(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6294569号
(P6294569)

(45)発行日　平成30年3月14日(2018.3.14)　(24)登録日　平成30年2月23日(2018.2.23)

(51)Int.Cl.　F I
G06F 13/10 (2006.01)　G06F 13/10　340A
G06F 3/06 (2006.01)　G06F 3/06　302A

請求項の数 15　(全 18 頁)

(21)出願番号　特願2017-524248(P2017-524248)
(86)(22)出願日　平成27年6月19日(2015.6.19)
(86)国際出願番号　PCT/JP2015/067676
(87)国際公開番号　WO2016/203629
(87)国際公開日　平成28年12月22日(2016.12.22)
審査請求日　平成29年8月10日(2017.8.10)

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　110000279
特許業務法人ウィルフォート国際特許事務所
(72)発明者　井澤　信介
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　杉本　定広
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　坂下　悠貴
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内

最終頁に続く

(54)【発明の名称】ストレージシステム及びキャッシュ制御方法

(57)【特許請求の範囲】
【請求項１】
ホストシステムに接続されたストレージシステムであって、
複数のストレージ装置を有し、
前記複数のストレージ装置の各々が、前記ホストシステムに接続されており、
前記複数のストレージ装置の各々が、記憶デバイスと、その記憶デバイスに対して入出力されるデータがキャッシュされるキャッシュメモリ領域とを有し、
各ストレージ装置が、そのストレージ装置内のキャッシュメモリ領域だけでなく、そのストレージ装置に接続されているいずれの他のストレージ装置内のキャッシュメモリ領域も管理しており、いずれの他のストレージ装置のキャッシュメモリ領域にデータを入出力することができ、
論理ボリュームのアドレスを指定したＩ／Ｏ（Input/Output）要求を前記複数のストレージ装置のうちのいずれかが受信した場合、前記複数のストレージ装置のうちの少なくとも１つが、対象Ｉ／Ｏパターンと接続形態とのうちの少なくとも１つに基づいて、前記受信したＩ／Ｏ要求に従うデータであるＩ／Ｏデータがキャッシュされるキャッシュ先ストレージ装置を決定し、
前記対象Ｉ／Ｏパターンは、複数のＩ／Ｏパターンのうち、前記受信したＩ／Ｏ要求に従うＩ／Ｏが属するＩ／Ｏパターンであり、
前記複数のＩ／Ｏパターンの各々は、前記論理ボリュームにおけるＩ／Ｏ先アドレスがランダムとなるかシーケンシャルとなるかに関するパターンであり、

前記接続形態は、Ｉ／Ｏ要求を受信するストレージ装置が確定しているか否かである、ストレージシステム。
【請求項２】
前記対象Ｉ／Ｏパターンが、ランダムライト及びランダムリードのうちのいずれかの場合、前記キャッシュ先ストレージ装置が、前記Ｉ／Ｏ要求を受信したストレージ装置である、
請求項１記載のストレージシステム。
【請求項３】
前記対象Ｉ／Ｏパターンが、シーケンシャルリードの場合、前記キャッシュ先ストレージ装置が、前記Ｉ／Ｏデータを格納している記憶デバイスを有するストレージ装置である、
請求項１記載のストレージシステム。
【請求項４】
前記接続形態が、Ｉ／Ｏ要求を受信するストレージ装置が確定していることを意味している場合、前記キャッシュ先ストレージ装置が、前記Ｉ／Ｏ要求を受信したストレージ装置であるホスト側ストレージ装置である、
請求項１記載のストレージシステム。
【請求項５】
前記ホスト側ストレージ装置が、複数のストレージコントローラを有し、
複数のストレージコントローラの各々が、前記ホストシステムに接続されており、
複数のストレージコントローラの各々が、キャッシュメモリ領域を有し、
前記接続形態が、Ｉ／Ｏ要求を受信するストレージコントローラが確定していることを意味している場合、前記Ｉ／Ｏデータがキャッシュされるキャッシュ先ストレージコントローラは、前記複数のストレージコントローラのうち、Ｉ／Ｏ要求を受信することが確定しているストレージコントローラである確定ストレージコントローラである、
請求項４記載のストレージシステム。
【請求項６】
Ｉ／Ｏ要求を受信するストレージ装置が確定していないことを前記接続形態が意味している場合、前記複数のストレージ装置のうちの少なくとも１つが、前記対象Ｉ／Ｏパターンに基づいて前記キャッシュ先ストレージ装置を決定する、
請求項４記載のストレージシステム。
【請求項７】
前記対象Ｉ／Ｏパターンが、ランダムライト及びランダムリードのうちのいずれかの場合、前記キャッシュ先ストレージ装置が、前記ホスト側ストレージ装置である、
請求項６記載のストレージシステム。
【請求項８】
前記ホスト側ストレージ装置が、複数のストレージコントローラを有し、
複数のストレージコントローラの各々が、前記ホストシステムに接続されており、
複数のストレージコントローラの各々が、キャッシュメモリ領域を有し、
前記接続形態が、Ｉ／Ｏ要求を受信するストレージコントローラが確定していることを意味している場合、前記Ｉ／Ｏデータがキャッシュされるキャッシュ先ストレージコントローラは、前記複数のストレージコントローラのうち、Ｉ／Ｏ要求を受信することが確定しているストレージコントローラである確定ストレージコントローラであり、
前記接続形態が、Ｉ／Ｏ要求を受信するストレージコントローラが確定していないことを意味しており、且つ、前記対象Ｉ／Ｏパターンが、ランダムライト及びランダムリードのうちのいずれかの場合、前記キャッシュ先ストレージコントローラは、前記複数のストレージコントローラからのいずれかである、
請求項７記載のストレージシステム。
【請求項９】
前記対象Ｉ／Ｏパターンが、シーケンシャルリードの場合、前記キャッシュ先ストレー

ジ装置が、前記Ｉ／Ｏデータを格納している記憶デバイスを有するストレージ装置であり、
前記キャッシュ先ストレージ装置のキャッシュメモリ領域に、シーケンシャルリードに従い読み出され得るデータが先読みされる、
請求項８記載のストレージシステム。
【請求項１０】
前記対象Ｉ／Ｏパターンが、ランダムライト、ランダムリード及びシーケンシャルリードのうちのいずれでもない場合、前記キャッシュ先ストレージ装置が、前記複数のストレージ装置のうちのいずれかである、
請求項９記載のストレージシステム。
【請求項１１】
前記複数のストレージ装置の各々が、ＡＬＵＡ（Asymmetric Logical Unit Access）を使用して前記ホストシステムと通信するようになっており、
前記確定ストレージコントローラは、前記複数のストレージコントローラと前記ホストシステム間の複数のパスのうちの最も優先度が高いパスに接続されているストレージコントローラである、
請求項１０記載のストレージシステム。
【請求項１２】
前記確定ストレージコントローラは、前記複数のストレージコントローラと前記ホストシステム間の複数のパスのうち唯一の通信可能状態パスに接続されているストレージコントローラである、
請求項１０記載のストレージシステム。
【請求項１３】
前記確定ストレージコントローラは、前記複数のストレージコントローラと前記ホストシステム間の複数のパスのうちＡＬＵＡ（Asymmetric Logical Unit Access）に従う優先度が最も高いパスに接続されているストレージコントローラである、
請求項１１記載のストレージシステム。
【請求項１４】
スケールアウト型のストレージシステムである請求項１０記載のストレージシステム。
【請求項１５】
それぞれホストシステムに接続された複数のストレージ装置を有するストレージシステムのキャッシュ制御方法であって、
前記複数のストレージ装置の各々が、記憶デバイスと、その記憶デバイスに対して入出力されるデータがキャッシュされるキャッシュメモリ領域とを有し、
各ストレージ装置が、そのストレージ装置内のキャッシュメモリ領域だけでなく、そのストレージ装置に接続されているいずれの他のストレージ装置内のキャッシュメモリ領域も管理しており、いずれの他のストレージ装置のキャッシュメモリ領域にデータを入出力することができ、
前記キャッシュ制御方法が、
論理ボリュームのアドレスを指定したＩ／Ｏ（Input/Output）要求を前記複数のストレージ装置のうちのいずれかにより受信し
対象Ｉ／Ｏパターンと接続形態とのうちの少なくとも１つに基づいて、前記受信したＩ／Ｏ要求に従うデータであるＩ／Ｏデータがキャッシュされるキャッシュ先ストレージ装置を決定し、
前記対象Ｉ／Ｏパターンは、複数のＩ／Ｏパターンのうち、前記受信したＩ／Ｏ要求に従うＩ／Ｏが属するＩ／Ｏパターンであり、
前記複数のＩ／Ｏパターンの各々は、前記論理ボリュームにおけるＩ／Ｏ先アドレスがランダムとなるかシーケンシャルとなるかに関するパターンであり、
前記接続形態は、Ｉ／Ｏ要求を受信するストレージ装置が確定しているか否かである、
キャッシュ制御方法。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、概して、ストレージシステムのキャッシュ制御に関する。
【背景技術】
【０００２】
近年、ソーシャルネットワーキングサービスや金融、医療及び交通といった社会インフラに関する膨大なデータを分析することにより、新たな価値を生み出すビッグデータ分析と呼ばれる技術が実用化されつつある。
【０００３】
ビッグデータ分析においては、社会インフラから収集される入力データ及び分析結果である出力データの容量が共に非常に大きく、時間と共に増加し続ける。ストレージシステムは、このような膨大なビックデータを安全に保管及び管理するプラットフォームとして、企業ＩＴシステムにおいて重要な位置づけとされている。
【０００４】
例えば、クラウドサービスを提供する企業は、導入コストを削減するために、サービスの初期稼働時は当面必要となる資源でストレージシステムを構築する。ストレージシステムとして、例えば、スケールアウト型ストレージシステムが採用される。すなわち、サービスが稼動し、資源の利用量が増加してきたら、ストレージ装置（ノード）の台数を増やすことで、システム全体の処理能力を向上させることが行われる。
【０００５】
ストレージシステムの処理能力を向上させる１つの手段として、ストレージシステムのキャッシュ制御が考えられる。キャッシュ制御に関して、例えば特許文献１の技術が知られている。特許文献１には、キャッシュ制御として、割り当てられるキャッシュ容量の制御が行われる。
【先行技術文献】
【特許文献】
【０００６】
【特許文献１】特開２０１０－２８６９２３号公報
【発明の概要】
【発明が解決しようとする課題】
【０００７】
一般に、スケールアウト型ストレージシステムにおけるストレージ装置間接続は、「疎結合」である。本明細書において、「疎結合」とは、一方のストレージ装置から他方のストレージ装置のキャッシュメモリ領域（ＣＭ領域）にデータを入出力できないストレージ装置間接続を意味する。疎結合のストレージ装置では、一方のストレージ装置は、自分のＣＭ領域を管理するが、他方のストレージ装置のＣＭ領域を管理しない。このため、一方のストレージ装置は、自分のＣＭ領域からデータの格納先キャッシュセグメント（ＣＭサブ領域の一例）を確保できるが、他のストレージ装置のＣＭ領域からキャッシュセグメントを確保できない。結果として、疎結合のストレージ装置間でＩ／Ｏデータ（Ｉ／Ｏ要求に従い入出力される対象のデータ）が転送された場合、そのＩ／Ｏデータは、転送元のストレージ装置のＣＭ領域にも転送先のストレージ装置のＣＭ領域にも格納されることになる。
【０００８】
そこで、スケールアウト型ストレージシステムにおけるストレージ装置間接続を、「密結合」にすることを検討する。本明細書において、「密結合」とは、一方のストレージ装置から他方のストレージ装置のＣＭ領域にＩ／Ｏデータを入出力できるストレージ装置間接続を意味する。密結合のストレージ装置では、一方のストレージ装置は、自分のＣＭ領

域だけでなく、他方のストレージ装置のＣＭ領域を管理する。このため、一方のストレージ装置は、自分のＣＭ領域と他のストレージ装置のＣＭ領域のいずれからもキャッシュセグメントを確保できる。結果として、密結合のストレージ装置間でＩ／Ｏデータが転送された場合、そのＩ／Ｏデータは、転送元のストレージ装置のＣＭ領域と転送先のストレージ装置のＣＭ領域とのいずれか一方にのみ格納されることになる。

【０００９】

このように、密結合では、１つのＩ／Ｏデータについて、キャッシュセグメントが確保されるＣＭ領域は１つでよい（キャッシュセグメントの二重化が行われることもあり得るが、その場合、メインのキャッシュセグメントが確保されるＣＭ領域は１つでよい）。

【００１０】

しかし、そのために、ＣＭ領域からホストへのＩ／Ｏデータがストレージシステムにおいて転送される回数が多くなり得る。以下、具体例を説明する。なお、課題の具体例の説明では、各ストレージ装置において、ストレージコントローラ（ＣＴＬ）が二重化されており、各ＣＴＬが、ＣＭ領域を有する。Ｉ／Ｏ要求を受信したストレージ装置を、「ホスト側ストレージ装置」と言い、ホスト側ストレージ装置の要素ＸＸを、「ホスト側ＸＸ」と言い、Ｉ／Ｏ要求を受信したホスト側ＣＴＬを「Ｉ／Ｏ受信ＣＴＬ」と言うことができる。また、Ｉ／Ｏデータを記憶しているＰＤＥＶ（物理記憶デバイス）を有するストレージ装置を、「ＰＤＥＶ側ストレージ装置」と言い、ＰＤＥＶ側ストレージ装置の要素ＸＸを、「ＰＤＥＶ側ＸＸ」と言うことができる。

【００１１】

（課題の具体例１（図８））

【００１２】

ホスト側ＣＴＬ５１０４Ａａが、ホスト５１０２からのリード要求を受信したときに、ＰＤＥＶ側ＣＴＬ５１０４Ｂａ内のＣＭ領域５２０５Ｂａからキャッシュセグメントを確保したとする。この場合、ＰＤＥＶ５１１０からＰＤＥＶ側ＣＭ領域５２０５Ｂａにリードデータが格納され、その後に２段転送が行われる。すなわち、リードデータが、ＰＤＥＶ側ＣＭ領域５２０５Ｂａからホスト側ＣＴＬ５１０４Ａａ内のバッファ領域５２０４Ａａに転送され（矢印９０１）、そのホスト側バッファ領域５２０４Ａａからホスト５１０２にリードデータが転送される（矢印９０２）。

【００１３】

その後に、ホスト側ＣＴＬ５１０４Ａａが同じリードデータのリード要求を受信しても、同様の２段転送が発生する。そのリードデータのＣＭ領域はＰＤＥＶ側ＣＭ領域５２０５Ｂａであり、且つ、ホスト側バッファ領域５２０４Ａａには、ＣＭ領域と異なり、転送されたリードデータが残らないためである。

【００１４】

このように、リードデータのＣＭ領域がＰＤＥＶ側であると、同一データのリード要求を受信する度に、ストレージ装置間データ通信を含んだ２段転送が生じる。このため、性能低下が起こり得る。２段転送を避けるために、リードデータのＣＭ領域をホスト側ＣＭ領域にすることが考えられるが、それは容易ではない。いずれのストレージ装置がホスト側ストレージ装置になるかは、リード要求を受信するまでわからないためである。言い換えれば、いずれのストレージ装置もホストからリード要求を受信し得るためである。

【００１５】

（課題の具体例２（図９））

【００１６】

同一ストレージ装置５１０１Ａ内のＣＴＬ５１０４Ａａ及び５１０４Ａｂのいずれも、ホスト５１０２からＩ／Ｏ要求を受信し得るとする。言い換えれば、ストレージ装置５１０１Ａがホスト５１０２にＡｃｔｉｖｅ／Ａｃｔｉｖｅで接続されているとする。また、ストレージ装置５１０１Ａが、少なくとも１つのリード要求の処理としてシーケンシャルリードを実行することを決定した場合、ＣＭ領域５２０５Ａａ又は５２０５Ａｂに、シーケンシャルリードによりやがて読み出されることになると予測されるデータを先読みする

（前もって読み出す）ようになっているとする。

【００１７】

一方のＣＴＬ５１０４Ａａが、ホスト５１０２からリード要求を受信して、ＣＭ領域５２０５Ａａからキャッシュセグメントを確保し、シーケンシャルリードの実行を決定したとする。この場合、ＰＤＥＶ５１１０からＰＤＥＶ側ＣＴＬ５１０４Ｂａのバッファ領域５２０４Ｂａにリードデータが転送され、そのリードデータが、そのＰＤＥＶ側バッファ領域５２０４Ｂａからホスト側ＣＭ領域５２０５Ａａ（確保されたキャッシュセグメント）に転送され、ホスト側ＣＭ領域５２０５Ａａからホスト５１０２へ転送される。その一連の処理に並行して（又はその一連の処理の後に）、シーケンシャルに読み出されると予測されるデータが、ＰＤＥＶ５１１０からＰＤＥＶ側バッファ領域５２０４Ｂａに転送され（矢印１００１）、そのデータが、そのＰＤＥＶ側バッファ領域５２０４Ｂａからホスト側ＣＭ領域５２０５Ａａに転送される（矢印１００２）。つまり、ホスト側ＣＭ領域５２０５Ａａにデータが先読みされる。

【００１８】

しかし、その後、先読みされたデータを指定したリード要求を、他方のＣＴＬ５１０４Ａｂが受信したとする。この場合、データが、ＣＭ領域５２０５ＡａからＣＴＬ５１０４Ａｂのバッファ領域５２０４Ａｂに転送され（矢印１００３）、そのデータが、バッファ領域５２０４Ａｂからホスト５１０２へ転送されることになる（矢印１００４）。

【００１９】

以上の通り、先のリード要求を受信した一方のＣＴＬ５１０４ＡａのＣＭ領域５２０５Ａａにデータが先読みされた後に、他方のＣＴＬ５１０４Ａｂがその先読みされたデータのリード要求を受信すると、ＣＴＬ５１０４Ａａ及び５１０４Ａｂ間のデータ転送が発生する。この結果、ＣＴＬ５１０４Ａａ及び５１０４Ａｂ間のパス５１０９Ａの帯域が消費されることになる。シーケンシャルリードが必要になる分析又はバックアップ等のユースケースでは、ＣＴＬ間帯域の消費が大きくなりがちであり、ＣＴＬ間のデータ転送を避けることが望ましい。

【００２０】

以上の課題は、特許文献１のようなキャッシュ容量制御を利用しても解決することはできない。この種の課題は、背景技術で述べた用途のストレージシステム以外の用途のストレージシステムについてもあり得る。

【課題を解決するための手段】

【００２１】

論理ボリュームのアドレスを指定したＩ／Ｏ（Input/Output）要求を複数のストレージ装置のうちのいずれかが受信した場合、対象Ｉ／Ｏパターンと接続形態とのうちの少なくとも１つに基づいて、受信したＩ／Ｏ要求に従うデータであるＩ／Ｏデータがキャッシュされるキャッシュ先ストレージ装置が決定される。キャッシュ先ストレージ装置のＣＭ領域にＩ／Ｏデータがキャッシュされることになる。対象Ｉ／Ｏパターンは、複数のＩ／Ｏパターンのうち、受信したＩ／Ｏ要求に従うＩ／Ｏが属するＩ／Ｏパターンである。複数のＩ／Ｏパターンの各々は、論理ボリュームにおけるＩ／Ｏ先アドレスがランダムとなるかシーケンシャルとなるかに関するパターンである。接続形態は、受信したＩ／Ｏ要求で指定されているアドレスと同一アドレスが指定されたＩ／Ｏ要求を受信するストレージ装置が確定しているか否かである。

【発明の効果】

【００２２】

対象Ｉ／Ｏパターン及び接続形態のうちの少なくとも１つに基づいてキャッシュ先ストレージ装置が決定される。このため、ＣＭ領域からホストへのＩ／Ｏデータがストレージシステムにおいて転送される回数を減らすことができる。

【図面の簡単な説明】

【００２３】

【図１】実施例に係る計算機システムの構成例を示す図である。

【図２】メモリの論理構成例を示す図である。
【図３】プログラム領域に格納されるプログラムの一例を示す図である。
【図４】共有メモリ領域に格納される情報の一例を示す図である。
【図５】ローカルメモリ領域に格納される情報の一例を示す図である
【図６】キャッシュセグメント属性管理テーブルの構成例を示す図である。
【図７】キャッシュセグメント管理テーブルの構成例を示す図である。
【図８】比較例の具体例１の概要図である。
【図９】比較例の具体例２の概要図である。
【図１０】キャッシュ先ＣＴＬ決定処理の流れを示すフローチャートである。
【図１１】確定ポートを有するＣＴＬがキャッシュ先ＣＴＬであることの利点の一例の説明図である。
【図１２】ＰＤＥＶ側ＣＴＬがキャッシュ先ＣＴＬであることの利点の一例の説明図である。
【発明を実施するための形態】
【００２４】
以下、図面を参照して一実施例を説明する。
【００２５】
以下の説明では、「ａｂｃテーブル」の表現にて情報を説明することがあるが、情報は、テーブル以外のデータ構成で表現されていてもよい。データ構成に依存しないことを示すために「ａｂｃテーブル」のうちの少なくとも１つを「ａｂｃ情報」と呼ぶことができる。また、以下の説明において、各テーブルの構成は一例であり、１つのテーブルは、２以上のテーブルに分割されてもよいし、２以上のテーブルの全部又は一部が１つのテーブルであってもよい。
【００２６】
また、以下の説明では、要素のＩＤは、要素の識別情報（例えば識別子）のことであり、識別情報は、文字、数字、記号又はそれらの組合せ等（例えば名前）で表現可能である。
【００２７】
また、以下の説明では、同種の要素を区別しないで説明する場合には、参照符号における共通符号（又は参照符号）を使用し、同種の要素を区別して説明する場合は、参照符号（又は要素のＩＤ）を使用することがある。
【００２８】
また、以下の説明では、「記憶部」は、メモリを含んだ１以上の記憶デバイスでよい。例えば、記憶部は、主記憶デバイス（典型的には揮発性のメモリ）及び補助記憶デバイス（典型的には不揮発性の記憶デバイス）のうちの少なくとも主記憶デバイスでよい。また、記憶部は、ＣＭ領域（例えばキャッシュメモリ又はその一部領域）とバッファ領域（例えばバッファメモリ又はその一部領域）とのうちの少なくとも１つを含んでもよい。
【００２９】
また、以下の説明では、「ＰＤＥＶ」は、物理的な記憶デバイスを意味し、典型的には、不揮発性の記憶デバイス（例えば補助記憶デバイス）でよい。ＰＤＥＶは、例えば、ＨＤＤ（Hard Disk Drive）又はＳＳＤ（Solid State Drive)でよい。
【００３０】
また、以下の説明では、「ＲＡＩＤ」は、Redundant Array of Independent (or Inexpensive) Disksの略である。ＲＡＩＤグループは、複数のＰＤＥＶで構成され、そのＲＡＩＤグループに関連付けられたＲＡＩＤレベルに従いデータを記憶する。ＲＡＩＤグループは、パリティグループと呼ばれてもよい。パリティグループは、例えば、パリティを格納するＲＡＩＤグループのことでよい。
【００３１】
また、以下の説明では、「プログラム」を主語として処理を説明する場合があるが、プログラムは、プロセッサ（例えばＣＰＵ（Central Processing Unit））によって実行

されることで、定められた処理を、適宜に記憶部（例えばメモリ）及び／又はインターフェースデバイス（例えば通信ポート）等を用いながら行うため、処理の主語がプロセッサとされてもよい。プログラムを主語として説明された処理は、プロセッサあるいはそのプロセッサを有する装置又はシステムが行う処理としてもよい。また、プロセッサは、処理の一部または全部を行うハードウェア回路を含んでもよい。プログラムは、プログラムソースから計算機のような装置にインストールされてもよい。プログラムソースは、例えば、プログラム配布サーバまたは計算機が読み取り可能な記憶メディアであってもよい。プログラムソースがプログラム配布サーバの場合、プログラム配布サーバはプロセッサ（例えばＣＰＵ）と記憶部を含み、記憶部はさらに配布プログラムと配布対象であるプログラムとを記憶してよい。そして、プログラム配布サーバのプロセッサが配布プログラムを実行することで、プログラム配布サーバのプロセッサは配布対象のプログラムを他の計算機に配布してよい。また、以下の説明において、２以上のプログラムが１つのプログラムとして実現されてもよいし、１つのプログラムが２以上のプログラムとして実現されてもよい。

【００３２】

また、以下の説明では、「ホストシステム」は、ストレージシステムにＩ／Ｏ要求を送信するシステムであり、インターフェースデバイスと、記憶部（例えばメモリ）と、それらに接続されたプロセッサとを有してよい。ホストシステムは、１以上のホスト計算機で構成されてよい。少なくとも１つのホスト計算機は、物理的な計算機でよく、ホストシステムは、物理的なホスト計算機に加えて仮想的なホスト計算機を含んでよい。

【００３３】

以下、Ｉ／Ｏパターン及び接続形態の少なくとも１つに応じてキャッシュ先ストレージ装置を決定するストレージシステムの例を説明する。なお、以下に説明する実施例は、請求の範囲にかかる発明を限定するものではなく、また実施例の中で説明されている特徴の組合せの全てが発明の解決手段に必須であるとは限らない。

【００３４】

図１は、実施例に係る計算機システムの構成例を示す図である。

【００３５】

計算機システムは、スケールアウト型ストレージシステムと、ホスト１０２と、それらを接続する外部ネットワーク１０３から構成される。スケールアウト型ストレージシステムは、複数のストレージ装置１０１を有する。ストレージ装置１０１間接続が密結合である。ホスト１０２は複数あってもよい。

【００３６】

複数のストレージ装置１０１の各々を、１つのストレージ装置１０１を例に取り説明する。

【００３７】

ストレージ装置１０１は、複数のＰＤＥＶ１１０と、コントローラ部５０とを有する。コントローラ部５０は、複数のストレージコントローラ（以下、ＣＴＬ）１０４を有する。複数のＣＴＬ１０４は、例えば二重化されたＣＴＬである。複数のＣＴＬ１０４の各々に、複数のＰＤＥＶ１１０に接続されている。

【００３８】

ＣＴＬ１０４は、メモリ１０６と、ＦＥ　Ｉ／Ｆ（フロントエンドインターフェース）１０７と、ＢＥ　Ｉ／Ｆ（バックエンドインターフェース）１０８と、ＨＣＡ（Host Channel Adapter）１１１とを、それらに接続されたプロセッサ（例えば、マイクロプロセッサ）１０５とを有する。ＣＴＬ１０４が有する各種要素の数は１以上でよい。

【００３９】

また、ＣＴＬ１０４はＣＴＬ間パス１０９を介して他のＣＴＬ１０４に接続されている。各ＣＴＬ１０４は、他のＣＴＬ１０４のメモリ１０６に、ＣＴＬ間パス１０９を介して、プロセッサ１０５（又は、図示されていないＤＭＡ（Direct Memory Access）回路等）によってアクセス可能である。

【００４０】
これ以降、説明の簡単のため、１つのストレージ装置１０１にＣＴＬ１０４は二つであるとする。しかし、ＣＴＬ１０４は、三つ以上でもよい。また、プロセッサ１０５がマルチコアプロセッサの場合、プロセッサ内のいくつかのコアをグループとして、論理的に複数のプロセッサとして管理してもよい。
【００４１】
ＦＥ　Ｉ／Ｆ１０７は、例えば、ＳＡＮ（Storage Area Network）などの外部ネットワーク１０３を通してホスト１０２に接続する。外部ネットワーク１０３は、データ通信ができるプロトコルであれば何でもよい。
【００４２】
ＣＴＬ１０４は、ＢＥ　Ｉ／Ｆ１０８を介し、ＰＤＥＶ１１０が接続される。ＰＤＥＶ１１０は、例えばＨＤＤ（Hard Disk Drive）、ＳＳＤ（Solid State Drive）、又はテープであってよい。
【００４３】
ＣＴＬ１０４は、論理ボリューム（以下、ＶＯＬ）をホスト１０２に提供する。ＶＯＬは、１以上のＰＤＥＶ１１０に基づく実体的なＶＯＬであってもよいし、１以上のＰＤＥＶ１１０に基づくプールから記憶領域が動的に割り当てられる仮想的なＶＯＬ（例えばThin Provisioningに従うＶＯＬ）であってもよい。複数のＰＤＥＶ１１０は、冗長化のためにＲＡＩＤグループを構成していてもよく、実体的なＶＯＬは、構成されたＲＡＩＤグループに基づくＶＯＬでもよい。ＶＯＬは、ホスト１０２に対して、ＬＵ（Logical Unit）として提供可能であり、ホスト１０２が指定するアドレスに対してライト要求及びリード要求を受け付けることが可能である。なお、各ＶＯＬには、Ｉ／Ｏ処理を担当するプロセッサ１０５が一つ設定されていてよい。また、ホスト１０２は、ホストシステムの一例である。
【００４４】
ストレージ装置１０１は、ＨＣＡ１１１と一以上のＳＷ（Ｓｗｉｔｃｈ）１１２を介し、一以上の他のストレージ装置１０１と接続される。
【００４５】
本実施例では、ストレージ装置１０１間接続は、密結合である。密結合のストレージ装置では、ストレージ装置１０１間の通信で使用される通信プロトコルと、ストレージ装置１０１内のデバイス（要素）間の通信で使用される通信プロトコルは、同じである。どちらの通信プロトコルも、例えば、ＰＣＩｅ（PCI-Express）である。一方、疎結合のストレージ装置では、ストレージ装置間の通信で使用される通信プロトコルと、ストレージ装置内のデバイス間の通信で使用される通信プロトコルは、異なる。前者の通信プロトコルは、例えば、ＦＣ（Fibre Channel）又はＩＰ（Internet Protocol）であり、後者の通信プロトコルは、例えば、ＰＣＩｅ（PCI-Express）である。なお、密結合及び疎結合の定義は、既に説明した通りである。
【００４６】
図２は、メモリ１０６の論理構成例を示す図である。
【００４７】
メモリ１０６には、プログラム領域２０１と、共有メモリ（ＳＭ）領域２０２と、ローカルメモリ（ＬＭ）領域２０３と、バッファ領域２０４と、キャッシュメモリ（ＣＭ）領域２０５とが確保されている。
【００４８】
プログラム領域２０１は、プロセッサ１０５が処理を実施するための各プログラムが格納されている領域である。
【００４９】
ＳＭ領域２０２は、全てのプロセッサからアクセスされる領域で、各種管理テーブルが格納される領域である。
【００５０】

ＬＭ領域２０３は、プロセッサ１０５毎に存在する領域であり、特定のプロセッサ１０５のみからアクセスされる領域である。

【００５１】

バッファ領域２０４及びＣＭ領域２０５は、ＦＥ　Ｉ／Ｆ１０７、ＢＥ　Ｉ／Ｆ１０８又はＣＴＬ間パス１０９等によるデータ転送の際に、一時的にデータが格納される領域である。なお、ＣＭ領域２０５は、複数のキャッシュセグメント（ＣＭサブ領域の一例）で構成され、キャッシュセグメント単位で領域が確保される。また、バッファ領域２０４から読み出されてデータは、バッファ領域２０４には残らない。一方、ＣＭ領域２０５から読み出されたデータは、ＣＭ領域２０５に残る。

【００５２】

図３は、プログラム領域２０１に格納されるプログラムの一例を示す図である。

【００５３】

プログラム領域２０１には、例えば、コマンド振分けプログラム３０１と、Ｉ／Ｏプログラム３０２と、新規セグメント割当てプログラム３０４と、フリーセグメント確保プログラム３０５と、フリーセグメント解放プログラム３０６と、ＣＴＬ間データ転送プログラム３０９と、ポート判定プログラム３１０と、Ｉ／Ｏパターン判定プログラム３１１とが格納される。

【００５４】

ホスト１０２からのリード／ライト要求は、コマンド振分けプログラム３０１により、リード／ライト対象のＶＯＬのＩ／Ｏ処理を担当するプロセッサ１０５に振り分けられる。次に、そのリード／ライト要求を受けたプロセッサ１０５（Ｉ／Ｏ処理を担当するプロセッサ１０５）が、Ｉ／Ｏプログラム３０２を実行することにより、そのリード／ライト要求に従い、対象領域のリード／ライト処理を実行する。その際、必要に応じて、ステージング（ＰＤＥＶ１１０からＣＭ領域２０５へのデータの読出し）及びデステージ（ＣＭ領域２０５からＰＤＥＶ１１０へのデータの書込み）のうちの少なくとも１つが、Ｉ／Ｏプログラム３０２により実行される。

【００５５】

Ｉ／Ｏプログラム３０２が、ポート判定プログラム３１０の判定結果（図１０のＳ１７０１の判定結果）と、Ｉ／Ｏパターン判定プログラム３１１の判定結果（図１０のＳ１７０２の判定結果）とに基づいて、キャッシュ先ＣＴＬ（新たにキャッシュセグメントが確保されるＣＭ領域２０５を有するＣＴＬ）１０４を決定する。そして、ＣＭ領域２０５からの新規キャッシュセグメントの割当て処理を新規セグメント割当てプログラム３０４が実行する。その割当て処理において、フリーセグメント確保プログラム３０５が、共有フリーキャッシュセグメントを確保することでフリーキャッシュセグメントを確保する。フリーセグメント解放プログラム３０６は、フリーキャッシュセグメントを解放することで、共有フリーキャッシュセグメントを増やす。

【００５６】

ＣＴＬ１０４間のデータ転送は、ＣＴＬ間データ転送プログラム３０９により実行される。

【００５７】

図４は、ＳＭ領域２０２に格納される情報の一例を示す図である。

【００５８】

ＳＭ領域２０２には、ＰＤＥＶテーブル５０１と、ＶＯＬテーブル５０２と、キャッシュセグメント属性管理テーブル５０３と、キャッシュセグメント管理テーブル５０４とが格納される。ＰＤＥＶテーブル５０１は、ストレージ装置１０１とＰＤＥＶ１１０との対応関係を示す。ＶＯＬテーブル５０２は、ＰＤＥＶ１１０とＶＯＬとの対応関係を示す。キャッシュセグメントは、キャッシュセグメント属性管理テーブル５０３及びキャッシュセグメント管理テーブル５０４を用いて管理される。ＳＭ領域２０２内の情報は、複数のＣＴＬ１０４のメモリ１０６にコピーされてもよい。

【００５９】

図５は、ＬＭ領域２０３に格納される情報の一例を示す図である。

【００６０】

ＬＭ領域２０３には、キャッシュセグメント属性管理テーブル（ローカル）６０１と、キャッシュセグメント管理テーブル（ローカル）６０２とが格納される。

【００６１】

図６は、キャッシュセグメント属性管理テーブル５０３の構成例を示す図である。

【００６２】

キャッシュセグメント属性管理テーブル５０３が有する各エントリ（レコード）は、ストレージ番号、ＣＴＬ番号、プロセッサ番号、及び、キャッシュセグメント属性を記憶する。

【００６３】

ストレージ番号は、ストレージ装置１０１の識別番号ある。ＣＴＬ番号は、ＣＴＬ１０４の識別番号である。プロセッサ番号は、プロセッサ１０５の識別番号である。キャッシュセグメント属性は、キャッシュセグメントの属性（クリーン、ダーティ、フリー及び共有フリーのいずれか）を示す。「クリーン」とは、ＰＤＥＶ１０１に格納済みのデータが記憶された確保済キャッシュセグメントを意味する（クリーンキャッシュセグメント内のデータを「クリーンデータ」と呼ぶことができる）。「ダーティ」は、ＰＤＥＶ１１０に未格納のデータを含んだ確保済キャッシュセグメントを意味する（ダーティキャッシュセグメント内のデータを「ダーティデータ」と呼ぶことができる）。「フリー」は、格納先キャッシュセグメントとして割り当てられる候補である確保済キャッシュセグメント（但しデータが論理的には存在しない）を意味する。「共有フリー」は、未確保のキャッシュセグメント（データが論理的に存在しない）ことを意味する。

【００６４】

キャッシュセグメント属性管理テーブル（ローカル）６０１は、キャッシュセグメント属性管理テーブル５０３と同じ構成であるため説明は省略する。

【００６５】

本実施例では、新規セグメント割当てプログラム３０４が、データをキャッシュ領域に格納する（キャッシュする）ために、Ｉ／Ｏを担当するプロセッサ１０５に対して、フリーキャッシュセグメントを、データ格納先として割り当てる。そのキャッシュセグメントの属性は、格納されたデータの状態に応じて、フリーからクリーン又はダーティに変化する。

【００６６】

図７は、キャッシュセグメント管理テーブル５０４の構成例を示す図である。

【００６７】

キャッシュセグメント管理テーブル５０４は、キャッシュセグメント毎にエントリを有する。各エントリは、キャッシュセグメント番号、ＶＯＬ番号、ＶＯＬ内セグメント番号、キャッシュセグメント属性、ダーティビットマップ、ＣＴＬ番号及びストレージ番号を記憶する。

【００６８】

キャッシュセグメント番号は、キャッシュセグメントの識別番号である。ＶＯＬ番号は、キャッシュセグメントに格納されるデータの格納先ＶＯＬの識別番号である。ＶＯＬセグメント番号は、キャッシュセグメントに格納されるデータの格納先ＶＯＬセグメント（ＶＯＬを構成する領域の一例）の識別番号である。ＶＯＬは、複数のＶＯＬセグメントで構成され、各ＶＯＬセグメントのサイズは、キャッシュセグメントのサイズと同じである。キャッシュセグメント属性は、キャッシュセグメントの属性（クリーン、ダーティ、フリー及び共有フリーのいずれであるか）を示す。ダーティビットマップは、キャッシュセグメント内のデータを構成する複数の各々がダーティか否かを示す。キャッシュセグメント内のデータが１つでもダーティビットを含んでいれば、そのデータはダーティデータである。ＣＴＬ番号は、キャッシュセグメントが属するＣＴＬの識別番号である。ストレージ番号は、キャッシュセグメントが属するストレージ装置の識別番号である。

【００６９】
キャッシュセグメント管理テーブル（ローカル）６０２は、キャッシュセグメント管理テーブル５０４と同じ構成であるため説明は省略する。
【００７０】
参照は、テーブル（ローカル）６０１及び６０２に対して実行され、更新は、テーブル５０３及び５０４と、テーブル（ローカル）６０１及び６０２の両方に対して実行される。但し、共有フリーキャッシュセグメントに関しては、ＬＭ領域２０３へのコピーが存在しないため、ＳＭ領域２０２へのアクセスが実行される。
【００７１】
さて、密結合のストレージ装置を含んだストレージシステムの比較例には、図８及び図９を参照して説明した具体例のような課題がある。
【００７２】
そこで、本実施例では、ＶＯＬのアドレスを指定したＩ／Ｏ要求（ライト／リード要求）を複数のストレージ装置１０１のうちのいずれかが受信した場合、キャッシュ先ＣＴＬ決定処理が行われる。キャッシュ先ＣＴＬ決定処理では、対象Ｉ／Ｏパターンと接続形態とのうちの少なくとも１つに基づいて、受信したＩ／Ｏ要求に従うデータであるＩ／Ｏデータがキャッシュされるキャッシュ先ＣＴＬ１０４が決定される。キャッシュ先ＣＴＬ１０４のＣＭ領域２０５にＩ／Ｏデータ（ライト／リードデータ）がキャッシュされることになる。対象Ｉ／Ｏパターンは、複数のＩ／Ｏパターンのうち、受信したＩ／Ｏ要求に従うＩ／Ｏが属するＩ／Ｏパターンである。複数のＩ／Ｏパターンの各々は、ＶＯＬにおけるＩ／Ｏ先アドレスがランダムとなるかシーケンシャルとなるかに関するパターンである。接続形態は、受信したＩ／Ｏ要求で指定されているアドレスと同一アドレスが指定されたＩ／Ｏ要求を受信するストレージ装置が確定しているか否かである。
【００７３】
図１０は、キャッシュ先ＣＴＬ決定処理の流れを示すフローチャートである。キャッシュ先ＣＴＬ決定処理は、ストレージシステムにおける複数のＣＴＬ１０４のうちの少なくとも１つにより実行される。本実施例では、Ｉ／Ｏ要求を受信したＣＴＬによりキャッシュ先ＣＴＬ決定処理が実行されるとする。
【００７４】
ポート判定プログラム３１０が、ホスト１０２からＩ／Ｏ要求を受信するポートが確定しているか否か（例えば、今回受信したＩ／Ｏ要求で指定されているアドレスと同一アドレスを指定したＩ／Ｏ要求を受信するポートが確定しているか否か）を判定する（Ｓ１７０１）。「ポート」は、ＣＴＬのＦＥ　Ｉ／Ｆ１０７が有するポートである。１つのＣＴＬ１０４の１以上のＦＥ　Ｉ／Ｆ１０７が、１以上のポートを有する。Ｉ／Ｏ要求を受信することが確定していると判断されたポートを「確定ポート」と言うことができる。
確定ポートの具体例は、図１１に示す通りである。すなわち、ホスト側ストレージ装置（（Ｉ／Ｏ要求を受信したストレージ装置）１０１ＡのＣＴＬ１０４Ａａ及び１０４Ａｂとホスト１０２間の複数のパスのうちＡＬＵＡ（Asymmetric Logical Unit Access）に従う優先度が最も高いパス１３０１に接続されているポートが確定ポートの一例である。或いは、ＣＴＬ１０４Ａａ及び１０４Ａｂとホスト１０２間の複数のパスのうち唯一の通信可能状態パス（例えばシングルパス）に接続されているポートが確定ポートの一例である。
【００７５】
Ｓ１７０１の判定結果が肯定の場合（Ｓ１７０１：ＹＥＳ）、新規セグメント割当てプログラム３０４が、テーブル５０３及び５０４を参照することで、図１１に示すように、確定ポートを有するＣＴＬ１０４Ａａ内のＣＭ領域２０５Ａａからフリーキャッシュセグメントを選択し、選択したフリーキャッシュセグメントを、当該Ｉ／Ｏ要求に従うＩ／Ｏ先ＶＯＬセグメントに割り当てる。これにより、そのキャッシュセグメントの属性が、フリーからダーティ又はクリーンに更新され、且つ、割当て先ＶＯＬセグメントの識別番号が、そのキャッシュセグメントに関連付けられる（図７のキャッシュセグメントテーブルが更新される）。以下、キャッシュセグメントの割当てに関しては、同様の処理が行われ

るため、キャッシュセグメントの割当ての説明を簡略化する。Ｓ１７０１：ＹＥＳの後のＳ１７０３によれば、確定ポートを有するＣＴＬ１０４Ａａ（すなわち、必ずＩ／Ｏ要求を受信することになるＣＴＬ１０４Ａａ）のＣＭ領域２０５Ａａがキャッシュ先となるので、図８を参照して説明したような無駄なデータ転送の発生を回避できる。例えば、今回受信したリード要求で指定されているアドレスと同一アドレスを指定したリード要求が受信された場合、そのリード要求を受信したＣＴＬ１０４Ａａ内のＣＭ領域２０５Ａａからホスト１０２へデータを転送でき、図８を参照して説明した２段転送は不要である。なお、キャッシュセグメントの二重化が行われる場合は、キャッシュ先とされたＣＴＬと同じストレージ装置内の他のＣＴＬからキャッシュセグメントが割り当てられてもよいし、ＳＷ１１２を介して接続された他のストレージ装置内のＣＴＬからキャッシュセグメントが割当てられてもよい。

【００７６】

Ｓ１７０１の判定結果が否定の場合（Ｓ１７０１：ＮＯ）、Ｉ／Ｏパターン判定プログラム３１１が、対象Ｉ／Ｏパターンを判定する。Ｉ／Ｏパターン判定よりもポート判定が優先される。この結果、Ｉ／Ｏパターン判定に関わらず、確定ポートがあれば、確定ポートを有するＣＴＬ１０４がキャッシュ先となる。これにより、上述した無駄なデータ転送の発生を回避できる。

【００７７】

Ｓ１７０２の判定結果、対象Ｉ／Ｏパターン（複数のＩ／Ｏパターンのうち、受信したＩ／Ｏ要求に従うＩ／Ｏが属するＩ／Ｏパターン）がランダムリード及びランダムライトのいずれかの場合、新規セグメント割当てプログラム３０４が、図１１に示すように、ホスト側ストレージ装置１０１Ａのいずれかの ＣＴＬ（例えば１０４Ａａ）のＣＭ領域（２０５Ａａ）が有するフリーキャッシュセグメントを、当該Ｉ／Ｏ要求に従うＩ／Ｏ先ＶＯＬセグメントに割り当てる（Ｓ１７０３）。つまり、ここでは、Ｓ１７０１：ＹＥＳの場合と異なり、Ｉ／Ｏ要求を受信したストレージ装置１０１から任意のＣＴＬ１０１が（例えばランダム又はラウンドロビンで）選択されてよい。ホスト側ストレージ装置１０１がキャッシュ先とされる理由は、そのストレージ装置１０１が提供したＶＯＬを指定したＩ／Ｏ要求はそのストレージ装置１０１が受信するためである。Ｉ／Ｏ要求を受信したストレージ装置１０１のうちの任意のＣＴＬ１０１がキャッシュ先とされる理由は、いずれのＣＴＬがＩ／Ｏ要求を受信するかわからないためである。

【００７８】

Ｓ１７０２の判定結果、対象Ｉ／Ｏパターンがシーケンシャルリードの場合、新規セグメント割当てプログラム３０４が、図１２に示すように、当該リード要求に従うリードデータが格納されているＰＤＥＶ（指定されたＶＯＬの基になっているＰＤＥＶ）１１０を有するストレージ装置（ＰＤＥＶ側ストレージ装置）１０１Ｂのうちのいずれかの ＣＴＬ（例えば１０４Ｂａ）内のＣＭ領域（２０５Ｂａ）が有するフリーキャッシュセグメントを、当該リード要求に従うリード元ＶＯＬセグメントに割り当てる（Ｓ１７０４）。Ｓ１７０４では、そのリード元ＶＯＬセグメントに連続する１以上のＶＯＬセグメントにも、それぞれ、ＰＤＥＶ側ストレージ装置１０１内の同じＣＴＬ１０４（又は別のＣＴＬ１０４）のＣＭ領域２０５から１以上のキャッシュセグメントを割り当てる。リードデータの他にシーケンシャルに読み出され得るデータが、ＰＤＥＶ側ストレージ装置１０１内のＣＭ領域２０５（割り当てられたキャッシュセグメント）に、例えばＩ／Ｏプログラム３０２により先読みされる。Ｓ１７０４によれば、シーケンシャルリードのようにデータが先読みされるＩ／Ｏパターンでは、ＰＤＥＶ側ＣＴＬ１０４Ｂａがキャッシュ先となるので、図９を参照して説明したような無駄なＣＴＬ間データ転送の発生を回避できる。すなわち、図１２に示すように、先読みされたデータのうちのいずれかをリード対象としたリード要求をホスト側ＣＴＬ１０４Ａａ及び１０４Ａｂのいずれが受信しても、リード対象のデータを、ＰＤＥＶ側ＣＭ領域２０５Ｂａから、ＳＷ１１２を介して、そのリード要求を受信したホスト側ＣＴＬ（例えば１０４Ａｂ）のバッファ領域（２０４Ａｂ）に転送できるからである。このため、ホスト側ストレージ装置１０１ＡでのＣＴＬ間データ転送が不

要である。

【００７９】

Ｓ１７０２の判定結果、対象Ｉ／Ｏパターンが、ランダムリード、ランダムライト及びシーケンシャルリードのいずれにも該当しない場合、新規セグメント割当てプログラム３０４が、任意のＣＴＬ１０４のＣＭ領域が有するフリーキャッシュセグメントを、当該Ｉ／Ｏ要求に従うＩ／Ｏ先ＶＯＬセグメントに割り当てる（Ｓ１７０５）。任意のＣＴＬ１０４は、ランダムに選択されてもよいし、ラウンドロビンで選択されてもよい。

【００８０】

本実施例によれば、／Ｏパターン及び接続形態の少なくとも１つに基づきキャッシュ先ＣＴＬが決定される。これにより、無駄なデータ転送を回避でき、以って、高性能化を図ることができる。

【００８１】

なお、本発明は、上記した実施例に限定されるものではなく、様々な変形例が含まれる。例えば、上記した実施例は本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。また、ある実施例の構成の一部を他の実施例の構成に置き換えることが可能であり、また、ある実施例の構成に他の実施例の構成を加えることも可能である。また、各実施例の構成の一部について、他の構成の追加・削除・置換をすることが可能である。

【００８２】

また、上記の各構成、機能、処理部、処理手段等は、それらの一部又は全部を、例えば集積回路で設計する等によりハードウェアで実現してもよい。また、上記の各構成、機能等は、プロセッサがそれぞれの機能を実現するプログラムを解釈し、実行することによりソフトウェアで実現してもよい。各機能を実現するプログラム、テーブル、ファイル等の情報は、メモリや、ハードディスク、ＳＳＤ等の記憶デバイス、または、ＩＣカード、ＳＤカード等の記憶デバイスに置くことができる。

【００８３】

また、制御線や情報線は説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。実際には殆ど全ての構成が相互に接続されていると考えてもよい。

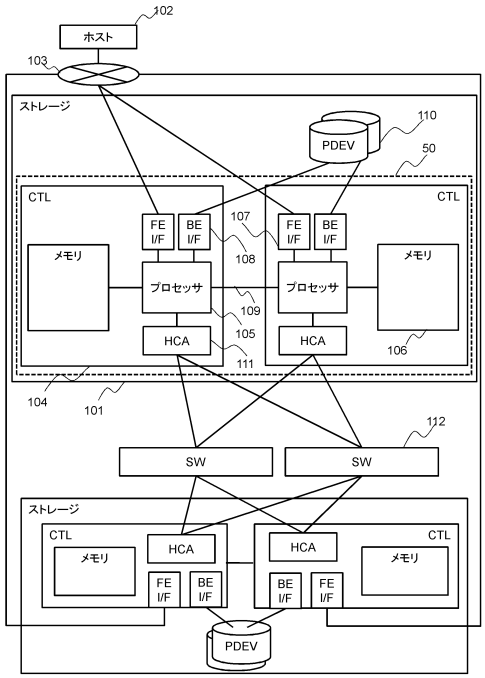
【符号の説明】

【００８４】

１０１…ストレージ装置

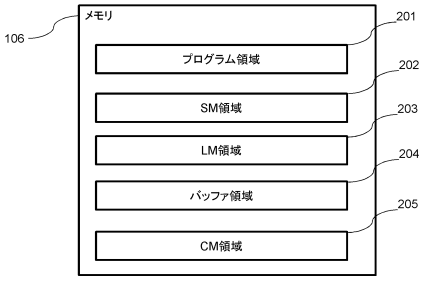
【図１】

FIG. 1



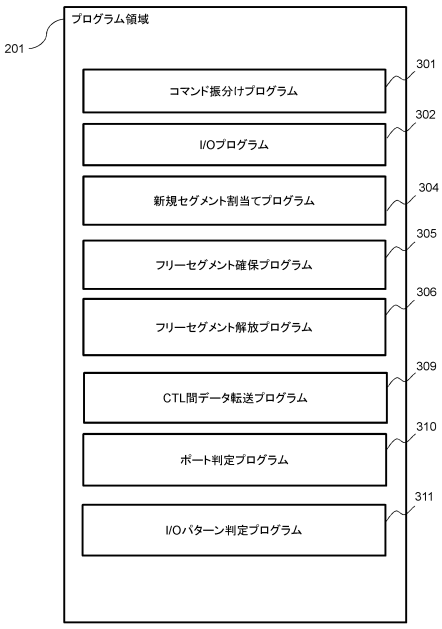
【図２】

FIG. 2



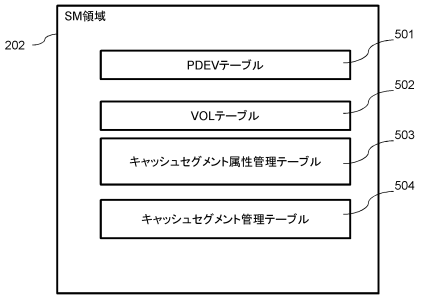
【図３】

FIG. 3



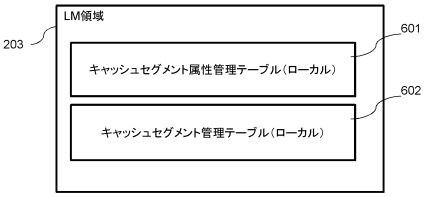
【図４】

FIG. 4



【図５】

FIG. 5



【図６】

FIG. 6

キャッシュセグメント属性管理テーブル
503

ストレージ番号	CTL番号	プロセッサ番号	キャッシュセグメント属性
1	1	1	クリーン
1	1	1	ダーティ
1	1	1	フリー
1	2	1	クリーン
1	2	1	ダーティ
2	2	NULL	共有フリー

【図７】

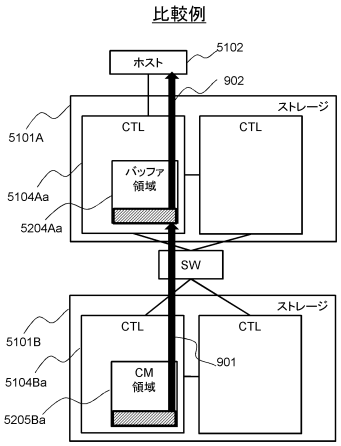
FIG. 7

キャッシュセグメント管理テーブル
504

キャッシュセグメント番号	VOL番号	VOL内セグメント番号	キャッシュセグメント属性	ダーティビットマップ	CTL番号	ストレージ番号
1	なし	なし	フリー	00000000	1	1
2	2	1	クリーン	00000000	1	2
3	1	1	クリーン	00000000	2	1
4	3	4	ダーティ	00100010	2	2
5	4	5	ダーティ	10010011	1	1

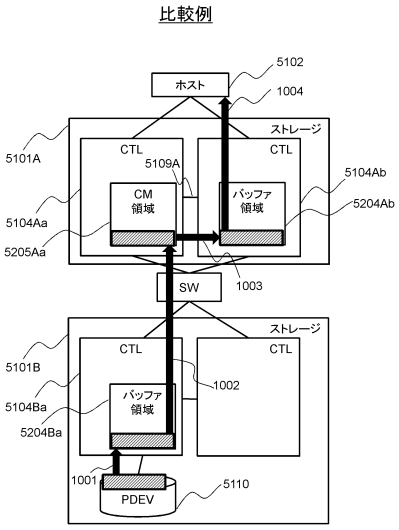
【図８】

FIG. 8



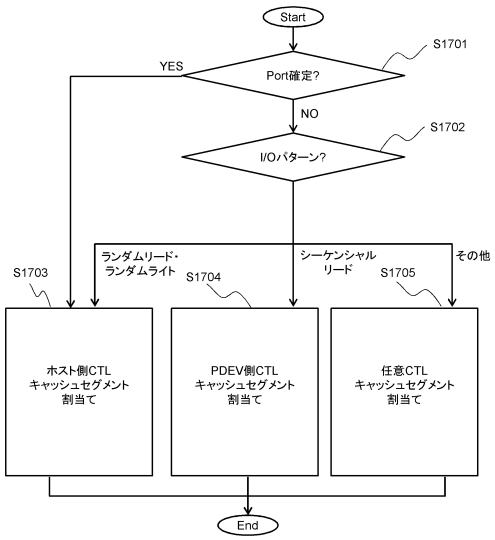
【図９】

FIG. 9



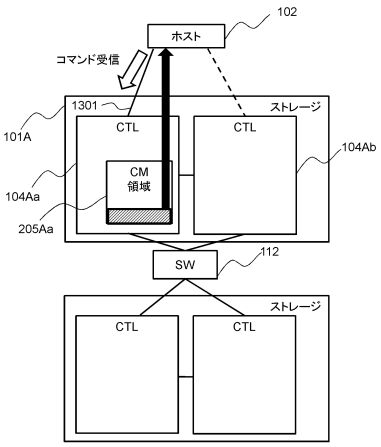
【図１０】

FIG. 10



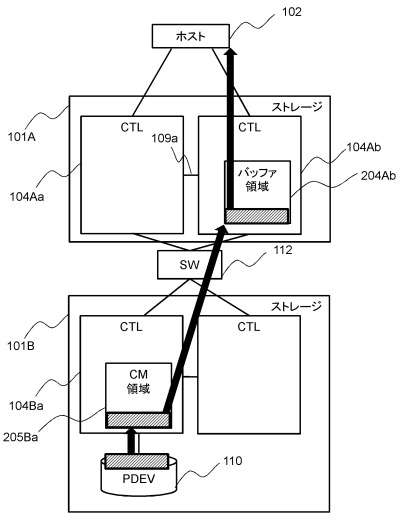
【図１１】

FIG. 11



【図１２】

FIG. 12



フロントページの続き

審査官　桜井　茂行

(56)参考文献　国際公開第２０１５／０５６３０１（WO，Ａ１）
特開２００５－２７５５２５（ＪＰ，Ａ）
米国特許出願公開第２００５／０２１６６９２（ＵＳ，Ａ１）
特開２００６－２５２０１９（ＪＰ，Ａ）
米国特許出願公開第２００６／０２１５６８２（ＵＳ，Ａ１）
特表２００７－５３７５４８（ＪＰ，Ａ）
国際公開第２００５／１１４４２８（WO，Ａ１）
特開２０００－１４８５８７（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１３／１０
Ｇ０６Ｆ　　３／０６