(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3697149号
(P3697149)
(45)発行日　平成17年9月21日(2005.9.21)　(24)登録日　平成17年7月8日(2005.7.8)

(51)Int.Cl.[7]	F I		
G06F 12/08	G06F	12/08	501F
G06F 3/06	G06F	12/08	509Z
G06F 12/12	G06F	12/08	551Z
	G06F	12/08	557
	G06F	3/06	301U

請求項の数 2　(全 14 頁)　最終頁に続く

(21)出願番号　特願2000-294735(P2000-294735)
(22)出願日　平成12年9月27日(2000.9.27)
(65)公開番号　特開2001-142778(P2001-142778A)
(43)公開日　平成13年5月25日(2001.5.25)
審査請求日　平成12年9月27日(2000.9.27)
(31)優先権主張番号　09/410499
(32)優先日　平成11年10月1日(1999.10.1)
(33)優先権主張国　米国(US)

前置審査

(73)特許権者　390009531
インターナショナル・ビジネス・マシーンズ・コーポレーション
ＩＮＴＥＲＮＡＴＩＯＮＡＬ　ＢＵＳＩＮＥＳＳ　ＭＡＳＣＨＩＮＥＳ　ＣＯＲＰＯＲＡＴＩＯＮ
アメリカ合衆国１０５０４　ニューヨーク州　アーモンク　ニュー　オーチャード　ロード
(74)代理人　100086243
弁理士　坂口　博
(74)代理人　100091568
弁理士　市位　嘉宏

最終頁に続く

(54)【発明の名称】キャッシュ・メモリを管理する方法

(57)【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でデータ・パスを介して接続され、かつ第１のキャッシュと第２のキャッシュとに論理的に分割されているキャッシュ・メモリを管理する方法であって、

前記第１のキャッシュがデータを含むセグメントを記憶し、前記第２のキャッシュが複数のセグメントからなるグループを記憶し、前記第１のキャッシュのためのセグメントＬＲＵ（最低使用頻度）リスト及び前記第２のキャッシュのためのグループＬＲＵリストが設けられており、

前記ホスト・コンピュータへの読み取り時に、

（ａ）前記ホスト・コンピュータからのデータにアクセスする要求に応えて、要求されたデータが前記セグメントＬＲＵリストにあるかどうかを判断するステップと、

（ｂ）前記ステップ（ａ）において、もしも前記要求されたデータが前記セグメントＬＲＵリストにあるならば、前記要求されたデータのコピーを前記第１のキャッシュから前記ホスト・コンピュータに転送するステップと、

（ｃ）前記ステップ（ａ）において、もしも前記要求されたデータが前記セグメントＬＲＵリストにないならば、前記要求されたデータが前記グループＬＲＵリストにあるかどうかを判断するステップと、

（ｄ）前記ステップ（ｃ）において、もしも前記要求されたデータが前記グループＬＲＵリストにあるならば、前記要求されたデータを含むセグメントのコピーを前記第２のキ

ャッシュから前記第１のキャッシュへ転送すると共に、前記セグメントＬＲＵリストを更新して、前記転送されたセグメントのアドレスを前記セグメントＬＲＵリストのＭＲＵ（最後に使用された）部分に入れるステップと、
（ｅ）前記ステップ（ｄ）に続いて、前記要求されたデータのコピーを前記第１のキャッシュから前記ホスト・コンピュータに転送するステップと、
（ｆ）前記ステップ（ｃ）において、もしも前記要求されたデータが前記グループＬＲＵリストにないならば、前記要求されたデータを含むグループのコピーを前記記憶装置から前記第２のキャッシュへ転送すると共に、前記グループＬＲＵリストを更新して、前記転送されたグループのアドレスを前記グループＬＲＵリストのＭＲＵ部分に入れるステップと、
（ｇ）前記ステップ（ｆ）に続いて、前記要求されたデータを含むセグメントのコピーを前記第２のキャッシュから前記第１のキャッシュに転送するステップと、
（ｈ）前記ステップ（ｇ）に続いて、前記要求されたデータのコピーを前記第１のキャッシュから前記ホスト・コンピュータに転送するステップと
を行い、
前記ホスト・コンピュータからのデータの書き込み時に、
（イ）通常はスリープ・モードになっており、前記ホスト・コンピュータにより修正され且つ前記第２のキャッシュ及び前記記憶装置へ転送する必要があるセグメントが前記第１のキャッシュ内に存在することに応答して起動するステップと、
（ロ）前記ステップ（イ）に続いて、前記セグメントＬＲＵリスト内の前記修正されたセグメントを見出すステップと、
（ハ）前記ステップ（ロ）に続いて、前記修正されたセグメントが前記グループＬＲＵリストにあるかどうかを判断するステップと、
（ニ）前記ステップ（ハ）において、前記修正されたセグメントが前記グループＬＲＵリストにあるならば、前記修正されたセグメントを前記第２のキャッシュに転送して既存のグループを修正するステップと、
（ホ）前記ステップ（ハ）において、前記修正されたセグメントが前記グループＬＲＵリストにないならば、前記修正されたセグメントのセグメント待ちフラグが０に設定されているかどうかを判断するステップと、
（ヘ）前記ステップ（ホ）において、前記セグメント待ちフラグが０に設定されているならば、該セグメント待ちフラグを１に設定するステップと、
（ト）前記ステップ（ヘ）に続いて、前記ステップ（イ）、（ロ）、（ハ）及び（ホ）を行い、該ステップ（ホ）において前記セグメント待ちフラグが１であることを検出して、前記修正されたセグメントを前記記憶装置に転送するステップと
を行うことを特徴とする、キャッシュ・メモリ管理方法。
【請求項２】
もしも前記ステップ（ａ）で要求されたデータが前記第１のキャッシュに格納されていないと判断したならば、最低使用頻度手順にもとづいて最低使用頻度セグメントを前記第１のキャッシュからデステージするステップと、
もしも前記ステップ（ａ）で要求されたデータが前記第２のキャッシュに格納されていないと判断したならば、最小使用頻度手順にもとづいて前記セグメントの最小使用頻度グループを前記第２のキャッシュからデステージするステップと、
をさらに有することを特徴とする請求項１に記載のキャッシュ・メモリ管理方法。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
この発明は、ディスクメモリ等の直接アクセス記憶装置とキャッシュメモリとを有する記憶システムに関する。特に、記憶システムへのデータ・アクセスに対するホスト・コンピュータ及びチャンネルの待ち時間を減らすことによってパフォーマンスを向上させる方法及びシステムに関する。

【０００２】
【従来の技術】
典型的なコンピュータ・システムは、ホスト・コンピュータと、直接アクセス記憶装置（ＤＡＳＤ）、例えば１枚以上の磁気ディスクとを含む。ホスト・コンピュータ上で実行されているアプリケーションは、データの読み取り及び書き出しのためにＤＡＳＤ内のアドレス位置にアクセスする。そのようなアクセスは、ディスク入出力（Ｉ／Ｏ）オペレーションとして知られている。ホスト・コンピュータは、ディスクＩ／Ｏ操作が該Ｉ／Ｏ操作の完了を実行中のアプリケーションに待たせるように、ＤＡＳＤよりもかなり速い速度で稼動する。結果は、ホスト・コンピュータの処理能力が妨げられるということである。このことを避けるため、別の高速キャッシュ・メモリが用いられ、アプリケーションによって最も頻繁に使用されるデータが格納される。
【０００３】
記憶システム・マネージャがディスクＩ／Ｏ操作の制御及びキャッシュ・メモリに対するアクセスの制御に使われる。一般に、データはレコード等のデータ対象物に編成される。実行中のアプリケーションがレコードを要求すると、記憶システム・マネージャは最初にキャッシュ・メモリ内のレコードを探索する。もし要求されたレコードがキャッシュ・メモリで見いだされるならば（すなわち、「ヒット」）記憶システム・マネージャは素早くそのレコードにアクセスするので時間を浪費するディスクＩ／Ｏ操作を実行する必要性がない。もし要求したレコードがキャッシュ・メモリで見いだされなければ（すなわち、「ミス」）、記憶システム・マネージャはディスクＩ／Ｏ操作を実行してＤＡＳＤから要求されたレコードを取得し、該要求されたレコードをキャッシュ・メモリに書き込む。
【０００４】
一般に、記憶システム・マネージャは、最低使用頻度（ＬＲＵ）手法によってキャッシュ・メモリ内のレコード保存を管理する。ＬＲＵ手法は、制御ブロックの連鎖又は待ち行列を使用する。それぞれの制御ブロックが（ａ）レコードのアドレス、（ｂ）連鎖内の次に続くレコードのアドレスを識別する順方向連鎖ポインタ、及び（ｃ）その連鎖内の先のレコードのアドレスを識別する逆方向連鎖ポインタを識別する。記憶システム・マネージャは、ＬＲＵレコード、例えば連鎖の先端を識別する第１のアンカー・ポインタを保持する。また、記憶システム・マネージャは最後に使用された（ＭＲＵ）レコード、例えば連鎖の末端を識別する第２のアンカー・ポインタを保持する。
【０００５】
キャッシュ・ヒットが起こるたびに、ヒット・レコードの制御ブロックがデキューされ、続いてＬＲＵ連鎖の末端でＭＲＵレコードとして待ち行列に入れられる。キャッシュ・ミスが起こるたびに、ＬＲＵ制御ブロックは連鎖の先端からデキューされる。新たに要求されたレコードはＤＡＳＤからキャッシュ・メモリ内の割り当てられたアドレスへステージされる。デキューされた制御ブロックは、ステージされたレコードと割り当てられたアドレスとの一致によってアップデートされ、さらにＭＲＵ制御ブロックとしてＬＲＵ連鎖の末端で待ち行列に入れられる。
【０００６】
記憶システムのキャッシュ・メモリを設計する上で、要求されたデータ・レコードがキャッシュ・メモリ内で見つかる可能性を増やすことにかなりの配慮が払われる。例えば、米国特許第５，７１７，８９３号は、グローバル・キャッシュと、各々が特定のタイプのデータ・レコードに割り当てられている複数のデステージング・ローカル・キャッシュとに分割されたキャッシュ・メモリを開示している。全てのタイプのデータ・レコードがローカル・キャッシュ又はディスク記憶システムからグローバル・キャッシュにデステージされる。ＬＲＵアルゴリズムにもとづいて、ＬＲＵデータ・レコードがグローバル・キャッシュからローカル・キャッシュへ降格される。この際、ローカル・キャッシュのデータ・タイプは降格されたＬＲＵレコードのデータ・タイプと一致する。ローカル・キャッシュが一杯になると、ＬＲＵは記憶システムにデステージされる。より頻繁に使用されるデータ・レコード・タイプに対してより多くのキャッシュを割り当てるように分割スキームが

設計されるので、キャッシュ・ヒット率が増加する。より頻繁に使用されるデータ・タイプに対してキャッシュが増加し、使用頻度の低いデータ・タイプに対してはキャッシュが同時に減少するように、区画の論理的かつダイナミックなサイズ変更を可能とすることも特徴の一つである。

【０００７】

他の従来のスキームは、キャッシュ内のデータ・レコードの複製を削除することによって、キャッシュ・ヒット率を増やす。このタイプの典型的なスキームは、米国特許第５，８０２，５７２号及び第５，６２７，９９０号に開示されている。

【０００８】

ＤＡＳＤシステムは、障害が生じた際にデータ回復を保証する幾何学的配置で構成された多数の小さな記憶装置モジュールを使用することで改善されている。改善がなされたそれらのシステムは、比較的安価な（又は独立した）複数のディスクからなる重複アレイ（ＲＡＩＤ）としてしばしば呼ばれる。そのような幾何学的配置のいくつかでは、データ・オブジェクトが複数のデータ部分に分割され、かつ各データ部分が複数のディスクの異なったディスク上に格納される。ＲＡＩＤレベル４として知られている一つの幾何学的配置では、ディスクの一つがデータ部分に対するパリティの保存専用となる。パリティは、障害が生じた際にデータ部分を再構築するために使われる。書き出し操作のために、この幾何学的配置は別々の２つの書き出しアクセスを必要とするもので、一つはデータ部分が格納されるディスクに対するアクセスであり、もう一つはパリティが格納されるディスクに対するアクセスである。

【０００９】

ＲＡＩＤレベル５として知られている別の幾何学的配置では、アレイ内の複数のアクティブなディスクをまたがってデータ及びパリティ情報の分配を行うためにディスクが分割される。各区画は一般にストライプと呼ばれる。ストライプに対するパリティ情報は通常一つのディスクに置かれ、データはストライプの残りのディスク上に置かれる。パリティ情報を含んでいるディスクはストライプ間で異なる。このことによって多数のストライプが並列処理可能となり、それによってステージ又はデステージされるデータのどちらかと言えば大きいチャンクを可能とする。

【００１０】

キャッシュ・ヒット率を高めるための上記のスキームは、どちらかと言うと小さなデータ・オブジェクト、例えばページ、テーブル等に関係している。また上記スキームは、より小さなページ・オブジェクトを大量に含むかなり大きなデータ・オブジェクト（例えばストライプ）を扱うためにＲＡＩＤシステムの能力を利用することはない。

【００１１】

【発明が解決しようとする課題】

したがって、キャッシュ・ヒットの確率が改善されたキャッシュ・メモリが求められている。特に、ＲＡＩＤ記憶装置が持つストライプ・アクセス能力を利用するキャッシュ・メモリが求められている。

【００１２】

【課題を解決するための手段】

本発明は、記憶システムからデータ・オブジェクトを必要とするアプリケーションを実行するホスト・コンピュータを使用する。記憶システムは、記憶装置（例えば、ディスク記憶装置）とキャッシュ・メモリとを有する。ホスト・コンピュータによって頻繁に使われるデータ・オブジェクトがキャッシュ・メモリに格納される。データ・オブジェクトは、データ・オブジェクトが論理的に配列された記憶装置にも格納される。キャッシュ・メモリは、第１のキャッシュと第２のキャッシュとに論理的に分割されている。

【００１３】

本発明の方法は、わずかな細分性(granularity)を持つセグメントを格納するための第１のキャッシュと、大きな細分性を持つセグメントからなるグループを格納するための第２のキャッシュとを使用する。ホスト・コンピュータがデータに対するアクセスを要求する

と、本発明の方法は要求されたデータが第１のキャッシュに格納されているかどうかを判断する。もし要求されたデータが第１のキャッシュに格納されていないならば、方法は要求されたデータが第２のキャッシュに格納されているかどうかを判断する。もし要求されたデータが第２のキャッシュに格納されていないならば、記憶装置に格納された複数のセグメントからなる一グループにアクセスする。この際、それらのセグメントの一つに要求されたデータが含まれている。次に、セグメントからなるグループを第２のキャッシュに格納し、要求されたデータが含まれる第１のセグメントを第１のキャッシュに格納する。次に、要求されたデータが第１のキャッシュからアクセスされる。

【００１４】

もし方法が第２のキャッシュではなく第１のキャッシュに要求されたデータが格納されていると判断するならば、要求されたデータを含むセグメントのコピーが第１のキャッシュに転送される。

【００１５】

本方法は、別々のＬＲＵ手順を用いて最小使用頻度セグメントを第１のキャッシュから、またセグメントからなるグループを第２のキャッシュからデステージして第１及び第２のキャッシュに格納されていない要求データに対して記憶域を割り当てる。

【００１６】

第１のキャッシュにセグメントを格納し、第２のキャッシュにグループを格納するとともに、キャッシュ・メモリを第１のキャッシュ及び第２のキャッシュへ論理的に分割することは、本発明の重要な特徴である。この特徴は、一グループ内の一つのデータ・オブジェクトを要求するアプリケーションが同一グループの別のデータ・オブジェクトも要求するが、同一のセグメントである必要性はない可能性を利用する。

【００１７】

本発明のキャッシュ・メモリ・システムは上記した方法の手順を含む多重細分性(multi-granular)キャッシュ・マネージャ・プログラムを使用する。

【００１８】

本発明にもとづくメモリ媒体は、キャッシュ・メモリを制御して上記した発明の方法の手順を実行する。

【００１９】

本発明の別の目的、利点、及び特徴は、図面と共に以下の説明を参照することによって理解することができよう。なお、同の一符号は同一の構成要素を示している。

【００２０】

【発明の実施の形態】

図１は、本発明に適用されるコンピュータ・システム１０の概略的構成を示すブロック図である。コンピュータ・システム１０は、ホスト・コンピュータ１２と記憶システム１４とを含む。記憶システム１４は、ホスト・アダプタ１６、記憶装置１８、及び本発明にもとづく多重細分性キャッシュ・メモリ３０を備える。多重細分性キャッシュ・メモリ３０はホスト・アダプタ１６とともに、ホスト・コンピュータ１２と記憶装置１８との間のデータ・パスに結合している。

【００２１】

ホスト・コンピュータ１２は、アプリケーションを実行する１種類以上のプロセッサを一般に有する。プロセッサは、並行して単一のアプリケーション、又は異なる時間又は同時に別々のアプリケーション、あるいはそれらの任意の組み合わせを実行してもよい。アプリケーションは、論理データ構造を構成するデータ・オブジェクトを使う。論理データ構造は一般にページ、テーブル等のレコードであってもよい。格納を目的として、複数のデータ・オブジェクトが一つのセグメントに論理的に配置され、また複数のセグメントが論理的に一つのグループに配置される。

【００２２】

ＲＡＩＤ－５アレイに関する好ましい実施形態例では、グループはストライプ（アレイ・シリンダとして当業者にしばしば知られている）に対応し、またセグメントはストライプ

（また、アレイ・トラックとしてしばしば知られている）の一つのディスクに対応する。しかし、本発明は、オブジェクト、セグメント、及びグループの論理データ構造をフレキシブルな方法で規定することが可能であると考える。また、当業者はそれらの論理データ構造に対応する代わりの物理的レイアウトが可能であることを理解するであろう。

【００２３】

ホスト・アダプタ１６は、好ましくは記憶システム１４とホスト・コンピュータ１２との間でデータ・レコードを交換する送信及び受信能力を持つ従来のファシリティであってもよい。

【００２４】

記憶装置１８は、好ましくは複数のディスク１９を有するディスク記憶装置であってもよい。好ましくは、記憶装置１８は多重細分性キャッシュ・メモリ３０がどちらかと言えば大きいチャンク（１つ以上のシリンダ）を操作するＲＡＩＤ能力を利用することができるように、ＲＡＩＤ記憶スキームでディスク１９を使用する。

【００２５】

本発明によれば、キャッシュ・メモリ３０は多重細分性マネージャ３６を含む。

本発明の方法によれば、マネージャ３６はキャッシュ・メモリ３０を論理的に第１のキャッシュ３２と第２のキャッシュ３４とに分割する。マネージャ３６は、ホスト・コンピュータ１２からのデータ・アクセス要求に応答して、より小さな細分性のデータ構造、例えば第１のキャッシュ３２に格納されるセグメントを生成する。マネージャ３６は、ホスト・コンピュータ１２からのデータ・アクセス要求に応答して、より大きな細分性のデータ構造、例えば第１のキャッシュ３４に格納されるセグメントを生成する。第１及び第２のキャッシュ３２、３４は、キャッシュ・メモリ３０の指定物理領域、又は論理領域であってもよい。

【００２６】

セグメントを第１のキャッシュに格納し、グループを第２のキャッシュ３４に格納するとともにキャッシュ・メモリ３０を第１のキャッシュ３２と第２のキャッシュ３４とに論理分割することは、本発明の重要な特徴の一つである。この特徴は、一グループ内の一つのデータ・オブジェクトを要求するアプリケーションが同一グループの別のデータ・オブジェクトも要求するが、同一のセグメントである必要性はない可能性を利用する。キャッシュ・ミス率は約４分の１に減少すると推測される。

【００２７】

マネージャ３６は、セグメントＬＲＵリスト３５と、第１のキャッシュ３２のためにセグメントＬＲＵリスト３５を管理する第１のキャッシュ・マネージャ３７とを含むもので、従来のＬＲＵ手順を利用する。また、マネージャ３６はグループＬＲＵリスト３８と、第２のキャッシュ３４のためにグループＬＲＵリスト３８を管理するグループ・マネージャ３９とを含むもので、従来のＬＲＵ手順を利用する。マネージャ３６は、さらに読み込み手順５０、セグメント書き出し手順６０、及びグループ書き出し手順８０を有する。

【００２８】

ここで、図２を参照する。読み込み手順５０は、ホスト・コンピュータ１２が読み込み操作のためにデータ・アクセスを要求する時に開示される。ステップ５１は、要求されたデータがセグメントＬＲＵリスト３５にあるかどうかを判断する。もし要求されたデータがセグメントＬＲＵリスト３５にあるならば、ステップ５７は要求されたデータのコピーを第１のキャッシュ３２からホスト・コンピュータ１２に転送する。

【００２９】

もし要求されたデータがセグメントＬＲＵリスト３５に含まれていない場合（第１のキャッシュ・ミス）、ステップ５２は要求されたデータがグループＬＲＵリスト３８にあるかどうかを判断する。もし要求されたデータがグループＬＲＵリスト３８に含まれている場合、ステップ５４は要求されたデータを含むセグメントのコピーを第２のキャッシュ３４から第１のキャッシュ３２へステージする。ステップ５４は、また第１のキャッシュ・マネージャと協力してセグメントＬＲＵリスト３５をアップデートし、そのようなステージ

されたセグメントのアドレスをリストのＭＲＵ部分に入れる。次に、ステップ５７は要求されたデータのコピーをホスト・コンピュータ１２に転送する。

【００３０】

もし要求されたデータがグループＬＲＵリスト３８にないならば（第２のキャッシュ・ミス）、ステップ５５が要求されたデータを含むグループのコピーをディスク記憶装置１８から第２のキャッシュ３４へステージする。また、ステップ５５は、第２のキャッシュ・マネージャと協力してグループＬＲＵリスト３８をアップデートし、そのようなステージされたグループのアドレスをリストのＭＲＵ部分に入れる。次に、ステップ５４は要求されたデータを含むセグメントを第２のキャッシュ３４から第１のキャッシュ３２へ上記したようにステージする。次に、ステップ５７は要求されたレコードのコピーをホスト・コンピュータ１２に転送する。

【００３１】

ここで、図３及び図４を参照しながら説明する。セグメント書き出し手順６０は、図３に示す同期プロセス６１と、図４に示す非同期プロセス７０とを含む。最初に図３を参照すると、同期プロセス６１はホスト・コンピュータ１２が書き出し操作のためにデータ・アクセスを要求する時に開始される。ステップ６２は、要求されたデータがすでにセグメントＬＲＵリスト３５に存在するかどうかを判断する。もしそうであるならば、ステップ６４は既存のデータを修飾するためにセグメントＬＲＵキャッシュ３２に書き込まれるデータを転送する。また、ステップ６４はセグメントＬＲＵリスト３５も更新する。さらにステップ６４はこのセグメントのセグメント待ちフラグを０にセットする。

【００３２】

もし要求データがセグメントＬＲＵリスト３５に存在しないとステップ６２よって判断されるばらば、ステップ６３は第１のキャッシュ３２に位置を割り当てる。次に、ステップ６４が上記したようにして実行される。この際、書き込まれるデータは既に存在する版を修正しない。

【００３３】

図４を参照する。非同期プロセス７０は通常はスリープ・モードになっており、第２のキャッシュ３４及び／又はディスク記憶装置１８にデステージされる必要がある第１のキャッシュ３２内に修飾されたセグメントが存在する場合にステップ７１によって起こされる。ステップ７１がデステージをする作業が存在すると判断した場合、ステップ７２はセグメントＬＲＵリスト３５に次の修飾されたセグメントを見つける。ステップ７３はこのセグメントがグループＬＲＵリスト３８にあるかどうかを判断する。もしなければ、ステップ７４はセグメント待ちフラグが０に設定されたかどうかを判断する。もしそうであるならば、ステップ７７はセグメント待ちフラグを１に設定する。次に、非同期プロセス７０は、セグメントＬＲＵリスト３５の次の修飾されたセグメントに進むか、又はもし作業のための現在の操作が完了するならば短時間の間スリープ・モードに入る。このことは、ディスク記憶装置１８にステージされる前に作られるセグメントに対する別の修飾を可能とする待ち機構を提供する。

【００３４】

次にステップ７１は再び非同期プロセス７０を起こし、またステップ７２からステップ７４までが繰り返される。それによって、ステップ７４はセグメント待ちフラグが０に設定されていないと判断する。次に、ステップ７６は修飾されたセグメントをディスク記憶装置１８にデステージする。そして、非同期プロセス７０はセグメントＬＲＵリスト３５の次の修飾されたセグメントに進むか、又は作業に対する現行の操作が完了したならば短時間の間スリープ・モードに入る。

【００３５】

もし修飾されたセグメントがグループＬＲＵリスト２８にあるとステップ７３によって判断されるならば、ステップ７３は既にあるグループを修飾するように、第２のキャッシュ３４へ修飾されたセグメントを転送する。また、ステップ７５はグループＬＲＵリスト３８を更新する。さらに、ステップ７５はこのグループに対するグループ待ちフラグを０に

設定する。次に、非同期プロセス７０がスリープ・モードに入る。

【００３６】

ここで、図５を参照する。グループ書き出し手順８０もまた非同期プロセスであり、グループをデステージする作業が存在する時にステップ８１で起こされる。グループ書き出し手順８０がステップ８１によって起こされると、ステップ８２はグループＬＲＵリスト３８に次の修飾されたグループを見つける。ステップ８３は、グループ待ちフラグが０に設定されているかどうかを判断する。もしそうであるならば、ステップ８４はグループ待ちフラグを１に設定する。次に、グループ書き出し手順８０は短時間の間、スリープ・モードに入る。このことは、ディスク記憶装置１８にステージされる前に作られるセグメントに対する別の修飾を可能とする待ち機構を提供する。

【００３７】

次に、ステップ８１は再びグループ書き出し手順８０を起こし、ステップ８２及び８３が繰り返される。それによって、やっとステップ８３はグループ待ちフラグが０に設定されていないと判断する。次に、ステップ８５は修飾されたセグメントをディスク記憶装置１８にデステージする。次に、グループ書き出し手順８０はスリープ・モードに入る。

【００３８】

以上、本発明をその好ましい実施の形態にもとづいて説明したが、本発明は特許請求の範囲で定義されるように、種々の変更及び修飾が本発明の精神及び範囲から逸脱することなく可能であることは当業者に容易に理解されよう。

【００３９】

まとめとして、本発明の構成に関して以下の事項を開示する。

（１）ホスト・コンピュータと記憶装置との間のデータ・パスに位置し、
かつ第１のキャッシュと第２のキャッシュとに論理的に分割されているキャッシュ・メモリを管理する方法であって、
（ａ）前記ホスト・コンピュータからのデータにアクセスする要求に応えて、要求されたデータが前記第１のキャッシュに格納されているかどうかを判断し、もし前記要求されたデータが第１のキャッシュに格納されていなければ、前記要求されたデータが第２のキャッシュに格納されているかどうかを判断するステップと、
（ｂ）前記要求されたデータが前記第２のキャッシュに格納されていないならば、前記記憶装置に格納されたセグメントからなるグループにアクセスし、前記セグメントのグループに含まれた第１のセグメントに前記要求されたデータが含まれるステップと、
（ｃ）前記第２のキャッシュに前記セグメントのグループを格納し、また前記第１のキャッシュに前記要求されたデータが含まれる前記第１のセグメントを格納するステップと、
（ｄ）前記ステップ（ａ）で前記要求されたデータが前記第１のキャッシュに格納されていると判断した場合、又は前記ステップ（ｃ）で前記第１のセグメントを前記第１のキャッシュに格納した時に前記第１のキャッシュからの前記要求されたデータにアクセスするステップと、
を有することを特徴とするキャッシュ・メモリ管理方法。

（２）（ｅ）もしステップ（ａ）が要求されたデータは第１のキャッシュではなく第２のキャッシュに格納されていると判断するならば、前記第２のキャッシュにおいて前記要求されたデータを含む第２のセグメントにアクセスし、前記第２のセグメントを前記第１のキャッシュに格納するステップと、
（ｆ）前記ステップ（ｅ）が前記第１のキャッシュを前記第２のセグメントに格納した時に、前記第１のキャッシュから前記要求されたデータにアクセスするステップと、
をさらに有することを特徴とする上記（１）に記載のキャッシュ・メモリ管理方法。

（３）（ｇ）もし前記ステップ（ａ）で要求されたデータが前記第１のキャッシュに格納されていないと判断したならば、最低使用頻度手順にもとづいて最低使用頻度セグメントを前記第１のキャッシュからデステージするステップと、
（ｈ）もし前記ステップ（ａ）で要求されたデータが前記第２のキャッシュに格納されていないと判断したならば、最小使用頻度手順にもとづいて前記セグメントの最小使用頻度

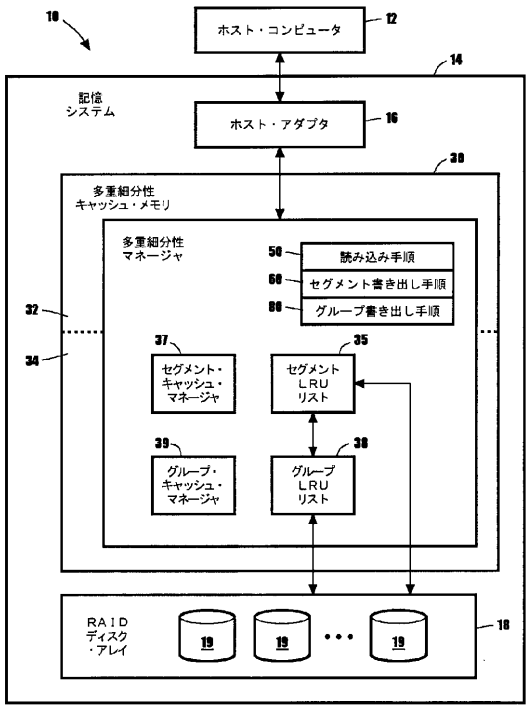
グループを前記第２のキャッシュからデステージするステップと、
をさらに有することを特徴とする上記（２）に記載のキャッシュ・メモリ管理方法。
（４）（ｉ）前記ステップ（ｇ）で使用される前記第１のキャッシュに対する第１の最小使用頻度リストを保持するステップと、
（ｊ）前記ステップ（ｈ）で使用される前記第２のキャッシュに対する第２の最小使用頻度リストを保持するステップと、
をさらに有することを特徴とする上記（２）に記載のキャッシュ・メモリ管理方法。
（５）（ｇ）もし前記アクセス要求が前記要求されたデータの修復バージョンを書き出す要求であるならば、前記修復バージョンを前記第１のキャッシュに転送するステップと、
（ｈ）前記記憶装置に対して前記第１のキャッシュから前記修復バージョンのコピーを直接的に、又は前記第２のキャッシュを介して間接的に転送するステップと、
をさらに有することを特徴とする上記（２）に記載のキャッシュ・メモリ管理方法。
（６）前記ステップ（ｈ）は、もし前記要求されたデータが前記第２のキャッシュになければ直接的に転送を行い、またもし要求されたデータが前記第２のキャッシュにあれば間接的に転送を行うことを特徴とする上記（５）に記載のキャッシュ・メモリ管理方法。
（７）前記記憶装置は、ＲＡＩＤ記憶装置であり、前記セグメントの各々は論理アレイ・トラックであり、また前記グループの各々は論理アレイ・シリンダであることを特徴とする上記（６）に記載のキャッシュ・メモリ管理方法。
（８）ホスト・コンピュータと記憶装置との間のデータ・パスに位置した多重細分性キャッシュ・メモリ・システムであって、
第１のキャッシュと第２のキャッシュとに論理的に分割されるキャッシュ・メモリと、
データにアクセスするために前記ホスト・コンピュータからの要求を処理するための多重細分性マネージャとを備え、さらに、
前記多重細分性マネージャは、
（ａ）前記ホスト・コンピュータからのデータにアクセスする要求に応えて、要求されたデータが前記第１のキャッシュに格納されているかどうかを判断し、もし前記要求されたデータが第１のキャッシュに格納されていなければ、前記要求されたデータが第２のキャッシュに格納されているかどうかを判断するステップと、
（ｂ）前記要求されたデータが前記第２のキャッシュに格納されていないならば、前記記憶装置に格納されたセグメントからなるグループにアクセスし、前記セグメントのグループに含まれた第１のセグメントに前記要求されたデータが含まれるステップと、
（ｃ）前記第２のキャッシュに前記セグメントのグループを格納し、また前記第１のキャッシュに前記要求されたデータが含まれる前記第１のセグメントを格納するステップと、
（ｄ）前記ステップ（ａ）で前記要求されたデータが前記第１のキャッシュに格納されていると判断した場合、又は前記ステップ（ｃ）で前記第１のセグメントを前記第１のキャッシュに格納した時に前記第１のキャッシュからの前記要求されたデータにアクセスするステップと、
を有する手順を実行することを特徴とする多重細分性キャッシュ・メモリ・システム。
（９）前記手順は、
（ｅ）もしステップ（ａ）が要求されたデータは第１のキャッシュではなく第２のキャッシュに格納されていると判断するならば、前記第２のキャッシュにおいて前記要求されたデータを含む第２のセグメントにアクセスし、前記第２のセグメントを前記第１のキャッシュに格納するステップと、
（ｆ）前記ステップ（ｅ）が前記第１のキャッシュを前記第２のセグメントに格納した時に、前記第１のキャッシュから前記要求されたデータにアクセスするステップと、
をさらに有することを特徴とする上記（８）に記載の多重細分性キャッシュ・メモリ・システム。
（１０）前記手順は、
（ｇ）もし前記ステップ（ａ）で要求されたデータが前記第１のキャッシュに格納されていないと判断したならば、最低使用頻度手順にもとづいて最低使用頻度セグメントを前記

第１のキャッシュからデステージするステップと、
（ｈ）もし前記ステップ（ａ）で要求されたデータが前記第２のキャッシュに格納されていないと判断したならば、最小使用頻度手順にもとづいて前記セグメントの最小使用頻度グループを前記第２のキャッシュからデステージするステップと、
をさらに有することを特徴とする上記（９）に記載の多重細分性キャッシュ・メモリ・システム。
（１１）前記手順は、
（ｉ）前記ステップ（ｇ）で使用される前記第１のキャッシュに対する第１の最小使用頻度リストを保持するステップと、
（ｊ）前記ステップ（ｈ）で使用される前記第２のキャッシュに対する第２の最小使用頻度リストを保持するステップと、
をさらに有することを特徴とする上記（１０）に記載の多重細分性キャッシュ・メモリ・システム。
（１２）前記手順は、
（ｇ）もし前記アクセス要求が前記要求されたデータの修復バージョンを書き出す要求であるならば、前記修復バージョンを前記第１のキャッシュに転送するステップと、
（ｈ）前記記憶装置に対して前記第１のキャッシュから前記修復バージョンのコピーを直接的に、又は前記第２のキャッシュを介して間接的に転送するステップと、
をさらに有することを特徴とする上記（９）に記載の多重細分性キャッシュ・メモリ・システム。
（１３）前記ステップ（ｈ）は、もし前記要求されたデータが前記第２のキャッシュになければ直接的に転送を行い、またもし要求されたデータが前記第２のキャッシュにあれば間接的に転送を行うことを特徴とする上記（１２）に記載の多重細分性キャッシュ・メモリ・システム。
（１４）前記記憶装置は、ＲＡＩＤ記憶装置であり、前記セグメントの各々は論理アレイ・トラックであり、また前記グループの各々は論理アレイ・シリンダであることを特徴とする上記（６）に記載の多重細分性キャッシュ・メモリ・システム。
（１５）ホスト・コンピュータと記憶装置との間のデータ・パスに位置した多重細分性キャッシュ・メモリ・システムを制御するためのメモリ媒体であって、
前記キャッシュ・メモリ・システムは、キャッシュ・マネージャと、第１のキャッシュと第２のキャッシュとに論理的に分割されるキャッシュ・メモリとを有し、
前記メモリ媒体は、
前記多重細分性キャッシュ・メモリを制御するための手段を有し、該制御は、
（ａ）前記ホスト・コンピュータからのデータにアクセスする要求に応えて、要求されたデータが前記第１のキャッシュに格納されているかどうかを判断し、もし前記要求されたデータが第１のキャッシュに格納されていなければ、前記要求されたデータが第２のキャッシュに格納されているかどうかを判断するステップと、
（ｂ）前記要求されたデータが前記第２のキャッシュに格納されていないならば、前記記憶装置に格納されたセグメントからなるグループにアクセスし、前記セグメントのグループに含まれた第１のセグメントに前記要求されたデータが含まれるステップと、
（ｃ）前記第２のキャッシュに前記セグメントのグループを格納し、また前記第１のキャッシュに前記要求されたデータが含まれる前記第１のセグメントを格納するステップと、
（ｄ）前記ステップ（ａ）で前記要求されたデータが前記第１のキャッシュに格納されていると判断した場合、又は前記ステップ（ｃ）で前記第１のセグメントを前記第１のキャッシュに格納した時に前記第１のキャッシュからの前記要求されたデータにアクセスするステップと、
を有する手順によることを特徴とするメモリ媒体。
（１６）前記手順は、
（ｅ）もしステップ（ａ）が要求されたデータは第１のキャッシュではなく第２のキャッシュに格納されていると判断するならば、前記第２のキャッシュにおいて前記要求された

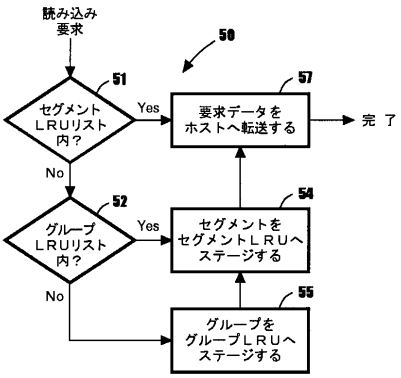
データを含む第２のセグメントにアクセスし、前記第２のセグメントを前記第１のキャッシュに格納するステップと、
（ｆ）前記ステップ（ｅ）が前記第１のキャッシュを前記第２のセグメントに格納した時に、前記第１のキャッシュから前記要求されたデータにアクセスするステップと、
をさらに有することを特徴とする上記（１５）に記載のメモリ媒体。
（１７）前記手順は、
（ｇ）もし前記ステップ（ａ）で要求されたデータが前記第１のキャッシュに格納されていないと判断したならば、最低使用頻度手順にもとづいて最低使用頻度セグメントを前記第１のキャッシュからデステージするステップと、
（ｈ）もし前記ステップ（ａ）で要求されたデータが前記第２のキャッシュに格納されていないと判断したならば、最小使用頻度手順にもとづいて前記セグメントの最小使用頻度グループを前記第２のキャッシュからデステージするステップと、
をさらに有することを特徴とする上記（１６）に記載のメモリ媒体。
（１８）前記手順は、
（ｉ）前記ステップ（ｇ）で使用される前記第１のキャッシュに対する第１の最小使用頻度リストを保持するステップと、
（ｊ）前記ステップ（ｈ）で使用される前記第２のキャッシュに対する第２の最小使用頻度リストを保持するステップと、
をさらに有することを特徴とする上記（１７）に記載のメモリ媒体。
（１９）前記手順は、
（ｇ）もし前記アクセス要求が前記要求されたデータの修復バージョンを書き出す要求であるならば、前記修復バージョンを前記第１のキャッシュに転送するステップと、
（ｈ）前記記憶装置に対して前記第１のキャッシュから前記修復バージョンのコピーを直接的に、又は前記第２のキャッシュを介して間接的に転送するステップと、
をさらに有することを特徴とする上記（１６）に記載のメモリ媒体。
（２０）前記ステップ（ｈ）は、もし前記要求されたデータが前記第２のキャッシュになければ直接的に転送を行い、またもし要求されたデータが前記第２のキャッシュにあれば間接的に転送を行うことを特徴とする上記（１９）に記載のメモリ媒体。
（２１）前記記憶装置は、ＲＡＩＤ記憶装置であり、前記セグメントの各々は論理アレイ・トラックであり、また前記グループの各々は論理アレイ・シリンダであることを特徴とする上記（２０）に記載のメモリ媒体。
【図面の簡単な説明】
【図１】　本発明にもとづく多重細分性キャッシュ・メモリを有するコンピュータ・システムの概略的構成を説明するためのブロック図である。
【図２】　本発明にもとづく多重細分性プログラムの読み込み手順を示すフローチャートである。
【図３】　本発明にもとづく多重細分性プログラムの書き出し手順を示すフローチャートである。
【図４】　本発明にもとづく多重細分性プログラムの書き出し手順を示すフローチャートである。
【図５】　本発明にもとづく多重細分性プログラムの書き出し手順を示すフローチャートである。
【符号の説明】
１０　　コンピュータ・システム
１２　　ホスト・コンピュータ
１４　　記憶システム
１６　　ホスト・アダプタ
１８　　記憶装置
１９　　ディスク
３０　　多重細分性キャッシュ・メモリ

３２　　第１のキャッシュ
３４　　第２のキャッシュ
３５　　セグメントＬＲＵリスト
３６　　多重細分性マネージャ
３７　　第１のキャッシュ・マネージャ
３８　　グループＬＲＵリスト
３９　　グループ・マネージャ
５０　　読み込み手順
６０　　セグメント書き出し手順
６１　　同期プロセス
７０　　非同期プロセス
８０　　グループ書き出し手順

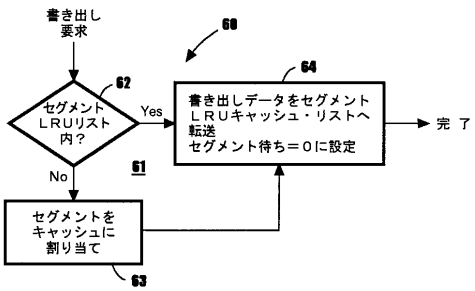
【図１】



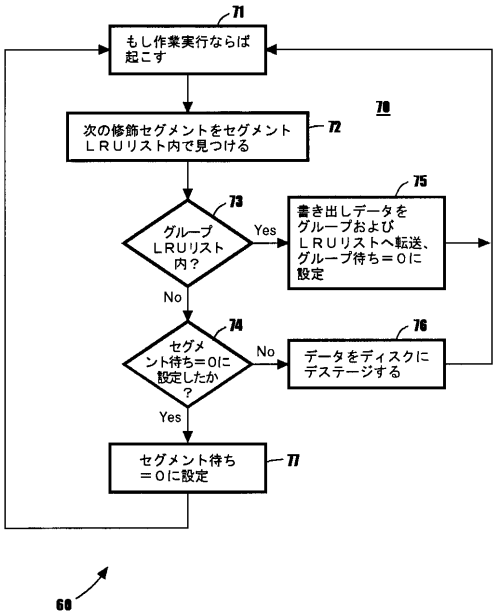
【図２】



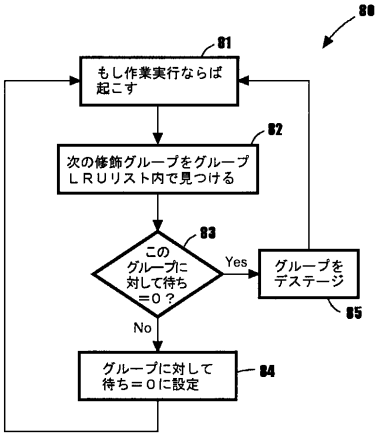
【図３】



【図４】


71
もし作業実行ならば起こす
70
次の修飾セグメントをセグメントＬＲＵリスト内で見つける
72
73
グループＬＲＵリスト内？
Yes
75
書き出しデータをグループおよびＬＲＵリストへ転送、グループ待ち＝０に設定
No
74
セグメント待ち＝０に設定したか？
No
76
データをディスクにデステージする
Yes
77
セグメント待ち＝０に設定
60

【図５】


80
81
もし作業実行ならば起こす
82
次の修飾グループをグループＬＲＵリスト内で見つける
83
このグループに対して待ち＝０？
Yes
グループをデステージ
85
No
グループに対して待ち＝０に設定
84

フロントページの続き

(51)Int.Cl.[7] FI
G06F 3/06 302A
G06F 3/06 540
G06F 12/12 557Z

(72)発明者 ブルース・マクナット
アメリカ合衆国95020 カリフォルニア州、 ギルロイ、 デニオ・アベニュー 95

審査官 清木 泰

(56)参考文献 特開平4－112253（JP，A）
特開平6－243003（JP，A）
特開平2－281350（JP，A）

(58)調査した分野(Int.Cl.[7]，DB名)
G06F12/08-12/12
G06F 3/06- 3/08