(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2009-43030
(P2009-43030A)
(43)公開日　平成21年2月26日(2009.2.26)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 12/00	(2006.01)	G06F	12/00	514A	5B005
G06F 12/08	(2006.01)	G06F	12/08	551B	5B082
		G06F	12/08	501Z	
		G06F	12/08	523E	
		G06F	12/08	557	

審査請求 未請求　請求項の数 9　ＯＬ　（全 27 頁）

(21)出願番号　特願2007-207357 (P2007-207357)
(22)出願日　平成19年8月9日 (2007.8.9)

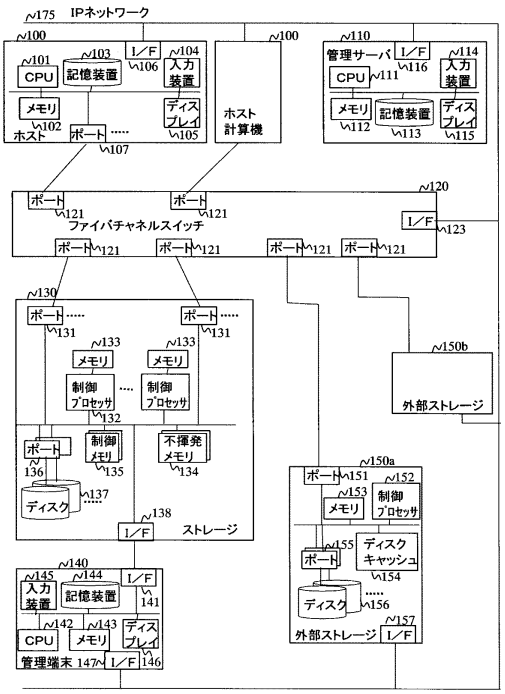
(71)出願人　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　100100310
弁理士　井上　学
(74)代理人　100098660
弁理士　戸田　裕二
(72)発明者　山本　康友
神奈川県川崎市麻生区王禅寺１０９９番地
株式会社日立製作所システム開発研究所内

Ｆターム(参考)　5B005 JJ23　MM11　UU12　UU42
5B082 CA01　FA12　JA06

(54)【発明の名称】ストレージシステム

(57)【要約】
【課題】ホスト計算機データを常置する媒体として不揮発メモリを搭載するストレージにおいて、常置領域のＩ／Ｏ処理に必要な一時領域を割り当てる不揮発メモリ容量を決定する。
【解決手段】当該ストレージが搭載するディスク装置や当該ストレージのストレージ仮想化機能により管理される外部のストレージ装置のディスク装置などのホスト計算機データ常置領域の容量やアクセス頻度などに応じて、常置領域全体に対して確保要な一時領域容量を算出し、当該容量の不揮発メモリを一時領域として定義し、ホスト計算機からのデータ入出力を処理する際に、ホスト計算機データを一時格納するために用いる。また、外部ストレージの増減設や不揮発メモリの増減設といった構成変更などのイベントに応じて、必要な不揮発メモリ一時領域容量を再算出し、必要に応じて、不揮発メモリ一時領域容量を増減するよう制御する。
【選択図】　図１

図1



【特許請求の範囲】
【請求項１】
計算機から送られるデータが格納される1つ以上の不揮発メモリと、
前記１つ以上の不揮発メモリに接続され、前記不揮発メモリの第１の領域を一時領域として割り当て、前記不揮発メモリの第２の領域を常置領域として割り当てるプロセッサと、を有するストレージシステム。
【請求項２】
請求項１記載のストレージシステムであって、
前記不揮発メモリを管理する管理サーバによって、前記不揮発メモリのうち、一時的にデータを保存するために必要な一時領域の容量を算出し、前記プロセッサは、当該容量分について一時領域として割り当てる、ストレージシステム。
【請求項３】
請求項２記載のストレージシステムであって、
前記不揮発メモリのうち、一時領域かつ常置領域として割当てられていない未割当領域の比率を前記管理サーバにより算出し、所定の比率以上の場合、前記プロセッサは、前記不揮発メモリのうち、一部の領域を常置領域となる物理デバイスとして使用するストレージシステム。
【請求項４】
請求項３記載のストレージシステムであって、
前記管理サーバから論理デバイスを定義する要求を受信した場合、前記プロセッサは、前記物理デバイスの一部を前記要求に従った論理デバイスとして定義し、前記管理サーバに処理完了を通知するストレージシステム。
【請求項５】
請求項４記載のストレージシステムであって、
前記ホスト計算機から前記論理デバイスに対して、Ｉ／Ｏ要求を受信した場合、前記プロセッサは、Ｉ／Ｏ要求の種別を判別し、ライト要求である場合、前記論理デバイスにデータを格納するストレージシステム。
【請求項６】
請求項４記載のストレージシステムであって、
前記管理サーバから前記一時領域の一部の割当を解除する要求を受領した場合、対象となる前記一時領域の一部に前記物理デバイスに書き込まれていないデータが存在する際、前記書き込まれていないデータを前記デバイスに書き込んだ後、前記一時領域の一部を未割当領域に変更するストレージシステム。
【請求項７】
請求項４記載のストレージシステムであって、
データを格納する複数のディスクを有し、
前記管理サーバからデバイスを移行又は入れ替える要求を受けた場合、前記制御プロセッサは、前記論理デバイスのデータを前記複数のディスクから構成される物理デバイスに移行又は入れ替えるストレージシステム。
【請求項８】
請求項４記載のストレージシステムであって、
前記管理サーバからデバイスを移行する要求を受けた場合、
前記管理サーバからデバイスを移行又は入れ替える要求を受けた場合、前記制御プロセッサは、前記論理デバイスのデータを他のストレージシステムにある外部デバイスに移行又は入れ替えるストレージシステム。
【請求項９】
請求項４記載のストレージシステムであって、
前記管理サーバから前記物理デバイスの割当を解除する要求を受けた場合、前記制御プロセッサは、前記物理デバイスを開放し、前記不揮発メモリに設定した常置領域との対応を解消するストレージシステム。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、計算機システムにおいて計算機によりアクセスされるストレージシステムの制御方法および管理方法に係り、特に、ストレージシステムに搭載した不揮発メモリの記憶領域の割り当てを行うための方法、および、記憶領域の管理方法に関する。
【背景技術】
【０００２】
文書電子化の規制緩和や手続きの電子化、インターネットビジネスの拡大に伴って、企業などで取り扱われる情報が爆発的に増大している。このような情報増大に加え、ディスク装置から別のディスク装置へのデータバックアップ（Ｄｉｓｋ－ｔｏ－Ｄｉｓｋ　Ｂａｃｋｕｐ）や監査対応などによる企業の業務活動記録（取引情報やメールなど）の長期保管などの顧客要求の高まりによって、ストレージ装置に格納するデータ量が飛躍的に増加し、ストレージシステムの大規模化を招いている。
【０００３】
これに伴い、企業情報システムでは、各部門／各システムのストレージ増強を図る一方で、複雑化するＩＴインフラストラクチャの管理の簡素化や効率化が求められている。特にストレージ装置の管理を簡素化し、データの価値に応じ最適なストレージを活用してトータルコストの最適化を図ることへの期待が高まっている。
【０００４】
大規模ストレージシステムの管理コストを低減する方法の一つとして、特許文献１に開示されるようなストレージ仮想化技術が挙げられる。特許文献１では、第一のストレージを一台以上の第二のストレージへ接続し、第二のストレージがホスト等上位に提供するデバイス（以下、論理デバイスと呼ぶ）を第一のストレージを介して第一のストレージの論理デバイスとしてホストへ提供するストレージ仮想化技術（以下、外部ストレージ接続方法と呼ぶ）が開示されている。第一のストレージはホストからの論理デバイスへの入出力要求を受信した際に、アクセス対象のデバイスが第二のストレージの論理デバイス若しくは第一ストレージ内のディスク装置などの物理デバイスのいずれに対応しているのかを判定し、判定結果に従って適当なアクセス先に入出力要求を振り分ける。
【０００５】
また、ストレージシステム設計を簡素化する方法の一つとして、特許文献２に開示されるような動的に容量を拡張可能なボリューム管理技術（以下、拡張ボリュームと呼ぶ）が挙げられる。特許文献２では、論理デバイス定義時に、ユーザから要求されたデバイス容量分の物理記憶領域(ＨＤＤなど)を固定的に割り当てるのではなく、ストレージ内にプール領域を設けて、当該論理デバイスへ更新アクセスが行われた際に、当該更新部分に対してのみ、プール領域から規定量の物理記憶領域を動的に追加していく。これにより、少ない物理記憶領域で大容量な論理デバイスを導入でき、ストレージ容量設計の簡易化が可能となる。
【０００６】
一方、上記のように大量データを格納する大規模ストレージシステムの導入・運用コストを低減する方法の一つとして、特許文献３に開示されるフラッシュメモリのような不揮発メモリを採用したストレージ実装技術が考えられる。特許文献３では、フラッシュメモリの欠点である、書き換え回数増加を契機としたライト失敗に対する対策について開示されている。特許文献３の技術などを用いることで、ＨＤＤに代えて又は加えてフラッシュメモリなどの不揮発メモリを搭載し、不揮発メモリで、ホストに書かれる又は読み出されるデータを記憶するストレージが考えられる。ホストデータを不揮発メモリへ格納することで、ディスク装置を搭載するストレージ装置と比較して、Ｉ／Ｏ処理高速化と消費電力低減を両立可能なストレージ装置を提供できる。これによって、導入システムの小規模化やシステムランニングコスト（消費電力量）の低減が実現可能となる。なお、現在主流の不揮発メモリであるフラッシュメモリは、計算機の主記憶として用いられるDRAMやS-RAM

などと比較して、低速かつ書き込み回数制限など、使い勝手に影響する特性を有しているため、ディスク装置代替など二次記憶装置として活用されている。

【０００７】

【特許文献１】特開２００５－１１２７７号公報

【特許文献２】特開２００３－１５９１５号公報

【特許文献３】特許３５０７１３２号

【発明の開示】

【発明が解決しようとする課題】

【０００８】

特許文献１に示す第一のストレージでは、外部ストレージである第二のストレージの入出力処理のためのキャッシュメモリを搭載する必要がある。キャッシュメモリとは、ストレージが搭載するディスク装置への入出力処理に際して、リード又はライト対象データを一時的に格納し、キャッシュヒットによるリードおよびライト処理の高速化を実現するもので、ほとんどのストレージシステムに搭載されている。本明細書では、キャッシュメモリのように、データを一時的に格納する記憶領域を「一時領域」、ディスク装置のように、データの最終格納先である記憶領域を「常置領域」と呼ぶこととする。

【０００９】

外部ストレージへの入出力を高速に実行するためには、外部ストレージの容量やアクセス特性（局所性など）に応じて、入出力処理に必要な容量のキャッシュメモリ（一時領域）を割り当てなければならない。

【００１０】

従来、このキャッシュメモリ（一時領域）にはDRAMやSRAMなどのメモリデバイスが採用され、高コスト高速小容量といった、ディスク装置（常置領域）とは全く異なる特性を持った記憶媒体として、ストレージシステムの高速化に寄与してきた。逆に、一時領域に求められる特性に応えるため、前述の第一のストレージでは、搭載するディスク装置（常置領域）とは別に、外部ストレージ（および内部に搭載するディスク装置）用の一時領域としてメモリデバイスを搭載せざるを得なかった。

【００１１】

ところが、前述の外部ストレージ接続手段を有する第一のストレージが、不揮発メモリを常置領域として用いるストレージである場合、外部ストレージの一時領域として、常置領域と同様、不揮発メモリを用いることが可能となる。一時領域／常置領域として共に不揮発メモリを用いることで、ディスク装置とメモリを共に実装する必要が無くなり、ハードウェアが単純化され、ハードウェアコストの低減などが可能となる。しかし、この実現には、不揮発メモリ上に割り当てる一時領域および常置領域の容量・場所を決定・管理可能な管理技術が必要となる。また、外部ストレージの増減設やアクセス特性の変化に応じて、外部ストレージの入出力処理要件を満たす一時領域割り当てへの変更が必要となる。

【００１２】

また、将来技術として研究が進められているMRAMなどの次世代不揮発メモリでは、DRAMなどに匹敵する高速性と書き込み制限の解消の実現が計画されており、新しい不揮発メモリの登場で新しいアーキテクチャが必要になる。

【００１３】

本発明の目的は、外部ストレージ接続手段を有し、常置領域として不揮発メモリを搭載したストレージと外部ストレージからなるストレージシステムにおいて、外部ストレージの入出力処理要件に応じた一時領域容量の決定、割り当て領域の選定および管理を可能とすることである。

【課題を解決するための手段】

【００１４】

管理サーバおよびストレージシステムは、ユーザもしくはアプリケーションプログラムからの指示、またはストレージデバイス増減設などの構成変更に応じて、ストレージ内に搭載した不揮発メモリに割り当てる一時領域、すなわち、ホストからのデータ入出力時に

一時的にデータを格納する記憶領域の容量を決定し、当該一時領域を不揮発メモリの特定領域へ割り当てる。

【００１５】

さらに、ディスク装置や不揮発メモリなどのデータ常置領域、すなわちホストデータを格納する記憶領域の容量変化やアクセス特性の変化に応じて、不揮発メモリに割り当てた一時領域の容量を増減させる。

【発明の効果】

【００１６】

本発明のストレージシステムによれば、仮想ストレージシステムの構成や被仮想化デバイスの容量やアクセス特性に応じて、一時領域容量を決定し、不揮発メモリを一時領域と常置領域それぞれに使い分けることが可能となる。

【発明を実施するための最良の形態】

【００１７】

本発明を実施するための最良の形態を説明する。

【００１８】

本発明の計算機システムとストレージ及びデバイス制御方法の実施例として、第一から第二の実施例を説明する。

【００１９】

第一の実施例は、ホスト計算機と第一のストレージと、ホスト計算機と第一のストレージの間に接続された第二のストレージと管理サーバとをネットワークで接続した計算機システムであって、第二のストレージが常置領域として不揮発メモリを搭載し、他のストレージのデバイスを自身のデバイスとして仮想統合可能なストレージ仮想化機能を搭載し、ユーザ又はアプリケーションの指示に応じて管理サーバと第二のストレージが一時領域の容量を決定し、当該容量の一時領域を不揮発メモリ上に確保し、割り当てる。併せて本発明でのデバイス定義処理とＩ／Ｏ処理について説明する。

【００２０】

第二の実施例は、第一の実施例と同様の構成で、第一のストレージデバイスや第二のストレージに搭載したディスク装置および不揮発メモリ上の常置領域の容量およびアクセス特性に応じて、割当が必要な一時領域容量を算出し、既割当容量との比較結果に応じて、一時領域割当要量を増減させるよう制御する。併せて本発明でのデバイス移行指示処理について説明する。

【実施例１】

【００２１】

まず図１から図１３と図１９と図２０を参照して第１の実施例を説明する。図１は本発明の第１の実施例が適用される計算機システムのハードウェア構成の一例を示す図である。

【００２２】

計算機システムは、１台以上のホスト計算機１００と、管理サーバ１１０と、ファイバチャネルスイッチ１２０と、ストレージ１３０と、管理端末１４０と、外部ストレージ１５０ａおよび１５０ｂ（総称して外部ストレージ１５０と呼ぶ）を有する。ホスト１００、ストレージ１３０および外部ストレージ１５０は各々ポート１０７，１３１，１５１を介してファイバチャネルスイッチ１２０のポート１２１に接続される。また、ホスト１００、ストレージ１３０、外部ストレージ１５０、ファイバチャネルスイッチ１２０は、各々インタフェース制御部（Ｉ／Ｆ）１０６、１３８，１５７，１２３からＩＰネットワーク１７５を介して管理サーバ１１０に接続され、管理サーバ１１０で動作する図示しないストレージ管理ソフトウェアによって統合管理される。なお、本実施例では、ストレージ１３０は管理端末１４０を介して管理サーバ１１０に接続する形態をとるものとするが、ストレージ１３０が直接ＩＰネットワークに接続される構成であっても良い。

【００２３】

ホスト計算機１００は、ＣＰＵ１０１やメモリ１０２などを有する計算機であり、ディ

スク装置や光磁気ディスク装置などの記憶装置１０３に格納されたオペレーティングシステムやアプリケーションプログラムなどのソフトウェアをメモリ１０２に読み上げ、ＣＰＵ１０１がメモリ１０２から読み出して実行することで、所定の機能を達成する。キーボードやマウスなどの入力装置１０４やディスプレイ１０５等の出力装置を具備し、入出力装置１０４がホスト管理者などからの入力を受け付け、出力装置がＣＰＵ１０１から指示された情報を出力する。又、ホスト１００はポート１０７以外に、ＩＰネットワーク１７５と接続するための一以上のインタフェース制御部１０６を有する。

【００２４】

管理サーバ１１０も、ＣＰＵ１１１やメモリ１１２を有する計算機であり、ディスク装置や光磁気ディスク装置等の記憶装置１１３に格納されたストレージ管理ソフトウェアなどをメモリ１１２に読み上げ、ＣＰＵ１１１がそれを読んで実行することで、計算機システム全体の運用・保守管理といった、所定の機能を達成する。ＣＰＵ１１１によってストレージ管理ソフトウェアが実行されると、管理サーバ１１０はインタフェース制御部１１６からＩＰネットワーク１７５を介して、計算機システム内の各機器から構成情報、リソース利用率、性能監視情報、障害ログなどを収集する。そして、管理サーバ１１０は収集したそれらの情報をディスプレイ１１５等の出力装置に出力してストレージ管理者に提示する。また、管理サーバは、キーボードやマウス等の入力装置１１４を介してストレージ管理者からの指示を受信し、受信した運用・保守指示をインタフェース制御部１１６を介して各機器に送信する。

【００２５】

ストレージ１３０は、１つ以上のポート１３１と、１つ以上の制御プロセッサ１３２と、各々制御プロセッサ１３２に接続される１つ以上のメモリ１３３と、１つ以上の不揮発メモリ１３４と、１つ以上の制御メモリ１３５と、１つ以上のポート１３６と、各々ポート１３６に接続される１台以上のディスク装置１３７と、インタフェース制御部１３８を内部結合網で接続した構成をとる。

【００２６】

制御プロセッサ１３２は、ポート１３１から受信した入出力要求についてアクセス対象デバイスを特定し、当該デバイスが対応するディスク装置１３７や外部ストレージ１５０内デバイスへの入出力要求を処理する。その際、制御プロセッサ１３２は受信した入出力要求に含まれるポートＩＤおよびＬＵＮ（Ｌｏｇｉｃａｌ Ｕｎｉｔ Ｎｕｍｂｅｒ）からアクセス対象のデバイスを特定する。なお本実施例では、ポート１３１としてＳＣＳＩ（Ｓｍａｌｌ Ｃｏｍｐｕｔｅｒ Ｓｙｓｔｅｍ Ｉｎｔｅｒｆａｃｅ）を上位プロトコルとしたファイバチャネルインタフェースに対応したポートを想定しているが、ＳＣＳＩを上位プロトコルとしたＩＰネットワークインタフェース（例えば、iSCSI）など、他のストレージ接続用ネットワークインタフェースに対応したポートであっても構わない。

【００２７】

本実施例のストレージ１３０は次のようなデバイス階層を有する。まず、ディスク装置１３７複数台または不揮発メモリ１３４常置領域の複数領域による記憶装置アレイが構成され、この記憶装置アレイは制御プロセッサ１３２によって１つの物理デバイスとして管理される。さらに、本発明の制御プロセッサ１３２では、ストレージ１３０内に搭載した物理デバイスに対して直接論理デバイスを割り当てる（即ち、制御プロセッサ１３２は物理デバイスと論理デバイスとを対応づける）。論理デバイスはストレージ１３０内で管理され、その番号はストレージ１３０毎に独立管理される。論理デバイスは、各ポート１３１に割り当てられたＬＵＮに対応付けられ、ストレージ１３０のデバイスとしてホスト計算機１００に提供される。即ち、ホスト計算機が認識するのはストレージ１３０の論理デバイスであり、ホスト計算機１００は論理デバイスに対応するポート１３１のＬＵＮを用いて、ストレージ１３０に格納されているデータにアクセスする。なお、本実施例では、制御プロセッサ１３２は外部ストレージ１５０の論理デバイスを外部デバイスとして管理し、ストレージ１３０のデバイスとして仮想化する外部ストレージ接続方法を有する。この外部ストレージ接続を用いて取り込んだ外部デバイスは、物理デバイスと同様に、その

１つ以上を同ストレージ１３０の論理デバイスと直接対応づけられる。なお、外部デバイスも各ストレージ１３０内で独立管理される。また、本実施例では、論理デバイスを拡張ボリューム技術、すなわち、更新アクセス部分に対して動的に物理記憶領域を追加割り当て可能な仮想ボリューム管理技術を適用することもできる。その場合、物理デバイス又は外部デバイスに対してプールボリュームを割り当て、さらにプールボリュームに対応させて拡張ボリュームを定義して、当該拡張ボリュームに論理デバイスを割り当てる。ここで、拡張ボリュームとは拡張ボリューム技術を用いた仮想的なデバイスであり、論理デバイスと１対１で対応する。プールボリュームとは拡張ボリュームに動的に割り当てる物理記憶領域を管理する仮想的なデバイスである。本実施例では、簡単のため、プールボリュームは１つの物理デバイス又は外部デバイスに対応させるが、１つのプールボリュームを複数の物理デバイスや外部デバイスで構成してもかまわない。以上のようなデバイス階層を実現するため、制御プロセッサ１３２は、論理デバイス、拡張ボリューム、プールボリューム、物理デバイス、ディスク装置１３７、不揮発メモリ１３４の常置領域、外部デバイス、外部ストレージ１５０の論理デバイス、それぞれのデバイスの間の対応関係を管理し、論理デバイスに対するアクセス要求をディスク装置１３７、不揮発メモリ１３４の常置領域、外部ストレージ装置１５０の論理デバイスへのアクセス要求へ変換し、適当なデバイスへ送信する処理を行う。尚、前述の通り、本実施例におけるストレージ１３０は、複数のディスク装置１３７または不揮発メモリ１３６常置領域の複数領域をまとめて１つまたは複数の物理デバイスを定義し（即ち複数のディスク装置１３７または不揮発メモリ１３６常置領域の複数領域をまとめて１つ又は複数の物理デバイスと対応付け）、１つの物理デバイスに１つの論理デバイス又は１つのプールボリュームを割り当てて、ホスト計算機１００に提供する。しかしもちろん、個々のディスク装置１３７または不揮発メモリ１３６常置領域の一領域を１つの物理デバイスおよび１つの論理デバイス又はプールボリュームとしてもよい。

【００２８】

また、制御プロセッサ１３２は、デバイスに対する入出力処理以外にも、データ複製やデータ再配置など、デバイス間のデータ連携を実現する各種処理を実行する。

【００２９】

また、制御プロセッサ１３２は、インタフェース制御部１３８を介して接続される管理端末１４０にストレージ管理者へ提示する構成情報を送信し、管理者から管理端末１４０に入力された保守・運用指示を管理端末１４０から受信してストレージ１３０の構成変更などを行う。

【００３０】

不揮発メモリ１３４は、一時領域と常置領域に区分して利用される。一時領域は、ホスト計算機１００からのアクセス要求に対する処理速度を高めるため、ディスク装置１３７から頻繁に読み出されるデータを予め格納しておいたり、ホスト計算機１００から受信したライトデータを一時的に格納したりする。尚、不揮発メモリ１３４一時領域を用いたライトアフタを行う場合、即ちホスト計算機１００から受信したライトデータが不揮発メモリ１３４一時領域に格納された後、ディスク装置１３７に実際に書き込まれる前にホスト計算機１００に対しライト要求に対する応答を返す場合には、不揮発メモリ１３４一時領域に格納されているライトデータがディスク装置１３７に書き込まれる前に消失することを防止するために、不揮発メモリ１３４を媒体障害への耐性向上のため二重化したりするなど、可用性を向上させておくことが望ましい。一方、常置領域は、ディスク装置１３７と同様にホスト計算機１００データを最終的に格納しておくために用いられる。よって、ビットコストおよびＩ／Ｏ処理性能の高速化のため、一時領域は単純に二重化することで小容量だが高速アクセスを実現し、常置領域はディスク装置１３７のようにＲＡＩＤ５などの冗長アレイ構成を採用することで、比較的低速アクセスだが大容量低コストな記憶領域を提供することが望ましい。なお、本発明では、簡単のため、不揮発メモリ１３４常置領域は規定サイズの小領域毎に識別番号を振られて管理され、常置領域の全領域に物理デバイスが定義されているとする。

【００３１】
図２０に、不揮発メモリ１３４の実装例を示す。不揮発メモリ１３４は、それぞれに複数のスイッチ２００１と複数のメモリパッケージ２００２を搭載する。メモリパッケージ２００２には複数の不揮発なメモリモジュール２００４と当該メモリモジュール２００４へのアクセスを制御するメモリコントローラ２００３が搭載される。メモリコントローラ２００３はスイッチ２００１と内部結合網を介してストレージ１３０内の他コンポーネントと結合し、メモリアクセスやデータ転送を実施する。
【００３２】
制御メモリ１３５は、前述のデバイス階層を実現するための各デバイスの属性やデバイス間の対応関係を管理するための制御情報や、不揮発メモリ１３４一時領域上に保持したディスク反映済み又は未反映データを管理するための制御情報などを格納する。制御メモリ１３５に格納されている制御情報が消失すると、ディスク装置１３７または不揮発メモリ１３４常置領域に格納されているデータへアクセスできなくなるため、制御メモリ１３５はバッテリバックアップなどにより不揮発化したり、媒体障害への耐性向上のため二重化したりするなど、高可用化のための構成を有することが望ましい。
【００３３】
ストレージ１３０内の各部位は第１図で示すとおり、内部結合網で接続され、各部位間でのデータ、制御情報、及び構成情報の送受信を実行する。この内部結合網により、制御プロセッサ１３２同士が、ストレージ１３０の構成情報を共有、管理することができる。なお、可用性向上の観点から内部結合網も多重化されていることが望ましい。
【００３４】
管理端末１４０は、ＣＰＵ１４２、メモリ１４３、記憶装置１４４、ストレージ１３０と接続されるインタフェース制御部１４１、ＩＰネットワーク１７５と接続されるインタフェース制御部１４７、ストレージ管理者からの入力を受け付ける入力装置１４５及びストレージ管理者にストレージ１３０の構成情報や管理情報を出力するディスプレイ１４６等の出力装置を有する。ＣＰＵ１４２は記憶装置１４４に格納されているストレージ管理プログラムをメモリ１４３に読み出して、これを実行することにより、構成情報の参照、構成変更の指示、特定機能の動作指示などを行い、ストレージ１３０の保守運用に関して、ストレージ管理者もしくは管理サーバ１１０とストレージ１３０間のインタフェースとなる。なお、管理端末１４０を省略して、ストレージ１３０を直接管理サーバ１１０へ接続し、ストレージ１３０を管理サーバ１１０で動作する管理ソフトウェアを用いて管理してもよい。
【００３５】
外部ストレージ１５０は、ファイバチャネルスイッチ１２０を介してストレージ１３０のポート１３１と接続される一又は複数のポート１５１、制御プロセッサ１５２、メモリ１５３、ディスクキャッシュ１５４、一又は複数のディスク装置１５６及び各々ディスク装置１５６に接続される一又は複数のポート１５５を有する。制御プロセッサ１５２はメモリ１５３に格納されているプログラムを実行することにより、ポート１５１から受信したディスク装置１５６への入出力要求を処理する。本実施形態では外部ストレージ１５０を、制御メモリを有さず、ストレージ１３０より小規模な構成のストレージとしたが、ストレージ１３０と同じ構成を有する同規模のストレージとしてもよい。
【００３６】
なお、本実施例では図１に示すようにストレージ装置１３０のポート１３１と外部ストレージ１５０のポート１５１とをファイバチャネルスイッチ１２０を介して接続するため、ホスト計算機１００から外部ストレージ１５０への直接アクセスを抑止するよう、ファイバチャネルスイッチ１２０のゾーニングを設定することが望ましい。また、ファイバチャネルスイッチ１２０を介さずにポート１３１とポート１５１を直接接続してもかまわない。
【００３７】
次に、本発明の実施例におけるストレージ１３０のソフトウェア構成について説明する

。図２はストレージ１３０および管理サーバ１１０の制御メモリおよびメモリに格納される、制御情報とストレージ制御処理のためのプログラムの一例を示したソフトウェア構成図である。

【００３８】

ストレージ１３０の構成管理情報としては、制御メモリ１３５に論理デバイス管理情報２０１とＬＵパス管理情報２０２と物理デバイス管理情報２０３と外部デバイス管理情報２０４とキャッシュ管理情報２０５と不揮発メモリ管理情報２０６と拡張ボリューム管理情報２０７とプールボリューム管理情報２０８とデバイス稼動情報２０９がある。これらの情報は情報消失を防ぐため本実施例では制御メモリ１３５に格納されている。これらの制御情報は制御プロセッサ１３２から参照・更新可能であるが、その際に相互接続網を介したアクセスが必要となる。よって、処理性能向上のため、各制御プロセッサ１３２で実行される処理に必要な制御情報の複製をメモリ１３３に保持してもよい。また、制御端末１４０や管理サーバ１１０にもストレージ１３０の構成情報は送信され保持されている。管理サーバ１１０もしくは管理端末１４０がストレージ管理ソフトウェアもしくはストレージ管理者からの指示を受けてストレージ１３０の構成を変更した場合や、ストレージ１３０内各部位が障害・自動交替などによって構成が変化した場合に、制御プロセッサ１３２のうちの一つが制御メモリ１３５内の該当構成情報を更新する。そして、制御プロセッサ１３２は制御情報が構成変更により更新された旨を、相互接続網を介して他制御プロセッサ１３２、管理端末１４０、管理サーバ１１０に通知し、最新情報を制御メモリ１３５から他部位のメモリへ取り込ませる。

【００３９】

また、外部ストレージ１５０は、ストレージ１３０と同様、自装置内部の論理デバイスや物理デバイスのデバイス管理やデータ管理のために、論理デバイス管理情報２５１、ＬＵパス管理情報２５２、物理デバイス管理情報２５３、キャッシュ管理情報２５４を保持する。これらの内容は、ストレージ１３０の同名情報と同じ目的で用いられる。

【００４０】

さらに、管理サーバ１１０は、ストレージ１３０や外部ストレージ１５０から収集したデバイス管理情報の複製や各ストレージの属性を示すストレージ管理情報２３２をメモリ１１２に保持している。これらの情報はデータ消失を避けるために、管理サーバ１１０に搭載した記憶装置１１３に保持してもよい。

【００４１】

次に各管理情報について説明する。第一に、論理デバイス管理情報２０１について説明する。図３に論理デバイス管理情報２０１の一例を示す。論理デバイス管理情報２０１は、各論理デバイスにつき、論理デバイス番号３０１からデータ移行中フラグ３１０までの情報組を保持する。論理デバイス番号３０１には、論理デバイスを識別するために論理デバイスに制御プロセッサ１３２が割り当てた番号が格納されている。サイズ３０２には、論理デバイス番号３０１により特定される論理デバイスの容量が格納されている。対応下位デバイス番号３０３には、当該論理デバイスに対応づけられている物理デバイス、外部デバイス、又は拡張ボリュームの番号、すなわちそれぞれの管理情報である物理デバイス管理情報２０３、外部デバイス管理情報２０４、又は拡張ボリューム管理情報２０７のエントリ番号が格納される。本実施例では論理デバイスと物理デバイス／外部デバイス／拡張ボリュームは１対１対応とするが、複数の物理デバイス又は外部デバイスを結合して一つの論理デバイスを形成する場合には、論理デバイス管理情報２０１に各論理デバイスが対応する物理／外部デバイスの番号リストとその数を格納するエントリが必要となる。なお、論理デバイスが未定義の場合、対応下位デバイス番号３０３には無効値が設定される。タイプ３０４には、当該論理デバイスのデバイスタイプ識別情報が格納される。ストレージ１３０は、キャッシュ上でのデータ管理単位やデバイス管理情報の格納形態（ディスク空間への管理情報格納有無や格納形態など）などの異なる複数のデバイスタイプの論理デバイスを定義することができ、各論理デバイスがどのデバイスタイプであるかをタイプ３０４で保持する。

【００４２】
デバイス状態３０５には、当該論理デバイスの状態を示す情報が設定される。状態としては、「アタッチド」、「デタッチド」、「未実装」、「ブロックド」が存在する。「アタッチド」は、当該論理デバイスが正常に稼動し、１つ以上のポート１３１にＬＵパス定義され、ホスト計算機１００からアクセスできる状態であることを示す。「デタッチド」は、当該論理デバイスは定義され、正常に稼動しているが、ＬＵパスが未定義であるなどの理由で、ホスト計算機１００からはアクセスできない状態にあることを示す。「未実装」は、当該論理デバイスが物理デバイス又は外部デバイス又は拡張ボリュームに対して定義されておらず、ホスト計算機１００からアクセスができない状態にあることを示す。「ブロックド」は、当該論理デバイスに障害が発生してホスト計算機１００からアクセスできないことを示す。デバイス状態３０５の初期値は「未実装」であり、論理デバイス定義処理により「デタッチド」に、更にＬＵパス定義処理により「アタッチド」に変更される。

【００４３】
エントリ３０７のポート番号には、当該論理デバイスが複数のポート１３１のうちどのポートにＬＵＮ定義されているかを表す情報、即ち当該論理デバイスにアクセスするために用いられるポート１３１の識別情報が設定される。ここでポート１３１の識別情報とは、各ポート１３１に割り振られているストレージ１３０内で一意な番号である。また、同エントリ３０７に格納されるターゲットＩＤとＬＵＮは、当該論理デバイスを識別するための識別子である。本実施例においては、論理デバイスを識別するための識別子として、ＳＣＳＩ上でホスト計算機１００からデバイスをアクセスする場合に用いられるＳＣＳＩ－ＩＤとＬＵＮが用いられる。エントリ３０７は、当該論理デバイスに対するＬＵパス定義が実行された際に情報が設定される。接続ホスト名３０８は、当該論理デバイスへのアクセスが許可されているホスト計算機１００を識別するホスト名である。ホスト名としては、ホスト計算機１００のポート１０７に付与されたＷＷＮ（Ｗｏｒｌｄ Ｗｉｄｅ Ｎａｍｅ）など、ホスト計算機１００もしくはポート１０７を一意に識別可能な値であれば何を用いてもよい。同ストレージ１３０には、このほかに、各ポート１３１のＷＷＮなどポートの属性に関する管理情報を保持する。当該エントリ３０８は、ストレージ管理者によって論理デバイス定義時に設定される。移行中の下位デバイス番号３０８には、当該論理デバイスが現在対応している物理デバイス／外部デバイス／拡張ボリュームから別の物理デバイス／外部デバイス／拡張ボリュームへ移行中である場合、移行先のデバイス番号が格納される。当該論理デバイスがデータ移行中の場合、データ移行中フラグ３１０に「ＯＮ」が格納され、データ移行進捗ポインタ３０９にデータ移行が完了した最後の領域を示すアドレス情報が格納される。なお、エントリ３０８と３０９はデータ移行中フラグ３１０が「ＯＮ」の時のみ有効となる。

【００４４】
第二に、ＬＵパス管理情報２０２について説明する。図４は、ＬＵパス管理情報２０２の一例である。ＬＵパス管理情報２０２は、ストレージ１３０内の各ポート１３１について、各ポートに定義されている有効なＬＵＮ分の情報を保持する。ターゲットＩＤ／ＬＵＮ４０２エントリには、ポート１３１に定義された（割り当てられた）ＬＵＮが格納される。対応論理デバイス番号４０３には、当該ＬＵＮが割り当てられている論理デバイスの番号が格納される。接続ホスト名４０４は、当該ポート１３１に定義されている当該ＬＵＮに対してアクセスを許可されているホスト計算機１００を示す情報が格納され、ホスト計算機１００を示す情報としては例えば上述のホスト計算機１００のポート１０７に付与されているＷＷＮが用いられる。尚、一つの論理デバイスに対して複数のポート１３１のＬＵＮが定義され（割り当てられ）ており、複数のポート１３１から当該論理デバイスにアクセスできる場合がある。この場合、当該複数のポート１３１のＬＵＮ各々に関するＬＵパス管理情報２０２の接続ホスト名４０４の和集合が、当該論理デバイスに関する論理デバイス管理情報２０１の接続ホスト名３０８に保持される。

【００４５】

第三に、物理デバイス管理情報２０３について説明する。物理デバイス管理情報２０３は、ストレージ１３０内の１台以上のディスク装置１３７または不揮発メモリ１３４常置領域の１つ以上の領域から構成される物理デバイスの管理に用いられる。図５は、物理デバイス管理情報２０３の一例である。各ストレージ１３０は、自装置内に存在する物理デバイス毎に、物理デバイス番号５０１からディスク／メモリ内サイズ、オフセット５０７の情報組を保持する。物理デバイス番号５０１には、物理デバイスを識別するための識別番号が登録される。サイズ５０２には、物理デバイス番号５０１により特定される物理デバイスの容量が格納されている。対応上位デバイス番号５０３には、当該物理デバイスが対応付けられている論理デバイス又はプールボリューム番号が当該論理デバイス又はプールボリュームが定義される契機に格納される。当該物理デバイスが論理デバイス又はプールボリュームへ割り当てられていない場合、エントリ５０３には無効値が設定される。デバイス状態５０４には、当該物理デバイスの状態を示す情報が設定される。状態としては、「アタッチド」、「デタッチド」、「未実装」、「ブロックド」が存在する。「アタッチド」は、当該物理デバイスが正常に稼動し、論理デバイス又はプールボリュームに割り当てられている状態を示す。「デタッチド」は、当該物理デバイスは定義され、正常に稼動しているが、論理デバイス又はプールボリュームに未割り当てであることを示す。「未実装」は、当該物理デバイス番号に対する物理デバイスがディスク装置１３７上に定義されていない状態にあることを示す。「ブロックド」は、当該物理デバイスに障害が発生してアクセスできない状態であることを示す。デバイス状態５０４の初期値は「未実装」であり、物理デバイス定義処理により「デタッチド」に、さらに論理デバイスが定義された契機で「アタッチド」に変更される。ＲＡＩＤ構成５０５には、当該物理デバイスが割り当てられたディスク装置１３７または不揮発メモリ１３４常置領域のＲＡＩＤレベル、データディスク装置／メモリ領域とパリティディスク装置/メモリ領域数、データ分割単位であるストライプのサイズなどＲＡＩＤ構成に関連する情報が保持される。ディスク／メモリ番号リスト５０６には、当該物理デバイスが割り当てられたＲＡＩＤを構成する複数のディスク装置１３７または不揮発メモリ１３４常置領域の複数領域の各々の識別番号が保持される。このディスク装置１３７および不揮発メモリ１３４常置領域の複数領域の識別番号は、ストレージ１３０内でディスク装置１３７および不揮発メモリ１３４常置領域の複数領域を識別するために付与した一意な値である。ディスク／メモリ内サイズ／オフセット５０７は、当該物理デバイスが各ディスク装置１３７または不揮発メモリ１３４常置領域の特定領域内のどの領域に割り当てられているかを示す情報である。本実施例では簡単のため、全物理デバイスについて、ＲＡＩＤを構成する各ディスク装置１３７または不揮発メモリ１３４常置領域の特定領域内のオフセットとサイズが同一であるとする。

【００４６】

第四に、外部デバイス管理情報２０４について説明する。外部デバイス管理情報２０４は、ストレージ１３０に接続された外部ストレージ１５０の論理デバイスを外部デバイスとして管理するために用いられる。図６は、外部デバイス管理情報２０４の一例である。ストレージ１３０は、ストレージ１３０で管理する外部デバイス毎に、外部デバイス番号６０１からターゲットポートＩＤ／ターゲットＩＤ／ＬＵＮリスト６０８までの情報組を保持する。外部デバイス番号６０１には、ストレージ１３０の制御プロセッサ１３２が当該外部デバイスに対して割り当てた、ストレージ１３０内で一意な値が格納される。サイズ６０２には、外部デバイス番号６０１により特定される外部デバイスの容量が格納される。対応上位デバイス番号６０３には、当該外部デバイスが対応付けられているストレージ１３０内の論理デバイス又はプールボリュームの番号が登録される。デバイス状態６０４には、物理デバイス管理情報２０３のデバイス状態５０４と同様に、当該外部デバイスの状態が設定される。なお、ストレージ１３０は初期状態では外部ストレージ１５０を接続していないため、デバイス状態６０４の初期値は「未実装」となる。ストレージ識別情報６０５には、当該外部デバイスを搭載する外部ストレージ１５０の識別情報を保持する。識別情報としては、同ストレージのベンダ識別情報と各ベンダが一意に割り振る製造シリアル番号の組み合わせ、などが考えられる。外部ストレージ内デバイス番号６０６には

、当該外部デバイスに対応する外部ストレージ１５０の論理デバイスについて、外部ストレージ１５０内で割り振られた識別番号、すなわち論理デバイス番号が格納される。イニシエータポート番号リスト６０７には、当該外部デバイスへアクセス可能なストレージ１３０のポート１３１の識別番号が登録される。複数のポート１３１から当該外部デバイスへアクセスできる場合には、複数のポート識別番号が登録される。ターゲットポートＩＤ／ターゲットＩＤ／ＬＵＮリスト６０８には、当該外部デバイスが外部ストレージ１５０の１つ以上のポート１５１にＬＵＮ定義されている場合、それらのポート１５１のポートＩＤおよび当該外部デバイスが割り当てられたターゲットＩＤ／ＬＵＮが１つ又は複数個保持される。なお、ストレージ装置１３０の制御プロセッサ１３２が外部デバイスにアクセスする場合（制御プロセッサがポート１３１から外部デバイスに対する入出力要求を送信する場合）には、当該外部デバイスが属する外部ストレージ１５０によって当該外部デバイスに割り当てられたターゲットＩＤ及びＬＵＮが、当該外部デバイスを識別するための情報として用いられる。

【００４７】

第五に、拡張ボリューム管理情報２０７について説明する。本実施の形態では、拡張ボリュームとは定義時には物理記憶領域を割り当てられず、ホスト計算機などからの更新アクセスに応じて、アクセス部分に対応する物理記憶領域としてプールボリュームの記憶領域を動的に追加割り当てしていく仮想ボリュームのことである。本情報は、この拡張ボリュームと論理デバイスおよびプールボリュームとの対応関係を管理するのに用いられる。

【００４８】

図７は、拡張ボリューム管理情報２０７の一例である。ストレージ１３０は、全ての拡張ボリューム毎に、拡張ボリューム番号７０１から実領域割当総量７０７までの情報組を保持する。拡張ボリューム番号７０１には、ストレージ１３０の制御プロセッサ１３２が当該拡張ボリュームに対して割り当てた、ストレージ１３０内で一意な値が格納される。サイズ７０２には、拡張ボリューム番号７０１により特定される拡張ボリュームのホスト計算機からの見掛けの容量が格納される。対応論理デバイス番号７０３には、当該拡張ボリュームが対応付けられているストレージ１３０内の論理デバイスの番号が登録される。ボリューム状態７０４には、物理デバイス管理情報２０３の同名の情報と同様に、当該拡張ボリュームの状態が設定される。なお、本実施例では、ボリューム状態７０４の初期値は「未実装」となる。対応プール番号リスト７０５には、当該拡張ボリュームの物理記憶領域を格納する全てのプールボリュームの番号が保持される。プールエントリ対応管理情報７０６には、拡張ボリュームがホスト計算機に提供する記憶空間と対応するプールボリューム内の物理記憶領域との対応を示す情報が格納される。拡張ボリュームが定義されたばかりの初期状態では、当該ボリュームには物理記憶領域が割り当てられていないため、当該ボリュームの全記憶領域は未割り当て状態である。ホスト計算機などから当該拡張ボリュームに対応する論理デバイスへの更新アクセスがあると、拡張ボリューム内のアクセス対象位置が算出され、該当する記憶領域エントリへのプールボリューム記憶領域の割り当て有無がチェックされる。アクセス対象記憶領域に対する物理記憶領域が未割り当ての場合、対応プール番号リスト７０５で示される１つ以上のプールボリュームから適当な記憶領域が選定され、拡張ボリュームのアクセス対象領域の物理記憶領域として割り当てられる。最後の実領域割当総量７０７には、当該拡張ボリュームに対して物理記憶領域を割り当てた総容量を格納する。

【００４９】

第六に、プールボリューム管理情報２０８について説明する。本発明では、プールボリュームとは拡張ボリュームの物理記憶領域として割り当てられる記憶領域を提供・管理する仮想的なボリュームのことである。本情報は、このプールボリュームと拡張ボリュームおよび物理／外部デバイスとの対応関係を管理するのに用いられる。

【００５０】

図８にプールボリューム管理情報２０８の一例を示す。ストレージ１３０は、全てのプールボリューム毎に、プールボリューム番号８０１から対応拡張ボリューム番号リスト８

０８までの情報組を保持する。プールボリューム番号８０１には、ストレージ１３０の制御プロセッサ１３２が当該プールボリュームに対して割り当てた、ストレージ１３０内で一意な値が格納される。サイズ８０２には、プールボリューム番号８０１により特定されるプールボリュームの容量が、空き容量８０３には、拡張ボリュームへ未割り当てな記憶領域の総容量が格納される。ボリューム状態８０４には、拡張ボリューム管理情報２０７の同名の情報と同様に、当該プールボリュームの状態が設定される。物理／外部デバイス番号８０５には、当該プールボリュームが割り当てられた物理／外部デバイスの番号が保持される。プールエントリサイズ８０６には、当該プールボリュームの記憶領域を拡張ボリュームの物理記憶領域として割り当てる際の割り当て単位（エントリ）の大きさを示す情報を格納する。拡張ボリュームの記憶領域は対応するプールボリュームのエントリサイズで複数のエントリに分割され、拡張ボリュームの各エントリとプールボリュームの各エントリとの対応状態が、プールエントリ割当状態管理情報８０７に保持される。対応拡張ボリューム番号リスト８０８には、当該プールボリュームの記憶領域を物理記憶領域として割り当てた１つ以上の拡張ボリュームの番号が保持される。

【００５１】

第七に、デバイス稼動情報２０９について説明する。本実施の形態では、各論理デバイス毎にホスト計算機１００からのＩ／Ｏ頻度やデータ長またはアクセス範囲の局所性・連続性などを採取する。本情報に基づいて、各論理デバイスに対応する物理デバイス又は外部デバイスへのキャッシュメモリ、すなわち不揮発メモリ１３４一時領域の必要量を算出する。

【００５２】

本実施例では、ストレージ１３０は上記の７つのデバイス／ボリューム管理情報を用いて、デバイス／ボリュームを管理する。当該ストレージ１３０の初期状態としては、当該装置が工場出荷時に既に各ディスク装置１３７および不揮発メモリ１３４の常置領域に物理デバイスが定義され、論理デバイス、外部デバイス、拡張ボリューム、プールボリュームはそれぞれ未定義である。また、ユーザ又はストレージ管理者はストレージ１３０を導入時に同ストレージ１３０に接続した外部ストレージ１５０の論理デバイスを外部デバイスとして定義し、これら物理デバイスおよび外部デバイス上に論理デバイスやプールボリュームを定義し、プールボリュームの場合は更に拡張ボリュームや論理デバイスを定義し、当該論理デバイスについて各ポート１３１にＬＵＮを定義する。

【００５３】

また、ストレージ１３０は制御メモリ１３５に、不揮発メモリ１３４および不揮発メモリ１３４に割り当てた一時領域を管理するために、不揮発メモリ管理情報２０６およびキャッシュ管理情報２０５を保持する。不揮発メモリ１３４管理情報は、搭載した不揮発メモリ１３４が提供するメモリ空間の各領域の割当状況、すなわち一時領域または常置領域のどちらかに割り当て済みかまたは未割当かを示す情報を保持する。キャッシュ管理情報２０５は、不揮発メモリ１３４一時領域の領域管理情報や、物理デバイスまたは外部デバイスのアドレス空間のどのデータがどの一時領域に格納されているかを示すヒットミス判定情報などで構成される。一時領域の管理方法として、例えば、一時領域を規定サイズの小領域（セグメント）に分割し、各セグメント毎に物理デバイスまたは外部デバイスのアドレス空間との対応情報を保持し、各セグメントが保持する情報がデバイスへ未反映（ダーティ）か、反映済み（クリーン）かを示す情報を保持するなどが考えられる。

【００５４】

次に、再び図２に戻り、ストレージ１３０、外部ストレージ１５０、管理サーバ１１０のメモリ１３３、１５３、１１２に格納されるプログラムついて説明する。各メモリには上述の制御情報の複製に加えて、これらの各部位内の制御プロセッサ１３２、１５２やＣＰＵ１１１で動作するプログラムが格納される。

【００５５】

本実施例では、ストレージ１３０の不揮発メモリ１３４一時領域の割当処理と、それに付随した物理デバイス割当処理、デバイス定義処理について説明する。また、後述する第

２の実施例では、当該一時領域の割当変更処理と、それに付随したデバイス移行処理について説明する。これらの処理は、ユーザ又はアプリケーションからの要求に対して、管理サーバ１１０、ストレージ１３０、外部ストレージ１５０が連携して実行する。

【００５６】

まず、不揮発メモリ１３４一時領域の割当処理に対応して、管理サーバ１１０には一時領域定義指示処理２４３がメモリ１１２に、ストレージ１３０には一時領域割当処理２２４および物理デバイス定義処理２２６がメモリ１３３にそれぞれ格納される。

【００５７】

図９は、管理サーバ１１０で実行される一時領域定義指示処理２４３の処理フローの一例を示す図である。管理サーバ１１０は、デバイス管理情報の複製２３１内の情報に基づいてストレージ１３０の不揮発メモリ１３４に割り当てなければならない一時領域の容量を算出する（ステップ９０１）。より具体的には、物理デバイス管理情報２０３や外部デバイス管理情報２０４から物理／外部デバイスの容量を算出し、またデバイス管理情報２０９から各論理デバイスに対応する物理／外部デバイスのアクセス特性（頻度、局所性など）を得て、それに基づいた一時領域必要量を算出する。例えば、アクセス頻度と局所性から一時領域割当容量に対するキャッシュヒット向上率を推定し、各論理デバイスに期待される性能が達成されるよう制御する方法などが考えられる。これらの一時領域の容量算出には様々なアルゴリズムが存在する。特定のアルゴリズムに基づいて一時領域必要量を算出したら、ストレージ１３０に対して、不揮発メモリ１３４への一時領域割当を指示する（ステップ９０２）。一時領域の割当が完了したら、次に不揮発メモリ１３４のうち、一時領域としても常置領域としても割り当てられていない領域(以下、未割当領域と呼ぶ）の量および比率を算出し、これらが規定レベルを超えている場合、超過分の不揮発メモリ１３４を常置領域として定義して物理デバイスを定義するよう、ストレージ１３０へ指示する（ステップ９０３）。この結果、超過分の不揮発メモリ１３４が物理デバイスとしてストレージ１３０が管理する空きデバイスとして算入される。

【００５８】

図１０は、ストレージ１３０で実行される一時領域割当処理２２４の処理フローの一例を示す図である。管理サーバ１１０からの一時領域割当指示を受信したストレージ１３０の制御プロセッサ１３２は、不揮発メモリ管理情報２０６に基づいて、指示された容量分の不揮発メモリ１３４未割当領域を選定し、当該未割当領域を一時領域へ変更する（ステップ１００１，１００２）。さらに、新規に割り当てた当該一時領域をセグメントに分割管理するよう、キャッシュ管理情報２０５へ登録する（ステップ１００３）。ここで、制御プロセッサ１３２は、制御メモリ１３５に対して、登録する。制御プロセッサ１３２がメモリ１３３にもキャッシュ管理情報２０５を有している場合、メモリ１３３に対しても登録する。

【００５９】

図１１は、ストレージ１３０で実行される物理デバイス定義処理２２６の処理フローの一例を示す図である。管理サーバ１１０からの不揮発メモリ１３４の未割当領域への物理デバイス定義指示を受信したストレージ１３０の制御プロセッサ１３２は、不揮発メモリ管理情報２０６に基づいて、指示された容量分の不揮発メモリ１３４の未割当領域を選定し、当該未割当領域を常置領域へ変更する（ステップ１１０１、１１０２）。さらに、新規に割り当てた当該常置領域へ物理デバイスを定義するよう、制御プロセッサ１３２は、物理デバイス管理情報２０３を設定する。具体的には、制御プロセッサ１３２は、特定の物理デバイスに対応する物理デバイス管理情報２０３へ、物理デバイスを定義する当該常置領域のサイズと、当該常置領域の識別番号および各領域内のオフセットおよびサイズをそれぞれ登録する。

【００６０】

以上の説明のように、管理サーバ１１０とストレージ１３０の連携によって、不揮発メモリ１３０の一時領域が定義される。次に、こうしたシステム環境におけるストレージ運用管理に伴う通常処理として、論理デバイスの定義処理とＩ／Ｏ処理を例に説明する。

【００６１】
まず論理デバイス定義処理に対応して、管理サーバ１１０にはデバイス定義指示処理２４１がメモリ１１２に、ストレージ１３０にはデバイス定義処理２２２がメモリ１３３にそれぞれ格納される。
【００６２】
図１２は、管理サーバ１１０で実行される、デバイス定義指示処理２４１の処理フローの一例を示す図である。管理サーバ１１０は、ＩＰネットワーク１７５とインタフェース制御部１１６を介してユーザまたはアプリケーションプログラムからストレージ１３０の論理デバイス定義要求を受信する（ステップ１２０１）。当該要求に含まれる情報としては、例えば、ストレージ１３０の識別情報、論理デバイス番号、アクセス元ＨＢＡ特定情報（ＷＷＮなど）、ポート特定情報（ポート１３１のポートＩＤ、ターゲットＩＤ、ＬＵＮのリストなど）、デバイス属性情報（通常ボリューム／拡張ボリューム、用途、など）などが考えられる。これらの情報から、まず管理サーバ１１０は、当該ストレージ１３０に設定された管理アクセス権設定などから、当該要求が当該ストレージ１３０の管理アクセス権を持ったユーザ又はホスト計算機１００上のアプリケーションプログラムからのものであることを確認して、当該要求の実行可否を判定する。続いて、管理サーバ１１０は、前記デバイス定義指示に含まれた情報から、当該論理デバイスを通常のボリュームとして定義するか、拡張ボリュームとして定義するかを判定する（ステップ１２０２）。当該論理デバイスを拡張ボリュームとして割り当てる場合、管理サーバ１１０は、物理／外部デバイスに対応している全てのプールボリュームの残容量、すなわち拡張ボリュームの物理記憶領域として未割り当てな領域の総容量が規定値以上あれば、既に定義されているプールボリュームのうちの一つに対応した拡張ボリュームを定義させて、当該拡張ボリュームへ当該論理デバイスを割り当てる旨をストレージ１３０へ指示する（ステップ１２０３，１２０４）。管理サーバ１１０は、容量が十分でなければ、物理／外部デバイスのうち、論理デバイスにもプールボリュームにも未割り当てなデバイス（以下、空きデバイス）をチェックし、該当デバイスがあれば新たなプールボリュームおよび拡張ボリュームを定義して、当該論理デバイスを割り当てる旨をストレージ１３０へ指示する（ステップ１２０５）。もし、空きデバイスが無い場合には、管理サーバ１１０は、当該論理デバイス定義処理のエラーを要求元へ報告する（ステップ１２０８）。一方、ステップ１２０２の判定で、当該論理デバイスを拡張ボリュームではなく、通常ボリュームとして定義する場合、空き物理／外部デバイスを検索し、割り当て対象デバイスを決定し、同デバイスへの当該論理デバイス割り当てをストレージ１３０へ指示する（ステップ１２０６）。管理サーバ１１０は、ストレージ１３０へ指示した後、要求元へ完了報告する（ステップ１２０７）。
【００６３】
図１３は、ストレージ１３０で実行される、デバイス定義処理２２２の処理フローの一例を示す図である。まず、ストレージ１３０は、管理サーバ１１０から論理デバイス定義要求を受信する（ステップ１３０１）。この際、管理サーバ１１０は、IPネットワーク１７５及び管理端末１４０を介して、ストレージ１３０に論理デバイス定義を送信する。制御プロセッサ１３２は、当該要求を制御メモリ１３５に格納する。当該要求に含まれる情報としては、管理サーバ１１０がデバイス定義指示処理２４１で受信した情報に加えて、通常ボリュームの場合は論理デバイス割り当て対象物理デバイス又は外部デバイス番号などが含まれる。一方、拡張ボリュームの場合は新規に割り当てる拡張ボリューム番号と拡張ボリューム割り当て対象プールボリューム番号か、又は新規に割り当てる拡張ボリューム番号／プールボリューム番号とプールボリューム割り当て対象物理デバイス又は外部デバイス番号などが含まれる。当該論理デバイスが拡張ボリュームの場合制御プロセッサ１３２は、必要ならば指定された物理デバイス又は外部デバイスに対してプールボリュームを定義し（ステップ１３０２、１３０３、１３０４）、プールボリュームに対して拡張ボリュームを定義する（ステップ１３０５）。ここで、プールボリューム定義として、具体的には、制御プロセッサ１３２は、指定されたプールボリューム番号に対応する情報エン

トリについて、割り当て対象の物理デバイス番号又は外部デバイス番号を物理デバイス番号又は、外部デバイス番号８０５に設定し、同デバイス管理情報を参照してサイズ８０２を設定し、空き容量８０３をサイズ８０２の値に、ボリューム状態８０４を「デタッチ」に、対応拡張ボリューム番号リスト８０８を無効値にそれぞれ初期化する。また、制御プロセッサ１３２は、プールエントリサイズ８０６について、ストレージ１３０で固定の値を用いてもよいし、ユーザからの論理デバイス割り当て指示に同サイズを示す情報を設定してもよい。また、制御プロセッサ１３２は、プールエントリ割当状態管理情報８０７を全エントリが拡張ボリュームに未割当となるよう初期化する。また、拡張ボリューム定義として、具体的には、指定された拡張ボリューム番号に対応する拡張ボリューム管理情報２０７の情報エントリと対応するプールボリュームのプールボリューム管理情報２０８の情報エントリについて、対応拡張ボリューム番号リスト８０８に当該拡張ボリューム番号を、ボリューム状態８０４に「アタッチ」をそれぞれ登録する。さらに、制御プロセッサ１３２は、サイズ７０２に当該拡張ボリュームに対応させる論理デバイスのサイズを設定し、対応論理デバイス番号７０３に無効値を、ボリューム状態７０４に「デタッチ」を、停止フラグに「Ｏｎ」を、最終起動／停止時間７０６に現在時刻を、対応プール番号リスト７０７に当該プールボリューム番号を、実領域割当総量に０をそれぞれ設定する。プールエントリ対応管理情報７０８には、全エントリが未割当状態となるよう初期化する。次に、制御プロセッサ１３２は、当該拡張ボリュームに対して当該論理デバイスを定義してから、指定されたポート１３１に対してＬＵＮパス定義を行い、処理完了を管理サーバ１１０へ報告する（ステップ１３０６、１３０７）。ここで、論理デバイス定義として具体的には、制御プロセッサ１３２は、対応する論理デバイス管理情報２０１と拡張ボリューム管理情報２０７について、対応論理デバイス番号７０３に当該論理デバイス番号を、ボリューム状態７０４に「アタッチ」をそれぞれ設定する。さらに、サイズ３０２とタイプ３０４に割当指示内容を、対応下位デバイス番号３０３に当該拡張ボリューム番号を、デバイス状態に「デタッチ」を、エントリ３０６から３０９に無効値を、データ移行中フラグ３１０に「Ｏｆｆ」をそれぞれ設定する。また、ＬＵＮパス定義として具体的には、制御プロセッサ１３２は、対応する論理デバイス管理情報２０１とＬＵパス管理情報２０２について、エントリ３０６、３０７および４０２、４０４に割当指示で指定されたポート情報や接続ホスト計算機１００情報をそれぞれ設定する。また、対応論理デバイス番号４０３に当該論理デバイス番号を設定する。一方、当該論理デバイスが通常ボリュームの場合は、指定された空き物理デバイス又は外部デバイスに対して当該論理デバイスを定義し、ＬＵＮパス定義を行い、処理完了を報告する（ステップ１３０２、１３０６，１３０７）。ここで、論理デバイス定義として具体的には、対応する論理デバイス管理情報２０１と物理デバイス管理情報２０２又は外部デバイス管理情報２０３に必要情報を設定する。

【００６４】

次にホスト計算機から論理デバイスに対するＩ／Ｏ処理に対応して、ストレージ１３０にＩ／Ｏ処理２２１がメモリ１３３に格納される。

【００６５】

図１９は、ストレージ１３０で実行される、Ｉ／Ｏ処理２２１の処理フローの一例を示す図である。ストレージ１３０は、ホスト計算機１００からのＩ／Ｏ要求を受信する（ステップ１９０１）。制御プロセッサ１３２は、Ｉ／Ｏ要求の種別によって処理を切り分ける（ステップ１９０２）。

【００６６】

ステップ１９０２で、制御プロセッサ１３２がリード要求であると判断した場合、制御プロセッサ１３２は、キャッシュ管理情報２０５に基づき、リード対象データのキャッシュヒットミスであるか否かを判定する（ステップ１９０３）。キャッシュミスの場合はキャッシュ管理情報２０５を操作して不揮発メモリ一時領域から空きセグメントを割り当てて（ステップ１９０４）、物理デバイスであるディスク１３７または外部ストレージ１５０からステージングしたデータを不揮発メモリ１３４に格納し（ステップ１９０５）、ホスト計算機へデータを転送する（ステップ１９０６）。キャッシュヒットの場合、制御プ

ロセッサ１３２は、不揮発メモリ１３４からデータをリードする。
【００６７】
また、ステップ１９０２で、制御プロセッサ１３２がＩ／Ｏ要求をライト要求であると判別し、キャッシュミスであるか否かを判別する（ステップ１９０８）。キャッシュミスの場合、制御プロセッサ１３２は、空きセグメントを割り当てる（ステップ１９０９）。ここで、キャッシュミスとは、論理デバイスのあるデータ領域に対して、不揮発メモリ１３４にセグメントが割り当てられていない状態をいう。その後、制御プロセッサ１３２は、ホスト計算機からのライトデータを当該セグメントに格納した後（ステップ１９１０）、物理デバイスまたは外部デバイスなどの常置領域へデステージする（ステップ１９１１）。一方、ステップ１９０８で、キャッシュヒットした場合、不揮発メモリ１３４に常置領域が存在することになるので、該当するアドレスにデータを格納する。
【００６８】
他方、ステップ１９０２で、制御プロセッサ１３２がライト要求でもリード要求でもないと判別した場合（例えば、センスコマンドなど）、制御プロセッサ１３２は、要求された処理を実行する（ステップ１９１２）。
【００６９】
最後に、処理を実行した制御プロセッサ１３２は、処理完了をホスト計算機１００へ報告し、処理を終了する（ステップ１９０７）。
【００７０】
本実施例は、上述のように、構成されるので、以下の効果を奏する。仮想ストレージシステムの構成や被仮想化デバイスの容量やアクセス特性に応じて、一時領域容量を決定し、不揮発メモリを一時領域と常置領域それぞれに使い分けることが可能となる。
【実施例２】
【００７１】
実施例１では、ストレージ１３０が管理する物理デバイスおよび外部デバイスの容量やアクセス特性などから一時領域の必要量を算出し、必要分の一時領域を割り当てる方法について説明した。実施例２では、外部ストレージの増減設や各デバイスへのアクセス特性の変化を感知し、一時領域割当量の見直しをかける方法について説明する。
【００７２】
第２の実施例は、図１から図８および図１４から図１８を用いて説明する。第１と第２の実施例は多くの共通点を持つため、両者の差異についてのみ述べる。
【００７３】
第２の実施例におけるソフトウェア構成は第１の実施例と同様に図２に示される。第２の実施例では、まず、不揮発メモリ１３４の一時領域の割当変更処理に対応して、管理サーバ１１０には一時領域変更指示処理２４４がメモリ１１２に、ストレージ１３０には一時領域割当解除処理２２５とデバイス移行処理２２３と物理デバイス定義解除処理２２７がメモリ１３３にそれぞれ格納される。
【００７４】
図１４は、管理サーバ１１０で実行される、一時領域変更指示処理２４４の処理フローの一例を示す図である。管理サーバ１１０のＣＰＵ１１１は、ある時点での不揮発メモリ１３４に既に割り当てられている一時領域の容量を、キャッシュ管理情報２０５や不揮発メモリ管理情報２０６に基づき算出する（ステップ１４０１）。更に、一時領域割当処理２２４のステップ１００１と同様の方法で、その時点の一時領域割当必要量を算出する（ステップ１４０２）。これらの値を比較し、両者の差が規定値を超えた場合、ＣＰＵ１１１は、一時領域の割当量を制御する。もし、不揮発メモリ１３４の既割当一時領域の容量が必要量に対して規定値以上の場合（ステップ１４０３）、解除候補として超過容量分の一時領域を選定する（ステップ１４０８）。一時領域の選定方法としては、未使用セグメントを多く含む一時領域を選定してもよいし、ダーティデータを格納したセグメントが少ない一時領域を選定してもよい。ＣＰＵ１１１は、選定した領域を特定する情報を付加して、ストレージ１３０に対して特定一時領域の割当解除を指示する（ステップ１４０９）

。なお、このとき、割当を解除する一時領域を選定せず、割当解除が必要な一時領域の容量のみを付加し、ストレージ１３０で割当を解除する一時領域を決めても良い。特定一時領域の割当が解除されたら、ＣＰＵ１１１は、未割当領域となった当該不揮発メモリ１３４の領域に対して常置領域への設定および物理デバイス定義をストレージ１３０へ指示する（ステップ１４１０）。もし、既割当一時領域の容量が必要量に対して規定値より少ない場合、超過容量分に相当する、不揮発メモリ１３４に定義された物理デバイスを選定し、当該物理デバイスに対応する論理デバイスをディスク装置１３７に定義された物理デバイス又は外部デバイスに移行するようストレージ１３０へ指示する（ステップ１４０５）。移行が完了したら、ＣＰＵ１１１は、当該物理デバイスの割当解除をストレージ１３０へ指示する（ステップ１４０６）。それが完了したら、管理サーバ１１０はストレージ１３０に対して当該未割当領域を一時領域へ割り当てるよう指示する（ステップ１４０７）。

【００７５】

図１５は、ストレージ１３０で実行される、一時領域割当解除処理２２５の処理フローの一例を示す図である。は、管理サーバ１１０からの一時領域割当解除指示を管理端末１４０を介して、受領する。その後、制御プロセッサ１３２は、必要に応じて割当解除対象となる不揮発メモリ１３４の一時領域を選定し（ステップ１５０１，１５０２）、割当解除対象となる一時領域について、キャッシュ管理情報２０５に基づき、ダーティデータの有無すなわち、常置領域へ未反映なホスト更新データを格納したセグメントの有無を確認し、存在すれば、当該未反映データ（ダーティデータ）を対応する物理デバイス又は外部デバイスへデステージする（ステップ１５０３）。対象となる一時領域の全てのダーティデータのデステージが完了したら、不揮発メモリ管理情報２０６の対応するエントリを未割当領域を示すよう変更し、対応するキャッシュ管理情報２０５を全て初期化する（ステップ１５０４）。

【００７６】

図１６は、ストレージ１３０で実行される、デバイス移行処理２２３の処理フローの一例を示す図である。制御プロセッサ１３２は、周期的に本処理を起動し、デバイス移行要求が登録されていないかをチェックする（ステップ１６０１）。先述の通り、デバイス移行要求は、不揮発メモリ１３４の一時領域を拡張するために常置領域に定義した物理デバイスに対応する論理デバイスを別の物理デバイスとなるディスク装置１３７又は外部デバイスに移行する際、登録される。また、後述するように、ストレージ１３０が構築する、不揮発メモリ１３４の常置領域、ディスク装置１３７や外部デバイスといった階層ストレージにおける、データ価値などに基づく、階層間データ移行が必要な際、登録される。デバイス移行要求が登録されていなければ、なにもせず本処理を終了する（ステップ１６０１）。要求が登録されていた場合、対象論理デバイスが拡張ボリュームか通常ボリュームかで処理を切り分ける（ステップ１６０２）。当該論理デバイスが通常ボリュームの場合、要求処理が空きデバイスへの移行の場合、移行元論理デバイスのデータを先頭アドレスから順次移行先物理デバイス又は外部デバイスへ移行し、全領域のデータ移行が完了したら、移行元論理デバイスを現在対応している物理デバイス又は外部デバイスから移行先空き物理デバイス又は外部デバイスへ対応づくように、デバイス管理情報を更新する（ステップ１６０６、１６０７）。また、要求処理が別の論理デバイスとの入替の場合、移行元／移行先論理デバイスのデータを先頭アドレスから順次不揮発メモリ１３４一時領域などの別の記憶装置に読み上げ、移行元論理デバイスのデータを移行先論理デバイスへ、移行先論理デバイスのデータを移行元論理デバイスへ入れ替えて、書き込んでいく。当該論理デバイス全領域についてデータ入替が完了したら、移行元論理デバイスと移行先論理デバイスについて、それぞれが対応する物理デバイス又は外部デバイスとの対応関係を入れ替えるよう、デバイス管理情報を更新する（ステップ１６０４、１６０５）。

【００７７】

一方、当該論理デバイスが拡張ボリュームの場合、要求処理が別のプールボリュームへの移行処理の場合、制御プロセッサ１３２は、当該論理デバイスの先頭エントリから順次

、対応する拡張ボリュームで実領域が割当済みのエントリを検索し、当該エントリのデータのみを移行先プールボリュームの空きエントリに追記していく。より詳細には、当該拡張ボリュームの対応プールボリュームに移行先プールボリュームを追加し、移行元プールボリュームのエントリから読み上げたデータを移行先プールボリュームの空きエントリを割り当てて書き込み、移行先ボリュームのエントリを当該拡張ボリュームの当該エントリに対応づくように管理情報を更新する。本処理を論理デバイス管理情報の進捗ポインタを用いながら順次実行していき、全領域について移行処理が完了したら、当該拡張ボリュームの対応プールボリュームリストから移行元プールボリューム（当初当該論理デバイスが対応していたプールボリューム）を削除する（ステップ１６０９、１６１０）。また、要求処理が別の論理デバイスとの入替指示の場合、制御プロセッサ１３２は、移行元／移行先論理デバイスの先頭エントリから順に実領域（プールボリュームのエントリ）が割り当てられているエントリを検索し、対象エントリについて相手側のプールボリュームの空きエントリを新規に割り当てて当該エントリのデータを移行していく。この移行処理過程において、移行先／移行元論理デバイスの移行実施エントリ数に大きな差がついてしまうと、一方のプールボリュームの空きエントリ数が枯渇する恐れがある。従って、移行実施エントリ数が移行先／移行元論理デバイスで大きく乖離しないよう、進捗を制御する必要がある。このように全領域の入替処理が完了したら、移行先／移行元論理デバイスとそれぞれのプールボリュームとの対応が入れ替わるようデバイス管理情報を更新する（ステップ１６１１から１６１６）。

【００７８】

図１７は、ストレージ１３０で実行される、物理デバイス定義解除処理２２７の処理フローの一例を示す図である。制御プロセッサ１３２は、指定された物理デバイスに対応する物理デバイス管理情報２０３を初期化し、不揮発メモリ１３４に設定した常置領域と未割当領域との対応を解消する（ステップ１７０１）。次に当該不揮発メモリ１３４の常置領域が未割当領域となるよう、不揮発メモリ管理情報２０６を更新し、処理を終了する（ステップ１７０２）。

【００７９】

以上の説明のように、管理サーバ１１０とストレージ１３０の連携によって、不揮発メモリ１３４の一時領域の割当状態が変更される。次に、こうしたシステム環境におけるストレージ運用管理に伴う通常処理として、論理デバイスの移行処理を例に説明する。

【００８０】

論理デバイス移行処理に対応して、管理サーバ１１０にはデバイス移行指示処理２４２がメモリ１１２に、ストレージ１３０にはデバイス移行処理２２３がメモリ１３３にそれぞれ格納される。このうち、これまでに説明していないデバイス移行指示処理２４２について説明する。

【００８１】

図１８は、管理サーバ１１０で実行される、デバイス移行指示処理２４２の処理フローの一例を示す図である。管理サーバ１１０は、論理デバイスの中から別の物理デバイス又は外部デバイスに対応する別の論理デバイスとの入替が必要かを判定する（ステップ１８０１）。具体的には、ＤＬＣＭ（データライフサイクル管理）のように、各論理デバイスに格納されたデータの価値、アクセス頻度やデータの保持期限などに基づいて、ストレージ１３０が管理する階層ストレージの各階層間でより好適なデータ配置となるよう、デバイス間でのデータ配置を再考する。判定の結果、入替が必要なデバイスが存在した場合（ステップ１８０２）、管理サーバ１１０は移行元および移行先デバイスを特定する情報を付加したデバイス入替をストレージ１３０に指示し（ステップ１８０３）、入替が完了した時点で当該デバイスの最新情報をストレージ１３０からメモリ１１２へ取り込み、デバイス管理情報の複製２３１を更新する（ステップ１８０４）。

【００８２】

ステップ１８０１において、例えば、デバイス単位のアクセス頻度により、データ配置

する際、制御プロセッサ１３２は、デバイス単位あたりのアクセス頻度を記録する。管理サーバ１１０のＣＰＵ１１１は、このアクセス頻度の情報を取得し、デバイス移行指示処理２４２に格納されたアルゴリズムに従って、データ再配置を指示する。

【００８３】

次に、データ再配置をデバイス単位ではなくファイル単位で実行する場合について、説明する。図２１は、管理サーバ１１０が有するアドレス変換テーブル２１００の一例を示す図である。ＣＰＵ１１１は、ホスト計算機から情報を取得することによって、アドレス変換テーブル２１０を作成することができる。アドレス変換テーブル２１０を使用することによって、ＣＰＵ１１１は、ファイル単位でデータを読み出すことや書込みすることができる。また、このアドレス変換テーブルは、オペレーティングシステム毎に作成される。

【００８４】

図２２は、ホスト計算機１００のＣＰＵ１０１が管理サーバ１１０の指示により、実行した結果、収集できる情報の一例を示す図である。ホスト計算機１００のＣＰＵ１０１は、各ファイルについて、ファイル名、ファイル生成日時、最終アクセス日時、ファイルサイズ、ファイルアドレスや格納先のＬＵ番号等をそれぞれ収集する。収集した後、ＣＰＵ１０１は、管理サーバ１１０に送信する。この情報のファイル生成日時や最終アクセス日時により、長期間アクセスされなかったファイルをデータ再配置することができる。　上記の実施例２の構成により、実施例１の効果に加えて、システム構成変更やアクセス特性変化に応じて、不揮発メモリの一時／常置領域の割り当てを動的に変更することが可能となる。

【００８５】

以上、本発明の好適な幾つかの実施形態を説明したが、これらは本発明の説明のための例示であって、本発明の範囲をこれらの実施形態にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。

【００８６】

例えば、オープンシステムの替わりにメインフレームシステムにも本発明を適用することができる。

【００８７】

また、例えば、管理サーバ１１０の機能をホスト計算機１００に組み込むこともできる。この場合、管理サーバ１１０が不要になり、ホスト計算機により、運用と管理の両方をすることができる。また、管理サーバ１１０は、ストレージ１３０に対して、指示する場合、管理サーバ１１０にファイバチャネルスイッチを介して、アウトバンドではなくインバンドで指示を実行することができる。

【００８８】

また、例えば、ＮＡＳアダプタをストレージシステムに内蔵することができる。図２３は、ＮＡＳアダプタを内蔵したストレージシステムの一例を示す図である。図１との相違点は、ストレージシステム１３０にＮＡＳアダプタ２３１２が内蔵されていることと、ＩＰスイッチ１２２が追加されていることである。ＮＡＳアダプタ２３１０のプロセッサ２３１１は、メモリ２３１２に格納されたオペレーティングシステム又はファイルシステムにより、ホスト計算機からのファイルアクセスをブロックアクセスに変換して、不揮発メモリ１３４、ディスク１３７や外部ストレージ１５０にデータを保存することができる。ＮＡＳアダプタ２３１０が内蔵されている場合、ファイル収集情報２２００は、プロセッサ２３１１が収集し、管理端末１４０を介して、管理サーバ１１０に送信される。

【図面の簡単な説明】

【００８９】

【図１】計算機システムのハードウェア構成の一例を示す図である。

【図２】ストレージのソフトウェア構成の一例を示す図である。

【図３】論理デバイス管理情報の一例を示す図である。

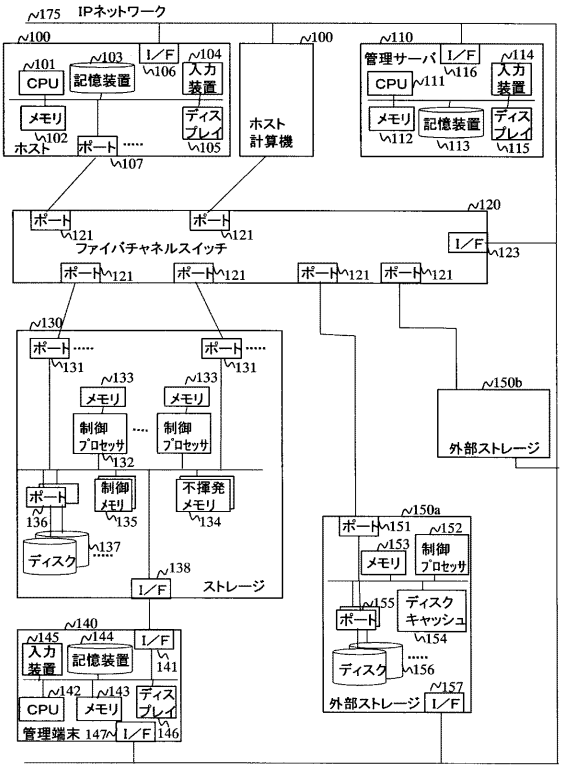
【図４】ＬＵパス管理情報の一例を示す図である。

【図５】物理デバイス管理情報の一例を示す図である。
【図６】外部デバイス管理情報の一例を示す図である。
【図７】拡張ボリューム管理情報の一例を示す図である。
【図８】プールボリューム管理情報の一例を示す図である。
【図９】実施例１において、管理サーバ１１０で実行される、一時領域定義指示処理２４３の処理フローの一例を示す図である。
【図１０】実施例１において、ストレージ１３０で実行される、一時領域定義処理２２４の処理フローの一例を示す図である。
【図１１】実施例１において、ストレージ１３０で実行される、物理デバイス定義処理２２６の処理フローの一例を示す図である。
【図１２】実施例１において、管理サーバ１１０で実行される、デバイス定義指示処理２４１の処理フローの一例を示す図である。
【図１３】実施例１において、ストレージ１３０で実行される、デバイス定義処理２２２の処理フローの一例を示す図である。
【図１４】実施例２において、管理サーバ１１０で実行される、一時領域変更指示処理２４４の処理フローの一例を示す図である。
【図１５】実施例２において、ストレージ１３０で実行される、一時領域割当解除処理２２５の処理フローの一例を示す図である。
【図１６】実施例２において、ストレージ１３０で実行される、デバイス移行処理２２３の処理フローの一例を示す図である。
【図１７】実施例２において、ストレージ１３０で実行される、物理デバイス定義解除処理２２７の処理フローの一例を示す図である。
【図１８】実施例２において、管理サーバ１１０で実行される、デバイス移行指示処理２４２の処理フローの一例を示す図である。
【図１９】実施例１および２において、ストレージ１３０で実行される、Ｉ／Ｏ処理処理２２１の処理フローの一例を示す図である。
【図２０】実施例１および２において、ストレージ１３０に搭載する不揮発メモリ１３４の構成の一例を示す図である。
【図２１】実施例２において、管理サーバが保持するアドレス変換テーブルの一例を示す図である。
【図２２】実施例２において、ホスト計算機１００が収集するファイル収集情報の一例を示す図である。
【図２３】実施例１及び２に共通する変形例での計算機システムのハードウェア構成の一例を示す図である。
【符号の説明】
【００９０】
１００…ホスト計算機
１１０…管理サーバ
１２０…ファイバチャネルスイッチ
１３０…ストレージ
１４０…管理端末
１５０…外部ストレージ
１７５…ＩＰネットワーク
２０１…論理デバイス管理情報
２０２…ＬＵパス管理情報
２０３…物理デバイス管理情報
２０４…外部デバイス管理情報
２０５…キャッシュ管理情報
２０６…不揮発メモリ管理情報
２０７…拡張ボリューム管理情報

２０８…プールボリューム管理情報
２０９…デバイス稼動情報
２２１…Ｉ／Ｏ処理
２２２…デバイス定義処理
２２３…デバイス移行処理
２２４…一時領域割当処理
２２５…一時領域割当解除処理
２２６…物理デバイス定義処理
２２７…物理デバイス定義解除処理
２６１…Ｉ／Ｏ処理
２４１…デバイス定義指示処理
２４２…デバイス移行指示処理
２４３…一時領域定義指示処理
２４４…一時領域変更指示処理

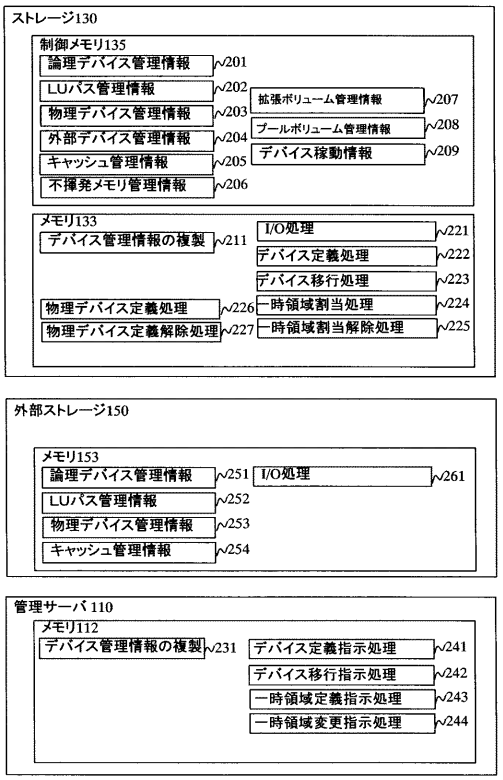
【図１】

図1



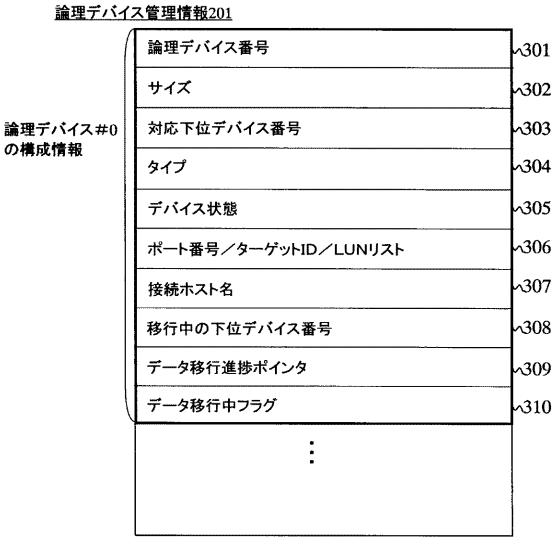
【図２】

図2



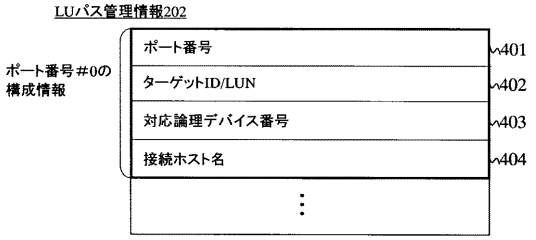
【図３】

図3



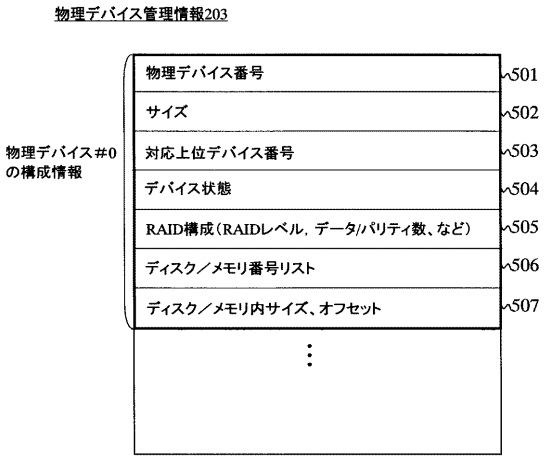
【図４】

図4



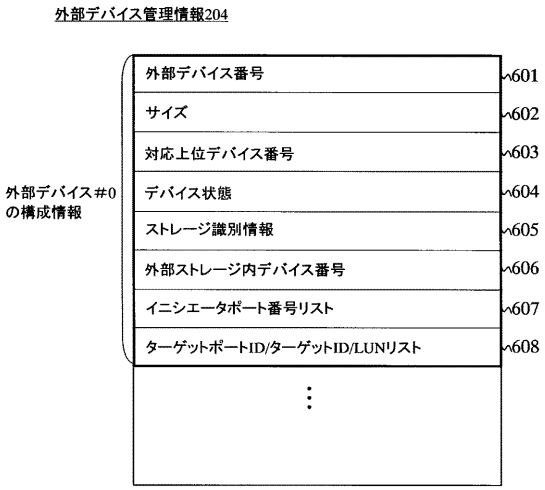
【図５】

図5



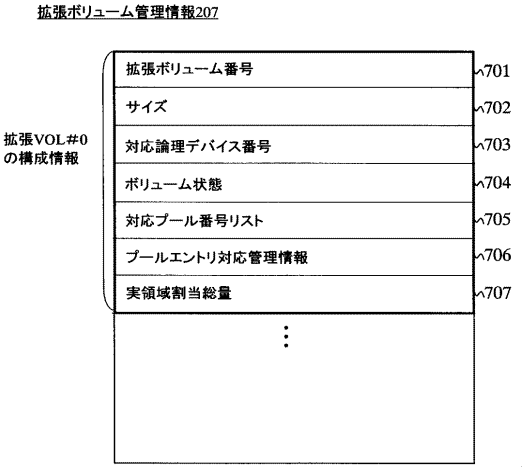
【図６】

図6



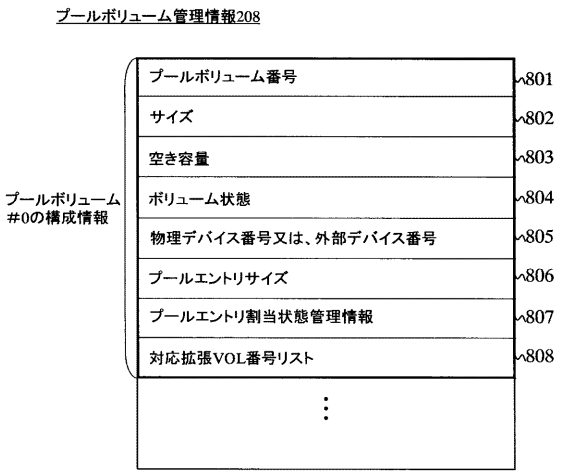
【図７】

図7



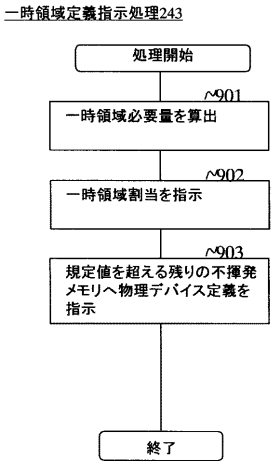
【図８】

図8



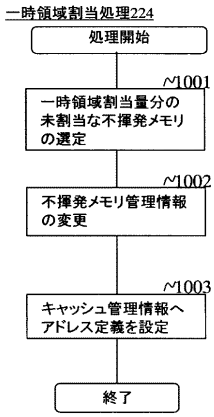
【図９】

図9



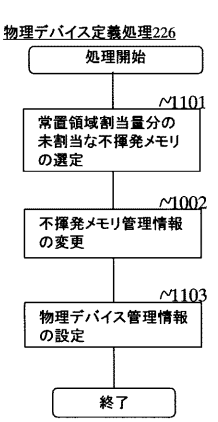
【図１０】

図10



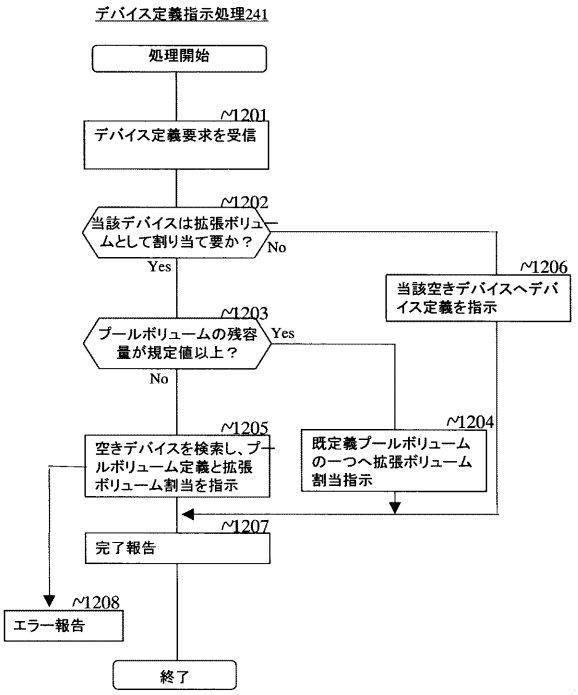
【図１１】

図11



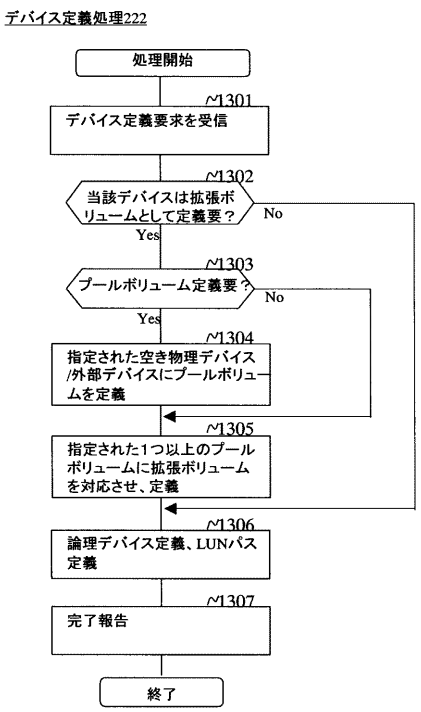
【図１２】

図12



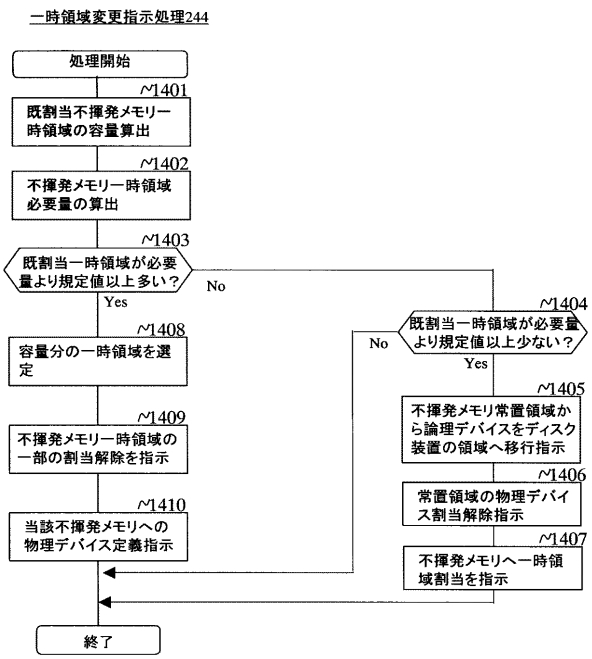
【図１３】

図13



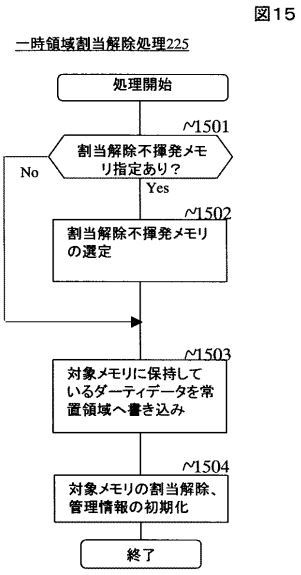
【図１４】

図14



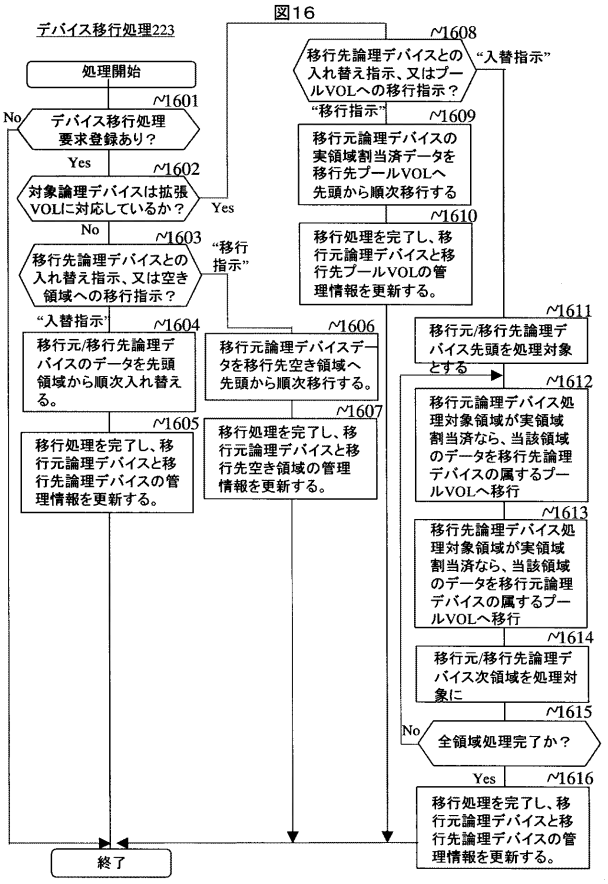
【図１５】

図15



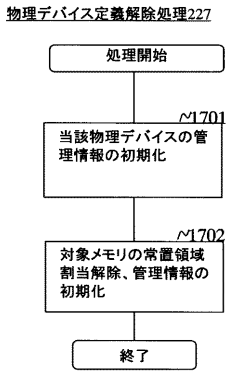
【図１６】

図16



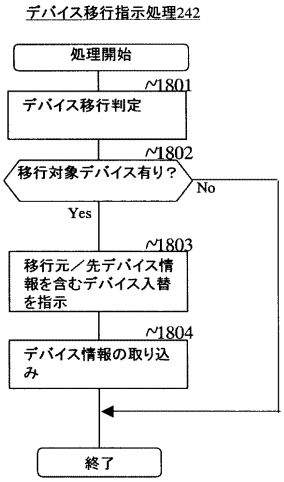
【図１７】

図17



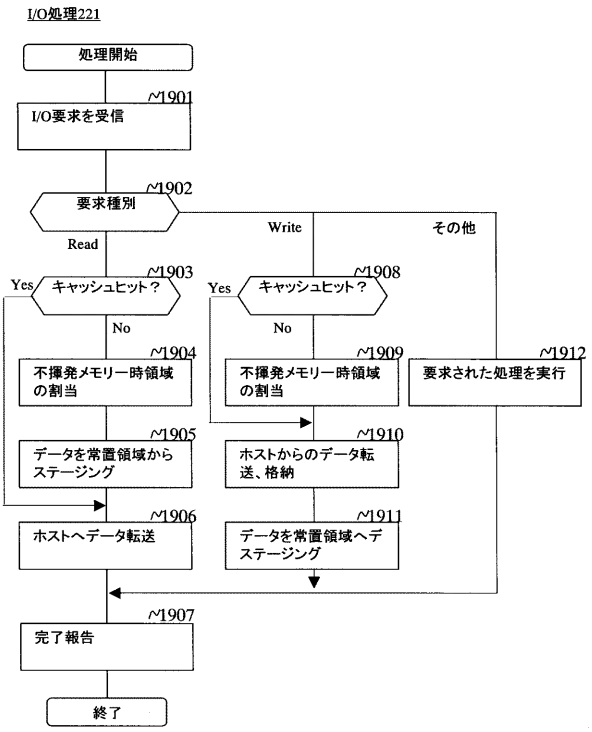
【図１８】

図18



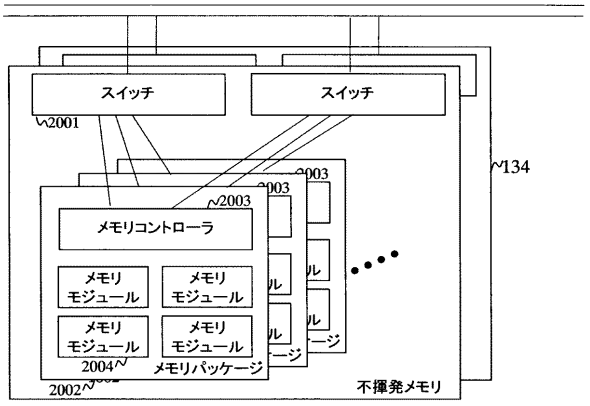
【図１９】

図19



【図２０】

図20



【図２１】

図21

アドレス変換テーブル2100

ファイルアドレス1	∿2110	LBA01	∿2120
ファイルアドレス2	∿2111	LBA02	∿2121
ファイルアドレス3	∿2112	LBA03	∿2122
ファイルアドレス4	∿2113	LBA04	∿2123
ファイルアドレス5	∿2114	LBA05	∿2124
ファイルアドレス6	∿2115	LBA06	∿2125
ファイルアドレス7	∿2116	LBA07	∿2126
⋮		⋮	

【図２２】

図22

ファイル収集情報2200

ファイル名	∿2220
ファイル生成日時	∿2221
最終アクセス日時	∿2222
ファイルサイズ	∿2223
ファイルアドレス	∿2224
格納先LU#	∿2225

【図２３】

図23

