新管理部に対応する前記各マイクロプロセッサが前記ローカル制御情報を更新する場合に、その対応する前記マイクロプロセッサに関連するカウンタ値が更新されるようになっている、さらに、前記対応する前記マイクロプロセッサ以外の他の前記マイクロプロセッサに関連する前記パージメッセージに基づいて、前記他の前記マイクロプロセッサに対応するカウンタ値が更新されるようになっている、

請求項 4 に記載の記憶制御装置。

20

【請求項 6】

前記各更新値記憶部に対応する前記各マイクロプロセッサが前記共有メモリ内の前記共有制御情報にアクセスした場合に、前記共有制御情報更新管理部で管理されている前記各カウンタ値が取得されて、それら取得された前記各カウンタ値が前記更新値記憶部に書きで記憶されるようになっている、

請求項 5 に記載の記憶制御装置。

【請求項 7】

前記共有メモリは、

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域（8A，1311）と、

30

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域（8B，1312）と、

を備えている、

請求項 1 に記載の記憶制御装置。

【請求項 8】

前記共有メモリは、

40

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域（8A，1311）と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域（8C，1313）と、
を備えている、

請求項 1 に記載の記憶制御装置。

【請求項 9】

前記共有メモリは、

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセ

50

ッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域（８Ｂ，１３１２）と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域（８Ｃ，１３１３）と、
を備えている、
請求項１に記載の記憶制御装置。

【請求項１０】

前記共有メモリは、

10

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域（８Ａ，１３１１）と、

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域（８Ｂ，１３１２）と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域（８Ｃ，１３１３）と、
を備えている、
請求項１に記載の記憶制御装置。

20

【請求項１１】

前記共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第１情報（Ｄ１，Ｄ１１）が記憶されるようになっており、

前記排他キャッシュ領域には、前記オーナーマイクロプロセッサが比較的高頻度に参照または更新する第２情報（Ｄ２，Ｄ１２）が記憶されるようになっており、

前記非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度で参照または更新される第３情報（Ｄ３，Ｄ１３）が記憶されるようになっている、
請求項１０に記載の記憶制御装置。

30

【請求項１２】

前記各マイクロプロセッサは、

前記共有キャッシュ領域に記憶されている情報を更新した場合には、前記更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを他の各マイクロプロセッサにそれぞれ送信し、

前記排他キャッシュ領域に記憶されている情報を更新する場合において、自分が前記オーナーマイクロプロセッサである場合は、前記パージメッセージを作成せず、自分が前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを前記オーナーマイクロプロセッサのみに送信し、

40

前記非キャッシュ領域に記憶されている情報を更新した場合には、前記パージメッセージを作成しないようになっている、
請求項１０または請求項１１のいずれかに記載の記憶制御装置。

【請求項１３】

前記各マイクロプロセッサは、前記共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合に、複数の更新を統合するよ

50

うにして前記ページメッセージを送信させる、
請求項 1 に記載の記憶制御装置。

【請求項 14】

前記各マイクロプロセッサは、
他の各マイクロプロセッサが正常に監視しているか否かを監視しており（S120）、
いずれかのマイクロプロセッサで障害が発生した場合には（S121）、前記共有制御情報更新管理部で管理されている各カウント値を取得して前記更新値記憶部に上書きで記憶し（S123）、

前記障害の発生した前記マイクロプロセッサについて、前記ローカル更新管理部で管理されているカウント値と前記更新値管理部に記憶されているカウント値と異なるか否かを判定し（S124）、

前記ローカル更新管理部で管理されているカウント値と前記更新値管理部に記憶されているカウント値とが異なる場合、前記ローカル更新管理部で管理されている前記障害の発生した前記マイクロプロセッサの前記カウント値を、前記更新値管理部内の対応するカウント値で上書きし（S125）、

前記ローカルメモリにコピーされているローカル制御情報をクリアさせる（S126）、
請求項 1 に記載の記憶制御装置。

【請求項 15】

共有制御情報を記憶する共有メモリ（4，131）と、複数のマイクロプロセッサ（6，121）と、前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリ（7，122）とを備える記憶制御装置の制御方法であって、

前記各ローカルメモリには、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされており、

前記共有制御情報を更新する場合には、更新元のマイクロプロセッサに対応する前記ローカルメモリ以外の他の前記各ローカルメモリに記憶されている前記ローカル制御情報が無効であることを示すページメッセージを作成して保存し、

前記共有制御情報を更新する一連の処理が完了した場合に、前記保存されている前記ページメッセージを前記他の前記各ローカルメモリに対応する前記各マイクロプロセッサに送信し、

前記ページメッセージを受信した前記他の前記各マイクロプロセッサは、前記無効とされた箇所の情報を使用する場合に、前記共有メモリにアクセスして取得するようになっている、

記憶制御装置の制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶制御装置及び記憶制御装置の制御方法に関する。

【背景技術】

【0002】

記憶制御装置は、ハードディスクドライブのような記憶装置を複数用いて、RAID（Redundant Arrays of Inexpensive Disks）に基づく記憶領域をホストコンピュータに提供する。記憶制御装置は、複数のマイクロプロセッサを備えている。各マイクロプロセッサは、共有メモリに記憶されているデータを共有する。マイクロプロセッサが共有メモリに直接アクセスしてデータを使用すると、処理時間がかかる。そこで、共有されるデータの一部を、マイクロプロセッサの近傍に配置されるローカルメモリにコピーして使用する構成が提案されている。

【0003】

【特許文献 1】JP - A - 2006 - 285778

10

20

30

40

50

【発明の開示】

【発明が解決しようとする課題】

【0004】

従来技術では、いずれかのマイクロプロセッサが共有メモリ内のデータを更新した場合に、他のローカルメモリにコピーされているデータ（つまり、他のローカルメモリにキャッシュされているデータ）を破棄させる必要がある。大元のデータである共有メモリ内のデータが変更された場合、各ローカルメモリにコピーされているデータは古くなり、使用できないためである。

【0005】

そこで、共有メモリ内のデータが更新された場合、そのデータ更新を行ったマイクロプロセッサから他の各マイクロプロセッサに向けて、ローカルメモリに記憶されたデータが使用できない旨を通知する。この通知を、ここではバージメッセージと呼ぶ。バージメッセージを受け取った他のマイクロプロセッサは、ローカルメモリにコピーされているデータが古くて使えないことを知る。従って、他のプロセッサは、共有メモリにアクセスして新しいデータを読み出し、ローカルメモリに記憶させる。

10

【0006】

各マイクロプロセッサが比較的高頻度で共有メモリ内のデータを更新すると、各マイクロプロセッサ間で多数のバージメッセージが交換される。従って、バージメッセージを作成したり、バージメッセージを送信したり、バージメッセージに基づいてローカルメモリ内のデータを破棄したり等の処理が多数行われる。

20

【0007】

つまり、従来技術では、各マイクロプロセッサ間で多数のバージメッセージが通信されるため、バージメッセージ関連の処理がボトルネックになり易い。そのため、記憶制御装置の処理性能を向上させるのが難しいという問題がある。

【0008】

そこで、本発明の目的は、複数のマイクロプロセッサ間で通信されるバージメッセージの量を低減することにより、処理性能を向上できるようにした記憶制御装置及び記憶制御装置の制御方法を提供することにある。本発明の他の目的は、共有メモリに記憶される情報を複数種類に分類し、各情報の性質に応じてバージメッセージの送信を制御することにより、処理性能を向上できるようにした記憶制御装置及び記憶制御装置の制御方法を提供することにある。本発明の他の目的は、後述する実施形態の記載から明らかになるであろう。

30

【課題を解決するための手段】

【0009】

上記課題を解決すべく、本発明の第1観点に従う記憶制御装置は、ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって、共有制御情報を記憶する共有メモリと、複数のマイクロプロセッサと、各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、共有メモリに記憶される共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、各マイクロプロセッサのうちの一つのマイクロプロセッサから他の各マイクロプロセッサに、または、特定の一つまたは複数のマイクロプロセッサに、向けて送信されるバージメッセージであって、ローカルメモリに記憶されているローカル制御情報が無効である旨を通知するためのバージメッセージを記憶するバージメッセージ記憶部と、各ローカル制御情報が共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、を備え、各マイクロプロセッサは、共有制御情報を更新する場合に、バージメッセージを作成してバージメッセージ記憶部に記憶させ、さらに、共有制御情報の更新と非同期で、バージメッセージ記憶部に記憶されたバージメッセージを送信させる。

40

【0010】

第2観点では、第1観点において、複数のマイクロプロセッサパッケージを備え、各マイクロプロセッサパッケージは、少なくとも一つのマイクロプロセッサと、少なくとも一

50

つのローカルメモリと、少なくとも一つのページメッセージ記憶部とを備えており、各マイクロプロセッサパッケージと共有メモリとは、スイッチを介して接続されており、

共有メモリには、共有制御情報に加えて、共有制御情報の更新状態を管理するための共有制御情報更新管理部が設けられており、

各ローカルメモリには、ローカル制御情報及びページメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部と、共有制御情報更新管理部で管理されている更新状態を示す値のコピーを記憶する更新値記憶部とが記憶されており、

共有制御情報更新管理部は、各マイクロプロセッサパッケージ内の各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、マイクロプロセッサにより共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応するカウンタ値が更新されるようになっており、

各ローカル更新管理部も、各マイクロプロセッサ識別情報毎にカウンタ値を対応付けて管理しており、さらに、各ローカル更新管理部に対応する各マイクロプロセッサがローカル制御情報を更新する場合に、その対応するマイクロプロセッサに関連するカウンタ値が更新されるようになっており、さらに、対応するマイクロプロセッサ以外の他のマイクロプロセッサに関連するページメッセージに基づいて、他のマイクロプロセッサに対応するカウンタ値が更新されるようになっており、

各更新値記憶部に対応する各マイクロプロセッサが共有メモリ内の共有制御情報にアクセスした場合に、共有制御情報更新管理部で管理されている各カウンタ値が取得されて、それら取得された各カウンタ値が更新値記憶部に上書きで記憶されるようになっており、

共有メモリは、各マイクロプロセッサが各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、オーナーマイクロプロセッサ内のローカルメモリにコピー可能であり、かつ、オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、各マイクロプロセッサのいずれもが各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、を備えており、

共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第1情報が記憶されるようになっており、

排他キャッシュ領域には、オーナーマイクロプロセッサが比較的高頻度に参照または更新する第2情報が記憶されるようになっており、

非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度に参照または更新される第3情報が記憶されるようになっており、

各マイクロプロセッサは、

共有キャッシュ領域に記憶されている情報を更新した場合には、更新のたびにページメッセージを作成してページメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、ページメッセージ記憶部に記憶された各ページメッセージを他の各マイクロプロセッサにそれぞれ送信し、

排他キャッシュ領域に記憶されている情報を更新する場合において、自分がオーナーマイクロプロセッサである場合は、ページメッセージを作成せず、自分がオーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびにページメッセージを作成してページメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、ページメッセージ記憶部に記憶された各ページメッセージをオーナーマイクロプロセッサのみに送信し、

非キャッシュ領域に記憶されている情報を更新した場合には、ページメッセージを作成しないようになっており、

さらに、各マイクロプロセッサは、共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合には、複数の更新を統合する

10

20

30

40

50

ようにしてパージメッセージを送信させるようになっている。

【 0 0 1 1 】

第 3 観点では、第 1 観点において、各マイクロプロセッサは、制御情報同期管理部に基づいて各ローカル制御情報が共有制御情報に同期していると判定される場合は、対応する各ローカルメモリからローカル制御情報を使用し、制御情報同期管理部に基づいて各ローカル制御情報が共有制御情報に同期していないと判定される場合は、共有メモリ内の共有制御情報を使用する。

【 0 0 1 2 】

第 4 観点では、第 1 観点において、共有メモリには、共有制御情報に加えて、共有制御情報の更新状態を管理するための共有制御情報更新管理部が設けられており、各ローカルメモリには、ローカル制御情報及びパージメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部と、共有制御情報更新管理部で管理されている更新状態を示す値のコピーを記憶する更新値記憶部とが記憶されており、制御情報同期管理部は、共有制御情報更新管理部と、各ローカル更新管理部と、各更新値管理部とを備えて構成される。

【 0 0 1 3 】

第 5 観点では、第 4 観点において、共有制御情報更新管理部は、各マイクロプロセッサパッケージ内の各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、マイクロプロセッサにより共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応するカウンタ値が更新されるようになっており、各ローカル更新管理部も、各マイクロプロセッサ識別情報毎にカウンタ値を対応付けて管理しており、さらに、各ローカル更新管理部に対応する各マイクロプロセッサがローカル制御情報を更新する場合に、その対応するマイクロプロセッサに関連するカウンタ値が更新されるようになっており、さらに、対応するマイクロプロセッサ以外の他のマイクロプロセッサに関連するパージメッセージに基づいて、他のマイクロプロセッサに対応するカウンタ値が更新されるようになっている。

【 0 0 1 4 】

第 6 観点では、第 5 観点において、各更新値記憶部に対応する各マイクロプロセッサが共有メモリ内の共有制御情報にアクセスした場合に、共有制御情報更新管理部で管理されている各カウンタ値が取得されて、それら取得された各カウンタ値が更新値記憶部に書きで記憶されるようになっている。

【 0 0 1 5 】

第 7 観点では、第 1 観点において、共有メモリは、各マイクロプロセッサが各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、オーナーマイクロプロセッサ内のローカルメモリにコピー可能であり、かつ、オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、を備えている。

【 0 0 1 6 】

第 8 観点では、第 1 観点において、共有メモリは、各マイクロプロセッサが各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、各マイクロプロセッサのいずれもが各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、を備えている。

【 0 0 1 7 】

第 9 観点では、第 1 観点において、共有メモリは、各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、オーナーマイクロプロセッサ内のローカルメモリにコピー可能であり、かつ、オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、各マイクロプロセッサのいずれもが各ローカルメモリにコピーできない情報を記憶するための非

10

20

30

40

50

キャッシュ領域と、を備えている。

【0018】

第10観点では、第1観点において、共有メモリは、各マイクロプロセッサが各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、オーナーマイクロプロセッサ内のローカルメモリにコピー可能であり、かつ、オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、各マイクロプロセッサのいずれもが各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、を備えている。

【0019】

第11観点では、第10観点において、共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第1情報が記憶されるようになっており、排他キャッシュ領域には、オーナーマイクロプロセッサが比較的高頻度に参照または更新する第2情報が記憶されるようになっており、非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度で参照または更新される第3情報が記憶されるようになっている。

【0020】

第12観点では、第10または第11観点のいずれかにおいて、各マイクロプロセッサは、共有キャッシュ領域に記憶されている情報を更新した場合には、更新のたびにパージメッセージを作成してパージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、パージメッセージ記憶部に記憶された各パージメッセージを他の各マイクロプロセッサにそれぞれ送信し、排他キャッシュ領域に記憶されている情報を更新する場合において、自分がオーナーマイクロプロセッサである場合は、パージメッセージを作成せず、自分がオーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびにパージメッセージを作成してパージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、パージメッセージ記憶部に記憶された各パージメッセージをオーナーマイクロプロセッサのみに送信し、非キャッシュ領域に記憶されている情報を更新した場合には、パージメッセージを作成しないようになっている。

【0021】

第13観点では、第1観点において、各マイクロプロセッサは、共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合に、複数の更新を統合するようにしてパージメッセージを送信させる。

【0022】

第14観点では、第1観点において、各マイクロプロセッサは、他の各マイクロプロセッサが正常に監視しているか否かを監視しており、いずれかのマイクロプロセッサで障害が発生した場合には、共有制御情報更新管理部で管理されている各カウント値を取得して更新値記憶部に上書きで記憶し、障害の発生したマイクロプロセッサについて、ローカル更新管理部で管理されているカウント値と更新値管理部に記憶されているカウント値と異なるか否かを判定し、ローカル更新管理部で管理されているカウント値と更新値管理部に記憶されているカウント値とが異なる場合、ローカル更新管理部で管理されている障害の発生したマイクロプロセッサのカウント値を、更新値管理部内の対応するカウント値で上書きし、ローカルメモリにコピーされているローカル制御情報をクリアさせる。

【0023】

第15観点に従う記憶制御装置の制御方法は、共有制御情報を記憶する共有メモリと、複数のマイクロプロセッサと、各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリとを備える記憶制御装置の制御方法であって、各ローカルメモリには、共有メモリに記憶される共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされており、共有制御情報を更新する場合には、更新元のマイクロプロセッサに対応するローカルメモリ以外の他の各ローカルメモリに記憶されているローカル制御情報が無効であることを示すパージメッセージを作成して保存し、共有制御情報を更新する一連の処理が

10

20

30

40

50

完了した場合に、保存されているパージメッセージを他の各ローカルメモリに対応する各マイクロプロセッサに送信し、パージメッセージを受信した他の各マイクロプロセッサは、無効とされた箇所の情報を使用する場合に、共有メモリにアクセスして取得するようになっている。

【 0 0 2 4 】

さらに、本発明の観点は明示された以外に種々組み合わせることも可能であり、そのような組合せは本発明の範囲に含まれる。

【 図面の簡単な説明 】

【 0 0 2 5 】

【 図 1 】 本発明の実施形態を示す概念図である。

10

【 図 2 】 記憶制御装置のブロック図である。

【 図 3 】 ローカルメモリ及び共有メモリの構成を模式的に示す図である。

【 図 4 】 パージバッファの構成を示す図である。

【 図 5 】 パージメッセージを統合する様子を示す図である。

【 図 6 】 各更新クロックの関係を示す図である。

【 図 7 】 クロックの大小を判定する方法を示す図である。

【 図 8 】 パージメッセージの作成及び送信を示すフローチャートである。

【 図 9 】 第 2 実施例に係る記憶制御装置の要部を示すブロック図である。

【 図 1 0 】 キャッシュ属性管理テーブルを示す図である。

【 図 1 1 】 共有メモリに記憶されるデータの例を示す図である。

20

【 図 1 2 】 共有メモリにデータを書き込む処理を示すフローチャートである。

【 図 1 3 】 図 1 2 に続くフローチャートである。

【 図 1 4 】 パージメッセージを送信する処理を示すフローチャートである。

【 図 1 5 】 共有メモリからデータを読み出す処理を示すフローチャートである。

【 図 1 6 】 第 3 実施例に係る記憶制御装置で実行される、マイクロプロセッサに障害が発生した場合の処理を示すフローチャートである。

【 図 1 7 】 障害から回復したマイクロプロセッサによって実行される処理を示すフローチャートである。

【 図 1 8 】 第 4 実施例に係る記憶制御装置の要部を示すフローチャートである。

【 符号の説明 】

30

【 0 0 2 6 】

1 : 記憶制御装置

3 : マイクロプロセッサパッケージ

4 : 共有メモリ

5 : スイッチ

6 : マイクロプロセッサ

7 : ローカルメモリ

8 A : 共有キャッシュ領域

8 B : 排他キャッシュ領域

8 C : 非キャッシュ領域

40

1 0 : 記憶制御装置

1 1 0 : フロントエンドパッケージ

1 2 0 : マイクロプロセッサパッケージ

1 2 1 : マイクロプロセッサ

1 2 2 : ローカルメモリ

1 3 0 : メモリパッケージ

1 3 1 : 共有メモリ

1 4 0 : バックエンドパッケージ

1 5 1 : 記憶装置

【 発明を実施するための最良の形態 】

50

【 0 0 2 7 】

図 1 は、本発明の実施形態の概要を示す説明図である。図 1 に示す構成は、本発明の範囲を限定するものではなく、図 1 に示す構成以外の構成も本発明の範囲に含まれる。本実施形態に係る記憶制御装置 1 は、ホストコンピュータ（以下、「ホスト」と略記）2 に接続されている。ホスト 2 としては、例えば、サーバコンピュータ、メインフレームコンピュータ、エンジニアリングワークステーション、パーソナルコンピュータ等のコンピュータ装置を挙げることができる。

【 0 0 2 8 】

図 1 は、記憶制御装置 1 の要部を示している。記憶制御装置 1 は、複数のマイクロプロセッサパッケージ 3 と、共有メモリ 4 と、スイッチ 5 とを備える。その他、後述する実施例に示すように、記憶制御装置 1 は、複数の記憶装置 151 を備えることもできる。

10

【 0 0 2 9 】

図 1 では、2 つのマイクロプロセッサパッケージ 3 を示すが、3 個以上のマイクロプロセッサを設けることもできる。各マイクロプロセッサパッケージ 3 は、例えば、マイクロプロセッサ 6 と、ローカルメモリ 7 とを備える。

【 0 0 3 0 】

マイクロプロセッサ 6 は、ホスト 2 からのコマンドを解釈して所定の処理を実行し、その処理結果をホスト 2 に返す。ローカルメモリ 7 は、例えば、ローカル制御情報 7 A と、ページバッファ 7 B と、ローカルメモリ更新管理部 7 C と、共有メモリ更新値記憶部 7 D とを備えている。図面では、ローカルメモリを「L M」と、共有メモリを「S M」と略して表示している。なお、実施例においても、ローカルメモリを L M、共有メモリを S M と略記する場合がある。

20

【 0 0 3 1 】

ローカル制御情報 7 A は、共有メモリ 4 に記憶されている制御情報の一部である。制御情報として、例えば、論理ボリュームを識別するための番号、R A I D 構成を管理するための情報、ジョブを管理するための情報等のように、コマンド処理及び／または記憶制御装置 1 の制御等に必要な情報を挙げることができる。

【 0 0 3 2 】

ページバッファ 7 B は、ページメッセージを記憶するための記憶領域である。本実施形態におけるページメッセージには、複数の役割がある。一つの役割は、ローカルメモリ 7 に記憶されているローカル制御情報 7 A が無効である旨を通知することである。この場合、ページメッセージとは、ローカルメモリ 7 にキャッシュされているローカル制御情報 7 A のページを要求するためのメッセージである。

30

【 0 0 3 3 】

他の役割は、マイクロプロセッサ 6 によって共有メモリ 4 内の制御情報が更新された回数（カウンタ値）を、他のマイクロプロセッサ 6 に通知することである。この場合、ページメッセージは、更新通知用メッセージとなる。ページメッセージ内の所定ビットを設定することにより、ページ要求メッセージまたは更新通知用メッセージのいずれか一方の役割を与えることができる。

【 0 0 3 4 】

各マイクロプロセッサ 6 が共有メモリ 4 内の制御情報を更新する度に、ページ要求のためのページメッセージが作成されて、ページバッファ 7 B に記憶される。なお、同一の領域、または、連続する領域について更新されたような場合、複数の更新を一つのページメッセージにまとめる。

40

【 0 0 3 5 】

共有メモリ 4 内の制御情報に関する一連の更新が完了すると、更新通知用のページメッセージが作成されてページバッファ 7 B に記憶される。更新通知用のページメッセージがページバッファ 7 B に記憶された後、他のマイクロプロセッサ 6 に向けて、ページバッファ 7 B 内の各ページメッセージが送信される。

【 0 0 3 6 】

50

LM更新管理部7Cは、マイクロプロセッサ6による制御情報の更新状態を管理するための管理部である。LM更新管理部7Cは、各マイクロプロセッサ6毎に、その更新回数(カウンタ値)を管理する。例えば、マイクロプロセッサ6(#0)が共有メモリ4内の制御情報を更新した場合、LM更新管理部7Cは、マイクロプロセッサ6(#0)のカウンタ値を「1」に設定する。さらに、マイクロプロセッサ6(#0)が共有メモリ4内の制御情報を更新した場合、カウンタ値は「2」に変更される。

【0037】

図1中右側に示す別のマイクロプロセッサ6(#1)が共有メモリ4内の制御情報を更新した場合、その更新の事実は上述のページメッセージを介して通知される。マイクロプロセッサ6(#0)は、マイクロプロセッサ6(#1)から更新通知用メッセージとしてのページメッセージを受領すると、LM更新管理部7C内のマイクロプロセッサ6(#1)のカウンタ値を更新させる。

10

【0038】

SM更新値記憶部7Dには、所定のタイミングで、共有メモリ4内のSM更新管理部9で管理されている各マイクロプロセッサ6のカウント値がコピーされる。所定のタイミングとしては、マイクロプロセッサ6が共有メモリ4にアクセスする場合を挙げることができる。例えば、ローカル制御情報7Aに含まれていない他の制御情報を使用する場合、マイクロプロセッサ6は、共有メモリ4にアクセスして所望の制御情報を取得する。その際に、SM更新管理部9で管理されている各カウンタ値が読み出されて、SM更新値記憶部7Dに記憶される。

20

【0039】

共有メモリ4は、スイッチ5を介して各マイクロプロセッサパッケージ3に接続されている。共有メモリ4は、制御情報記憶領域8と、SM更新管理部9とを備える。制御情報記憶領域8は、制御情報を記憶するための領域である。制御情報記憶領域8に記憶される制御情報は、各マイクロプロセッサ6により共有される。

【0040】

制御情報記憶領域8には、共有キャッシュ領域8Aと、排他キャッシュ領域8Bと、非キャッシュ領域8Cとが含まれている。共有キャッシュ領域8Aには、各マイクロプロセッサ6によって共有され、かつ、各マイクロプロセッサ6によりキャッシュされる第1制御情報D1が記憶される。排他キャッシュ領域8Bには、オーナーとなる特定のマイクロプロセッサ6のみがキャッシュでき、他のマイクロプロセッサ6はキャッシュすることのできない第2制御情報D2が記憶される。非キャッシュ領域8Cには、いずれのマイクロプロセッサ6もキャッシュすることのできない第3制御情報D3が記憶される。

30

【0041】

第1制御情報D1としては、例えば、記憶制御装置1の構成を管理する情報、プログラム製品の設定情報等のように、更新頻度が比較的少なく、かつ、複数のマイクロプロセッサ6から比較的高頻度で参照され得る情報が該当する。

【0042】

第2制御情報D2としては、例えば、論理ボリュームを管理する情報、リモートコピーの差分データを管理する情報等のように、特定のマイクロプロセッサ6が比較的高頻度で更新するような情報が該当する。ここでは、その特定のマイクロプロセッサをオーナーマイクロプロセッサと呼ぶ。オーナーマイクロプロセッサ以外のマイクロプロセッサを、非オーナーマイクロプロセッサ、または、その他のマイクロプロセッサと呼ぶ。理解のために、オーナーマイクロプロセッサを6(#0)とし、その他のマイクロプロセッサを6(#1)とする。

40

【0043】

第3制御情報D3としては、例えば、記憶制御装置1内で実行される各種ジョブを管理する情報、障害に関する情報等のように、更新頻度が比較的少ない情報が該当する。但し、第1 - 第3制御情報の例は、上記に限られない。記憶制御装置1の制御方式等に応じて変化し得る。

50

【 0 0 4 4 】

本実施形態では、制御情報の使われ方等に基づいて制御情報を複数種類に分類し、制御用の属性を与えて共有メモリ4に記憶させる。共有キャッシュ領域8A内の第1制御情報D1は、各マイクロプロセッサ6がキャッシュ可能である。各マイクロプロセッサ6は、第1制御情報D1について対等であり、各マイクロプロセッサ6は、第1制御情報D1をそれぞれのローカルメモリ7にコピーして使用する。各マイクロプロセッサ6が第1制御情報D1をローカルメモリ7にコピーして使用するため（つまり、第1制御情報D1をキャッシュするため）、各マイクロプロセッサ6間でパージメッセージが交換される。

【 0 0 4 5 】

排他キャッシュ領域8B内の第2制御情報D2は、オーナーマイクロプロセッサ6（#0）のみがローカルメモリ7にコピー可能であり（つまり、オーナーマイクロプロセッサ6（#0）のみがキャッシュ可能であり）、その他のマイクロプロセッサ6（#1）は第2制御情報D2をローカルメモリ7にコピーすることはできない。第2制御情報D2の使用を希望するその他のマイクロプロセッサ6（#1）は、共有メモリ4にアクセスすることにより、第2制御情報D2を使用する。

【 0 0 4 6 】

第2制御情報D2をローカルメモリ7にコピーできるのはオーナーマイクロプロセッサ6（#0）だけであるから、オーナーマイクロプロセッサ6（#0）は、共有メモリ4内の第2制御情報D2を更新した場合でも、その他のマイクロプロセッサ6（#1）にパージメッセージを送信する必要はない。その他のマイクロプロセッサ6（#1）が共有メモリ4内の第2制御情報D2を更新した場合、その他のマイクロプロセッサ6（#1）からオーナーマイクロプロセッサ6（#0）にパージメッセージが送信される。

【 0 0 4 7 】

非キャッシュ領域8C内の第3制御情報D3は、いずれのマイクロプロセッサ6もローカルメモリ7にコピーすることができない。各マイクロプロセッサ6は、第3制御情報D3を必要とする場合、共有メモリ4に直接アクセスして使用する。従って、第3制御情報D3については、各マイクロプロセッサ6間でパージメッセージは交換されない。

【 0 0 4 8 】

SM更新管理部9は、各マイクロプロセッサ6によって共有メモリ4内の各制御情報D1，D2が更新された回数を管理する。非キャッシュ領域8C内の第3制御情報D3については、更新回数を管理する必要はないが、管理してもよい。

【 0 0 4 9 】

SM更新管理部9は、各マイクロプロセッサ6を識別するためのマイクロプロセッサ番号（MP#）と、各マイクロプロセッサ6による更新回数（COUNT）とを対応付けて管理する。なお、各LM更新管理部7C及び各SM更新値記憶部7Dも、マイクロプロセッサ番号と更新回数とを対応付けて管理する。つまり、7C，7D，9は、それぞれ同一の配列を有する。しかし、更新回数が更新されるタイミング等が異なる。ここでは、更新回数をカウンタ値とも呼ぶ。

【 0 0 5 0 】

第1制御情報D1に着目して、パージ制御を説明する。以下、ローカルメモリ7，パージバッファ7B、LM更新管理部7C、SM更新値記憶部7Dに、（#0）または（#1）を添えることにより、一方のマイクロプロセッサ6（#0）側の構成であるのか、それとも他方のマイクロプロセッサ6（#1）側の構成であるのかを区別する。

【 0 0 5 1 】

一方のマイクロプロセッサ6（#0）が共有メモリ4内の第1制御情報D1を更新すると、一方のマイクロプロセッサ6（#0）に対応するLM更新管理部7C（#0）において、一方のマイクロプロセッサ6（#0）のカウンタ値が1つインクリメントされる。さらに、一方のマイクロプロセッサ6（#0）による第1制御情報D1の更新について、パージ要求用のパージメッセージが作成される。作成されたパージ要求用のパージメッセージは、パージバッファ7B（#0）に記憶される。

10

20

30

40

50

【 0 0 5 2 】

さらに、一方のマイクロプロセッサ 6 (# 0) が共有メモリ 4 内の第 1 制御情報 D 1 を更新する場合、一方のマイクロプロセッサ 6 (# 0) は、S M 更新管理部 9 内の一方のマイクロプロセッサ 6 (# 0) のカウンタ値を 1 つインクリメントさせる。

【 0 0 5 3 】

一方のマイクロプロセッサ 6 (# 0) は、例えば、ホスト 2 からのコマンドに対応して、共有メモリ 4 内の第 1 制御情報 D 1 を複数回更新する。複数の更新のそれぞれについて、パージ要求用のパージメッセージが作成され、保存される。なお、一つにまとめることができる場合、複数のパージメッセージが一つのパージメッセージに統合される。

【 0 0 5 4 】

一連の更新が完了すると、更新通知用のパージメッセージが作成される。その更新通知用のパージメッセージには、L M 更新管理部 7 C (# 0) に記憶されている、一方のマイクロプロセッサ 6 (# 0) のマイクロプロセッサ番号及びカウンタ値が含まれる。

【 0 0 5 5 】

更新通知用のパージメッセージがパージバッファ 7 B (# 0) に記憶されて、所定の契機が訪れると、パージバッファ 7 B (# 0) に記憶された各パージ要求用のパージメッセージと一つの更新通知用パージメッセージとが、他方のマイクロプロセッサ 6 (# 1) に送信される。パージメッセージを送信する契機としては、例えば、共有メモリ 4 内の制御情報への一連の更新が完了し、プログラムが明示的にコールした場合、または、パージバッファ 7 B に保存されるパージメッセージの数が予め設定される上限値に達した場合等が挙げられる。

【 0 0 5 6 】

他方のマイクロプロセッサ 6 (# 1) は、一方のマイクロプロセッサ 6 (# 0) からパージメッセージを受領すると、パージ要求用のパージメッセージに基づいて、ローカルメモリ 7 (# 1) 内に記憶されている第 1 制御情報 D 1 を無効とする。

【 0 0 5 7 】

さらに、他方のマイクロプロセッサ 6 (# 1) は、一方のマイクロプロセッサ 6 (# 0) から受信した更新通知用パージメッセージに基づいて、L M 更新管理部 7 C (# 1) 内の、一方のマイクロプロセッサ 6 (# 0) のカウンタ値を更新させる。

【 0 0 5 8 】

このように、更新元のマイクロプロセッサ 6 (# 0) が共有メモリ 4 内の制御情報を更新する場合、更新元のマイクロプロセッサ 6 (# 0) に関するカウンタ値は、更新元のマイクロプロセッサ 6 (# 0) 側の L M 更新管理部 7 C (# 0) と S M 更新管理部 9 との両方で、ほぼ同時期に更新される。

【 0 0 5 9 】

更新元のマイクロプロセッサ 6 (# 0) 以外の他のマイクロプロセッサ 6 (# 1) では、更新元マイクロプロセッサ 6 (# 0) から受信する更新通知用パージメッセージに基づいて、L M 更新管理部 7 C (# 1) 内の更新元マイクロプロセッサ 6 (# 0) のカウンタ値を更新させる。

【 0 0 6 0 】

S M 更新値記憶部 7 D には、マイクロプロセッサ 6 が共有メモリ 4 にアクセスした場合に S M 更新管理部 9 から読み出される各カウンタ値がコピーされる。上述の通り、更新元マイクロプロセッサ 6 (# 0) に関する S M 更新管理部 9 内のカウンタ値は、更新元マイクロプロセッサ 6 (# 0) による更新処理に同期して更新される。

【 0 0 6 1 】

従って、更新元マイクロプロセッサ 6 (# 0) 以外の他のマイクロプロセッサ 6 (# 1) は、S M 更新値記憶部 7 D (# 1) 内の全カウンタ値と L M 更新値管理部 7 C (# 1) 内の全カウンタ値とを比較することにより、ローカルメモリ 7 (# 1) 内に記憶されている第 1 制御情報 D 1 が有効であるか否かを判定できる。ここで、共有メモリ 4 と S M 更新値記憶部 7 D との区別を明らかにするために、S M 更新値をターゲット (T G T) と示す

10

20

30

40

50

場合がある。

【 0 0 6 2 】

LM更新管理部7C(# 1)内の全カウンタ値COUNT(LM # 1)がSM更新値記憶部7D(# 1)内の全カウンタCOUNT(TGT # 1)値と等しいか、あるいは、それよりも大きい場合(COUNT(LM # 1) > COUNT(TGT # 1))、ローカルメモリ7(# 1)に記憶されている第1制御情報D1は、有効であると判定できる。

【 0 0 6 3 】

逆に、COUNT(LM # 1) < COUNT(TGT # 1)の場合、更新元マイクロプロセッサ6(# 0)から未だ届いていないパージメッセージが存在すると判断できる。従って、この場合、ローカルメモリ7(# 1)内に記憶されている第1制御情報D1は古くなっている可能性があり、使用することはできない。

10

【 0 0 6 4 】

共有キャッシュ領域8Aに記憶される第1制御情報D1についてパージ制御の動作を述べたが、排他キャッシュ領域8Bに記憶される第2制御情報D2についても、同様に理解できる。

【 0 0 6 5 】

但し、第2制御情報D2の場合、オーナーマイクロプロセッサ6(# 0)のみがキャッシュ可能である。従って、オーナーマイクロプロセッサ6(# 0)が共有メモリ4内の第2制御情報D2を更新した場合でも、その他のマイクロプロセッサ6(# 1)にパージ要求用パージメッセージを送信する必要はない。つまり、オーナーマイクロプロセッサ6(# 0)は、第2制御情報D2に関するパージメッセージを作成したり、それを保存したりする必要はない。

20

【 0 0 6 6 】

その他のマイクロプロセッサ6(# 1)は、共有メモリ4内の第2制御情報D2に直接アクセスして更新することができる。第2制御情報D2は、オーナーマイクロプロセッサ6(# 0)のローカルメモリ7(# 0)にコピーされている。従って、その他のマイクロプロセッサ6(# 1)が第2制御情報D2を更新した場合、その他のマイクロプロセッサ6(# 1)からオーナーマイクロプロセッサ(# 0)に向けて、パージ要求用パージメッセージを送信する。

【 0 0 6 7 】

このように構成される本実施形態によれば、以下の効果を奏する。第1に、本実施形態では、共有メモリ4内の制御情報への一連の更新が完了した後で、パージ要求用パージメッセージを各マイクロプロセッサ6間で送受信させる。従って、パージ要求用パージメッセージの通信頻度を少なくでき、パージ処理に要する負荷を低減できる。この結果、マイクロプロセッサ6の演算資源等をホスト2からのコマンドを処理するために用いることができ、記憶制御装置1の処理性能を高めることができる。

30

【 0 0 6 8 】

第2に、本実施形態では、制御情報の同一箇所または連続する箇所が更新された場合、それら複数の更新を一つのパージ要求用パージメッセージにまとめる。従って、パージ要求用パージメッセージの通信頻度を低減でき、記憶制御装置1の処理性能を高めることができる。

40

【 0 0 6 9 】

第3に、本実施形態では、LM更新管理部7C内のカウンタ値とSM更新値記憶部7D内のカウンタ値とを比較することにより、ローカルメモリ7にキャッシュされている制御情報の有効性を判断できる。従って、古い制御情報を誤って使用することを防止し、記憶制御装置の動作の一貫性を保つことができる。

【 0 0 7 0 】

第4に、本実施形態では、共有メモリ4に記憶される制御情報の属性として、共有キャッシュと、排他キャッシュ及び非キャッシュとを用意しており、それら属性に応じてパージ制御を行う。従って、パージメッセージの作成及び送信に関する処理を少なくすること

50

ができ、各マイクロプロセッサ 6 の演算資源等をより一層有効に利用できる。以下、本実施形態についてさらに詳細に説明する。

【実施例 1】

【0071】

図 2 は、記憶制御装置 10 を含む情報処理システムの全体を示す。本実施例では、制御情報を分類しない場合について述べる。先に図 1 との対応関係を説明する。記憶制御装置 10 は記憶制御装置 1 に、ホスト 30 はホスト 2 に、共有メモリ 131 は共有メモリ 4 に、マイクロプロセッサ 121 はマイクロプロセッサ 6 に、ローカルメモリ 122 はローカルメモリ 7 に、対応する。図 3 に示す制御情報 D10L は制御情報 7A に、図 3 に示すページバッファ T10 はページバッファ 7B に、図 3 に示す LM 更新クロック T11 は LM 更新管理部 7C に、図 3 に示すターゲット更新クロック T12 は SM 更新値記憶部 7D に、図 3 に示す SM 更新クロック T13 は、SM 更新管理部 9 に、対応する。図 9 の第 1 制御情報 D11 は第 1 制御情報 D1 に、図 9 の第 2 制御情報 D12 は第 2 制御情報 D2 に、図 9 の第 3 制御情報 D13 は第 3 制御情報 D3 に、対応する。

10

【0072】

記憶制御装置 10 は、例えば、フロントエンドパッケージ 110 (図中、FEPK110) と、マイクロプロセッサパッケージ 120 (図中、MPPK120) と、メモリパッケージ 130 (図中、メモリ PK130) と、バックエンドパッケージ 140 (図中、BEPK140) と、スイッチ 160 と、サービスプロセッサ 170 (図中、SV P170) とを備える。

20

【0073】

フロントエンドパッケージ 110 は、ホスト 30 との通信を担当するための制御基板である。フロントエンドパッケージ 110 は、複数の通信ポート 111 を備える。通信ポート 111 は、通信ネットワーク 41 を介してホスト 30 に接続される。通信ネットワーク 41 としては、例えば、FC__SAN (Fibre Channel_Storage Area Network) や IP __SAN (Internet Protocol_SAN) を用いることができる。

【0074】

FC__SAN の場合、ホスト 30 とフロントエンドパッケージ 110 とは、ファイバチャネルプロトコルに従ってデータ通信を行う。ホスト 30 がメインフレームの場合は、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、データ通信が行われる。IP __SAN の場合、ホスト 30 とフロントエンドパッケージ 110 とは、TCP / IP (Transmission Control Protocol / Internet Protocol) 等のプロトコルに従って、データ通信を行う。

30

【0075】

バックエンドパッケージ 140 は、記憶装置 151 との通信を担当するための制御基板である。バックエンドパッケージ 140 は、複数の通信ポート 141 を備えており、各通信ポート 141 は各記憶装置 151 に接続されている。

【0076】

40

メモリパッケージ 130 は、共有メモリ領域 131 と、キャッシュメモリ領域 132 とを備える。共有メモリ領域 131 には、例えば、ホスト 30 から受信したコマンドや制御の各種情報等が記憶される。キャッシュメモリ領域 132 には、例えば、ユーザデータやキャッシュメモリ領域 132 を管理するためのテーブル等が記憶される。以下の説明では、キャッシュメモリ領域 132 をキャッシュメモリ 132 と呼ぶ。

【0077】

マイクロプロセッサパッケージ 120 は、モジュール 100 の動作を制御するための制御基板である。マイクロプロセッサパッケージ 120 は、マイクロプロセッサ 121 とローカルメモリ 122 とを備える。

【0078】

50

マイクロプロセッサパッケージ 120 は、ホスト 30 から発行されたコマンドを実行し、その実行結果をホスト 30 に送信する。フロントエンドパッケージ 110 がリードコマンドを受信した場合、マイクロプロセッサパッケージ 120 は、要求されたデータをキャッシュメモリ 132 または記憶装置 151 から取得し、ホスト 30 に送信させる。

【0079】

フロントエンドパッケージ 110 がライトコマンドを受信した場合、マイクロプロセッサパッケージ 120 は、ライトデータをキャッシュメモリ 132 に書き込み、ホスト 30 に処理完了を通知する。キャッシュメモリ 132 に書き込まれたデータは、その後、記憶装置 151 に書き込まれる。マイクロプロセッサパッケージ 120 の構成については、図 3 でさらに詳述する。

【0080】

記憶装置 151 は、ハードディスクドライブまたはフラッシュメモリデバイス等の記憶装置から構成される。複数の記憶装置 151 を一つのグループ 150 にまとめることができる。このグループ 150 は、例えば、RAID グループまたはパリティグループと呼ばれる。そして、グループ化された記憶領域には、一つまたは複数の論理ボリューム 152 が形成される。

【0081】

記憶装置 151 としては、例えば、ハードディスクデバイス、半導体メモリデバイス、光ディスクデバイス、光磁気ディスクデバイス、磁気テープデバイス、フレキシブルディスクデバイス等のデータを読み書き可能な種々のデバイスを利用可能である。

【0082】

記憶装置 151 としてハードディスクデバイスを用いる場合、例えば、FC (Fibre Channel) ディスク、SCSI (Small Computer System Interface) ディスク、SATA ディスク、ATA (AT Attachment) ディスク、SAS (Serial Attached SCSI) ディスク等を用いることができる。

【0083】

記憶装置 151 として半導体メモリデバイスを用いる場合、例えば、フラッシュメモリ、FeRAM (Ferroelectric Random Access Memory)、MRAM (Magnetoresistive Random Access Memory)、相変化メモリ (Ovonic Unified Memory)、RRAM (Resistance RAM) 等の種々のメモリデバイスを利用可能である。なお、記憶デバイスの種類は、上記のものに限定されず、将来製品化される他の種類の記憶デバイスを利用することもできるであろう。

【0084】

スイッチ 160 は、各パッケージ 110, 120, 130, 140 を接続するための回路である。フロントエンドパッケージ 110、バックエンドパッケージ 140、マイクロプロセッサパッケージ 130 は、スイッチ 160 を介して、メモリパッケージ 130 にアクセスする。

【0085】

SV P 170 は、記憶制御装置 10 内の各種情報を収集して管理端末 20 に提供したり、管理端末 20 から入力された設定値等を共有メモリ 131 に記憶させたりするための制御回路である。SV P 170 は、例えば、LAN (Local Area Network) 等の通信ネットワーク 42 を介して、管理端末 20 に接続されている。

【0086】

図 3 は、マイクロプロセッサパッケージ 120 の構成を模式的に示す図である。ローカルメモリ 122 には、制御情報 D10L と、ページバッファ T10 と、LM 更新クロック T11 と、ターゲット更新クロック T12 とが記憶される。

【0087】

制御情報 D10L は、共有メモリ 131 内の制御情報 D10 の一部をコピーしたものである。ページバッファ T10 は、ページメッセージを記憶するものである。ページメッセージについては、図 4 で後述する。LM 更新クロック T11 は、マイクロプロセッサ 12

10

20

30

40

50

1 による更新をローカルメモリ側で管理するためのものである。ターゲット更新クロック T 1 2 には、所定のタイミングで、共有メモリ 1 3 1 内の S M 更新クロック T 1 3 の各カウンタ値がコピーされる。

【 0 0 8 8 】

共有メモリ 1 3 1 には、制御情報記憶領域 1 3 1 0 と S M 更新クロック T 1 3 とが設けられる。制御情報記憶領域 1 3 1 0 には、制御情報 D 1 0 が記憶される。S M 更新クロック T 1 3 は、各マイクロプロセッサ 1 2 1 により更新された回数を示すカウンタ値を、各マイクロプロセッサ 1 2 1 を識別するためのマイクロプロセッサ番号毎にそれぞれ管理している。各 L M 更新クロック T 1 1 及び各ターゲット更新クロック T 1 2 も、マイクロプロセッサ番号とカウンタ値とを対応付けて管理する。つまり、各更新クロック T 1 1 , T 1 2 , T 1 3 は、同一の構造を有する。

10

【 0 0 8 9 】

図 4 は、ページバッファの構成を示す図である。ページバッファ T 1 0 は、例えば、メッセージタイプ欄 C 1 0 0 と、第 1 データ欄 C 1 0 1 と、第 2 データ欄 C 1 0 2 とを備えている。

【 0 0 9 0 】

メッセージタイプ欄 C 1 0 0 には、ページメッセージが、ページ要求用メッセージであるか、それとも、更新通知用メッセージであるかを示すタイプが格納される。メッセージタイプ 0 は、ページ要求用ページメッセージであることを示す。

【 0 0 9 1 】

20

ページ要求用ページメッセージの場合、第 1 データ欄 C 1 0 1 には、更新されたデータのサイズが格納される。例えば、データサイズとしてライン数が設定される。1 ラインは、例えば 6 4 バイトである。ページ要求用ページメッセージの場合、第 2 データ欄 C 1 0 2 には、共有メモリ 1 3 1 内の記憶先アドレス（開始アドレス）が格納される。つまり、ページ要求用ページメッセージを用いれば、第 2 データ欄 C 1 0 2 で示される開始アドレスから、第 1 データ欄 C 1 0 1 で示されるサイズまでのデータ（制御情報 D 1 0 ）が更新された事がわかる。

【 0 0 9 2 】

メッセージタイプ 1 は、更新通知用ページメッセージであることを示す。更新通知用ページメッセージの場合、第 1 データ欄 C 1 0 1 には、その更新通知用ページメッセージを発行する発行元マイクロプロセッサ 1 2 1 のマイクロプロセッサ番号（M P #）が格納される。第 2 データ欄 C 1 0 2 には、発行元マイクロプロセッサ 1 2 1 側で管理されている L M 更新クロック T 1 1 内の各カウンタ値のうち、発行元マイクロプロセッサ 1 2 1 のカウンタ値が格納される。

30

【 0 0 9 3 】

更新元（後の発行元）マイクロプロセッサ 1 2 1 が共有メモリ 1 3 1 内の制御情報を更新するたびに、メッセージタイプ 0 のページメッセージが作成されて、ページバッファ T 1 0 に登録される。一連の更新が完了すると、メッセージタイプ 1 のページメッセージが作成されて、ページバッファ T 1 0 に登録される。

【 0 0 9 4 】

40

図 5 は、複数のページ要求用ページメッセージを一つのページメッセージに統合する様子を示す図である。一方のページメッセージ M S G 1 の更新箇所と他方のページメッセージ M S G 2 の更新箇所とが、同一、または、連続、または、重複する場合、それらページメッセージ M S G 1 , 2 を一つのページメッセージ M S G 3 に統合して、ページバッファ T 1 0 に記憶させる。これにより、ページメッセージの数を低減して、ページメッセージの通信頻度を少なくできる。

【 0 0 9 5 】

図 6 は、L M 更新クロック T 1 1 とターゲット更新クロック T 1 2 及び S M 更新クロック T 1 3 の関係を示す図である。L M 更新クロック T 1 1 には、各マイクロプロセッサパッケージ 1 2 0 側で管理される各カウンタ値が記憶される。

50

【0096】

LM更新クロックT11を管理するマイクロプロセッサ121のカウント値は、そのマイクロプロセッサ121による制御情報D10の更新とリアルタイムで更新される。例えば、マイクロプロセッサ121(#0)の管理するLM更新クロックT11の場合、マイクロプロセッサ121(#0)のカウント値は、更新処理とほぼ同時にLM更新クロックT11に記録される。他のマイクロプロセッサ121(#1-#3)の各カウント値は、他のマイクロプロセッサ121(#1-#3)から発行される更新通知用パージメッセージに基づいて更新される。

【0097】

なお、図2、図7等ではマイクロプロセッサ121を2個示し、図4ではマイクロプロセッサ121が少なくとも8個ある場合を示し、図6ではマイクロプロセッサ121が4個ある場合を示しているが、説明の便宜上の例示に過ぎない。本発明は、複数のマイクロプロセッサ121が共有メモリ131を共有する構成の場合に有用である。混乱を避けるために、第1実施例では、マイクロプロセッサ121が2個の場合を中心に述べる。

【0098】

図6に示すターゲット更新クロックT12は、上述の通り、SM更新クロックT13のコピーである。例えば、マイクロプロセッサ121が、ローカルメモリ7にキャッシュされている制御情報D10L以外の制御情報を使用する場合(キャッシュミス時)、マイクロプロセッサ121は共有メモリ131内の制御情報D10にアクセスする。マイクロプロセッサ121が共有メモリ131にアクセスした場合に、共有メモリ131内のSM更新クロックT13の各カウント値が読み出されて、ターゲット更新クロックT12にコピーされる。

【0099】

図6に示すSM更新クロックT13は、各マイクロプロセッサ121による制御情報D10の更新回数を管理する。上述の通り、更新元マイクロプロセッサ121で管理されているLM更新クロックT11内の更新元マイクロプロセッサ121のカウント値と、SM更新クロックT13内の更新元マイクロプロセッサ121のカウント値とは、ほぼ同時期に更新される。そして、他の各マイクロプロセッサ121が共有メモリ131にアクセスした場合、SM更新管理クロックT13内の各カウント値が、他の各マイクロプロセッサ121で管理されているターゲット更新クロックT12にコピーされる。

【0100】

図7は、更新クロックの大小を判定する方法を模式的に示す図である。上述の通り、各更新クロックT11、T12、T13は、マイクロプロセッサ番号とカウント値とを対応付ける、同一構造を有する。そして、LM更新クロックT11に記憶されている各カウント値と、ターゲット更新クロックT12に記憶されている各カウント値との大小関係に基づいて、ローカルメモリ122にコピーされている制御情報D10Lが有効であるか否かが判定される。

【0101】

各更新クロックの大小関係は、次のように判定される。図7に示す更新クロックC1-C3を例に挙げて説明する。更新クロックC1と更新クロックC2との関係のように、更新クロックC2の各カウント値が更新クロックC1の対応する各カウント値以上の場合に、C2=C1であると判定する。なお、同一マイクロプロセッサ番号について、カウント値の大小を比較する。

【0102】

同様に、更新クロックC1と更新クロックC3との関係に着目すると、更新クロックC3の各カウント値は、更新クロックC1の対応する各カウント値以上であるため、C3=C1と判定される。

【0103】

更新クロックC2と更新クロックC3の関係に着目すると、マイクロプロセッサ番号#0の場合、更新クロックC2のカウント値の方が更新クロックC3の対応するカウント値

10

20

30

40

50

よりも大きい。しかし、別のマイクロプロセッサ番号 # 1 の場合、更新クロック C 2 のカウンタ値は、更新クロック C 3 の対応するカウンタ値よりも小さい。従って、更新クロック C 2 と更新クロック C 3 の関係は、不定であり、いずれが大でいずれが小であるか区別することはできない。

【 0 1 0 4 】

図 8 は、パージ制御の動作を示すフローチャートである。共有メモリ 1 3 1 のアドレス X には、制御情報 D 1 0 の一部としてのデータ " 0 " が記憶されている (S 1 0)。マイクロプロセッサ 1 2 1 (# 0) は、共有メモリ 1 3 1 のアドレス X からデータ " 0 " を読み出す (S 1 1)。読み出されたデータ " 0 " は、ローカルメモリ 1 2 2 (# 0) に記憶される。

10

【 0 1 0 5 】

この時点における、LM 更新クロック (LM - CLK) T 1 1 (# 0) の値は、(0 , 0) である。先の " 0 " は、マイクロプロセッサ 1 2 1 (# 0) のカウンタ値であり、後の " 0 " は、マイクロプロセッサ 1 2 1 (# 1) のカウンタ値である。ターゲット更新クロック (TGT - CLK) T 1 2 (# 0) の値も (0 , 0) である。

【 0 1 0 6 】

マイクロプロセッサ 1 2 1 (# 0) が再びアドレス X のデータを参照しようとする場合には、ローカルメモリ 1 2 2 (# 0) にキャッシュされているデータ " 0 " が使用される (S 1 2)。LM 更新クロックのカウンタ値とターゲット更新クロックのカウンタ値とを比較すると、LM - CLK TGT - CLK の関係が成り立つため、ローカルメモリ 1 2 2 (# 0) に記憶されているデータ " 0 " を使用できる。

20

【 0 1 0 7 】

別のマイクロプロセッサ 1 2 1 (# 1) がアドレス X にデータ " 8 " を書き込む場合を説明する (S 1 3)。その更新元マイクロプロセッサ 1 2 1 (# 1) が管理する LM 更新クロック T 1 1 (# 1) 内の更新元マイクロプロセッサ 1 2 1 (# 1) のカウンタ値は、1 つだけインクリメントされる。従って、LM - CLK (LM 更新クロック) は、(0 , 0) から (0 , 1) に変化する。共有メモリ 1 3 1 内の SM 更新クロック (SM - CLK) T 1 3 のカウンタ値も、(0 , 0) から (0 , 1) に変化する (S 1 4)。

【 0 1 0 8 】

マイクロプロセッサ 1 2 1 (# 0) が共有メモリ 1 3 1 の別アドレス Y から制御情報 D 1 0 の一部としてのデータを読み出す場合 (S 1 5)、マイクロプロセッサ 1 2 1 (# 0) は、SM 更新クロック T 1 3 の各カウンタ値 (0 , 1) も取得する。これにより、マイクロプロセッサ 1 2 1 (# 0) が管理しているターゲット更新クロック T 1 3 のカウンタ値は、(0 , 0) から (0 , 1) に変化する (S 1 6)。

30

【 0 1 0 9 】

別のマイクロプロセッサ 1 2 1 (# 1) が共有メモリ 1 3 1 のアドレス X にデータ " 1 0 " を書き込む場合を説明する (S 1 7)。その更新元マイクロプロセッサ 1 2 1 が管理している LM 更新クロック T 1 1 (# 1) のカウンタ値は、(0 , 1) から (0 , 2) に変化する。2 回目の更新だからである。更新元マイクロプロセッサ 1 2 1 (# 1) がアドレス X にデータ " 1 0 " を書き込む際に、SM 更新クロック T 1 3 のカウンタ値は (0 , 1) から (0 , 2) に更新される (S 1 8)。なお、ターゲット更新クロック T 1 3 (# 1) のカウンタ値は、(0 , 0) のまま変化していない。

40

【 0 1 1 0 】

マイクロプロセッサ 1 2 1 (# 0) が、アドレス X のデータを再び参照しようとする場合、ローカルメモリ 1 2 2 (# 0) にキャッシュされているデータ " 0 " が有効であるか否か (最新であるか否か) を判定する。マイクロプロセッサ 1 2 1 (# 0) は、LM 更新クロック T 1 1 (# 0) のカウンタ値 (0 , 0) とターゲット更新クロック T 1 3 (# 0) のカウンタ値 (0 , 1) とを比較する。

【 0 1 1 1 】

LM 更新クロック T 1 1 (# 0) のカウンタ値 (0 , 0) は、ターゲット更新クロック

50

T 1 3 (# 0) のカウンタ値 (0 , 1) よりも小さいため、ローカルメモリ 1 2 2 (# 0) にキャッシュされているデータ " 0 " は古くて使用できないと判定される。つまり、別のマイクロプロセッサ 1 2 1 (# 1) による更新を示すパージメッセージがマイクロプロセッサ 1 2 1 (# 0) に届いていないと判断される。

【 0 1 1 2 】

そこで、マイクロプロセッサ 1 2 1 (# 0) は、共有メモリ 1 3 1 にアクセスして、アドレス X から最新データ " 1 0 " を読み出す (S 1 9)。読み出された最新データ " 1 0 " は、ローカルメモリ 1 2 2 (# 0) に記憶される。マイクロプロセッサ 1 2 1 (# 0) が共有メモリ 1 3 1 のアドレス X からデータ " 1 0 " を読み出す際に、S M 更新クロック T 1 3 のカウンタ値 (0 , 2) も読み出される。これにより、マイクロプロセッサ 1 2 1 (# 0) が管理しているターゲット更新クロック T 1 3 (# 0) のカウンタ値は、(0 , 1) から (0 , 2) に更新される (S 2 0)。

10

【 0 1 1 3 】

別のマイクロプロセッサ 1 2 1 (# 1) は、一連の更新処理 (S 1 3 , S 1 7) が完了すると、更新通知用パージメッセージを作成してパージバッファ T 1 0 (# 1) に登録する。その後、パージバッファ T 1 0 (# 1) に登録された全てのパージメッセージを、マイクロプロセッサ 1 2 1 (# 0) に送信する (S 2 1)。マイクロプロセッサ 1 2 1 (# 1) は、共有メモリ 1 3 1 内のデータを 2 回更新しているが (S 1 3 , S 1 7)、最終的に 1 つのパージ要求用パージメッセージに統合される (Type=0, line=1, adr=X)。

【 0 1 1 4 】

20

マイクロプロセッサ 1 2 1 (# 0) は、マイクロプロセッサ 1 2 1 (# 1) からのパージメッセージに基づいて、L M 更新クロック T 1 1 (# 0) のカウンタ値を (0 , 0) から (0 , 2) に更新させる (S 2 2)。

【 0 1 1 5 】

本実施例によれば、共有メモリ 1 3 1 内の制御情報 D 1 0 への一連の更新が完了した後で、パージ要求用パージメッセージを発行させる。従って、パージ要求用パージメッセージの通信頻度を少なくでき、パージ処理に要する負荷を低減できる。この結果、記憶制御装置 1 0 の処理性能を高めることができる。

【 0 1 1 6 】

本実施例では、制御情報への複数回の更新を一つのパージ要求用パージメッセージに統合させる。従って、パージ要求用パージメッセージの通信頻度を低減でき、記憶制御装置 1 0 の処理性能を高めることができる。

30

【 0 1 1 7 】

本実施例では、L M 更新クロック T 1 1 のカウンタ値とターゲット更新クロック T 1 2 のカウンタ値とを比較することにより、ローカルメモリ 1 2 2 にキャッシュされている制御情報 D 1 0 L の有効性を判断できる。従って、古い制御情報を誤って使用することを防止し、記憶制御装置の動作の一貫性を高めることができる。

【 実施例 2 】

【 0 1 1 8 】

図 9 - 図 1 5 に基づいて第 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 実施例の変形例に相当する。従って、以下の各実施例では、第 1 実施例との相違点を中心に説明する。本実施例では、制御情報を、その技術的性質に基づいて複数種類に分類し、それぞれの種類に適した制御を実行する。つまり、本実施例では、制御情報に所定のキャッシュ属性を設定して管理する。

40

【 0 1 1 9 】

図 9 は、本実施例による記憶制御装置 1 0 の要部を示す図である。本実施例の共有メモリ 1 3 1 に着目すると、制御情報記憶領域 1 3 1 0 には、さらに 3 つの領域 1 3 1 1 , 1 3 1 2 , 1 3 1 3 が設定されている。

【 0 1 2 0 】

第 1 の領域 1 3 1 1 は、共有キャッシュ領域である。共有キャッシュ領域 1 3 1 1 には

50

、第 1 制御情報 D 1 1 が記憶される。第 1 制御情報 D 1 1 は、各マイクロプロセッサ 1 2 1 によって共有され、各ローカルメモリ 1 2 2 にそれぞれキャッシュされる。

【 0 1 2 1 】

図 1 1 に示すように、第 1 制御情報 D 1 1 としては、例えば、記憶制御装置 1 0 の構成を管理する情報 D 1 1 0、プログラム製品の設定情報 D 1 1 1 等のように、更新頻度が比較的少なく、かつ、複数のマイクロプロセッサ 1 2 1 から比較的高頻度で参照され得る情報が該当する。

【 0 1 2 2 】

第 2 の領域 1 3 1 2 は、排他キャッシュ領域である。排他キャッシュ領域 1 3 1 2 には、第 2 制御情報 D 1 2 が記憶される。第 2 制御情報 D 1 2 は、オーナーとなる特定のマイクロプロセッサ 1 2 1 のみがキャッシュでき、他のマイクロプロセッサ 1 2 1 はキャッシュすることができない。他のマイクロプロセッサ 1 2 1 は、共有メモリ 1 3 1 に直接アクセスして第 2 制御情報 D 1 2 を使用する。

10

【 0 1 2 3 】

図中では、一つの排他キャッシュ領域 1 3 1 2 のみを示すが、排他キャッシュ領域 1 3 1 2 は、各マイクロプロセッサ 1 2 1 毎に設けることもできる。つまり、各マイクロプロセッサ 1 2 1 がオーナーとなる排他キャッシュ領域をそれぞれ設定できる。

【 0 1 2 4 】

第 2 制御情報 D 1 2 としては、図 1 1 に示すように、例えば、論理ボリューム 1 5 2 を管理する情報 D 1 2 0、リモートコピーの差分データを管理する情報 D 1 2 1 等のように、特定のマイクロプロセッサ 1 2 1 が比較的高頻度で更新する情報が該当する。

20

【 0 1 2 5 】

第 3 の領域 1 3 1 3 は、非キャッシュ領域である。非キャッシュ領域 1 3 1 3 には、第 3 制御情報 D 1 3 が記憶される。第 3 制御情報 D 1 3 は、いずれのマイクロプロセッサ 1 2 1 もキャッシュできない。

【 0 1 2 6 】

第 3 制御情報 D 1 3 としては、図 1 1 に示すように、例えば、記憶制御装置 1 0 内で実行される各種ジョブを管理する情報 D 1 3 0、障害に関する管理情報等 D 1 3 1 のように、更新頻度が比較的少ない情報が該当する。但し、第 1 - 第 3 制御情報の例は、上記に限られない。

30

【 0 1 2 7 】

上述のキャッシュ属性（共有キャッシュ、排他キャッシュ、非キャッシュ）は、キャッシュ属性管理テーブル T 1 4 によって管理される。キャッシュ属性管理テーブル T 1 4 は、共有メモリ 1 3 1 に記憶されている。各ローカルメモリ 1 2 2 には、共有メモリ 1 3 1 に記憶されたキャッシュ属性管理テーブル T 1 4 のコピーが設けられる。

【 0 1 2 8 】

図 1 0 の上部は、キャッシュ属性管理テーブル T 1 4 を示す図である。キャッシュ属性管理テーブル T 1 4 は、例えば、S M アドレス（共有メモリ）の上位を示す欄 C 1 4 0 と、エリア属性欄 C 1 4 1 とを備える。図 1 0 の下部は、エリア属性の関係を模式的に示す図である。

40

【 0 1 2 9 】

本実施例では、S M アドレスの上位ビット毎に S M のキャッシュ属性を設定するようになっている。図 1 0 に示す例では、S M アドレスを 3 2 ビットとすると、0x00000000-0x0001FFFF が非キャッシュ、0x00020000-0x0002FFFF が共有キャッシュ、0x00030000-0x0003FFFF が排他キャッシュ（オーナー M P は # 0 ）となる。共有メモリ 1 3 1 にリードまたはライトする場合、図 1 0 上部に示すテーブル T 1 4 を参照して、アクセスしようとするアドレスのエリア属性を判断する。各制御情報 D 1 1 , D 1 2 , D 1 3 は、ひとかたまりずつの連続領域として存在するのではなく、テーブル T 1 4 で決定されているアドレス範囲ごとに自由に設定できる。

【 0 1 3 0 】

50

共有キャッシュの属性を有する第1制御情報D11は、複数のマイクロプロセッサ121によりキャッシュされる。排他キャッシュの属性を有する第2制御情報D12の場合、オーナーマイクロプロセッサは第2制御情報D12をキャッシュ可能であるが、オーナーマイクロプロセッサ以外の他のマイクロプロセッサは、第2制御情報D12をキャッシュすることができない。非キャッシュの属性を有する第3制御情報D13は、いずれのマイクロプロセッサ121もキャッシュすることができない。

【0131】

キャッシュの可否とパージメッセージの送信可否とは密接に関係する。第1制御情報D11は、全てのマイクロプロセッサ121にキャッシュが許可されている。各マイクロプロセッサ121は、共有メモリ131内の第1制御情報D11を更新すると、パージメッセージを作成し、所定の時点で他の各マイクロプロセッサ121に送信する。

10

【0132】

排他キャッシュの場合を説明する。オーナーマイクロプロセッサは、共有メモリ131内の第2制御情報D12を更新した場合でも、パージメッセージを作成したり、パージメッセージを他の各マイクロプロセッサに送信したりする必要はない。第2制御情報D12をキャッシュできるのは、オーナーマイクロプロセッサのみだからである。

【0133】

これに対し、オーナーマイクロプロセッサ以外の他のマイクロプロセッサ121が共有メモリ131内の第2制御情報D12を更新した場合、他のマイクロプロセッサ121はパージメッセージを作成して、オーナーマイクロプロセッサにのみ送信する。オーナーマイクロプロセッサは、共有メモリ131内の第2制御情報D12をローカルメモリ122にコピーして使用しているためである。

20

【0134】

非キャッシュの場合を説明する。各マイクロプロセッサ121は、非キャッシュの属性を有する第3制御情報D13をローカルメモリ122にコピーして使用することができないようになっている。従って、共有メモリ131内の第3制御情報D13が更新された場合でも、パージメッセージを作成して送信する必要はない。

【0135】

図12は、共有メモリ131にデータ(制御情報)を書き込む場合の処理を示すフローチャートである。マイクロプロセッサ121は、ライトアドレスに基づいてキャッシュ属性管理テーブルT14を参照する(S51)。共有メモリ131内のデータを更新しようとするマイクロプロセッサを、更新元マイクロプロセッサと呼ぶ。

30

【0136】

マイクロプロセッサ121は、ライトアクセスが共有キャッシュ領域1311に含まれるか否かを判定する(S52)。共有キャッシュ領域1311への書込みである場合(S52: YES)、つまり、第1制御情報D11を更新する場合、マイクロプロセッサ121は、自分の管理しているLM更新クロックT11内の対応するカウンタ値を1つインクリメントさせる(S53)。即ち、更新元のマイクロプロセッサ121は、更新元マイクロプロセッサ121が管理するLM更新クロックT11内の更新元マイクロプロセッサ121のカウンタ値を1つ増加させる。

40

【0137】

さらに、更新元マイクロプロセッサ121は、更新元マイクロプロセッサ121が管理しているパージバッファT10にライトアクセス及びライトデータのサイズを登録する(S54)。

【0138】

更新元マイクロプロセッサ121は、共有メモリ131に新しいデータを書込み(S55)、SM更新クロックT13内の更新元マイクロプロセッサ121のカウンタ値を1つインクリメントさせる(S56)。

【0139】

更新元マイクロプロセッサ121は、更新対象のデータ(更新対象の制御情報D11)

50

がローカルメモリ 1 2 2 にコピーされているか否かを判定する (S 5 7)。更新対象データがローカルメモリ 1 2 2 にコピーされている場合 (S57:YES)、つまり、キャッシュヒットの場合 (S57:YES)、更新元マイクロプロセッサ 1 2 1 は、ローカルメモリ 1 2 2 にコピーされている制御情報を更新させる (S 5 8)。

【 0 1 4 0 】

キャッシュミスの場合 (S57:NO)、本処理を終了する。なお、キャッシュミス時に、共有メモリ 1 3 1 内で更新された制御情報 D 1 1 を読み出して、ローカルメモリ 1 2 2 にコピーさせる構成でもよい。その場合、更新元マイクロプロセッサ 1 2 1 が管理しているターゲット更新クロック T 1 2 のカウンタ値は、S M 更新クロック T 1 3 のカウンタ値で更新される。

10

【 0 1 4 1 】

そのライトアドレスが共有キャッシュ領域内に存在しない場合 (S52:NO)、更新元マイクロプロセッサ 1 2 1 は、そのライトアドレスが更新元マイクロプロセッサ 1 2 1 がオーナーである排他キャッシュ領域 1 3 1 2 内に存在するか否かを判定する (S 5 9)。

【 0 1 4 2 】

更新元マイクロプロセッサ 1 2 1 がオーナーである排他キャッシュ領域 1 3 1 2 内の第 2 制御情報 D 1 2 を更新する場合 (S59:YES)、パージメッセージを作成したり保存したりする必要はない。そこで、S 5 3 , S 5 4 をスキップして S 5 5 に移る。

【 0 1 4 3 】

そのライトアドレスが、更新元マイクロプロセッサ 1 2 1 がオーナーである排他キャッシュ領域 1 3 1 2 内に存在しない場合 (S59:NO)、更新元マイクロプロセッサ 1 2 1 は、そのライトアドレスが更新元マイクロプロセッサ 1 2 1 がオーナーではない排他キャッシュ領域内に存在するか否かを判定する (S 6 0)。

20

【 0 1 4 4 】

更新元マイクロプロセッサ 1 2 1 がオーナーではない排他キャッシュ領域内の第 2 制御情報 D 1 2 を更新する場合 (S60:YES)、更新元マイクロプロセッサ 1 2 1 は、「オーナーマイクロプロセッサ以外の他のマイクロプロセッサ」となる。

【 0 1 4 5 】

更新元マイクロプロセッサ 1 2 1 は、更新元マイクロプロセッサ 1 2 1 が管理している L M 更新クロック T 1 1 内の更新元マイクロプロセッサ 1 2 1 のカウンタ値を 1 つインクリメントさせる (S 6 1)。さらに、更新元マイクロプロセッサ 1 2 1 は、パージバッファ T 1 0 に、ライトアドレス及びライトデータのサイズを登録する (S 6 2)。

30

【 0 1 4 6 】

更新元マイクロプロセッサ 1 2 1 は、共有メモリ 1 3 1 内の第 2 制御情報 D 1 2 を更新し (S 6 3)、S M 更新クロック T 1 3 内の更新元マイクロプロセッサ 1 2 1 のカウンタ値を 1 つインクリメントさせる (S 6 4)。

【 0 1 4 7 】

そのライトアドレスが、更新元マイクロプロセッサ 1 2 1 がオーナーマイクロプロセッサではない排他キャッシュ領域内にも存在しない場合 (S60:NO)、そのライトアドレスは、非キャッシュ領域 1 3 1 3 内に存在する。従って、更新元マイクロプロセッサ 1 2 1 は、S 6 1 , S 6 2 をスキップして S 6 3 に移る。非キャッシュ領域 1 3 1 3 に属する第 3 制御情報 D 1 3 は、いずれのマイクロプロセッサ 1 2 1 もキャッシュできないため、パージメッセージを作成したり保存したりする必要はない。

40

【 0 1 4 8 】

図 1 4 は、パージメッセージを他のマイクロプロセッサ 1 2 1 に送信する処理を示すフローチャートである。ここでは、パージメッセージを送信するマイクロプロセッサを発行元マイクロプロセッサと呼ぶ。

【 0 1 4 9 】

発行元マイクロプロセッサ 1 2 1 は、パージメッセージの送信契機が到来したか否かを判定する (S 7 0)。例えば、共有メモリ 1 3 内に記憶されている制御情報について一連

50

の更新が全て完了した場合に、その更新に関するプログラムからの明示的な指示によって、バージョンメッセージが送信される。または、例えば、バージョンバッファ T 1 0 に保存されたバージョンメッセージ数が所定の上限値に達した場合に、バージョンメッセージが送信される。その他の送信契機を採用してもよい。

【 0 1 5 0 】

送信契機が到来すると (S70:YES)、発行元マイクロプロセッサ 1 2 1 は、更新通知用バージョンメッセージを作成する (S 7 1)。つまり、発行元マイクロプロセッサ 1 2 1 は、バージョンバッファ T 1 0 に、発行元マイクロプロセッサ 1 2 1 のマイクロプロセッサ番号と、LM更新クロック T 1 1 内の発行元マイクロプロセッサ 1 2 1 のカウンタ値とを、登録する (S 7 1)。

10

【 0 1 5 1 】

そして、発行元マイクロプロセッサ 1 2 1 は、バージョンバッファ T 1 0 内の各バージョンメッセージを、他の各マイクロプロセッサ 1 2 1 に向けて送信する (S 7 2)。但し、第 2 制御情報 D 1 2 に関するバージョンメッセージの場合、発行元マイクロプロセッサ 1 2 1 が、オーナーマイクロプロセッサ以外の他のマイクロプロセッサのときは、オーナーマイクロプロセッサにのみバージョンメッセージを送信する。

【 0 1 5 2 】

図 1 5 は、共有メモリ 1 3 1 内からデータ (制御情報) を読み出す場合の処理を示すフローチャートである。データを読み出すマイクロプロセッサを、ここでは、読み出し元マイクロプロセッサと呼ぶ。

20

【 0 1 5 3 】

読み出し元マイクロプロセッサ 1 2 1 は、リードアドレスに基づいてキャッシュ属性管理テーブル T 1 4 を参照する (S 1 0 0)。読み出し元マイクロプロセッサ 1 2 1 は、そのリードアドレスが共有キャッシュ領域 1 3 1 1 または排他キャッシュ領域 (オーナー) のいずれかに存在するか否かを判定する (S 1 0 1)。

【 0 1 5 4 】

ここで、排他キャッシュ領域 (オーナー) とは、マイクロプロセッサがオーナーである排他キャッシュ領域であることを示す。これに対し、排他キャッシュ領域 (その他) とは、マイクロプロセッサがオーナーでは無い排他キャッシュ領域であることを示す。

【 0 1 5 5 】

30

リードアドレスが共有キャッシュ領域 1 3 1 1 または排他キャッシュ領域 (オーナー) 1 3 1 2 のいずれかに含まれる場合 (S101:YES)、読み出し元マイクロプロセッサ 1 2 1 は、他の各マイクロプロセッサ 1 2 1 からのバージョンメッセージが届いているか否かを確認する (S 1 0 2)。

【 0 1 5 6 】

読み出し元マイクロプロセッサ 1 2 1 は、バージョンメッセージに基づいて、ローカルメモリ 1 2 2 内のキャッシュデータ (ローカルメモリ 1 2 2 内にコピーされた制御情報) の一部または全部を無効にする (S 1 0 3)。

【 0 1 5 7 】

さらに、読み出し元マイクロプロセッサ 1 2 1 は、他の各マイクロプロセッサ 1 2 1 から届けられた更新通知用バージョンメッセージに基づいて、LM更新クロック T 1 1 内のカウンタ値を更新させる (S 1 0 4)。

40

【 0 1 5 8 】

読み出し元マイクロプロセッサ 1 2 1 は、LM更新クロック T 1 1 内の各カウンタ値がターゲット更新クロック T 1 2 内の対応する各カウンタ値以上であるか否かを判定する (S 1 0 5)。LM更新クロック T 1 1 ターゲット更新クロック T 1 2 の場合 (S105:YES)、ローカルメモリ 1 2 2 にコピーされている制御情報は有効であると判断される。

【 0 1 5 9 】

読み出し元マイクロプロセッサ 1 2 1 は、ローカルメモリ 1 2 2 内に所望のデータが存在するか否かを確認する (S 1 0 6)。所望のデータがローカルメモリ 1 2 2 内に存在す

50

る場合 (S106:YES)、そのデータは有効であるから、読み出し元マイクロプロセッサ 1 2 1 は、ローカルメモリ 1 2 2 から所望のデータを読み出して使用する (S 1 0 7)。

【0 1 6 0】

L M 更新クロック T 1 1 < ターゲット更新クロック T 1 2 の場合 (S105:NO)、または、有効とされたデータがローカルメモリ 1 2 2 に記憶されていなかった場合 (S106:NO) のいずれかの場合、S 1 0 8 に移る。

【0 1 6 1】

読み出し元マイクロプロセッサ 1 2 1 は、ローカルメモリ 1 2 2 に記憶領域を確保し (S 1 0 8)、所望のデータを共有メモリ 1 3 1 から読み出して使用する (S 1 0 9)。読み出し元マイクロプロセッサ 1 2 1 は、S M 更新クロック T 1 3 の各カウンタ値を取得し、それらのカウンタ値でターゲット更新クロック T 1 2 内の各カウンタ値を更新させる (S 1 1 0)。

【0 1 6 2】

なお、リードアドレスが排他キャッシュ領域 (その他) または非キャッシュ領域に存在する場合 (S101:NO)、S 1 0 9 に移って、共有メモリ 1 3 1 から所望のデータを直接読み出す。

【0 1 6 3】

このように構成される本実施例も第 1 実施例と同様の効果を奏する。さらに本実施例では、共有メモリ 1 3 1 内に記憶される制御情報の属性として、共有キャッシュと、排他キャッシュ及び非キャッシュとを用意し、それら属性に応じてパージ制御を行う。従って、パージメッセージの作成及び送信に関する処理を少なくでき、各マイクロプロセッサ 1 2 1 の演算資源等をより一層有効に利用できる。

【実施例 3】

【0 1 6 4】

図 1 6 , 図 1 7 に基づいて第 3 実施例を説明する。本実施例では、マイクロプロセッサ 1 2 1 に障害が発生した場合の対応方法を説明する。上述の通り、本発明は、パージメッセージの通信量を低減するべく、共有メモリ 1 3 1 内の制御情報の更新時期とパージメッセージの送信時期とを異ならせている。

【0 1 6 5】

従って、例えば、何らかの障害がマイクロプロセッサ 1 2 1 に生じた場合、通知すべきパージメッセージが各マイクロプロセッサ 1 2 1 に通知されることなく、マイクロプロセッサ 1 2 1 が機能を停止する可能性がある。

【0 1 6 6】

そこで、本実施例では、障害発生時に以下の処理を実行する。障害の発生しているマイクロプロセッサを障害マイクロプロセッサと呼ぶ場合がある。最初に、マイクロプロセッサ 1 2 1 は、他の各マイクロプロセッサ 1 2 1 の状態をチェックする (S 1 2 0)。例えば、各マイクロプロセッサ 1 2 1 は、自身の状態を所定の記憶領域に書き込む。従って、その所定の記憶領域に書き込まれた値を参照することにより、他の各マイクロプロセッサ 1 2 1 の状態を知ることができる。

【0 1 6 7】

マイクロプロセッサ 1 2 1 は、障害マイクロプロセッサが検出されたか否かを判定する (S 1 2 1)。障害マイクロプロセッサが検出された場合 (S121:YES)、マイクロプロセッサ 1 2 1 は、他の各マイクロプロセッサからのパージメッセージを受信する (S 1 2 2)。さらに、マイクロプロセッサ 1 2 1 は、S M 更新クロック T 1 3 のカウンタ値を読み出して、ターゲット更新クロック T 1 2 にコピーさせる (S 1 2 3)。

【0 1 6 8】

マイクロプロセッサ 1 2 1 は、障害マイクロプロセッサのカウンタ値について、L M 更新クロック T 1 1 内の値とターゲット更新クロック T 1 2 内の値とが異なるか否かを判定する (S 1 2 4)。つまり、マイクロプロセッサ 1 2 1 は、L M 更新クロック T 1 1 内の障害マイクロプロセッサのカウンタ値とターゲット更新クロック T 1 2 内の障害マイクロ

10

20

30

40

50

プロセッサのカウンタ値とが相違するか否かを判定する (S 1 2 4)。

【0 1 6 9】

L M更新クロック T 1 1 内のカウンタ値とターゲット更新クロック T 1 2 内のカウンタ値とが相違する場合 (S124:YES)、マイクロプロセッサ 1 2 1 は、L M更新クロック T 1 1 内の障害マイクロプロセッサのカウンタ値を、ターゲット更新クロック T 1 2 内の障害マイクロプロセッサのカウンタ値で上書きする (S 1 2 5)。

【0 1 7 0】

L M更新クロック T 1 1 内のカウンタ値とターゲット更新クロック T 1 2 内のカウンタ値とが相違する場合 (S124:YES)、障害マイクロプロセッサは、送信すべきパージメッセージを送信する前に、停止したと判定されるためである。共有メモリ 1 3 1 内の制御情報を書き換える際に、S M更新クロック T 1 3 内のカウンタ値は更新される。従って、障害マイクロプロセッサがパージメッセージの送信前に停止した場合でも、S M更新クロック T 1 3 内のカウンタ値は信頼することができる。

【0 1 7 1】

マイクロプロセッサ 1 2 1 は、自身が管理するローカルメモリ 1 2 2 内のキャッシュデータ (制御情報のコピーデータ) を全て破棄し、本処理を終了する (S 1 2 6)。送信すべきパージメッセージを送信するよりも前に、障害マイクロプロセッサが停止したことは判明するが、制御情報のどの部分が更新されたのかを確認できないため、いったんキャッシュデータを全てクリアさせる。

【0 1 7 2】

図 1 7 は、障害マイクロプロセッサが障害から回復した場合の処理を示す。例えば、障害マイクロプロセッサを含むマイクロプロセッサパッケージ 1 2 0 が正常なマイクロプロセッサパッケージ 1 2 0 に交換された場合に、図 1 7 のフローチャートが実行される。なお、記憶制御装置 1 0 内にマイクロプロセッサパッケージ 1 2 0 を追加する場合にも、図 1 7 の処理を応用可能である。

【0 1 7 3】

障害から回復したマイクロプロセッサを、回復済マイクロプロセッサと呼ぶ。回復済マイクロプロセッサ 1 2 1 は、L M更新クロック T 1 1 内の回復済マイクロプロセッサのカウンタ値を、ターゲット更新クロック T 1 2 内の回復済マイクロプロセッサのカウンタ値で上書きする (S 1 3 1)。これにより、回復済マイクロプロセッサが管理する L M更新クロック T 1 1 内の、回復済マイクロプロセッサのカウンタ値を正しい値に戻すことができる。

【0 1 7 4】

このように構成される本実施例も第 1 実施例と同様の効果を奏する。さらに、本実施例では、いずれかのマイクロプロセッサ 1 2 1 に障害が発生した場合でも適切に対応することができ、記憶制御装置 1 0 の信頼性が向上する。

【実施例 4】

【0 1 7 5】

図 1 8 に基づいて第 4 実施例を説明する。図 1 8 は、本実施例による記憶制御装置の要部を示すブロック図である。本実施例のマイクロプロセッサパッケージ 1 2 0 A は、パッケージメモリ 1 2 3 を備えている。

【0 1 7 6】

マイクロプロセッサパッケージ 1 2 0 A は、複数のサブパッケージ 1 2 4 と、複数のサブパッケージ 1 2 4 によって共有されるパッケージメモリ 1 2 3 とを備える。各サブパッケージ 1 2 4 は、マイクロプロセッサ 1 2 1 及びローカルメモリ 1 2 2 を備える。

【0 1 7 7】

パッケージメモリ 1 2 3 には、制御情報 D 1 0 と、L M更新クロック T 1 1 及びターゲット更新クロック T 1 2 が記憶される。パージバッファ T 1 0 は、各ローカルメモリ 1 2 2 内に設けられており、パッケージメモリ 1 2 3 内には設けられていない。各マイクロプロセッサ 1 2 1 のアクセス局所性を利用してパージメッセージの量を低減させるために、

10

20

30

40

50

ページバッファ 110 はローカルメモリ 122 内に設ける。なお、パッケージメモリ 123 としては、比較的高速にアクセス可能なメモリ装置が使用される。

【0178】

マイクロプロセッサパッケージ 120 A 内の各マイクロプロセッサ 121 は、パッケージメモリ 123 内に記憶されている制御情報を使用する。同一のマイクロプロセッサパッケージ 120 A 内に設けられている各マイクロプロセッサ 121 は、同一のパッケージメモリ 123 を共有しているため、マイクロプロセッサパッケージ 120 A 内でページメッセージを交換する必要はない。

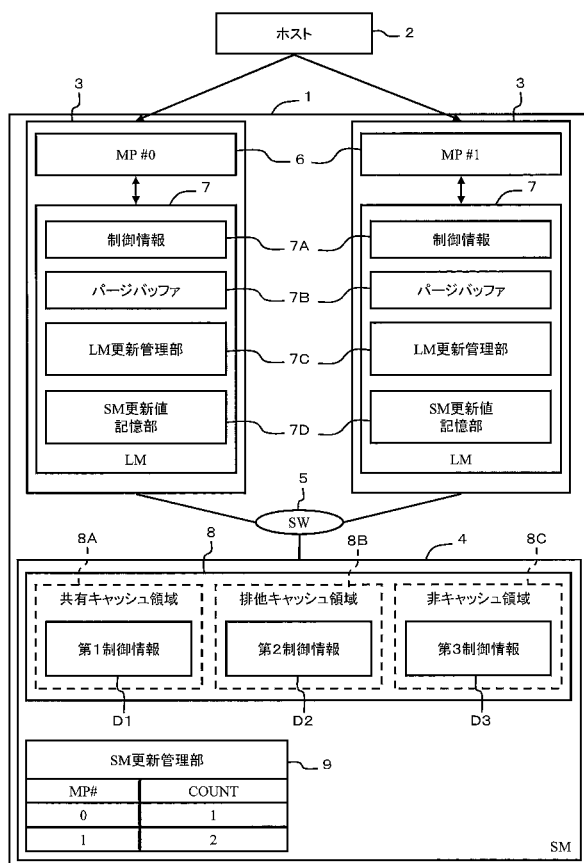
【0179】

本実施例では、マイクロプロセッサパッケージ 120 A 間だけでページメッセージ通信が行われる。従って、第 1 実施例よりもページメッセージの通信頻度及び通信量を少なくすることができる。

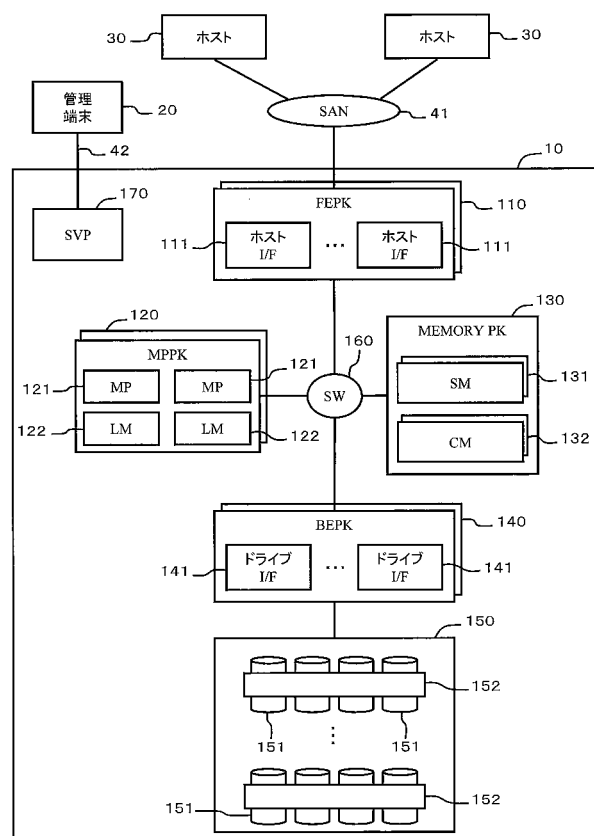
なお、本発明は、上述した実施例に限定されない。当業者であれば、本発明の範囲内で、種々の追加や変更等を行うことができる。例えば、各実施例を適宜組み合わせる構成としてもよい。

10

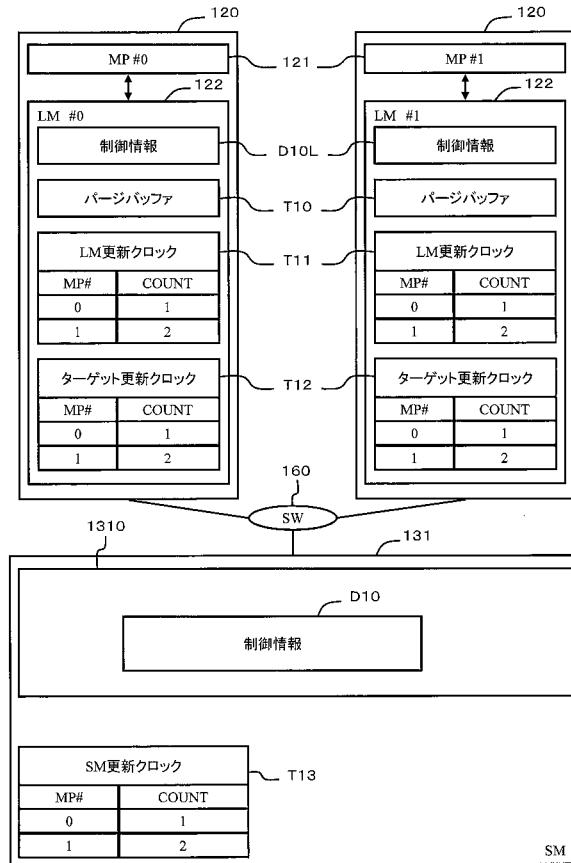
【図 1】



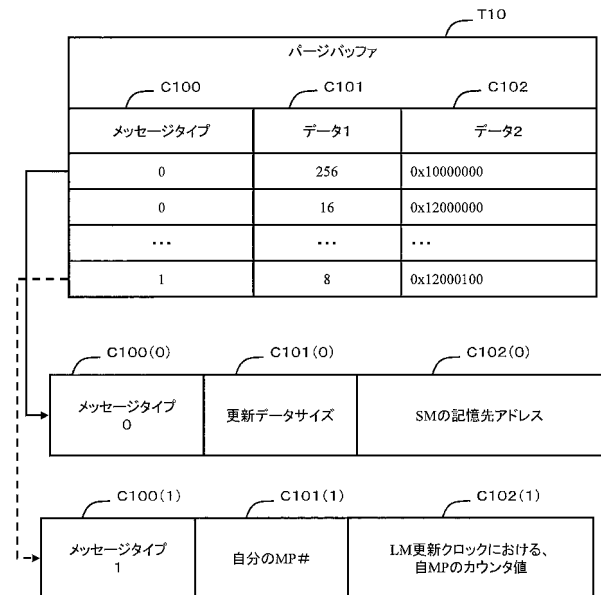
【図 2】



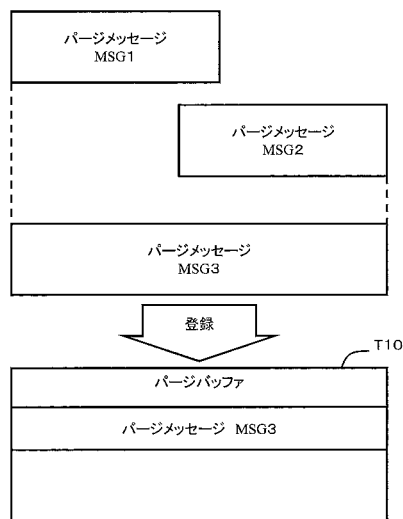
【図 3】



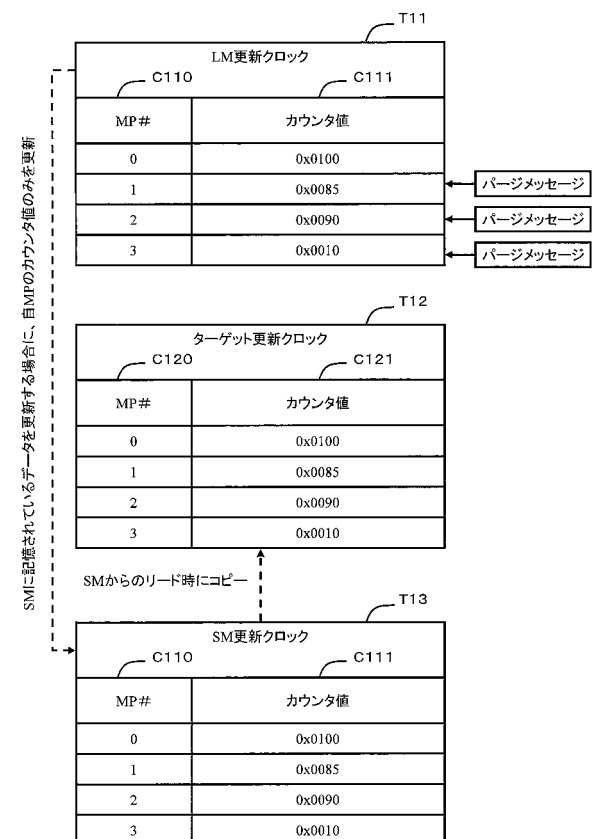
【図 4】



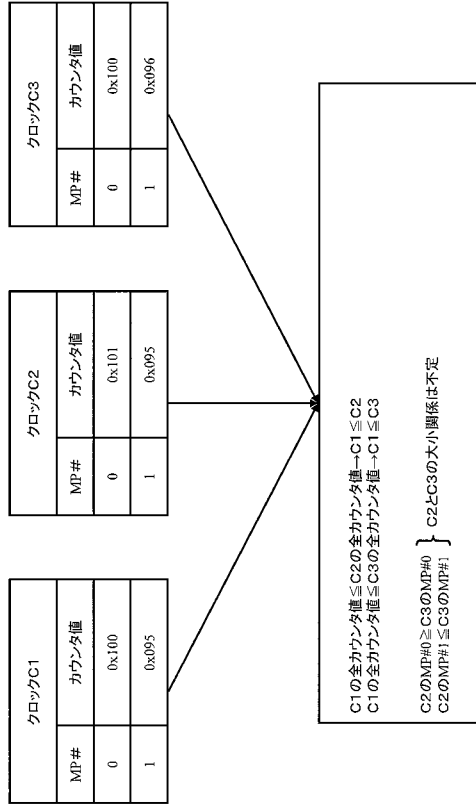
【図 5】



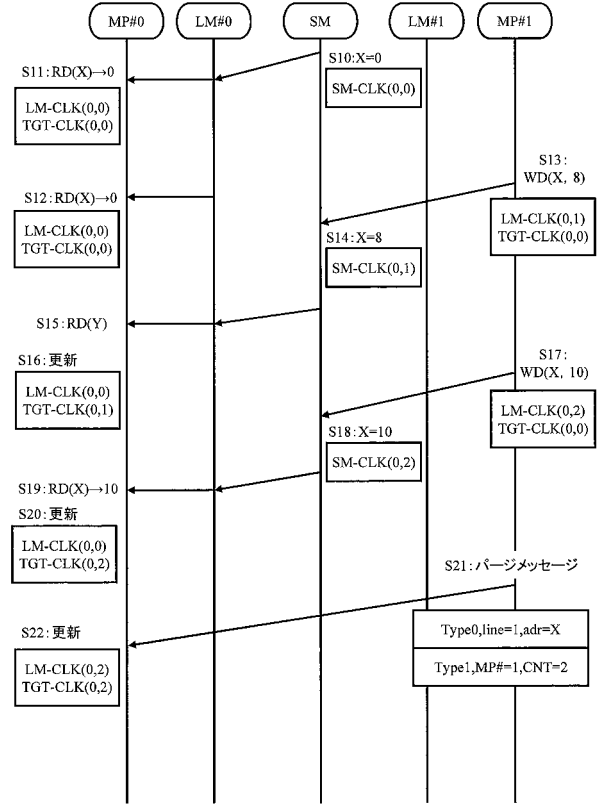
【図 6】



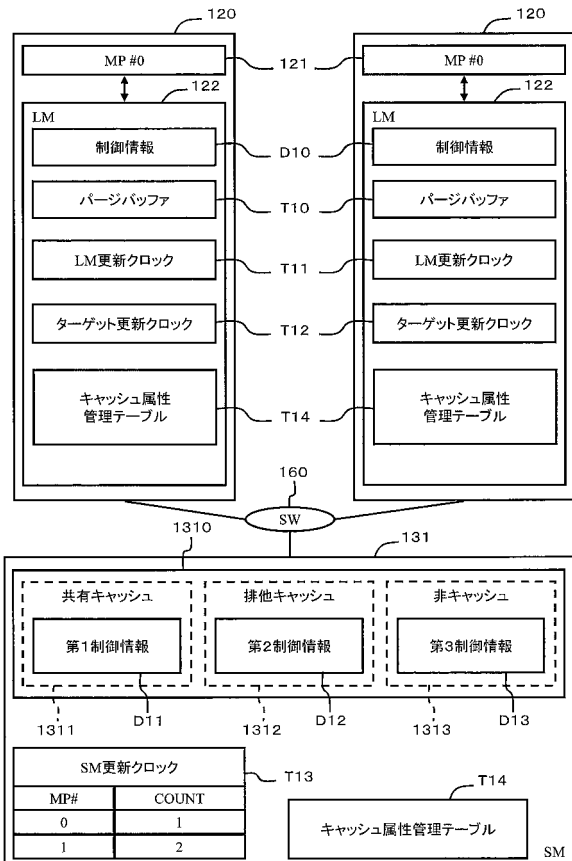
【図 7】



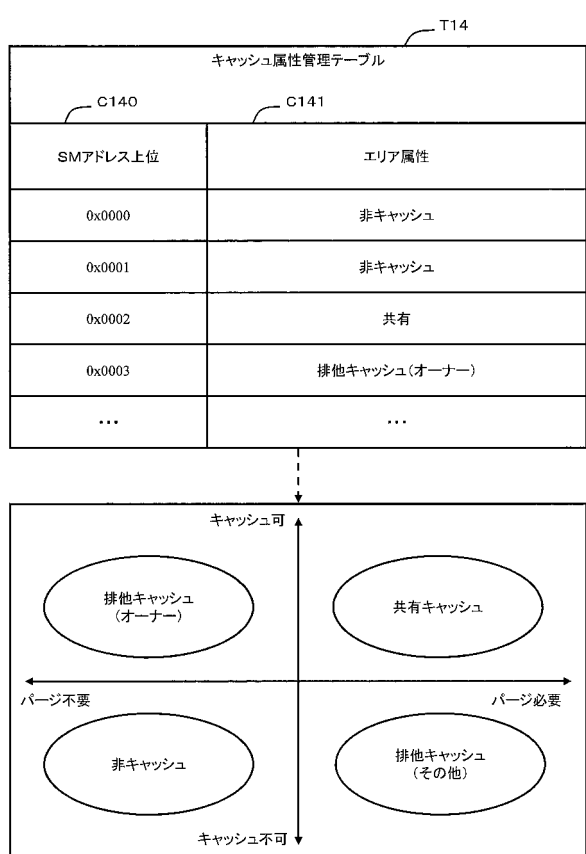
【図 8】



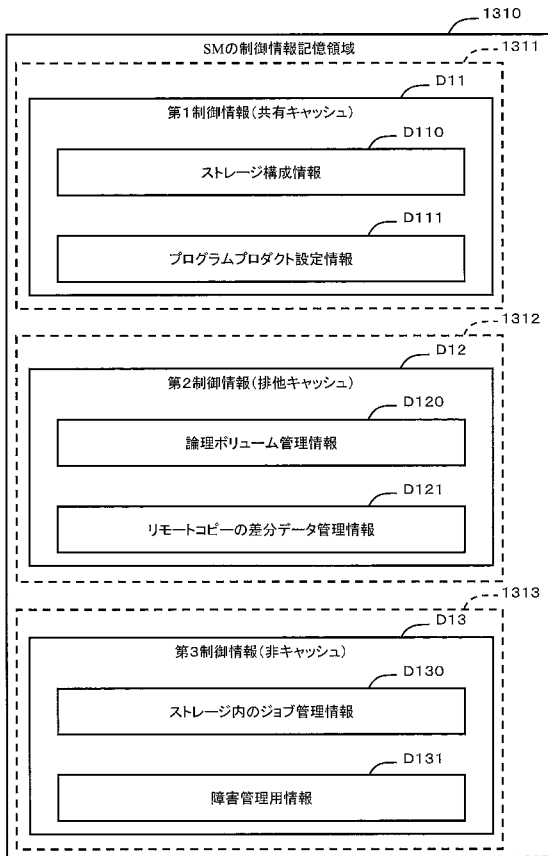
【図 9】



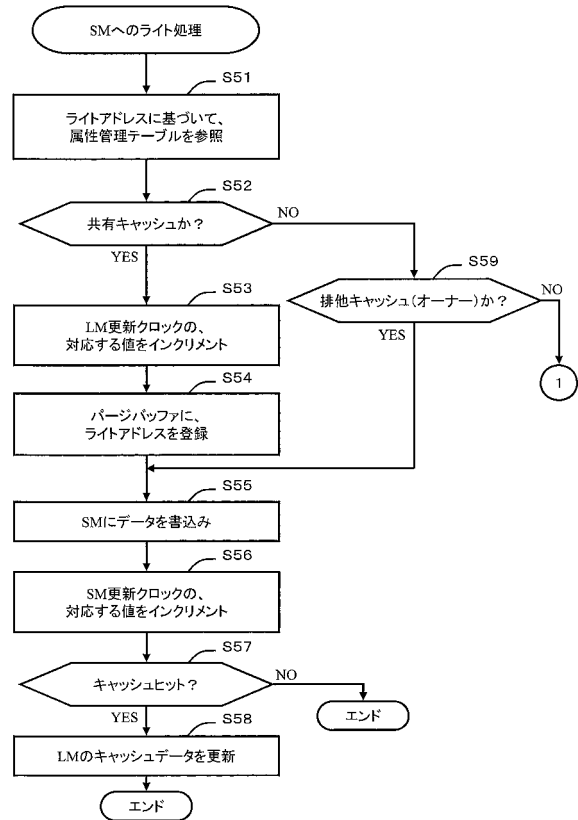
【図 10】



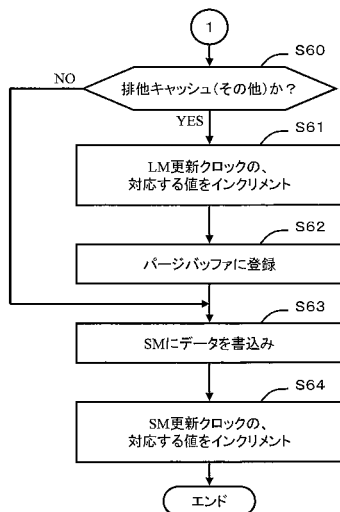
【図 1 1】



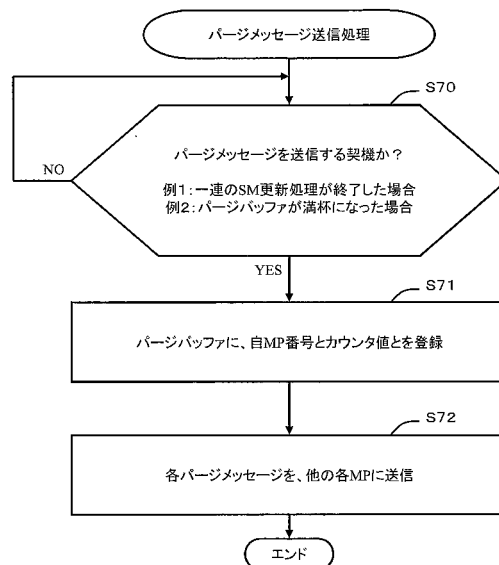
【図 1 2】



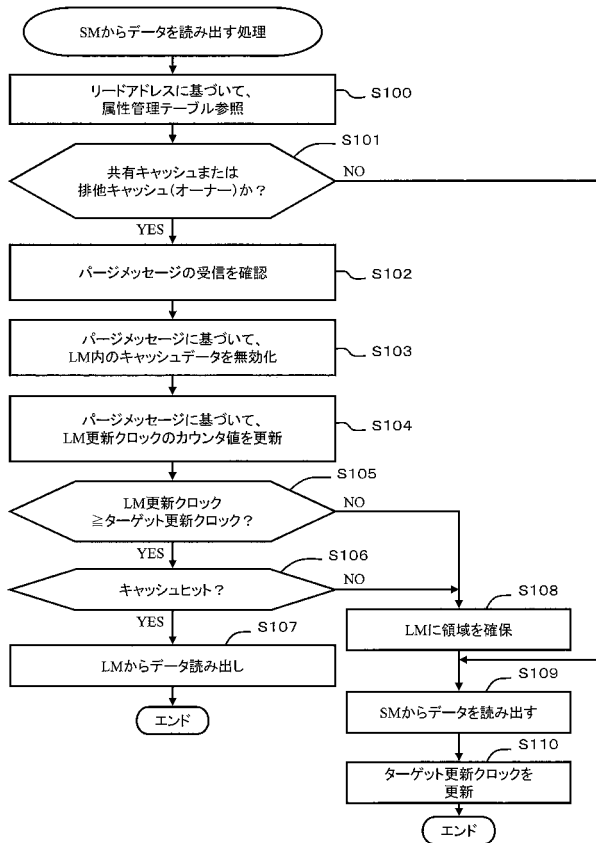
【図 1 3】



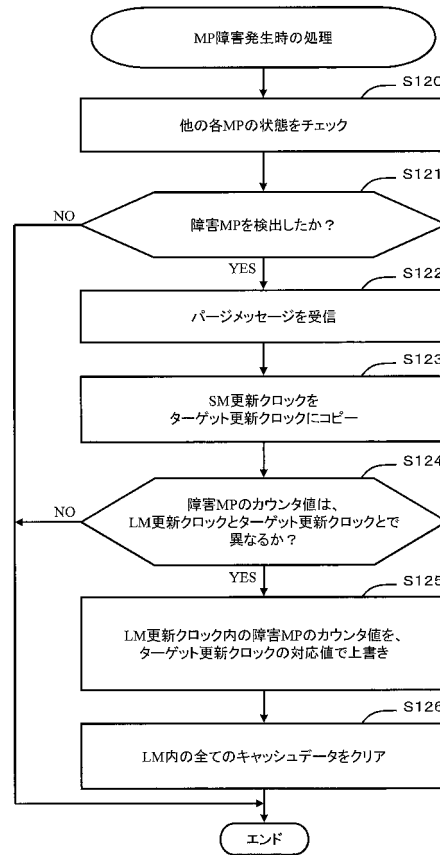
【図 1 4】



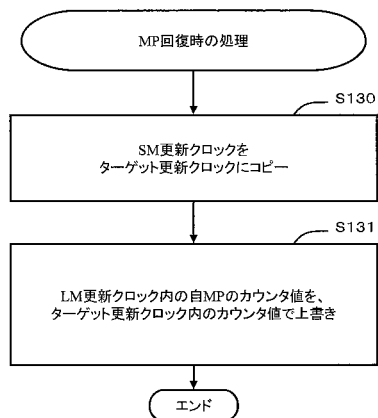
【図 15】



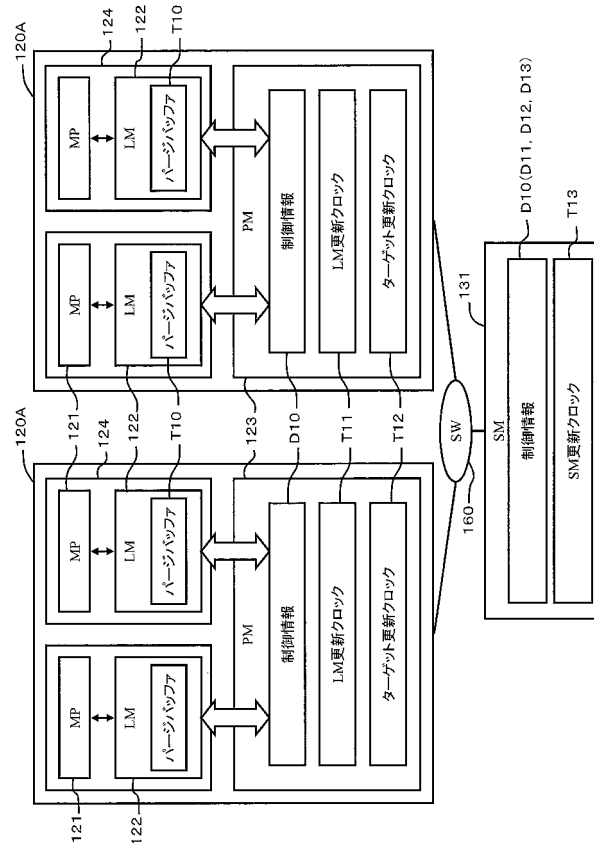
【図 16】



【図 17】



【図 18】



【手続補正書】

【提出日】平成23年4月6日(2011.4.6)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって

、

共有制御情報を記憶する共有メモリと、

複数のマイクロプロセッサと、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、

を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させるようになっており、

複数のマイクロプロセッサパッケージを備え、前記各マイクロプロセッサパッケージは、少なくとも一つの前記マイクロプロセッサと、少なくとも一つの前記ローカルメモリと、少なくとも一つの前記パージメッセージ記憶部とを備えており、

前記各マイクロプロセッサパッケージと前記共有メモリとは、スイッチを介して接続されており、

前記共有メモリには、前記共有制御情報に加えて、前記共有制御情報の更新状態を管理するための共有制御情報更新管理部が設けられており、

前記各ローカルメモリには、前記ローカル制御情報及び前記パージメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部と、前記共有制御情報更新管理部で管理されている前記更新状態を示す値のコピーを記憶する更新値記憶部とが記憶されており、

前記共有制御情報更新管理部は、前記各マイクロプロセッサパッケージ内の前記各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、前記各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、前記マイクロプロセッサにより前記共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応する前記カウンタ値が更新されるようになっており、

前記各ローカル更新管理部も、前記各マイクロプロセッサ識別情報毎に前記カウンタ値を対応付けて管理しており、さらに、前記各ローカル更新管理部に対応する前記各マイクロプロセッサが前記ローカル制御情報を更新する場合に、その対応する前記マイクロプロセッサに関連するカウンタ値が更新されるようになっており、さらに、前記対応する前記マイクロプロセッサ以外の他の前記マイクロプロセッサに関連する前記パージメッセージ

に基づいて、前記他の前記マイクロプロセッサに対応するカウンタ値が更新されるようになっており、

前記各更新値記憶部に対応する前記各マイクロプロセッサが前記共有メモリ内の前記共有制御情報にアクセスした場合に、前記共有制御情報更新管理部で管理されている前記各カウンタ値が取得されて、それら取得された前記各カウンタ値が前記更新値記憶部に書きで記憶されるようになっており、

前記共有メモリは、前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、を備えており、

前記共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第1情報が記憶されるようになっており、

前記排他キャッシュ領域には、前記オーナーマイクロプロセッサが比較的高頻度に参照または更新する第2情報が記憶されるようになっており、

前記非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度で参照または更新される第3情報が記憶されるようになっており、

前記各マイクロプロセッサは、

前記共有キャッシュ領域に記憶されている情報を更新した場合には、前記更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを他の各マイクロプロセッサにそれぞれ送信し、

前記排他キャッシュ領域に記憶されている情報を更新する場合において、自分が前記オーナーマイクロプロセッサである場合は、前記パージメッセージを作成せず、自分が前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを前記オーナーマイクロプロセッサのみに送信し、

前記非キャッシュ領域に記憶されている情報を更新した場合には、前記パージメッセージを作成しないようになっており、

さらに、前記各マイクロプロセッサは、前記共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合には、複数の更新を統合するようにして前記パージメッセージを送信させるようになっている、

記憶制御装置。

【請求項2】

ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって

、
共有制御情報を記憶する共有メモリと、
複数のマイクロプロセッサと、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制

御情報同期管理部と、
を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させるようになっており、

前記共有メモリは、

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、

を備えている、

記憶制御装置。

【請求項 3】

ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって

、

共有制御情報を記憶する共有メモリと、

複数のマイクロプロセッサと、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、

を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させるようになっており、

前記共有メモリは、

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、

を備えている、

記憶制御装置。

【請求項 4】

ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって

、

共有制御情報を記憶する共有メモリと、

複数のマイクロプロセッサと、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、

を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させるようになっており、

前記共有メモリは、

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、

を備えている、

記憶制御装置。

【請求項5】

ホストコンピュータと記憶装置との間のデータ入出力を制御する記憶制御装置であって

、

共有制御情報を記憶する共有メモリと、

複数のマイクロプロセッサと、

前記各マイクロプロセッサにそれぞれ対応して設けられるローカルメモリであって、前記共有メモリに記憶される前記共有制御情報のうち少なくとも一部がローカル制御情報としてコピーされる各ローカルメモリと、

前記各マイクロプロセッサのうちの一つのマイクロプロセッサから他の前記各マイクロプロセッサに、または、特定の一つまたは複数の前記マイクロプロセッサに、向けて送信されるパージメッセージであって、前記ローカルメモリに記憶されている前記ローカル制御情報が無効である旨を通知するためのパージメッセージを記憶するパージメッセージ記憶部と、

前記各ローカル制御情報が前記共有制御情報に同期しているか否かを管理するための制御情報同期管理部と、

を備え、

前記各マイクロプロセッサは、前記共有制御情報を更新する場合に、前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、さらに、前記共有制御情報の更新と非同期で、前記パージメッセージ記憶部に記憶された前記パージメッセージを送信させるようになっており、

前記共有メモリは、

前記各マイクロプロセッサが前記各ローカルメモリにコピー可能な情報を記憶するための共有キャッシュ領域と、

前記各マイクロプロセッサのうちオーナーとして設定されるオーナーマイクロプロセッサのみが、前記オーナーマイクロプロセッサ内の前記ローカルメモリにコピー可能であり、かつ、前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサはコピーできない情報を記憶するための排他キャッシュ領域と、

前記各マイクロプロセッサのいずれもが前記各ローカルメモリにコピーできない情報を記憶するための非キャッシュ領域と、

を備えている、
記憶制御装置。

【請求項 6】

前記共有キャッシュ領域には、比較的更新頻度が少ないが複数のマイクロプロセッサから比較的高頻度でアクセスされ得る第 1 情報が記憶されるようになっており、

前記排他キャッシュ領域には、前記オーナーマイクロプロセッサが比較的高頻度に参照または更新する第 2 情報が記憶されるようになっており、

前記非キャッシュ領域には、複数のマイクロプロセッサから比較的高頻度で参照または更新される第 3 情報が記憶されるようになっている、
請求項 5 に記載の記憶制御装置。

【請求項 7】

前記各マイクロプロセッサは、

前記共有キャッシュ領域に記憶されている情報を更新した場合には、前記更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを他の各マイクロプロセッサにそれぞれ送信し、

前記排他キャッシュ領域に記憶されている情報を更新する場合において、自分が前記オーナーマイクロプロセッサである場合は、前記パージメッセージを作成せず、自分が前記オーナーマイクロプロセッサ以外の他のマイクロプロセッサである場合は、更新のたびに前記パージメッセージを作成して前記パージメッセージ記憶部に記憶させ、一連の更新処理が完了した後で、前記パージメッセージ記憶部に記憶された前記各パージメッセージを前記オーナーマイクロプロセッサのみに送信し、

前記非キャッシュ領域に記憶されている情報を更新した場合には、前記パージメッセージを作成しないようになっている、

請求項 5 または請求項 6 のいずれかに記載の記憶制御装置。

【請求項 8】

前記各マイクロプロセッサは、前記制御情報同期管理部に基づいて前記各ローカル制御情報が前記共有制御情報に同期していると判定される場合は、対応する前記各ローカルメモリから前記ローカル制御情報を使用し、前記制御情報同期管理部に基づいて前記各ローカル制御情報が前記共有制御情報に同期していないと判定される場合は、前記共有メモリ内の前記共有制御情報を使用する、

請求項 1 ~ 7 のいずれかに記載の記憶制御装置。

【請求項 9】

前記共有メモリには、前記共有制御情報に加えて、前記共有制御情報の更新状態を管理するための共有制御情報更新管理部が設けられており、

前記各ローカルメモリには、前記ローカル制御情報及び前記パージメッセージ記憶部に加えて、ローカル制御情報の更新状態を管理するためのローカル更新管理部と、前記共有制御情報更新管理部で管理されている前記更新状態を示す値のコピーを記憶する更新値記憶部とが記憶されてあり、

前記制御情報同期管理部は、前記共有制御情報更新管理部と、前記各ローカル更新管理部と、前記各更新値管理部とを備えて構成される、

請求項 1 ~ 7 のいずれかに記載の記憶制御装置。

【請求項 10】

前記共有制御情報更新管理部は、前記各マイクロプロセッサパッケージ内の前記各マイクロプロセッサを識別するためのマイクロプロセッサ識別情報と、前記各マイクロプロセッサにより更新された回数を示すカウンタ値とを対応付けて管理しており、さらに、前記マイクロプロセッサにより前記共有制御情報が更新される場合に、更新元のマイクロプロセッサによって当該更新元のマイクロプロセッサに対応する前記カウンタ値が更新されるようになっており、

前記各ローカル更新管理部も、前記各マイクロプロセッサ識別情報毎に前記カウンタ値

を対応付けて管理しており、さらに、前記各ローカル更新管理部に対応する前記各マイクロプロセッサが前記ローカル制御情報を更新する場合に、その対応する前記マイクロプロセッサに関連するカウンタ値が更新されるようになっており、さらに、前記対応する前記マイクロプロセッサ以外の他の前記マイクロプロセッサに関連する前記パージメッセージに基づいて、前記他の前記マイクロプロセッサに対応するカウンタ値が更新されるようになっている、

請求項 9 に記載の記憶制御装置。

【請求項 11】

前記各更新値記憶部に対応する前記各マイクロプロセッサが前記共有メモリ内の前記共有制御情報にアクセスした場合に、前記共有制御情報更新管理部で管理されている前記各カウンタ値が取得されて、それら取得された前記各カウンタ値が前記更新値記憶部に上書きで記憶されるようになっている、

請求項 10 に記載の記憶制御装置。

【請求項 12】

前記各マイクロプロセッサは、前記共有制御情報のうち同一箇所の更新、または、連続する箇所の更新、または、重複する箇所の更新がされた場合に、複数の更新を統合するようにして前記パージメッセージを送信させる、

請求項 1 ~ 7 のいずれかに記載の記憶制御装置。

【請求項 13】

前記各マイクロプロセッサは、

他の各マイクロプロセッサが正常に監視しているか否かを監視しており、

いずれかのマイクロプロセッサで障害が発生した場合には、前記共有制御情報更新管理部で管理されている各カウンタ値を取得して前記更新値記憶部に上書きで記憶し、

前記障害の発生した前記マイクロプロセッサについて、前記ローカル更新管理部で管理されているカウンタ値と前記更新値管理部に記憶されているカウンタ値と異なるか否かを判定し、

前記ローカル更新管理部で管理されているカウンタ値と前記更新値管理部に記憶されているカウンタ値とが異なる場合、前記ローカル更新管理部で管理されている前記障害の発生した前記マイクロプロセッサの前記カウンタ値を、前記更新値管理部内の対応するカウンタ値で上書きし、

前記ローカルメモリにコピーされているローカル制御情報をクリアさせる、

請求項 1 ~ 7 のいずれかに記載の記憶制御装置。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2009/000645

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F3/06 G06F12/08		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 577 772 A1 (HITACHI LTD [JP]) 21 September 2005 (2005-09-21) abstract paragraphs [0008], [0046], [0054] - [0056], [0064], [0087], [0088] figures 1-4	1-6, 13-15
Y A	US 5 506 967 A (BARAJAS SAUL [US] ET AL) 9 April 1996 (1996-04-09) abstract column 2, line 30 - line 53 column 3, line 33 - column 4, line 2 column 5, line 11 - line 30 figures 1,2,3 -/-	1-6, 13-15 7-12
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search 23 July 2009		Date of mailing of the international search report 30/07/2009
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Mandato, Davide

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2009/000645

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 818 794 A2 (HITACHI LTD [JP]) 15 August 2007 (2007-08-15) abstract figure 1 paragraph [0013]	1-6, 13-15
Y	US 2003/110157 A1 (MAKI NOBUHIRO [JP] ET AL) 12 June 2003 (2003-06-12) figure 1	1-6, 13-15
A	US 5 353 428 A (SHIBATA MASABUMI [JP]) 4 October 1994 (1994-10-04) claims 3,8	7-12
A	EP 1 376 348 A2 (FUJITSU LTD [JP]) 2 January 2004 (2004-01-02) abstract figures 1,7,11,12	7-12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/JP2009/000645

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 1577772	A1	21-09-2005	CN 1670652 A	21-09-2005
			CN 101013385 A	08-08-2007
			CN 101308531 A	19-11-2008
			JP 2005267008 A	29-09-2005
			US 2008282043 A1	13-11-2008
			US 2006206683 A1	14-09-2006
			US 2005210217 A1	22-09-2005
US 5506967	A	09-04-1996	NONE	
EP 1818794	A2	15-08-2007	JP 2007207007 A	16-08-2007
			US 2007180188 A1	02-08-2007
US 2003110157	A1	12-06-2003	NONE	
US 5353428	A	04-10-1994	NONE	
EP 1376348	A2	02-01-2004	US 2004055000 A1	18-03-2004

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

１．ＲＲＡＭ

(72)発明者 阪口 治

東京都品川区東品川四丁目１２番７号 株式会社日立ソリューションズ内

Fターム(参考) 5B005 KK02 MM11 PP03 PP11

5B014 EB04

5B065 CA11 CH01