

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5124103号
(P5124103)

(45) 発行日 平成25年1月23日(2013.1.23)

(24) 登録日 平成24年11月2日(2012.11.2)

(51) Int. Cl.

F I

G06F 13/10	(2006.01)	G06F 13/10	340A
G06F 3/06	(2006.01)	G06F 3/06	301J
G06F 12/08	(2006.01)	G06F 3/06	301K
		G06F 12/08	501E
		G06F 12/08	557

請求項の数 20 (全 73 頁)

(21) 出願番号 特願2006-136869 (P2006-136869)
 (22) 出願日 平成18年5月16日(2006.5.16)
 (65) 公開番号 特開2007-310495 (P2007-310495A)
 (43) 公開日 平成19年11月29日(2007.11.29)
 審査請求日 平成21年1月30日(2009.1.30)

(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 100093861
 弁理士 大賀 真司
 (72) 発明者 川村 俊二
 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
 (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

最終頁に続く

(54) 【発明の名称】 計算機システム

(57) 【特許請求の範囲】

【請求項1】

ホスト計算機と、

外部ボリュームを有する外部ストレージシステムと、

前記外部ボリュームを仮想化してなる第一の拡張デバイスと、前記ホスト計算機から認識できるように中間記憶階層を介して前記第一の拡張デバイスを仮想化してなる第一の論理ユニットと、前記第一の論理ユニットに読み書きされるデータを一時的に格納するディスクキャッシュとを有する第一のストレージシステムと、

前記ホスト計算機及び前記外部ストレージシステムに接続可能な第二のストレージシステムと、を備え、

前記第一のストレージシステムは、前記第一の論理ユニットを前記第一のストレージシステムから前記第二のストレージシステムに移動させる指示を受けると、前記第一のストレージシステムの動作モードを、前記ディスクキャッシュに格納されたデータを外部ボリュームにも書き込むキャッシュスルーモードに設定し、

前記第二のストレージシステムは、前記第一の論理ユニットを前記第一のストレージシステムから前記第二のストレージシステムに移動させる指示を受けると、前記外部ボリュームを仮想化する第二の拡張デバイスを前記第二のストレージシステム内に作成し、前記第一の論理ユニットを仮想化する第三の拡張デバイスを前記第二のストレージシステム内に作成し、前記第一の論理ユニットと前記第三の拡張デバイスを接続するパスを定義し、前記ホスト計算機から認識できるように前記第二のストレージシステム内の中間記憶階層

を介して前記第三の拡張デバイスを仮想化してなる第二の論理ユニットを作成し、前記第二の論理ユニットと前記ホスト計算機とを接続するパスを、前記第一の論理ユニットと前記ホスト計算機とを接続するパスの交替パスとして設定し、

前記ホスト計算機は、前記交替パスを経由して、前記第二の論理ユニットから前記第二のストレージシステム内の中間記憶階層、前記第一及び前記第二のストレージシステムを接続するパス、前記第一の論理ユニット、及び前記第一のストレージシステム内の中間記憶階層を介して前記外部ボリュームにアクセスし、

前記第一のストレージシステムの前記ディスクキャッシュに一時的に格納されているデータの全てが前記外部ボリュームに書き込まれると、前記第二のストレージシステムは、前記第三の拡張デバイスと前記第二の論理ユニットとの対応関係を解除して、前記第二の拡張デバイスを前記第二の論理ユニットに対応付ける、

ことを特徴とする、計算機システム。

【請求項 2】

前記第一のストレージシステム内の中間記憶階層は、前記第一の拡張デバイスを仮想化するための仮想デバイスを含み、前記第一の論理ユニットは、前記仮想デバイスの一部を仮想化したものである、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 3】

前記第一のストレージシステムは、前記外部ボリュームを構成する記憶領域のうち前記仮想デバイスの一部に対応する記憶領域へのアクセスが禁止される、

ことを特徴とする、請求項 2 に記載の計算機システム。

【請求項 4】

前記第一のストレージシステム内の中間記憶階層は、前記第一の拡張デバイスを仮想化するための論理デバイスを含み、前記第一の論理ユニットは、複数の論理デバイスを一つにまとめて仮想化したものである、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 5】

前記第一のストレージシステム内の中間記憶階層は、前記第一の拡張デバイスの記憶領域に R A I D 構成を適用することにより、前記拡張デバイスを仮想化する仮想デバイスを含み、前記第一の論理ユニットは、前記仮想デバイスの一部を仮想化したものである、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 6】

前記外部ボリュームは、前記ホスト計算機が読み書きするユーザデータを格納する第一の記憶領域と、前記第一の論理ユニットの管理情報を格納する第一の記憶領域を含み、

前記第一のストレージシステム内の中間記憶階層は、前記第一の拡張デバイスを仮想化するための論理デバイスを含み、

前記論理デバイスは、前記外部ボリュームの前記第一の記憶領域を仮想化する第一の記憶領域と、前記外部ボリュームの前記第二の記憶領域を仮想化する第二の記憶領域とを含み、

前記第一の論理ユニットは、前記論理デバイスの前記第一の記憶領域を仮想化したものである、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 7】

前記第二のストレージシステムは、前記外部ボリュームの前記第二の記憶領域に格納されている前記管理情報を前記第二の論理ユニットの属性情報として設定する、

ことを特徴とする、請求項 6 に記載の計算機システム。

【請求項 8】

複数の第一の論理ユニットを前記第一のストレージシステムから前記第二のストレージシステムに移動させる、

ことを特徴とする、請求項 1 に記載の計算機システム。

10

20

30

40

50

【請求項 9】

前記第一の論理ユニットの一部を前記第一のストレージシステムから前記第二のストレージシステムに移動させる、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 10】

前記第一の論理ユニットは、複数のファイルシステムを格納しており、前記論理ユニットの一部は、前記複数のファイルシステムのうち何れか一つ以上のファイルシステムである、

ことを特徴とする、請求項 9 に記載の計算機システム。

【請求項 11】

前記外部ストレージシステムは、前記外部ボリュームを構成する記憶領域のうち前記第一の論理ユニットの一部に対応する記憶領域については、前記第二のストレージシステムからのアクセスを排他的に受け付け、前記外部ボリュームを構成する記憶領域のうち前記第一の論理ユニットの一部を除く部分に対応する記憶領域については、前記第一のストレージシステムからのアクセスを排他的に受け付ける、

ことを特徴とする、請求項 9 に記載の計算機システム。

【請求項 12】

前記ホスト計算機は、前記外部ボリュームを構成する各記憶領域についてアクセス権を有するストレージシステムに接続するためにパス切り替えを実行する、

ことを特徴とする、請求項 11 に記載の計算機システム。

【請求項 13】

前記第二のストレージシステムが前記第三の拡張デバイスに書き込むデータを暗号化する機能はオフに設定されており、前記第一のストレージシステムが前記第一の拡張デバイスに書き込むデータを暗号化する機能はオンに設定されている、

ことを特徴とする、請求項 1 に記載の計算機システム。

【請求項 14】

ホスト計算機と、

外部ボリュームを有する外部ストレージシステムと、

前記外部ボリュームを仮想化してなる第一の拡張デバイスと、前記ホスト計算機から認識できるように中間記憶階層を介して前記第一の拡張デバイスを仮想化してなる第一の論理ユニットと、前記第一の論理ユニットに読み書きされるデータを一時的に格納するディスクキャッシュとを有する複数の第一のストレージシステムと、

前記ホスト計算機及び前記外部ストレージシステムに接続可能な第二のストレージシステムと、

前記第一及び前記第二のストレージシステムに接続する管理サーバと、を備え、

前記管理サーバは、前記複数の第一のストレージシステムのうち障害又は性能低下が生じたストレージシステムを移動元ストレージシステムとして設定し、前記移動元ストレージシステムが有する前記第一の論理ユニットを移動元論理ユニットとして設定し、前記第二のストレージシステムを移動先ストレージシステムとして設定し、

前記移動元ストレージシステムは、前記第一のストレージシステムの動作モードをディスクキャッシュに格納されたデータを外部ボリュームにも書き込むキャッシュスルーモードに設定し、

前記移動先ストレージシステムは、前記外部ボリュームを仮想化する第二の拡張デバイスを前記移動先ストレージシステム内に作成し、前記移動元論理ユニットを仮想化する第三の拡張デバイスを前記移動先ストレージシステム内に作成し、前記移動元論理ユニットと前記第三の拡張デバイスを接続するパスを定義し、前記ホスト計算機から認識できるように前記移動先ストレージシステム内の中間記憶階層を介して前記第三の拡張デバイスを仮想化してなる移動先論理ユニットを作成し、前記移動先論理ユニットと前記ホスト計算機とを接続するパスを、前記移動元論理ユニットと前記ホスト計算機とを接続するパスの交替パスとして設定し、

10

20

30

40

50

前記ホスト計算機は、前記交替パスを経由して、前記移動先論理ユニットから前記移動先ストレージシステム内の中間記憶階層、前記移動先ストレージシステムと前記移動元ストレージシステムとを接続するパス、前記移動元論理ユニット、及び前記移動元ストレージシステム内の中間記憶階層を介して前記外部ボリュームにアクセスし、

前記ディスクキャッシュに一時的に格納されているデータの全てが前記外部ボリュームに書き込まれると、前記第二のストレージシステムは、前記第三の拡張デバイスと前記移動先論理ユニットとの対応関係を解除して、前記第二の拡張デバイスを前記移動先論理ユニットに対応付ける、

ことを特徴とする、計算機システム。

【請求項 15】

ホスト計算機と、

外部ボリュームを有する外部ストレージシステムと、

前記外部ボリュームを仮想化してなる第一の拡張デバイスと、前記ホスト計算機から認識できるように中間記憶階層を介して前記第一の拡張デバイスを仮想化してなる第一の論理ユニットと、前記第一の論理ユニットに読み書きされるデータを一時的に格納するディスクキャッシュとを有する第一のストレージシステムと、

前記外部ボリュームを仮想化してなる第二の拡張デバイスと、前記ホスト計算機から認識できるように中間記憶階層を介して前記第二の拡張デバイスを仮想化してなる第二の論理ユニットとを有する第二のストレージシステムと、を備え、

前記第一のストレージシステムは、前記第一の論理ユニットを前記第一のストレージシステムから前記第二のストレージシステムに移動させる指示を受けると、前記第一のストレージシステムの動作モードを、前記ディスクキャッシュに格納されたデータを外部ボリュームにも書き込むキャッシュスルーモードに設定し、

前記第一の拡張デバイスは、前記第二の論理ユニットを仮想化したものであり、

前記第二の論理ユニットと前記ホスト計算機とを接続するパスは、前記第一の論理ユニットと前記ホスト計算機とを接続するパスの交替パスとして設定されており、

前記ホスト計算機は、前記第一の論理ユニットから前記第一のストレージシステム内の中間記憶階層、前記第一及び前記第二のストレージシステムを接続するパス、前記第二の論理ユニット、及び前記第二のストレージシステム内の中間記憶階層を介して前記外部ボリュームにアクセスし、

前記第一のストレージシステムの前記ディスクキャッシュに一時的に格納されているデータの全てが前記外部ボリュームに書き込まれると、前記ホスト計算機は、前記交替パスを経由して、前記第二の論理ユニットにアクセスする、

ことを特徴とする、計算機システム。

【請求項 16】

ホスト計算機と、

前記ホスト計算機に利用されるデータを格納する外部ボリュームを有する外部ストレージシステムと、

前記外部ボリュームと前記第一のストレージシステムとの間の第一のパスを介して前記外部ボリュームを仮想化してなる第一の拡張デバイスを有し、前記第一の拡張デバイスに対応する第一の論理ユニットを前記ホスト計算機に提供する第一のストレージシステムと、

前記外部ボリュームと前記第二のストレージシステムとの間の第二のパスを介して前記外部ボリュームを仮想化してなる第二の拡張デバイスを有し、前記前記第二の拡張デバイスに対応する第二の論理ユニットを前記ホスト計算機に提供する第二のストレージシステムと、

を備え、

前記第一のストレージシステムは、

前記第一のストレージシステムと前記第二のストレージシステムの間の第三のパスを介して前記第二の論理ユニットを仮想化してなる第三の拡張デバイスをさらに備え、

10

20

30

40

50

前記第一のストレージシステムが前記第一の論理ユニットへの書き込みデータを受信すると、前記第一のストレージシステムは、キャッシュスルーモードにより、前記第三の拡張デバイス、前記第三のパス、前記第二の拡張デバイス、前記第二のパスを介して前記外部ボリュームにアクセスして前記外部ボリュームにデータを格納し、

前記ホスト計算機は、

前記第一のストレージシステムに障害が発生した場合に、前記第一の論理ユニットに代えて前記第二の論理ユニットを利用して前記外部ボリュームに格納されているデータへのアクセスに切り替える

ことを特徴とする、計算機システム。

【請求項 17】

前記第一のストレージシステムが前記第一の論理ユニットへのリードアクセスを受信すると、前記第一のストレージシステムは、前記第一の拡張デバイス及び前記第一のパスを介して前記外部ボリュームからデータを読み込む

ことを特徴とする、請求項 16 に記載の計算機システム。

【請求項 18】

前記ホスト計算機と、

前記ホスト計算機に利用されるデータを格納している外部ボリュームを有する前記外部ストレージシステムと、

前記外部ボリュームと前記第一のストレージシステムとの間の第一のパスを介して前記外部ボリュームを仮想化してなる第一の拡張デバイスを有し、前記第一の拡張デバイスに対応する第一の論理ユニットを前記ホスト計算機に提供し、前記ホスト計算機からの前記第一の論理ユニットへのアクセスを制御する前記第一ストレージシステムと、

外部ボリュームと前記第二のストレージシステムとの間の第二のパスを介して前記外部ボリュームを仮想化してなる第二の拡張デバイスを有し、前記前記第二の拡張デバイスに対応する第二の論理ユニットを前記ホスト計算機に提供し、前記ホスト計算機からの前記第二論理ユニットへのアクセスを制御する前記第二ストレージシステムと、

を備え、

前記第一ストレージシステムは、

前記第一のストレージシステムと前記第二のストレージシステムとの間の第三のパスを介して前記第二の論理ユニットを仮想化してなる第三の拡張デバイスをさらに備え、

前記第一のストレージシステムにのみが前記外部ボリュームへのアクセス権限を有し、前記第一のストレージシステムは、前記第一の論理ユニットへの書き込みデータを受信すると、前記第一のストレージシステムは、キャッシュスルーモードにより、前記第三の拡張デバイス及び前記第三のパスを介してアクセスして前記第二の論理ユニットにデータを格納し、

前記第一のストレージシステムが前記第一の論理ユニットから前記外部ボリュームへのダーティデータをデステージした後で、前記第二のストレージシステムは、前記第二の論理ユニットに格納されているデータを削除する

前記ホスト計算機は、

前記第一のストレージシステムに障害が発生した場合に、前記第一の論理ユニットに代えて前記第二の論理ユニットを利用して前記外部ボリュームに格納されているデータへのアクセスに切り替える

ことを特徴とする、計算機システム。

【請求項 19】

前記第二のストレージシステムは、前記外部ボリュームへのアクセス権限を取得して、前記外部ボリュームに前記第二論理ユニットに格納されているダーティデータをデステージする

ことを特徴とする、請求項 18 に記載の計算機システム

【請求項 20】

前記第二の論理ユニットは、同期コピーを形成するコピーペアとして前記第一の論理ユ

10

20

30

40

50

ニットに対応すると関連付けられる

ことを特徴とする、請求項 18 に記載の計算機システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はストレージ仮想化技術を適用したストレージシステムを有する計算機システムに関する。

【背景技術】

【0002】

インターネットビジネスの拡大、手続きの電子化など、急速に情報システムが進展する中、コンピュータが送受信するデータ量は、飛躍的に増加している。このようなデータ量の飛躍的増加に加え、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、ストレージに格納するデータ量も飛躍的に増加を続けている。これに伴い、企業情報システムにおいて、各部門/各システムのストレージの増強を図る一方で、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技術に対する期待が高まっている。

【0003】

大規模ストレージを有するシステムの管理コストを低減する方法として、特開平2005-011277号公報に開示されているようなストレージ仮想化技術を挙げることができる。特開平2005-011277号公報では、ホスト計算機に接続するストレージシステム(以下、第一のストレージシステムと称する。)を一つ以上の外部ストレージシステム(以下、第二のストレージシステムと称する。)に接続し、第二のストレージシステムが有するデバイス(以下、外部デバイスと称する。)をあたかも第一のストレージシステム内の論理デバイスであるかのように仮想化してホスト計算機に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技術とも称する)が開示されている。第一のストレージシステムは、ホスト計算機からの論理デバイスへの入出力要求を受信すると、アクセス先のデバイスが第二のストレージシステム内の外部デバイスに対応しているのか、或いは第一のストレージシステムが搭載するディスク装置などの物理デバイス(内部デバイス)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。

【0004】

特開平2005-011277号公報に示す外部ストレージ接続技術を利用したストレージシステムを用いると、性能、信頼性、価格などの属性の異なる複数のストレージシステムを統合したシステムを構築できる。例えば、外部ストレージ接続技術を利用した高コスト・高機能・高信頼な第一のストレージシステムと、低コスト・低機能・低信頼な第二のストレージシステムとを接続し、記憶領域を論理的に階層化することによって、データの更新日時やデータの価値に応じた最適なデータ配置が可能となる。階層化された記憶領域を有するストレージシステムを用いることで、監査対応などを目的として、日々の業務で発生する取引情報やメールなどの大量の情報を、各情報が持つ価値に応じて最適なコストで長期間保存することが可能となる。

【0005】

ところで、ストレージシステムの寿命を超えた数十年の長期にわたって大容量のデータを保存するには、データの保存期間中に計算機システムを構成する機器を交換する作業が必要になる。特開平2005-011277号公報では、外部ストレージ接続技術を適用した第一のストレージシステム(以下、移動元ストレージと称する。)を他の第一のストレージシステム(以下、移動先ストレージと称する。)に交換するために、移動元ストレージが有する論理ユニット(以下、移動元論理ユニットと称する。)内のデータを、移動先ストレージが有する論理ユニット(以下、移動先論理ユニットと称する。)へコピーし

10

20

30

40

50

て移動することが開示されている。

【特許文献1】特開平2005-011277号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

しかし、移動元論理ユニットから移動先論理ユニットへのコピー処理によって、移動元ストレージが行うデバイスの入出力処理にスループット低下などの悪影響を及ぼすことが懸念される。更に、移動元論理ユニット内のデータを全て移動先論理ユニットにコピーする必要があるため、移動元ストレージから移動先ストレージへのデバイス引継ぎに多くの時間を要するという問題が生じる。

10

【0007】

また、ストレージ階層の複雑化を背景として、移動元論理ユニットと外部デバイスとが一对一に対応している場合のみならず、両者が一对一に対応していない場合も想定されるところである。特に、移動元論理ユニットと外部デバイスとが一对一に対応していない場合においても、ホスト計算機が認識する記憶領域の単位である論理ユニット単位でのデータ移動を可能にすることが望ましい。

【0008】

本発明の目的は、ストレージ仮想化技術を適用した計算機システムにおいて、計算機システムの構成変更に伴う論理ユニット単位でのデータ移動を、ストレージシステムの性能低下を抑制しつつ、可能とすることにある。

20

【0009】

本発明の他の目的は、ストレージ仮想化技術を適用した計算機システムにおいて、外部ボリュームを複数のストレージシステムによって共有することにある。

【課題を解決するための手段】

【0010】

上記の課題を解決するため、本発明の計算機システムは、ホスト計算機と、外部ボリュームを有する外部ストレージシステムと、外部ボリュームを仮想化してなる第一の拡張デバイスと、第一の拡張デバイスを仮想化してなる、ホスト計算機から認識可能な第一の論理ユニットとを有する第一のストレージシステムと、外部ボリュームを仮想化してなる第二の拡張デバイスと、第二の拡張デバイスを仮想化してなる、ホスト計算機から認識可能な第二の論理ユニットとを有する第二のストレージシステムと、を備える。

30

【0011】

ここで、例えば、外部ストレージシステムは、外部ボリュームを構成する各記憶領域について、第一又は第二のストレージシステムのうち何れか一方のみの排他的アクセスを制御する。

【0012】

ホスト計算機は、例えば、外部ボリュームを構成する各記憶領域についてアクセス権を有するストレージシステムに接続するためにパス切り替えを実行する。

【0013】

外部ボリュームは、第一及び第二のストレージシステムの双方からアクセスされる共用領域を含むものであってもよい。

40

【0014】

本発明の他の側面に関する計算機システムは、ホスト計算機と、外部ボリュームを有する外部ストレージシステムと、外部ボリュームを仮想化してなる第一の拡張デバイスと、ホスト計算機から認識できるように中間記憶階層を介して第一の拡張デバイスを仮想化してなる第一の論理ユニットとを有する第一のストレージシステムと、ホスト計算機及び外部ストレージシステムに接続可能な第二のストレージシステムと、を備える。第二のストレージシステムは、第一の論理ユニットを第一のストレージシステムから第二のストレージシステムに移動させる指示を受けると、外部ボリュームを仮想化する第二の拡張デバイスを第二のストレージシステム内に作成し、第一の論理ユニットを仮想化する第三の拡張デ

50

バイスを第二のストレージシステム内に作成し、第一の論理ユニットと第三の拡張デバイスを接続するパスを定義し、ホスト計算機から認識できるように第二のストレージシステム内の中間記憶階層を介して第三の拡張デバイスを仮想化してなる第二の論理ユニットを作成し、第二の論理ユニットとホスト計算機とを接続するパスを、第一の論理ユニットとホスト計算機とを接続するパスの交替パスとして設定する。ホスト計算機は、交替パスを経由して、第二の論理ユニットから第二のストレージシステム内の中間記憶階層、第一及び第二のストレージシステムを接続するパス、第一の論理ユニット、及び第一のストレージシステム内の中間記憶階層を介して外部ボリュームにアクセスする。ホスト計算機から第二の論理ユニットに書き込まれたデータが全て外部ボリュームに書き込まれると、第二のストレージシステムは、第三の拡張デバイスと第二の論理ユニットとの対応関係を解除して、第二の拡張デバイスを第二の論理ユニットに対応付ける。

10

【0015】

第一のストレージシステムは、第一の論理ボリュームに読み書きされるデータを一時的に格納するディスクキャッシュを更に備えており、第一のストレージシステムは、第一の論理ユニットを第一のストレージシステムから第二のストレージシステムに移動させる指示を受けると、第一のストレージシステムの動作モードをキャッシュスルーモードに設定する。

【0016】

第一のストレージシステム内の中間記憶階層は、第一の拡張デバイスを仮想化するための仮想デバイスを含み、第一の論理ユニットは、仮想デバイスの一部を仮想化したものである。

20

【0017】

第一のストレージシステムは、外部ボリュームを構成する記憶領域のうち仮想デバイスの一部に対応する記憶領域へのアクセスが禁止される。

【0018】

第一のストレージシステム内の中間記憶階層は、第一の拡張デバイスを仮想化するための論理デバイスを含み、第一の論理ユニットは、複数の論理デバイスを一つにまとめて仮想化したものである。

【0019】

第一のストレージシステム内の中間記憶階層は、第一の拡張デバイスの記憶領域に RAID 構成を適用することにより、拡張デバイスを仮想化する仮想デバイスを含み、第一の論理ユニットは、仮想デバイスの一部を仮想化したものである。

30

【0020】

外部ボリュームは、ホスト計算機が読み書きするユーザデータを格納する第一の記憶領域と、第一の論理ユニットの管理情報を格納する第一の記憶領域を含み、第一のストレージシステム内の中間記憶階層は、第一の拡張デバイスを仮想化するための論理デバイスを含み、論理デバイスは、外部ボリュームの第一の記憶領域を仮想化する第一の記憶領域と、外部ボリュームの第二の記憶領域を仮想化する第二の記憶領域とを含み、第一の論理ユニットは、論理デバイスの第一の記憶領域を仮想化したものである。

【0021】

第二のストレージシステムは、外部ボリュームの第二の記憶領域に格納されている管理情報を第二の論理ユニットの属性情報として設定する。

40

【0022】

本発明の計算機システムは、複数の第一の論理ユニットを第一のストレージシステムから第二のストレージシステムに移動させる機能を備えてもよい。

【0023】

本発明の計算機システムは、第一の論理ユニットの一部を第一のストレージシステムから第二のストレージシステムに移動させる機能を備えてもよい。

【0024】

本発明の計算機システムにおいて、第一の論理ユニットは、複数のファイルシステムを

50

格納しており、論理ユニットの一部は、複数のファイルシステムのうち何れか一つ以上のファイルシステムである。

【0025】

本発明の計算機システムにおいて、外部ストレージシステムは、外部ボリュームを構成する記憶領域のうち第一の論理ユニットの一部に対応する記憶領域については、第二のストレージシステムからのアクセスを排他的に受け付け、外部ボリュームを構成する記憶領域のうち第一の論理ユニットの一部を除く部分に対応する記憶領域については、第一のストレージシステムからのアクセスを排他的に受け付ける。

【0026】

本発明の計算機システムにおいて、ホスト計算機は、外部ボリュームを構成する各記憶領域についてアクセス権を有するストレージシステムに接続するためにパス切り替えを実行する。

【0027】

本発明の計算機システムにおいて、第二のストレージシステムが第三の拡張デバイスに書き込むデータを暗号化する機能はオフに設定されており、第一のストレージシステムが第一の拡張デバイスに書き込むデータを暗号化する機能はオンに設定されている。

【0028】

本発明の計算機システムは、ホスト計算機と、外部ボリュームを有する外部ストレージシステムと、外部ボリュームを仮想化してなる第一の拡張デバイスと、ホスト計算機から認識できるように中間記憶階層を介して第一の拡張デバイスを仮想化してなる第一の論理ユニットとを有する複数の第一のストレージシステムと、ホスト計算機及び外部ストレージシステムに接続可能な第二のストレージシステムと、第一及び第二のストレージシステムに接続する管理サーバと、を備える。管理サーバは、複数の第一のストレージシステムのうち障害又は性能低下が生じたストレージシステムを移動元ストレージシステムとして設定し、移動元ストレージシステムが有する第一の論理ユニットを移動元論理ユニットとして設定し、第二のストレージシステムを移動先ストレージシステムとして設定する。移動先ストレージシステムは、外部ボリュームを仮想化する第二の拡張デバイスを移動先ストレージシステム内に作成し、移動元論理ユニットを仮想化する第三の拡張デバイスを移動先ストレージシステム内に作成し、移動元論理ユニットと第三の拡張デバイスを接続するパスを定義し、ホスト計算機から認識できるように移動先ストレージシステム内の中間記憶階層を介して第三の拡張デバイスを仮想化してなる移動先論理ユニットを作成し、移動先論理ユニットとホスト計算機とを接続するパスを、移動元論理ユニットとホスト計算機とを接続するパスの交替パスとして設定する。ホスト計算機は、交替パスを経由して、移動先論理ユニットから移動先ストレージシステム内の中間記憶階層、移動先ストレージシステムと移動元ストレージシステムとを接続するパス、移動元論理ユニット、及び移動元ストレージシステム内の中間記憶階層を介して外部ボリュームにアクセスする。ホスト計算機から移動先論理ユニットに書き込まれたデータが全て外部ボリュームに書き込まれると、第二のストレージシステムは、第三の拡張デバイスと移動先論理ユニットとの対応関係を解除して、第二の拡張デバイスを移動先論理ユニットに対応付ける。

【0029】

本発明の計算機システムは、ホスト計算機と、外部ボリュームを有する外部ストレージシステムと、外部ボリュームを仮想化してなる第一の拡張デバイスと、ホスト計算機から認識できるように中間記憶階層を介して第一の拡張デバイスを仮想化してなる第一の論理ユニットとを有する第一のストレージシステムと、外部ボリュームを仮想化してなる第二の拡張デバイスと、ホスト計算機から認識できるように中間記憶階層を介して第二の拡張デバイスを仮想化してなる第二の論理ユニットとを有する第二のストレージシステムと、を備える。第一の拡張デバイスは、第二の論理ユニットを仮想化したものである。第二の論理ユニットとホスト計算機とを接続するパスは、第一の論理ユニットとホスト計算機とを接続するパスの交替パスとして設定されている。ホスト計算機は、第一の論理ユニットから第一のストレージシステム内の中間記憶階層、第一及び第二のストレージシステムを

10

20

30

40

50

接続するバス、第二の論理ユニット、及び第二のストレージシステム内の中間記憶階層を介して外部ボリュームにアクセスする。ホスト計算機から第一の論理ユニットに書き込まれたデータが全て外部ボリュームに書き込まれると、ホスト計算機は、交替バスを經由して、第二の論理ユニットにアクセスする。

【発明の効果】

【0030】

本発明によれば、ストレージ仮想化技術を適用した計算機システムにおいて、計算機システムの構成変更に伴う論理ユニット単位でのデータ移動を、ストレージシステムの性能低下を抑制しつつ、可能とすることができる。

【0031】

また本発明によれば、ストレージ仮想化技術を適用した計算機システムにおいて、外部ボリュームを複数のストレージシステムによって共有することができる。

【発明を実施するための最良の形態】

【0032】

以下、各図を参照しながら本発明の実施形態について説明する。まず、図1を参照しながら本実施形態に関わる論理ユニット移動処理の原理を説明する。本実施形態に関わる計算機システム500は、ホスト計算機100と、複数のストレージシステム130a, 130bと、外部ストレージシステム150とを備える。

【0033】

ストレージシステム130aは、一つ以上の論理ユニット(LU)201aを備える。この論理ユニット201aには、LUN(Logical Unit Number)がアサインされる。論理ユニット201aは、論理的な記憶領域として、ホスト計算機100に提供される。ホスト計算機100は、論理ユニット201aへのデータの書込みや、論理ユニット201aからのデータの読み出しを行うことができる。論理ユニット201aの下層には、複数の中間記憶階層(LDEV/VDEV/EDEV)が設けられている。この中間記憶階層は、論理ユニット201aと外部ボリューム301との間を対応つけるための論理的な記憶資源である。ホスト計算機100からのアクセスは、論理ユニット201aから中間記憶階層を介して外部ボリューム301に伝えられる。

【0034】

ストレージシステム130a内の中間記憶階層には、例えば、論理デバイス(LDEV)202a、仮想デバイス(VDEV)203a、及び拡張デバイス(EDEV)206aを含めることができる。これらの各デバイスは必ずしも必須ではなく、一部のデバイスを省略してもよい。また、これらの各デバイスに加えて他のデバイス(例えば、後述するデバイスグループ(DEVGr)など)を更に含めてもよい(図3参照)。

【0035】

尚、図3に示すように、ストレージシステム130が物理デバイス(PDEV)205を有する場合には、物理デバイス205と拡張デバイス206とは、共に同一の中間記憶階層に属する。

【0036】

説明の便宜上、論理ユニットが属する記憶階層を「LDEV階層」と称する場合がある。同様に、論理デバイスが属する記憶階層を「LDEV階層」と称し、仮想デバイスが属する記憶階層を「VDEV階層」と称し、物理デバイス及び/又は拡張デバイスが属する記憶階層を「PDEV/EDEV階層」と称する場合がある。

【0037】

ストレージシステム130aは、バス401を介してホスト計算機100に接続するための一つ以上のポート131aと、ポート131aに割り当てられる論理ユニット201aと、論理ユニット201aにマッピングされる論理デバイス202aと、論理デバイス202aにマッピングされる仮想デバイス203aと、仮想デバイス203aにマッピングされる拡張デバイス206aとを備えている。ここで、「マッピング」とは、あるデバイスのアドレス空間と、他のデバイスのアドレス空間とを対応付けることをいう。本明細

10

20

30

40

50

書において、あるデバイスを他のデバイスにマッピングすることは、あるデバイスを他のデバイスに対応付ける（又は割り当てる）ことと同義である。

【 0 0 3 8 】

ストレージシステム 1 3 0 a は、パス 4 0 2 を介して外部ストレージシステム 1 5 0 に接続している。外部ストレージシステム 1 5 0 は、ディスクドライブ等の実記憶領域を有する物理デバイス上に定義された外部ボリューム 3 0 1 を有する。

【 0 0 3 9 】

ここで、拡張デバイス 2 0 6 a は、ストレージシステム 1 3 0 a の外部に存在する記憶資源である外部ボリューム 3 0 1 を仮想化したものである。つまり、拡張デバイス 2 0 6 a は外部記憶資源を仮想化してなる記憶階層である。

10

【 0 0 4 0 】

仮想デバイス 2 0 3 a は、上位記憶階層（例えば、L U / L D E V）と下位記憶階層（例えば、D E V G r / E D E V）とを接続する記憶階層である。下位記憶階層が物理デバイスである場合には、仮想デバイス 2 0 3 a は、例えば、複数の物理デバイスのそれぞれが提供する記憶領域を R A I D 構成してなる記憶階層である。一方、下位記憶階層が拡張デバイス 2 0 6 a である場合には、仮想デバイス 2 0 3 a は、例えば、複数の拡張デバイス 2 0 6 a のそれぞれの記憶領域の全部又は一部を集合してなる記憶領域、或いは拡張デバイス 2 0 6 a の記憶領域の一部を抜き出してなる記憶領域である。

【 0 0 4 1 】

論理デバイス 2 0 2 a は、例えば、複数の仮想デバイス 2 0 3 a のそれぞれの記憶領域の全部又は一部を集合してなる記憶領域、或いは仮想デバイス 2 0 3 a の記憶領域の一部を抜き出してなる記憶領域である。ホスト計算機 1 0 0 がオープン系システムの場合、ホスト計算機 1 0 0 は、論理デバイス 2 0 2 a を一つの物理的なデバイスとして認識し、L U N や論理ブロックアドレスを指定することにより、所望の論理デバイス 2 0 2 a にアクセスする。ホスト計算機 1 0 0 がメインフレーム系システムの場合、ホスト計算機 1 0 0 は、論理デバイス 2 0 2 a を直接認識する。

20

【 0 0 4 2 】

論理ユニット 2 0 1 a は、外部ボリューム 3 0 1 を仮想化したものであって、且つホスト計算機 1 0 0 が認識する論理的な記憶領域である。例えば、ホスト計算機 1 0 0 が U N I X（登録商標）系システムである場合には、論理ユニット 2 0 1 a は、デバイスファイル（Device File）に対応付けられる。一方、ホスト計算機 1 0 0 が W i n d o w s（登録商標）系システムである場合には、論理ユニット 2 0 1 a は、ドライブレター（ドライブ名）に対応付けられる。論理ユニット 2 0 1 a には、固有の L U N（Logical Unit Number）がアサインされる。

30

【 0 0 4 3 】

論理ユニット移動処理を説明するための便宜上、論理ユニット 2 0 1 a と論理デバイス 2 0 2 a との対応関係は一對一とし、論理デバイス 2 0 2 a と仮想デバイス 2 0 3 a との対応関係は一對一とし、仮想デバイス 2 0 3 a と拡張デバイス 2 0 6 a との対応関係は一對一とし、拡張デバイス 2 0 6 a と外部ボリューム 3 0 1 との対応関係は一對一とする。かかる対応関係の下では、ホスト計算機 1 0 0 が認識する論理ユニット 2 0 1 a の記憶容量は、外部ボリューム 3 0 1 の記憶容量に等しい。

40

【 0 0 4 4 】

一方、ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 a と同様に階層化された複数の記憶階層を有している。ストレージシステム 1 3 0 b は、パス 4 0 3 を介してホスト計算機 1 0 0 に接続するための一つ以上のポート 1 3 1 b と、ポート 1 3 1 b に割り当てられる論理ユニット 2 0 1 b と、論理ユニット 2 0 1 b にマッピングされる論理デバイス 2 0 2 b と、論理デバイス 2 0 2 b にマッピングされる仮想デバイス 2 0 3 b と、仮想デバイス 2 0 3 b にマッピングされる拡張デバイス 2 0 6 c、2 0 6 b とを備えている。

【 0 0 4 5 】

50

論理ユニット移動処理を説明するための便宜上、論理ユニット201bと論理デバイス202bとの対応関係は一對一とし、論理デバイス202bと仮想デバイス203bとの対応関係は一對一とし、仮想デバイス203bと拡張デバイス206cとの対応関係は一對一とし、仮想デバイス203bと拡張デバイス206bとの対応関係は一對一とし、拡張デバイス206cと論理ユニット201aとの対応関係は一對一とし、拡張デバイス206bと外部ボリューム301との対応関係は一對一とする。また、拡張デバイス206cと拡張デバイス206bとは、同一の記憶容量を有しており、拡張デバイス206cと拡張デバイス206bとのうち何れか一方が択一的に仮想デバイス203bにマッピングされるものとする。

【0046】

次に、ストレージシステム130a(移動元ストレージ)内の論理ユニット201a(移動元論理ユニット)にあるデータをストレージシステム130b(移動先ストレージ)内の論理ユニット201b(移動先論理ユニット)に移動する処理について概説する。論理ユニット移動処理は、下記の(1)~(9)の処理手順を含む。

【0047】

(1)ストレージシステム130bは、外部ボリューム301と、ストレージシステム130bとの間を外部接続するためのパス405を定義し、外部ボリューム301を仮想化するための拡張デバイス206bをストレージシステム130b内に作成する。

【0048】

(2)ストレージシステム130bは、外部ボリューム301の記憶容量と同一の記憶容量を有する仮想デバイス203bをストレージシステム130b内に作成する。

【0049】

(3)ストレージシステム130bは、論理ユニット201aとストレージシステム130bとの間を外部接続するためのパス404を定義し、論理ユニット201aを仮想化するための拡張デバイス206cをストレージシステム130b内に作成する。

【0050】

(4)ストレージシステム130bは、拡張デバイス206cを仮想デバイス203bにマッピングする。

【0051】

(5)ストレージシステム130bは、論理デバイス202aの論理構成と同一の論理構成を有する論理デバイス202bをストレージシステム130b内に作成する。仮想デバイス203bは、論理デバイス202bにマッピングされる。

【0052】

(6)ストレージシステム130bは、論理ユニット201aの論理構成と同一の論理構成を有する論理ユニット201bをストレージシステム130b内に作成し、ホスト計算機100と論理ユニット201bとを接続するためのパス403を定義する。論理デバイス202bは、論理ユニット201bにマッピングされる。

【0053】

(7)ホスト計算機100は、外部ボリューム301にアクセスするための経路として、パス401からパス403に切り替える。このとき、パス401を削除する等して、ホスト計算機100から論理ユニット201aへのアクセスを禁止する。尚、パス401からパス403へのパス切り替えは、ストレージシステム130a, 130b, 又は管理サーバ110等が行っても良い。また、必ずしもパス401を削除する必要はなく、例えば、パス401をパス403の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス401の状態を無効としたまま、パス401を残しても良い。

【0054】

上記の処理により、論理ユニット201bは、論理ユニット201aを仮想化したものであるだけでなく、外部ボリューム301をも仮想化したものでもあるので、論理ユニット移動処理中における、ホスト計算機100から論理ユニット201aへのデータ入出力経路は、パス403 ストレージシステム130b内の記憶階層(LU/LDEV/VD

10

20

30

40

50

EV/EDEV) パス404 ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV) パス402 外部ボリューム301となる。

【0055】

尚、論理ユニット移動処理に参与するストレージシステム130a, 130bに暗号化機能が搭載されている場合、二つのストレージシステム130a, 130bの暗号化機能がオンに設定されていると、外部ボリューム301に書き込まれたデータを復号できなくなる場合がある。このような場合には、ストレージシステム130bの暗号化機能をオフに設定し、ストレージシステム130aの暗号化機能をオンに設定する。つまり、ストレージシステム130bによって、拡張デバイス206cに書き込まれるデータは暗号化されず、ストレージシステム130aによって拡張デバイス206aに書き込まれるデータは暗号化されない。暗号化機能をこのように設定することで、外部ボリューム301に書き込まれるデータが暗号化される回数を一回のみにすることができるので、外部ボリューム301に書き込まれたデータを復号することができる。

10

【0056】

(8) ストレージシステム130aは、ストレージシステム130aのディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム301にデステージするためにストレージシステム130aの動作モードをキャッシュスルーモード(ライトスルーモード)に設定する。ダーティデータとは、ディスクキャッシュから外部ボリューム301にデステージされていない未反映データをいう。キャッシュスルーモードとは、ホスト計算機100からのデータ書き込み要求を受けてデータをディスクキャッシュに格納し、更にデバイス(内部デバイス又は外部デバイス)にもそのデータを書き込んでから、ホスト計算機100にデータ更新完了報告をする動作をいう。尚、ストレージシステム130aは、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部(例えば、ストレージシステム130b、ホスト計算機100、又は後述する管理サーバ110など)からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

20

【0057】

(9) ストレージシステム130aのディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム301にデステージできたならば、ストレージシステム130bは、拡張デバイス206cと仮想デバイス203bとの間のマッピング関係を解除し、拡張デバイス206bを仮想デバイス203bにマッピングする。更に、ストレージシステム130bは、パス404を削除する。

30

【0058】

かかる方法によれば、ホスト計算機100から論理ユニット201aへのデータ入出力処理を継続させたまま論理ユニット201aを移動させることができる。これにより、ホスト計算機100からのI/O要求に起因する複数のストレージシステム130間の負荷変動の発生に応じて外部ボリュームへのI/O処理担当を複数のストレージシステム130a, 130b間で切り替えることで、計算機システム500の負荷分散を実現できる。しかも、ストレージシステム130a, 130bのうち移動元ストレージの動作モードをキャッシュスルーモードに設定し、移動先ストレージの動作モードをライトアフトモードに設定することによりディスクキャッシュを用いた非同期ライト処理が可能となるので、論理ユニット移動処理に伴うストレージシステム130aの性能低下を抑制できる。

40

【0059】

尚、移動先のストレージシステム130bの動作モードをキャッシュスルーモードに設定しても、論理ユニット201aを移動させることはできるが、移動元のストレージシステム130a内に蓄積されているダーティデータの全てを外部ボリューム301にデステージしない限り、論理ユニット201aの移動処理は完了しない。そこで、論理ユニット201aの移動処理時間を短くするには、移動元のストレージシステム130aの動作モードをキャッシュスルーモードに設定するのが好ましい。

【0060】

また、図1に示すシステム構成では、それぞれのストレージシステム130a, 130

50

bは、外部ボリューム301を仮想化してなる拡張デバイス206を有するが、内部デバイス(物理デバイス)を有しない。本発明に関わる論理ユニット移動処理は、図1に示すようなシステム構成への適用に限られるものではなく、例えば、外部ボリューム仮想化機能を有しないストレージシステム130a, 130bが内蔵するディスク装置等の物理デバイスによって提供される論理ユニットを移動させる処理にも適用できる。

【0061】

上記の論理ユニット移動処理の説明においては、説明の便宜上、それぞれの中間記憶階層相互間のマッピング関係を一対一とする場合を例示したが、それぞれの中間記憶階層相互間のマッピング関係は必ずしも一対一の関係に限られるものではない。例えば、それぞれの中間記憶階層相互間のマッピング関係として、例えば、CVS (Custom Volume Size)、LUSE (LU Size Expansion)、外部ボリュームRAID、VMA (Volume Management Area) 付きボリューム、メインフレームボリューム、DEV連結、DEV離散、AOU (Allocation on Use)、スナップショットなどの各種のバリエーションが考えられる。詳細については、後述の実施例において説明する(図4乃至図22参照)。

【実施例】

【0062】

図2は本実施例に関わる計算機システム500のネットワーク構成を示す。

計算機システム500は、複数のホスト計算機100a, 100b、一つ以上の管理サーバ110、一つ以上のファイバチャネルスイッチ120、複数のストレージシステム130a, 130b、複数の管理端末140a, 140b、及び複数の外部ストレージシステム150a, 150bを備える。以下の説明では、ホスト計算機100a, 100bを区別しないときには、ホスト計算機100と総称する。ストレージシステム130a, 130bを区別しないときには、ストレージシステム130と総称する。管理端末140a, 140bを区別しないときには、管理端末140と総称する。外部ストレージシステム150a, 150bを区別しないときには、ストレージシステム150と総称する。

【0063】

ホスト計算機100、ストレージシステム130、及び外部ストレージシステム150は、それぞれポート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、ホスト計算機100、ストレージシステム130、外部ストレージシステム150、及びファイバチャネルスイッチ120は、それぞれインタフェース制御部106, 138, 157, 123から、IPネットワーク175を介して管理サーバ110に接続される。管理サーバ110は、これらのネットワークノード(ホスト計算機100、ストレージシステム130、外部ストレージシステム150、及びファイバチャネルスイッチ120)を管理する。

【0064】

尚、本実施例では、ストレージシステム130は、管理端末140を介して管理サーバ110に接続するネットワーク構成を有するものとするが、ストレージシステム130がIPネットワーク175に直接接続するネットワーク構成であってもよい。

【0065】

ホスト計算機100は、CPU101、メモリ102、記憶装置103、入力装置104、ディスプレイ105、インタフェース制御部106、及びポート107を有する。記憶装置103は、例えば、ディスク装置、光磁気ディスク装置等である。ホスト計算機100は、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によりそのソフトウェアを読み出して実行させることで、情報処理を実行する。入力装置104は、キーボード又はマウス等である。ホスト計算機100は、入出力装置104を介してホスト管理者などからの入力を受け付け、ディスプレイ105に情報処理結果等を表示する。インタフェース制御部106は、IPネットワーク175に接続するためのLANアダプタ等である。ポート107は、ストレージシステム130に接続するためのホストバスアダプタ等である。

10

20

30

40

50

【 0 0 6 6 】

管理サーバ 1 1 0 は、CPU 1 1 1、メモリ 1 1 2、記憶装置 1 1 3、入力装置 1 1 4、ディスプレイ 1 1 5、及びインタフェース制御部 1 1 6 を有する。記憶装置 1 1 3 は、例えば、ディスク装置、光磁気ディスク装置等である。管理サーバ 1 1 0 は、記憶装置 1 1 3 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 1 1 2 に読み上げ、CPU 1 1 1 によりそのストレージ管理ソフトウェアを読み出して実行することにより、計算機システム 5 0 0 全体を保守管理する。CPU 1 1 1 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 1 1 0 は、インタフェース制御部 1 1 6 から IP ネットワーク 1 7 5 を介して、計算機システム 5 0 0 内の各ネットワークノードから構成情報、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 1 1 0 は、収集したそれらの情報をディスプレイ 1 1 5 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 1 1 0 は、キーボードやマウス等の入力装置 1 1 4 を介してストレージ管理者から保守管理に関する指示を受け付け、その指示をネットワークノードに送信する。

10

【 0 0 6 7 】

ストレージシステム 1 3 0 は、ストレージネットワークに接続するための一つ以上のポート 1 3 1、ディスク装置 1 3 7 へのデータ入出力処理を制御するための一つ以上の制御プロセッサ 1 3 2、それぞれの制御プロセッサ 1 3 2 のワークエリアとして機能する一つ以上のメモリ 1 3 3、ディスク装置 1 3 7 に入出力されるデータを一時的に格納するための一つ以上のディスクキャッシュ 1 3 4、ストレージシステム 1 3 0 の構成情報等を格納するための一つ以上の制御メモリ 1 3 5、ディスク装置 1 3 7 に接続するためのインタフェースとして機能する一つ以上のポート 1 3 6、データを格納するためのディスク装置 1 3 7 を有する。

20

【 0 0 6 8 】

制御プロセッサ 1 3 2 は、ポート 1 3 1 を介してホスト計算機 1 0 0 から受信した入出力要求に含まれる情報（ポート ID、及び LUN）に基づいて、アクセス先の論理ユニットを特定する。アクセス先の論理ユニットがディスク装置 1 3 7（内部デバイス）に対応しているならば、制御プロセッサ 1 3 2 は、ディスク装置 1 3 7 へのデータ入出力を制御する。アクセス先の論理ユニットが外部ストレージシステム 1 5 0（外部デバイス）に対応しているならば、制御プロセッサ 1 3 2 は、外部ストレージシステム 1 5 0 へのデータ入出力を制御する。

30

【 0 0 6 9 】

尚、本実施例では、ポート 1 3 1 として、SCSI（Small Computer System Interface）を上位プロトコルとするファイバチャネルインタフェースに対応するポートを想定しているが、ポート 1 3 1 は、SCSI を上位プロトコルとする IP ネットワークインタフェースに対応するポートであってもよい。

【 0 0 7 0 】

ストレージシステム 1 3 0 は、次のような記憶階層を有する。複数のディスク装置 1 3 7 によるディスクアレイが構成され、このディスクアレイは、制御プロセッサ 1 3 2 によって物理デバイスとして管理される。更に、制御プロセッサ 1 3 2 は、ストレージシステム 1 3 0 内に搭載した物理デバイスに論理デバイスを割り当てる（つまり、制御プロセッサ 1 3 2 は、物理デバイスと論理デバイスとを対応付ける。）。論理デバイスは、ストレージシステム 1 3 0 内で管理され、その LDEV 番号は、ストレージシステム 1 3 0 毎に独立に管理される。

40

【 0 0 7 1 】

論理デバイスは、各ポート 1 3 1 に割り当てられた LUN に対応付けられ、ストレージシステム 1 3 0 内のデバイスとしてホスト計算機 1 0 0 に提供される。即ち、ホスト計算機 1 0 0 が認識するのは、ストレージシステム 1 3 0 内の論理デバイスであり、ホスト計算機 1 0 0 は、ポート 1 3 1 に割り当てられた論理デバイスを識別するための LUN を用いて、ストレージシステム 1 3 0 に格納されているデータにアクセスする。

50

【 0 0 7 2 】

尚、本実施例では、制御プロセッサ 1 3 2 は、外部ストレージ 1 5 0 内の論理デバイスを外部デバイスとして管理し、ストレージシステム 1 3 0 内のデバイス（内部デバイス）として仮想化する。つまり、ストレージシステム 1 3 0 によって管理される外部デバイスは、仮想デバイスである。外部ストレージ接続技術を用いて、ストレージシステム 1 3 0 内に取り込まれた外部デバイスは、ストレージシステム 1 3 0 内の物理デバイスと同様にストレージシステム 1 3 0 内の論理デバイスとして管理される。

【 0 0 7 3 】

以上のような記憶階層を実現するため、制御プロセッサ 1 3 2 は、ストレージシステム 1 3 0 内における論理デバイス、物理デバイス、及びディスク装置 1 3 7 と、外部ストレージシステム 1 5 0 内における論理デバイス、及びディスク装置 1 5 6 との対応関係を管理する。そして、制御プロセッサ 1 3 2 は、ホスト計算機 1 0 0 からの I / O 要求を、ストレージシステム 1 3 0 内のディスク装置 1 3 7 へのアクセス要求、又は外部ストレージシステム 1 5 0 内のディスク装置 1 5 6 へのアクセス要求に変換する。

10

【 0 0 7 4 】

尚、前述の通り、本実施例におけるストレージシステム 1 3 0 は、複数のディスク装置 1 3 7 の記憶領域をまとめて一つ以上の物理デバイスを定義し（つまり、複数のディスク装置 1 3 7 の記憶領域をまとめて一つ以上の物理デバイスに対応付ける。）、一つの物理デバイスに一つの論理デバイスを割り当てて、その記憶領域をホスト計算機 1 0 0 に提供する。個々のディスク装置 1 3 7 をそれぞれ一つの物理デバイスに対応付けてもよい。

20

【 0 0 7 5 】

制御プロセッサ 1 3 2 は、デバイスへのデータ入出力処理だけでなく、ボリューム間のデータ複製やデータ再配置などを制御する。制御プロセッサ 1 3 2 は、インタフェース制御部 1 3 8 を介して管理端末 1 4 0 に接続しており、ストレージ管理者によって管理端末 1 4 0 に入力された構成変更指示を受信し、ストレージシステム 1 3 0 の構成変更を行うこともできる。

【 0 0 7 6 】

ディスクキャッシュ 1 3 4 は、ホスト計算機 1 0 0 とディスク装置 1 3 7 との間で入出力されるデータを一時的に格納する。ストレージシステム 1 3 0 がホスト計算機 1 0 0 から受信したデータをディスクキャッシュ 1 3 4 に格納し、ライト完了報告をホスト計算機 1 0 0 に返信してから、その後、データをディスク装置 1 3 7 にデステージする処理は、ライトアフタ処理と称される。

30

【 0 0 7 7 】

ストレージシステム 1 3 0 がライトアフタ処理を行うように構成されている場合、ディスクキャッシュ 1 3 4 に格納されているデータがディスク装置 1 3 7 に書き込まれる前に消失することを防止するために、ディスクキャッシュ 1 3 4 を電源バックアップしたり、或いはディスクキャッシュ 1 3 4 を二重化したりする等して、耐障害性を向上させるのが好ましい。

【 0 0 7 8 】

制御メモリ 1 3 5 には、ストレージシステム 1 3 0 の構成情報（例えば、デバイス間の対応関係を管理するためのマッピング情報、キャッシュスルーモードの設定情報など）が格納されている。制御メモリ 1 3 5 に格納されている構成情報が消失すると、制御プロセッサ 1 3 2 は、ディスク装置 1 3 7 へのアクセスできなくなるので、制御メモリ 1 3 5 を電源バックアップしたり、或いは制御メモリ 1 3 5 を二重化したりする等して、耐障害性を向上させるのが好ましい。

40

【 0 0 7 9 】

管理端末 1 4 0 は、インタフェース制御部 1 4 1、CPU 1 4 2、メモリ 1 4 3、記憶装置 1 4 4、入力装置 1 4 5、ディスプレイ 1 4 6、及びインタフェース制御部 1 4 7 を有する。インタフェース制御部 1 4 1 は、ストレージシステム 1 3 0 に接続するための通信インタフェースである。記憶装置 1 4 4 は、例えば、ディスク装置又は光磁気ディスク

50

装置等である。入力装置 145 は、キーボード又はマウス等である。ディスプレイ 146 は、ストレージ管理用のユーザインターフェース環境を提供するものであり、例えば、ストレージシステム 130 の構成情報や管理情報などを表示する。CPU 142 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 に読み出して実行することにより、構成情報の参照、構成変更の変更、特定機能の動作指示などを行う。

【0080】

尚、本実例では、管理端末 140 を介さずに、ストレージシステム 130 を管理サーバ 110 に直接接続し、管理サーバ 110 上で動作する管理ソフトウェアによって、ストレージシステム 130 を管理するネットワーク構成を採用してもよい。

【0081】

外部ストレージシステム 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 に接続するための一つ以上のポート 151、ディスク装置 156 へのデータ入出力処理を制御するための一つ以上の制御プロセッサ 152、それぞれの制御プロセッサ 152 のワークエリアとして機能する一つ以上のメモリ 153、ディスク装置 156 に入出力されるデータを一時的に格納するための一つ以上のディスクキャッシュ 154、制御プロセッサ 152 がディスク装置 156 に接続するためのインタフェースとして機能する一つ以上のポート 155、データを格納するための一つ以上のディスクドライブ 156、IP ネットワークに接続するためのインタフェースとして機能するインタフェース制御部 157 を有する。

【0082】

尚、本実施例では、ストレージシステム 130 に装備されているポート 131 と、外部ストレージシステム 150 に装備されているポート 151 とをファイバチャネルスイッチ 120 を介して接続するネットワーク構成を採用しているため、ホスト計算機 100 からの外部ストレージシステム 150 への直接的なアクセスを抑制するために、ファイバチャネルスイッチ 120 にゾーニングを設定することが望ましい。ストレージネットワーク上に必ずしもファイバチャネルスイッチ 120 を設置する必要はなく、例えば、ストレージシステムに装備されているポート 131 と、外部ストレージシステム 150 に装備されているポート 151 とを直接接続するようなネットワーク構成を採用してもよい。

【0083】

図 3 は本実施例に関わるストレージシステム 130 内に構築される記憶階層を示す。図 1 乃至図 2 に示す符号と同一符号の論理資源又は物理資源は、同一のものを示すものとしてその詳細な説明を省略する。

【0084】

ホスト計算機 100 は、リンクマネージャ 108、及び RAID マネージャ 109 を備える。リンクマネージャ 108 は、ストレージシステム 130 へのアクセス経路を切り替えるためのソフトウェアプログラムである。RAID マネージャ 109 は、ストレージシステム 130 の記憶資源を管理するためのストレージ管理ソフトウェアであり、ストレージシステム 130 に各種の指示（例えば、論理ユニット移動指示）を与える。

【0085】

ストレージシステム 130 に装備されているあるポート 131 は、ストレージネットワーク 510 を介してホスト計算機 100 に接続する一方、他のポート 131 は、ストレージネットワーク 510 を介して外部ストレージシステム 150 に接続している。ストレージネットワーク 510 は、一つ以上のファイバチャネルスイッチ 120 を含む FC - SAN (Fibre Channel Storage Area Network) 等のネットワークである。

【0086】

ストレージシステム 130 は、論理ユニット 201、論理デバイス 202、仮想デバイス 203、デバイスグループ 204、物理デバイス 205、及び拡張デバイス 206 から成る記憶階層を有する。デバイスグループ 204 は、例えば、複数の物理デバイス 205 又は拡張デバイス 206 のうち何れか一方又は両者のそれぞれの記憶領域の全部又は一部を集合してなる記憶領域、或いは物理デバイス 205 又は拡張デバイス 206 のうち何れ

10

20

30

40

50

か一方又は両者の記憶領域の一部を抜き出してなる記憶領域である。

【 0 0 8 7 】

尚、物理デバイス 2 0 5 は、ストレージシステム 1 3 0 が有するディスク装置 1 3 7 の記憶領域（内部デバイス）を仮想化したものであるのに対し、拡張デバイス 2 0 6 は、外部ストレージシステム 1 5 0 が有するディスク装置 1 5 6 の記憶領域（外部デバイス）を仮想化したものである点で異なるが、両者は共に実デバイスを仮想化したものである点では共通なので、同一階層（最下層）に属している。

【 0 0 8 8 】

ストレージシステム 1 3 0 は、ポート 1 3 1 に割り当てられる C M D（Command Device）2 0 7 を備える。C M D 2 0 7 は、ホスト計算機 1 0 0 とストレージシステム 1 3 0 との間でコマンドやステータスを受け渡すために用いられる論理ユニットである。ホスト計算機 1 0 0 からストレージシステム 1 3 0 に送信されるコマンドは、C M D 2 0 7 に書き込まれる。ストレージシステム 1 3 0 は、C M D 2 0 7 に書き込まれるコマンドに対応する処理を実行し、その実行結果をステータスとして C M D 2 0 7 に書き込む。ホスト計算機 1 0 0 は、C M D 2 0 7 に書き込まれるステータスを読み出して確認し、次に実行すべきコマンドを C M D 2 0 7 に書き込む。このようにして、ホスト計算機 1 0 0 は、ストレージシステム 1 3 0 に各種の指示を与えることができる。

【 0 0 8 9 】

ホスト計算機 1 0 0 がストレージシステム 1 3 0 に与える指示内容としては、例えば、論理ユニット単位でのデバイス移動指示（論理ユニット移動指示）又は論理デバイス単位でのデバイス移動指示（論理デバイス移動指示）がある。論理ユニット移動処理の概要については、上述した通りである。移動元の論理ユニット 2 0 1 が L U S E 構成を有している場合には、ホスト計算機 1 0 0 はストレージシステム 1 3 0 に論理デバイス移動指示を与えることができる。L U S E 構成を有する論理ユニット 2 0 1 では、複数の論理デバイス 2 0 2 が一つの論理ユニット 2 0 1 にマッピングされているので（図 5 参照）、複数の論理デバイス 2 0 2 のうち一つの論理デバイス 2 0 2（例えば、論理ユニット 2 0 1 の先頭アドレスに割り当てられている論理デバイス 2 0 2）を指定することで、その論理ユニット 2 0 1 にマッピングされている他の全ての論理デバイス 2 0 2 も移動対象として指定することもできる。

【 0 0 9 0 】

尚、ストレージシステム 1 3 0 は、所定の条件（例えば、ストレージシステム 1 3 0 に障害が発生した場合、又はストレージシステム 1 3 0 の性能が低下した場合など）が満たされたことを契機として、自発的に（ホスト計算機 1 0 0 からの指示を待つことなく）、論理ユニット移動処理を実行してもよい。ストレージシステム 1 3 0 に障害が発生した場合又はストレージシステム 1 3 0 の性能が低下した場合を契機として、論理ユニットを移動させる処理の概要については、後述する。

【 0 0 9 1 】

また、管理サーバ 1 1 0 が I P ネットワーク 1 7 5 及び管理端末 1 4 0 を経由して、ストレージシステム 1 3 0 に論理ユニット移動指示を与えてもよい。つまり、ホスト計算機 1 0 0 からインバンド経由（ストレージネットワーク経由）でストレージシステム 1 3 0 に論理ユニット移動指示を与えてもよく、或いは管理サーバ 1 1 0 からアウトバンド経由（I P ネットワーク経由）でストレージシステム 1 3 0 に論理ユニット移動指示を与えてもよい。

【 0 0 9 2 】

次に、図 4 乃至図 2 2 を参照しながら各記憶階層相互間のマッピング関係のバリエーションについて説明を加える。

【 0 0 9 3 】

(1) C V S

図 4 に示すように一つの仮想デバイス 2 0 3 を複数の V D E V 領域 2 0 3 - 1 , 2 0 3 - 2 , 2 0 3 - 3 に分割し、それぞれの V D E V 領域 2 0 3 - 1 , 2 0 3 - 2 , 2 0 3 -

10

20

30

40

50

3 に一つの論理デバイス 202 を割り当てるマッピング関係を C V S と称する。それぞれの V D E V 領域 203 - 1 , 203 - 2 , 203 - 3 のサイズは、必ずしも同一である必要はない。

【 0094 】

(2) L U S E

図 5 に示すように複数の論理デバイス 202 の記憶領域を集めて一つの論理ユニット 201 に割り当てるマッピング関係を L U S E と称する。

【 0095 】

(3) 外部ボリューム R A I D

図 6 に示すように R A I D 構成された拡張デバイス 206 をデバイスグループ 204 に割り当てるマッピング関係を外部ボリューム R A I D と称する。例えば、R A I D 1 によって R A I D 構成された複数の拡張デバイス 206 のうち一方の拡張デバイス 206 は、データディスクに対応し、他方の拡張デバイス 206 は、ミラーディスクに対応する。

【 0096 】

(4) V M A 付きボリューム

図 7 は V M A 付きボリュームを示す。V M A 付きボリュームは、オープンシステム系のシステムに用いられる論理デバイス 202 である。この論理デバイス 202 は、ユーザデータを格納するためのユーザ領域 202 - 1 と、管理情報を格納するための管理領域 202 - 2 とを含む。管理情報には、例えば、V M A 付きボリュームのアクセス属性が含まれる。アクセス属性は、リード/ライト可能、リードオンリ、リード/ライト不可、リードキャパシティゼロ、インクエリ (Inquiry) 抑止、セカンダリボリュームディセーブル (S -VOL disable) などの情報を含む。ユーザ領域 202 - 1 は、ホスト計算機 100 が認識することが可能であり、しかも論理ユニット 201 に割り当てられるのに対して、管理領域 202 - 2 は、ホスト計算機 100 が認識することはできず、しかも論理ユニット 201 に割り当てられることもない。

【 0097 】

(5) メインフレームボリューム

図 8 はメインフレームボリュームを示す。メインフレームボリュームは、メインフレーム系のシステムに用いられる論理デバイス 202 である。この論理デバイス 202 は、ユーザデータを格納するためのユーザ領域 202 - 1 と、管理情報を格納するための管理領域 202 - 3 とを含む。管理情報には、例えば、メインフレームボリュームのアクセス属性などが含まれる。ユーザ領域 202 - 1 は、ホスト計算機 100 が認識することが可能であり、しかも論理ユニット 201 に割り当てられるのに対して、管理領域 202 - 3 は、ホスト計算機 100 が認識することはできず、しかも論理ユニット 201 に割り当てられることもない。

【 0098 】

(6) V D E V 連結

図 9 に示すように複数の仮想デバイス 203 の記憶領域を集めて一つの論理デバイス 202 に割り当てるマッピング関係を V D E V 連結と称する。

【 0099 】

(7) V D E V 離散 (基本形)

図 10 に示すように複数のデバイスグループ 204 の記憶領域を集めて一つの仮想デバイス 203 に割り当てるマッピング関係を V D E V 離散と称する。V D E V 離散のマッピング関係には、後述する変形例 1 ~ 2 が含まれる。

【 0100 】

(8) V D E V 離散 (変形例 1)

V D E V 離散の一つの変形例として、例えば、図 11 に示すように一つのデバイスグループ 204 を複数の記憶領域に分割し、分割されたそれぞれの記憶領域に一つの仮想デバイス 203 を割り当ててもよい。

【 0101 】

10

20

30

40

50

(9) V D E V 離散 (変形例 2)

V D E V 離散の他の一つの変形例として、例えば、図 1 2 に示すように複数のデバイスグループ 2 0 4 のそれぞれを複数の記憶領域に分割し、分割されたそれぞれの記憶領域を複数の仮想デバイス 2 0 3 に割り当ててもよい。

【 0 1 0 2 】

(1 0) D E V G r / E D E V

デバイスグループ 2 0 4 と拡張デバイス 2 0 6 との間のマッピング関係については、例えば、図 1 3 に示すように、複数の拡張デバイス 2 0 6 のそれぞれの記憶領域の全部又は一部を集めて一つのデバイスグループ 2 0 4 に割り当ててもよい。拡張デバイス 2 0 6 の中には、デバイスグループ 2 0 4 に割り当てられない記憶領域を含むものが存在してもよい。また例えば、図 1 4 に示すように、複数の拡張デバイス 2 0 6 のそれぞれの記憶領域の全部又は一部を集めて複数のデバイスグループ 2 0 4 に割り当ててもよい。拡張デバイス 2 0 6 の中には、何れのデバイスグループ 2 0 4 にも割り当てられない記憶領域を含むものが存在してもよい。また例えば、図 1 5 に示すように、複数のデバイスグループ 2 0 4 に割り当てられる拡張デバイス 2 0 6 や、一部のデバイスグループ 2 0 4 にのみ割り当てられる拡張デバイス 2 0 6 が存在してもよい。

【 0 1 0 3 】

(1 1) 異種メディア混載 R A I D

図 1 6 に示すように、一つのデバイスグループ 2 0 4 の記憶領域に異種デバイス (例えば、物理デバイス 2 0 5 及び拡張デバイス 2 0 6) が混在した形態で割り当ててもよい。また例えば、図 1 7 に示すように、一つのデバイスグループ 2 0 4 の記憶領域に異種ディスクドライブ (例えば、F C ディスクドライブ 5 2 0 及び S A T A ディスクドライブ 5 3 0) が混在した形態で割り当ててもよい。また例えば、図 1 8 乃至図 2 0 に示すように一つのデバイスグループ 2 0 4 の記憶領域にフラッシュメモリ 5 4 0 及びディスクドライブ 5 5 0 が混在した形態で割り当ててもよい。

【 0 1 0 4 】

図 1 8 では、複数のフラッシュメモリ 5 4 0 と、複数のディスクドライブ 5 5 0 とによって R A I D 0 / 1 構成のデバイスグループ 2 0 4 が構成されている。複数のフラッシュメモリ 5 4 0 によってデータボリュームが構成され、複数のディスクドライブ 5 5 0 によってミラーボリュームが構成されている。このようなドライブ構成においては、ランダムリード時には、フラッシュメモリ 5 4 0 からデータが読み出される。シーケンシャルリード時には、フラッシュメモリ 5 4 0 とディスクドライブ 5 5 0 のそれぞれからデータが読み出される。ライト時には、フラッシュメモリ 5 4 0 とディスクドライブ 5 5 0 のそれぞれにデータが書き込まれる。このようなドライブ構成によればフラッシュメモリ 5 4 0 の高速性を活かしつつ、フラッシュメモリ 5 4 0 のみの R A I D よりも低コストである。

【 0 1 0 5 】

図 1 9 では、三つのフラッシュメモリ 5 4 0 と一つのディスクドライブ 5 5 0 とによって R A I D 4 構成のデバイスグループ 2 0 4 が構成されている。フラッシュメモリ 5 4 0 は、データを格納し、ディスクドライブ 5 5 0 は、パリティデータを格納する。このようなドライブ構成においては、書き込み寿命のあるフラッシュメモリ 5 4 0 を、アクセス数の多いパリティドライブに使用しないことにより、フラッシュメモリ 5 4 0 の使用期間を延ばすことができる。

【 0 1 0 6 】

図 2 0 では、一つのフラッシュメモリ 5 4 0 と三つのディスクドライブ 5 5 0 とによって R A I D 4 構成のデバイスグループ 2 0 4 が構成されている。フラッシュメモリ 5 4 0 は、データを格納し、ディスクドライブ 5 5 0 は、パリティデータを格納する。このようなドライブ構成においては、アクセス数の多いパリティドライブとして、フラッシュメモリ 5 4 0 を使用することにより、書き込み性能を向上させることができる。

【 0 1 0 7 】

(1 2) A O U

10

20

30

40

50

図 2 1 はストレージシステム 1 3 0 内における、A O U と称される記憶階層相互間のマッピング関係を示す。L U 階層には、論理ユニット 2 0 1 が属している。L D E V 階層には、A O U ボリューム 2 0 2 c と複数のプールボリューム 2 0 2 d が属している。V D E V 階層には、複数のプールボリューム 2 0 3 c が属している。P D E V / E D E V 階層には、複数の拡張デバイス 2 0 6 e が属している。ホスト計算機 1 0 0 が認識する論理ユニット 2 0 1 には、A O U ボリューム 2 0 2 c が割り当てられている。A O U ボリューム 2 0 2 c には、複数のプールボリューム 2 0 3 c のそれぞれの記憶領域の一部が割り当てられている。A O U ボリューム 2 0 2 c には、プールボリューム 2 0 3 c の記憶領域を動的に割り当てることが可能であり、A O U ボリューム 2 0 2 c の容量を自在に増減することができる。複数のプールボリューム 2 0 3 c のそれぞれの記憶領域には、複数の拡張デバイス 2 0 6 e のそれぞれの記憶領域が割り当てられている。尚、P D E V / E D E V 階層には、拡張デバイス 2 0 6 e に替えて物理デバイスを配置してもよい。

10

【 0 1 0 8 】

(1 3) スナップショット

図 2 2 はスナップショット機能を備えるストレージシステム 1 3 0 内における記憶階層相互間のマッピング関係を示す。L U 階層には、論理ユニット 2 0 1 が属している。L D E V 階層には、プライマリボリューム 2 0 2 e 、セカンダリボリューム 2 0 2 f 、複数のプールボリューム 2 0 2 g が属している。V D E V 階層には、仮想デバイス 2 0 3 d 、複数のプールボリューム 2 0 3 e が属している。P D E V / E D E V 階層には、拡張デバイス 2 0 6 f , 2 0 6 g が属している。論理ユニット 2 0 1 には、論理デバイス 2 0 2 e が割り当てられている。論理デバイス 2 0 2 e には、仮想デバイス 2 0 3 d が割り当てられている。仮想デバイス 2 0 3 d には、拡張デバイス 2 0 6 f が割り当てられている。セカンダリボリューム 2 0 2 f は、ある時刻にプライマリボリューム 2 0 2 e に格納されているデータと、ある時刻以降にプライマリボリューム 2 0 2 e からプールボリューム 2 0 3 e に退避されたデータとから、ある時刻におけるプライマリボリューム 2 0 2 e のデータイメージを復元するための仮想的な論理ボリュームである。

20

【 0 1 0 9 】

次に、図 2 3 を参照しながらホストグループについて説明を加える。ホストグループ番号 # 0 のホストグループに属する複数のホスト計算機 1 0 0 a , 1 0 0 c は、何れも、ポート 1 3 1 の番号としてポート番号 # 0 、及び論理ユニット 2 0 1 の番号として L U N # 0 を指定することにより、論理デバイス番号 # 0 の論理デバイス 2 0 2 h にアクセスできる。一方、ホストグループ番号 # 1 のホストグループに属するホスト計算機 1 0 0 b は、ポート 1 3 1 の番号としてポート番号 # 0 、及び論理ユニット 2 0 1 の番号として L U N # 0 を指定することにより、論理デバイス番号 # 1 の論理デバイス 2 0 2 i にアクセスできる。このように、あるホスト計算機 1 0 0 が同一のポート番号及び同一の論理ユニット番号を指定したとしても、そのホスト計算機 1 0 0 が所属するホストグループ番号によっては、アクセスできる論理デバイス 2 0 2 が異なる。

30

【 0 1 1 0 】

本実施形態に関わる論理ユニット移動処理においては、あるホストグループに属する一つのホスト計算機 1 0 0 が利用可能な論理ユニット 2 0 1 を他のストレージシステム 1 3 0 に移動させる場合には、そのホストグループに属する他のホスト計算機 1 0 0 についても同様の論理ユニット移動処理を実行させるものとする。

40

【 0 1 1 1 】

次に、図 2 4 乃至図 2 5 を参照しながら交替パスについて説明を加える。図 2 4 に示すようにホスト計算機 1 0 0 からストレージシステム 1 3 0 内のある論理デバイス 2 0 2 j に接続するための複数のパス 4 0 6 , 4 0 7 が存在する場合にその複数のパス 4 0 6 , 4 0 7 のうち何れか一つのパス 4 0 6 を常用系パスとすると、他のパス 4 0 7 は交替パスとなる。交替パスは、待機系パスと称されることもある。ホスト計算機 1 0 0 は、ポート 1 3 1 の番号としてポート番号 # 0 、及び論理ユニット 2 0 1 c の番号として L U N # 0 を指定することにより、パス 4 0 6 を介して論理デバイス 2 0 2 j にアクセスできる。また

50

ホスト計算機 100 は、ポート 131 の番号としてポート番号 # 1、及び論理ユニット 201 d の番号として LUN # 0 を指定することにより、パス 407 を介して論理デバイス 202 j にアクセスできる。リンクマネージャ 108 は、論理デバイス 202 j にアクセスするためのパスとして、パス 406、407 のうち何れか一方を選択する処理を実行する。論理ユニット 201 c、201 d は、何れも同一の論理デバイス 202 j を仮想化したものであるため、一方の論理ユニット（例えば、論理ユニット 201 c）を他のストレージシステム 130 に移動させる場合には、他方の論理ユニット（例えば、論理ユニット 201 d）も他のストレージシステム 130 に移動させる必要がある。

【0112】

図 25 に示すようにストレージシステム 130 から外部ストレージシステム 150 内のある論理デバイス 202 k に接続するための複数のパス 408、409 が存在する場合にその複数のパス 408、409 のうち何れか一つのパス 408 を常用系パスとすると、他のパス 409 は交替パスとなる。ストレージシステム 130 は、ポート 151 の番号としてポート番号 # 0、及び論理ユニット 201 e の番号として LUN # 0 を指定することにより、パス 408 を介して論理デバイス 202 k にアクセスできる。また、ストレージシステム 130 は、ポート 151 の番号としてポート番号 # 1、及び論理ユニット 201 f の番号として LUN # 0 を指定することにより、パス 409 を介して論理デバイス 202 k にアクセスできる。論理ユニット 201 e、201 f は、何れも同一の論理デバイス 202 k を仮想化したものであるため、一方の論理ユニット（例えば、論理ユニット 201 e）を他のストレージシステム 130 に移動させる場合には、他方の論理ユニット（例

【0113】

次に、図 26 乃至図 33 を参照しながら論理ユニット移動処理に用いられる各種の情報について説明を加える。

【0114】

図 26 はホストグループ管理情報 601 を示す。ホストグループ管理情報 601 は、ホストグループに関する情報を保持する。ホストグループ管理情報 601 は、エントリ 701 からエントリ 705 までの情報を保持する。

【0115】

エントリ 701 には、計算機システム 500 内でホストグループを識別するための識別情報として、ホストグループ番号が格納される。

【0116】

エントリ 702 には、ホストグループ番号により特定されるホストグループに属するホスト計算機 100 の台数が格納される。

【0117】

エントリ 703 には、ホストグループ番号により特定されるホストグループに属する全てのホスト計算機 100 のホスト名のリストが格納される。ホスト名としては、ホスト計算機 100 を一意に識別できる情報であれば、特に限定されるものではなく、例えば、ポート 107 にアサインされた WWN (World Wide Name) を用いることができる。

【0118】

エントリ 704 には、ホストグループ番号により特定されるホストグループに属するホスト計算機 100 からアクセス可能な論理ユニットの個数が格納される。

【0119】

エントリ 705 には、後述する LU パス管理情報 602 がリスト形式で格納される。

【0120】

図 27 は LU パス管理情報 602 を示す。LU パス管理情報 602 は、ストレージシステム 130 内の各ポート 131 について、各ポート 130 に定義されている有効な LUN の個数分の情報を保持する。LU パス管理情報 602 は、エントリ 801 からエントリ 805 までの情報を保持する。

【0121】

10

20

30

40

50

エントリ 801 には、ストレージシステム 130 内のポート 131 を一意に識別するためのポート番号が格納される。

【0122】

エントリ 802 には、論理ユニット 201 を識別するための情報（例えば、ターゲット ID 又は LUN）が格納される。ターゲット ID 及び LUN は、論理ユニット 201 を識別するための識別子である。例えば、論理ユニット 201 を識別するための識別子として、ホスト計算機 100 から SCSI プロトコル上でデバイスをアクセスする場合に用いられる SCSI - ID と LUN とが用いられる。

【0123】

エントリ 803 には、エントリ 802 に格納されている LUN が割り当てられる論理デバイスの番号（LDEV 番号）が格納される。論理ユニット 201 が LUSE 構成を有している場合における、論理ユニット 201 に割り当てられている複数の論理デバイス 202 のうち代表的な論理デバイス 202 の番号が格納される。

10

【0124】

エントリ 804 には、エントリ 803 に格納されている LDEV 番号により特定される論理デバイスにアクセス可能なホスト計算機 100 の台数が格納される。

【0125】

エントリ 805 には、エントリ 803 に格納されている LDEV 番号により特定される論理デバイスにアクセス可能な全てのホスト計算機 100 のホスト名のリストが格納される。

20

【0126】

図 28 は論理デバイス管理情報 603 を示す。論理デバイス管理情報 603 は、各論理デバイス 202 につきエントリ 901 から 918 までの情報を保持する。

【0127】

エントリ 901 には論理デバイス 202 を識別するための LDEV 番号が格納される。

【0128】

エントリ 902 には、LDEV 番号により特定される論理デバイス 202 の容量（サイズ）が格納される。

【0129】

エントリ 903 には、LDEV 番号により特定される論理デバイス 202 の LDEV タイプが格納される。ストレージシステム 130 には、ディスク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）の異なる、複数のデバイスタイプの論理デバイス 202 を定義することができる。LDEV タイプは、各論理デバイス 202 がどのデバイスタイプであるかを示す。

30

【0130】

エントリ 904 には、LDEV 番号により特定される論理デバイス 202 の LDEV 状態が格納される。LDEV 状態は、論理デバイス 202 のデバイス状態を示す。LDEV 状態には、「オンライン」、「オフライン」、「未実装」、「ブロックド」がある。「オンライン」は、論理デバイス 202 が正常に稼動し、1 つ以上のポート 131 に LU パスが定義され、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。「オフライン」は、論理デバイス 202 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、論理デバイス 202 に物理デバイス又は外部デバイスが割り当てられていない等の理由でホスト計算機 100 からアクセスできない状態にあることを示す。「ブロックド」は、論理デバイス 202 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状態を示す。エントリ 603 d の初期値は、「未実装」であるが、論理デバイス定義処理により、「オフライン」に更新され、更に LU パス定義処理により、「オンライン」に更新される。

40

【0131】

エントリ 905 には、LDEV 番号により特定される論理デバイス 202 の LUSE 状

50

態が格納される。LUSE状態は、論理ユニット201がLUSE構成を有しているか否かを示すとともに、論理ユニット201がLUSE構成を有している場合における、論理ユニット201に割り当てられている複数の論理デバイス202のうち代表的な論理デバイスのLDEV番号を示す。

【0132】

エントリ906には、連結LDEV数が格納される。連結LDEV数は、論理ユニット201がLUSE構成を有している場合における、論理ユニット201に割り当てられている論理デバイス202の個数を示す。

【0133】

エントリ907には、連結LDEV番号リストが格納される。連結LDEV番号リストは、論理ユニット201がLUSE構成を有している場合における、論理ユニット201に割り当てられている論理デバイス202のLDEV番号のリストを示す。

10

【0134】

エントリ908には、LUPAS数が格納される。LUPASとは、ホスト計算機100が論理ユニット201に接続するためのパスである。LUPAS数は、LUPASの本数を示す。

【0135】

エントリ909には、それぞれのホストグループの番号、ホストグループに属するホスト計算機100からアクセス可能なポート131を識別するためのポート番号、ポート131に割り当てられる論理ユニット201を識別するための情報(ターゲットID又はLUN)がリスト形式で格納される。

20

【0136】

エントリ910には、それぞれのホストグループに属するホスト計算機100の台数が格納される。

【0137】

エントリ911には、複数のホスト計算機100による一つの論理ユニット201へのアクセスが許可されている場合における、ホスト計算機100のホスト名がリスト形式で格納される。

【0138】

エントリ912には、論理デバイス202がVDEV連結構成を有している場合における、論理デバイス202に割り当てられている仮想デバイス203の個数が格納される。

30

【0139】

エントリ913には、論理デバイス202と仮想デバイス203とが一対一に対応していない場合における、仮想デバイス203の番号(VDEV番号)がリスト形式で格納される。

【0140】

エントリ914には、論理デバイス202と仮想デバイス203との間のマッピング関係を示す情報が格納される。例えば、エントリ914には、論理デバイス202がCVS構成を有している場合における、仮想デバイス203上の論理デバイス202のオフセット位置や、論理デバイス202がVDEV連結構成を有している場合における、論理デバイス202上の仮想デバイス203のオフセット位置などが格納される。

40

【0141】

エントリ915には、論理デバイス202が同期コピー中又は移行中の場合における、コピー先又は移行先の論理デバイス202の番号が格納される。論理デバイス202が同期コピー中でもなく移行中でもない場合には、エントリ915には無効値が格納される。

【0142】

エントリ916には、論理デバイス202がデータ移行中である場合における、データ移行が完了した最終アドレスを示す情報(以下、データ移行進捗ポイントと呼ぶ)が格納される。データ移行進捗ポイントは、ホスト計算機100から論理デバイス202へのI/O要求をストレージシステム130が受領した際に、アクセス先の記憶領域が移行済み

50

であるのか或いは移行未完了であるのかを判定し、処理を選択する際に用いられる。

【0143】

エントリ917には、論理デバイス202が同期コピー中か、データ移行中か、また両状態のどちらにも該当しないか、を示す情報（以下、同期コピー/データ移行中フラグと称する。）が格納される。

【0144】

エントリ918には、暗号化フラグが格納される。暗号化フラグとは、物理デバイス205又は拡張デバイス206に書き込まれるデータを暗号化する否かを示すフラグ情報である。暗号化フラグがオンに設定されると、ストレージシステム130は、物理デバイス205又は拡張デバイス206に書き込まれるデータを暗号化する。

10

【0145】

図29は仮想デバイス管理情報604を示す。仮想デバイス管理情報604は、各仮想デバイス203につきエントリ1001から1010までの情報を保持する。

【0146】

エントリ1001には、仮想デバイス203を識別するためのVDEV番号が格納される。

【0147】

エントリ1002には、VDEV番号により特定される仮想デバイス203の容量（サイズ）が格納される。

【0148】

エントリ1003には、VDEV番号により特定される仮想デバイス203のVDEV状態が格納される。VDEV状態は、仮想デバイス203のデバイス状態を示す。VDEV状態には、「オンライン」、「オフライン」、「未実装」、「ブロックド」がある。「オンライン」は、仮想デバイス203が正常に稼動し、かつホスト計算機100から仮想デバイス203へのパスが定義され、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。「オフライン」は、仮想デバイス203は定義され、正常に稼動しているが、仮想デバイス203への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、仮想デバイス203に物理デバイス又は外部デバイスが割り当てられていない等の理由でホスト計算機100からアクセスできない状態にあることを示す。「ブロックド」は、仮想デバイス203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状態を示す。エントリ1003の初期値は、「未実装」であるが、仮想デバイス定義処理により、「オフライン」に更新され、更にパス定義処理により「オンライン」に更新される。

20

30

【0149】

エントリ1004には、VDEV番号により特定される仮想デバイス203のRAID構成が格納される。RAID構成には、例えば、RAIDレベル、データパリティ数、ドライブ構成、ストライプサイズなどがある。

【0150】

エントリ1005には、論理デバイス202がCVS構成を有している場合における、一つの仮想デバイス203から割り当てられる論理デバイス202の個数が格納される。

40

【0151】

エントリ1006には、VDEV番号により特定される仮想デバイス203が割り当てられている論理デバイス202の番号がリスト形式で格納される。

【0152】

エントリ1007には、論理デバイス202と仮想デバイス203との間のマッピング関係を示す情報がリスト形式で格納される。例えば、エントリ1007には、論理デバイス202がCVS構成を有している場合における、仮想デバイス203上の論理デバイス202のオフセット位置や、論理デバイス202がVDEV連結構成を有している場合における、論理デバイス202上の仮想デバイス203のオフセット位置等が格納される。

【0153】

50

エントリ 1 0 0 8 には、仮想デバイス 2 0 3 に割り当てられるデバイスグループ 2 0 4 の個数が格納される。例えば、エントリ 1 0 0 8 には、仮想デバイス 2 0 3 が V D E V 離散の構成を有している場合における、仮想デバイス 2 0 3 に割り当てられるデバイスグループ 2 0 4 の個数が格納される。また例えば、エントリ 1 0 0 8 は、論理ユニット移動処理中において、複数のデバイスグループ 2 0 4 を一つの仮想デバイス 2 0 3 に割り当てるために用いられる。

【 0 1 5 4 】

エントリ 1 0 0 9 には、仮想デバイス 2 0 3 に割り当てられるデバイスグループ 2 0 4 の番号がリスト形式で格納される。

【 0 1 5 5 】

エントリ 1 0 1 0 には、仮想デバイス 2 0 3 とデバイスグループ 2 0 4 との間のマッピング関係を示す情報がリスト形式で格納される。

【 0 1 5 6 】

図 3 0 はデバイスグループ管理情報 6 0 5 を示す。デバイスグループ管理情報 6 0 5 は各デバイスグループ 2 0 4 につきエントリ 1 1 0 1 から 1 1 1 0 までの情報を保持する。

【 0 1 5 7 】

エントリ 1 1 0 1 には、デバイスグループ 2 0 4 を識別するためのデバイスグループ番号が格納される。

【 0 1 5 8 】

エントリ 1 1 0 2 には、デバイスグループ番号により特定されるデバイスグループ 2 0 4 の容量 (サイズ) が格納される。

【 0 1 5 9 】

エントリ 1 1 0 3 には、デバイスグループ番号により特定されるデバイスグループ 2 0 4 の D E V G r 状態が格納される。 D E V G r 状態は、デバイスグループ 2 0 4 のデバイス状態を示す。 D E V G r 状態には、「オンライン」、「オフライン」、「未実装」、「ブロックド」がある。「オンライン」は、デバイスグループ 2 0 4 が正常に稼動し、かつホスト計算機 1 0 0 からデバイスグループ 2 0 4 へのパスが定義され、ホスト計算機 1 0 0 からアクセスできる状態であることを示す。「オフライン」は、デバイスグループ 2 0 4 は定義され正常に稼動しているが、デバイスグループ 2 0 4 へのパスが未定義であるなどの理由でホスト計算機 1 0 0 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、デバイスグループ 2 0 4 に物理デバイス又は外部デバイスが割り当てられていない等の理由でホスト計算機 1 0 0 からアクセスできない状態にあることを示す。「ブロックド」は、デバイスグループ 2 0 4 に障害が発生してホスト計算機 1 0 0 からアクセスできない状態を示す。エントリ 1 1 0 3 の初期値は、「未実装」であるが、デバイスグループ定義処理により、「オフライン」に更新され、更にパス定義処理により「オンライン」に更新される。

【 0 1 6 0 】

エントリ 1 1 0 4 には、デバイスグループ番号により特定されるデバイスグループ 2 0 4 の R A I D 構成が格納される。 R A I D 構成には、例えば、 R A I D レベル、データパリティ数、ドライブ構成、ストライプサイズなどがある。但し、論理ユニット移動処理中の場合のように、一つの拡張デバイス 2 0 6 を一つのデバイスグループ 2 0 4 に割り当てられる場合には、デバイスグループ 2 0 4 には、 R A I D 構成は適用されない。

【 0 1 6 1 】

エントリ 1 1 0 5 には、デバイスグループ 2 0 4 に割り当てられている仮想デバイス 2 0 3 の個数が格納される。 L U S E 構成を有する論理ユニット 2 0 1 を移動させるときには、一つのデバイスグループ 2 0 4 に複数の仮想デバイス 2 0 3 を割り当てることがあるので、このような場合には、エントリ 1 1 0 5 の値が更新される。

【 0 1 6 2 】

エントリ 1 1 0 6 には、デバイスグループ番号により特定されるデバイスグループ 2 0 4 に割り当てられている仮想デバイス 2 0 3 の番号がリスト形式で格納される。

10

20

30

40

50

【 0 1 6 3 】

エントリ 1 1 0 7 には、デバイスグループ 2 0 4 と仮想デバイス 2 0 3 との間のマッピング関係を示す情報がリスト形式で格納される。

【 0 1 6 4 】

エントリ 1 1 0 8 には、デバイスグループ 2 0 4 に割り当てられるデバイス（物理デバイス 2 0 5 又は拡張デバイス 2 0 6 ）の個数が格納される。

【 0 1 6 5 】

エントリ 1 1 0 9 には、デバイスグループ 2 0 4 に割り当てられるデバイス（物理デバイス 2 0 5 又は拡張デバイス 2 0 6 ）の番号がリスト形式で格納される。

【 0 1 6 6 】

エントリ 1 1 1 0 には、デバイスグループ 2 0 4 とデバイス（物理デバイス 2 0 5 又は拡張デバイス 2 0 6 ）との間のマッピング関係を示す情報がリスト形式で格納される。例えば、エントリ 1 1 1 0 には、図 1 3 乃至図 1 5 に示すように、拡張デバイス 2 0 6 の一部分のみがデバイスグループ 2 0 4 に割り当てられているような場合における、拡張デバイス 2 0 6 の先頭からのアドレスと容量などが格納される。

【 0 1 6 7 】

図 3 1 は V D E V - D E V G r マッピング情報 6 0 6 を示す。V D E V - D E V G r マッピング情報 6 0 6 は、エントリ 1 2 0 1 からエントリ 1 2 0 7 までの情報を保持する。

【 0 1 6 8 】

エントリ 1 2 0 1 には、V D E V 領域を識別するための番号が格納される。V D E V 領域とは、仮想デバイス 2 0 3 の一部の領域をいう。

【 0 1 6 9 】

エントリ 1 2 0 2 には、V D E V 領域に対応付けられているデバイスグループ 2 0 4 の個数が格納される。

【 0 1 7 0 】

エントリ 1 2 0 3 には、V D E V 領域に対応付けられているデバイスグループ 2 0 4 の番号がリスト形式で格納される。

【 0 1 7 1 】

エントリ 1 2 0 4 には、有効フラグがリスト形式で格納される。有効フラグとは、論理ユニット 2 0 1 に割り当てられる仮想デバイス 2 0 3 にミラーリング又はパリティ生成等の処理を実行するか否かを示すフラグ情報である。論理ユニット移行処理中では、有効フラグはオフに設定され、移行元の論理ユニット 2 0 1 に割り当てられる仮想デバイス 2 0 3 に対するミラーリング又はパリティ生成等の処理は実施されない。

【 0 1 7 2 】

エントリ 1 2 0 5 には、デバイスグループ 2 0 4 に対応付けられている V D E V 領域のオフセットアドレスがリスト形式で格納される。

【 0 1 7 3 】

エントリ 1 2 0 6 には、それぞれの V D E V 領域について、キャッシュスルーリードの動作モードで動作するか否かを示す情報がリスト形式で格納される。

【 0 1 7 4 】

エントリ 1 2 0 7 には、それぞれの V D E V 領域について、キャッシュスルーライトの動作モードで動作するか否かを示す情報がリスト形式で格納される。

【 0 1 7 5 】

図 3 2 は物理デバイス管理情報 6 0 7 を示す。物理デバイス管理情報 6 0 7 は、ストレージシステム 1 3 0 内に定義される各物理デバイス 2 0 5 につきエントリ 1 3 0 1 からエントリ 1 3 0 6 までの情報を保持する。

【 0 1 7 6 】

エントリ 1 3 0 1 には、物理デバイス 2 0 5 を識別するための物理デバイス番号が格納される。

【 0 1 7 7 】

10

20

30

40

50

エントリ 1302 には、物理デバイス番号により特定される物理デバイス 205 の容量 (サイズ) が格納される。

【0178】

エントリ 1303 には、物理デバイス番号により特定される物理デバイス 205 のデバイス状態が格納される。デバイス状態としては、「オンライン」、「オフライン」、「未実装」、「ブロックド」がある。「オンライン」は、物理デバイス 205 が正常に稼働し、且つその物理デバイス 205 が論理デバイス 202 に割り当てられている状態を示す。「オフライン」は、物理デバイス 205 は定義され、正常に稼働しているが、その物理デバイス 205 は論理デバイス 202 に割り当てられていないことを示す。「未実装」は、物理デバイス 205 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、物理デバイス 205 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。エントリ 1303 の初期値は「未実装」に設定されており、物理デバイス定義処理により「オフライン」に更新され、物理デバイス 205 が論理デバイス 202 に定義された契機で「オンライン」に更新される。

10

【0179】

エントリ 1304 には、一つの物理デバイス 205 の記憶領域が部分的に複数のデバイスグループ 204 に割り当てられているような場合 (図 15 参照) に一つの物理デバイス 205 に割り当てられているデバイスグループ 204 の個数が格納される。

【0180】

エントリ 1305 には、物理デバイス 205 に割り当てられるデバイスグループ 204 の番号がリスト形式で格納される。

20

【0181】

エントリ 1306 には、物理デバイス 205 とデバイスグループ 204 との間のマッピング関係を示す情報が格納される。

【0182】

図 33 は拡張デバイス管理情報 608 を示す。拡張デバイス管理情報 608 は、外部ストレージシステム 150 が有する外部ボリューム 301 をストレージシステム 130 内の拡張デバイス 206 (内部ボリューム) として仮想化するためのものである。拡張デバイス管理情報 608 は、各外部ボリューム 301 につきエントリ 1401 からエントリ 1411 までの情報を保持する。

30

【0183】

エントリ 1401 には、拡張デバイス 206 を識別するための EDEV 番号が格納される。EDEV 番号は、ストレージシステム 130 内で外部ボリューム 301 を一意に識別するための識別番号である。

【0184】

エントリ 1402 には、EDEV 番号により特定される拡張デバイス 206 の容量 (サイズ) が格納される。

【0185】

エントリ 1403 には、EDEV 番号により特定される拡張デバイス 206 のデバイス状態が格納される。デバイス状態としては、「オンライン」、「オフライン」、「未実装」、「ブロックド」がある。「オンライン」は、拡張デバイス 206 が正常に稼働し、且つその拡張デバイス 206 が論理デバイス 202 に割り当てられている状態を示す。「オフライン」は、拡張デバイス 206 は定義され、正常に稼働しているが、その拡張デバイス 206 は論理デバイス 202 に割り当てられていないことを示す。「未実装」は、拡張デバイス 206 が定義されていない状態にあることを示す。「ブロックド」は、拡張デバイス 206 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。エントリ 1403 の初期値は「未実装」に設定されており、拡張デバイス定義処理により「オフライン」に更新され、拡張デバイス 206 が論理デバイス 202 に定義された契機で「オンライン」に更新される。

40

【0186】

50

エントリ 1404 には、一つの拡張デバイス 206 の記憶領域が部分的に複数のデバイスグループ 204 に割り当てられているような場合（図 15 参照）に一つの拡張デバイス 206 に割り当てられているデバイスグループ 204 の個数が格納される。

【0187】

エントリ 1405 には、拡張デバイス 206 に割り当てられるデバイスグループ 204 の番号がリスト形式で格納される。

【0188】

エントリ 1406 には、拡張デバイス 206 とデバイスグループ 204 との間のマッピング関係を示す情報が格納される。

【0189】

エントリ 1407 には、拡張デバイス 206 を仮想化する外部ストレージシステム 150 を識別するためのストレージ識別情報が格納される。ストレージ識別情報としては、例えば、外部ストレージシステム 150 を製造販売するベンダを識別するためのベンダ識別情報と各ベンダが一意に割り振るシリアル番号との組み合わせ等が考えられる。

【0190】

エントリ 1408 には、外部ストレージシステム 150 が管理する外部ボリューム 301 の LDEV 番号が格納される。

【0191】

エントリ 1409 には、ストレージシステム 130 内の複数のポート 131 のうち外部ストレージシステム 150 に接続するためのイニシエータポートを識別するためのポート番号が格納される。

【0192】

エントリ 1410 には、外部ボリューム 301 が外部ストレージシステム 150 内の一つ以上のポート 151 に定義されている場合における、それらのポート 151 を識別するためのポート ID、及び外部ボリューム 301 を識別するためのターゲット ID 及び LUN がリスト形式で格納される。

【0193】

エントリ 1411 には、アクセス制御情報が格納される。アクセス制御情報は、外部ボリューム 301 が複数のストレージシステム 130 によって共有されている場合において、外部ボリューム 301 の各領域についてどのストレージシステム 130 がアクセス権限を有しているかを管理するための情報である。

【0194】

図 34 は計算機システム 500 のソフトウェア構成を示す。ストレージシステム 130 内の制御メモリ 135 には、各種の構成管理情報（ホストグループ管理情報 601、LUパス管理情報 602、論理デバイス管理情報 603、仮想デバイス管理情報 604、デバイスグループ管理情報 605、VDEV - DEVGr マッピング情報 606、物理デバイス管理情報 607、拡張デバイス管理情報 608、及びデバイス機能管理情報 609）が格納されている。

【0195】

デバイス機能管理情報 609 は、各論理デバイス 202 に設定される各種の属性を示す情報を保持する。論理デバイス 202 の属性情報の例としては、論理デバイス 202 へのアクセスを特定のホスト計算機 100 にのみに制限するアクセス制限情報、論理デバイス 202 へのリードアクセス又はライトアクセスを抑止するアクセス属性情報、論理デバイス 202 に格納されるデータへの暗号化適用の有無やデータの暗号化及び復号化に用いられる鍵情報などの暗号化情報を挙げるができる。

【0196】

ストレージシステム 130 内のメモリ 133 には、上述した各種の構成管理情報が格納される他、ストレージ I/O 処理プログラム 221、及び論理ユニット移動処理プログラム 222 が格納されている。ストレージ I/O 処理プログラム 221 が実行するストレージ I/O 処理の詳細、及び論理ユニット移動処理プログラム 222 が実行する論理ユニッ

10

20

30

40

50

ト移動処理プログラム 2 2 2 の詳細については、後述する。

【 0 1 9 7 】

管理サーバ 1 1 0 内のメモリ 1 1 2 には、上述した各種の構成管理情報が格納される他、論理ユニット移動指示処理プログラム 2 4 1、及びストレージ負荷監視プログラム 2 4 2 が格納されている。論理ユニット移動指示処理プログラム 2 4 1 が実行する L U 移動指示処理の詳細、及びストレージ負荷監視プログラム 2 4 2 が実行する処理の詳細については、後述する。

【 0 1 9 8 】

ホスト計算機 1 0 0 内のメモリ 1 0 2 には、デバイスパス管理情報 2 5 1、及びホスト I / O 処理プログラム 2 6 1 が格納されている。ホスト計算機 1 0 0 は、ストレージシステム 1 3 0 又は外部ストレージシステム 1 5 0 が提供する論理デバイスを認識し、論理デバイスをデバイスファイルに対応付けて管理する。デバイスパス管理情報 2 5 1 には、論理デバイスとデバイスファイルとの対応関係を示す情報が含まれている。例えば、ストレージシステム 1 3 0 が有する複数のポート 1 3 1 に同一の論理デバイスが割り当てられている場合には、同一の論理デバイスから複数のデバイスファイルが生成される。ホスト計算機 1 0 0 は、同一の論理デバイスに対応付けられている複数のデバイスファイルを一元管理し、論理デバイスへの入出力負荷を分散したり、或いはネットワーク障害時におけるパス切り替えを制御したりする。

【 0 1 9 9 】

尚、管理サーバ 1 1 0 又は管理端末 1 4 0 からストレージ管理ソフトウェア又はストレージ管理者からの指示を受けて、ストレージシステム 1 3 0 の構成を変更した場合や、或いはストレージシステム 1 3 0 内の各部位が障害又は自動交替等によって構成が変更した場合に、ストレージシステム 1 3 0 内の一つの制御プロセッサ 1 3 2 は、制御メモリ 1 3 5 内の構成管理情報を更新する。更に、制御プロセッサ 1 3 2 は、ストレージシステム 1 3 0 の構成変更により構成管理情報が更新された旨を他の制御プロセッサ 1 3 2、管理端末 1 4 0、及び管理サーバ 1 1 0 に通知し、他の制御メモリ 1 3 5 に格納されている全ての構成管理情報を最新の情報に更新させる。

【 0 2 0 0 】

次に、図 3 5 を参照しながら C V S 構成を有する論理ユニットの移動処理の概要について説明する。

【 0 2 0 1 】

ストレージシステム 1 3 0 a は、論理ユニット 2 0 1 a - 1、2 0 1 a - 2、論理デバイス 2 0 2 a - 1、2 0 2 a - 2、仮想デバイス 2 0 3 a、及び拡張デバイス 2 0 6 a を備える。論理ユニット 2 0 1 a - 1 は、パス 4 0 1 - 1 を介してホスト計算機 1 0 0 に接続されている。論理ユニット 2 0 1 a - 2 は、パス 4 0 1 - 2 を介してホスト計算機 1 0 0 に接続されている。仮想デバイス 2 0 3 a は、複数の記憶領域 2 0 3 a - 1、2 0 3 a - 2 に分割されている。一方の記憶領域 2 0 3 a - 1 は、論理デバイス 2 0 2 a - 1 を介して論理ユニット 2 0 2 a - 1 にマッピングされており、他方の記憶領域 2 0 3 a - 1 は、論理デバイス 2 0 2 a - 1 を介して論理ユニット 2 0 2 a - 1 にマッピングされている。それぞれの論理ユニット 2 0 1 a - 1、2 0 1 a - 2 は、仮想デバイス 2 0 3 a の一部を仮想化したものであり、C V S 構成を有している。

【 0 2 0 2 】

拡張デバイス 2 0 6 a は、外部ボリューム 3 0 1 を仮想化したものであり、複数の記憶領域 2 0 6 a - 1、2 0 6 a - 2 に分割されている。一方の記憶領域 2 0 6 a - 1 は、記憶領域 2 0 3 a - 1 にマッピングされており、他方の記憶領域 2 0 6 a - 2 は、記憶領域 2 0 3 a - 2 にマッピングされている。

【 0 2 0 3 】

ストレージシステム 1 3 0 a (移動元ストレージ) 内の論理ユニット 2 0 1 a - 2 (移動元論理ユニット) にあるデータをストレージシステム 1 3 0 b (移動先ストレージ) 内の論理ユニット 2 0 1 b (移動先論理ユニット) に移動する処理について概説する。論理

10

20

30

40

50

ユニット移動処理は、下記の(1)～(9)の処理手順を含む。

【0204】

(1) ストレージシステム130bは、外部ボリューム301と、ストレージシステム130bとの間を外部接続するためのパス405を定義し、外部ボリューム301を仮想化する拡張デバイス206bをストレージシステム130b内に作成する。拡張デバイス206bは、複数の記憶領域206b-1, 206b-2に分割されている。拡張デバイス206bの記憶容量は、外部ボリューム301の記憶量に等しい。記憶領域206b-2の記憶容量は、論理ユニット201a-2の記憶容量に等しい。

【0205】

(2) ストレージシステム130bは、外部ボリューム301の記憶容量と同一の記憶容量を有する仮想デバイス203bをストレージシステム130b内に作成する。仮想デバイス203bは、複数の記憶領域203b-1, 203b-2に分割されている。記憶領域203b-2の記憶容量は、論理ユニット201a-2の記憶容量に等しい。

10

【0206】

(3) ストレージシステム130bは、論理ユニット201a-2とストレージシステム130bとの間を外部接続するためのパス404を定義し、論理ユニット201a-2を仮想化するための拡張デバイス206cをストレージシステム130b内に作成する。

【0207】

(4) ストレージシステム130bは、拡張デバイス206cを仮想デバイス203b(より詳細には、記憶領域203b-2)にマッピングする。

20

【0208】

(5) ストレージシステム130bは、論理デバイス202a-2の論理構成と同一の論理構成を有する論理デバイス202bをストレージシステム130b内に作成する。仮想デバイス203b(より詳細には、記憶領域203b-2)は、論理デバイス202bにマッピングされる。

【0209】

(6) ストレージシステム130bは、論理ユニット201a-2の論理構成と同一の論理構成を有する論理ユニット201bをストレージシステム130b内に作成し、ホスト計算機100と論理ユニット201bとを接続するためのパス403を定義する。仮想デバイス202bは、論理ユニット201bにマッピングされる。

30

【0210】

(7) ホスト計算機100は、外部ボリューム301にアクセスするための経路として、パス401-2からパス403に切り替える。このとき、パス401-2を削除する等して、ホスト計算機100から論理ユニット201a-2へのアクセスを禁止する。尚、パス401-2からパス403へのパス切り替えは、ストレージシステム130a, 130b, 又は管理サーバ110等が行っても良い。また、必ずしもパス401-2を削除する必要はなく、例えば、パス401-2をパス403の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス401-2の状態を無効としたまま、パス401-2を残しても良い。

【0211】

上記の処理により、論理ユニット201bは、論理ユニット201a-2を仮想化したものであるだけでなく、外部ボリューム301の一部をも仮想化したものであるので、論理ユニット移動処理中における、ホスト計算機100から論理ユニット201a-2へのデータ入出力経路は、パス403 ストレージシステム130b内の記憶階層(LU/LDEV/VDEV/EDEV) パス404 ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV) パス402 外部ボリューム301となる。

40

【0212】

(8) ストレージシステム130aは、ストレージシステム130aのディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム301にデステージするため

50

にストレージシステム 130 a の動作モードをキャッシュスルーモードに設定する。尚、ストレージシステム 130 a は、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部（例えば、ストレージシステム 130 b、ホスト計算機 100、又は管理サーバ 110 など）からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

【0213】

(9) ストレージシステム 130 a のディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム 301 にデステージできたならば、ストレージシステム 130 b は、拡張デバイス 206 c と仮想デバイス 203 b との間のマッピング関係を解除し、拡張デバイス 206 b を仮想デバイス 203 b にマッピングする。更に、ストレージシステム 130 b は、パス 404 を削除する。

10

【0214】

尚、CVS構成を有する論理ユニットの移動処理では、複数のストレージシステム 130 a, 130 b によって外部ボリューム 301 の一部が共有されるので、外部ボリューム 301 に対する排他制御が必要になる。外部ボリューム 301 に対する排他制御は、外部ストレージシステム 150 が実行してもよく、或いは複数のストレージシステム 130 a, 130 b が連携しながら実行してもよい。

【0215】

次に、図 36 を参照しながら LUSE 構成を有する論理ユニットの移動処理の概要について説明する。

20

【0216】

ストレージシステム 130 a は、論理ユニット 201 a, 論理デバイス 202 a - 1, 202 a - 2, 仮想デバイス 203 a - 1, 203 a - 2, 及び拡張デバイス 206 a - 1, 206 a - 2 を備える。論理ユニット 201 a は、パス 401 を介してホスト計算機 100 に接続されている。論理ユニット 201 a は、複数の論理デバイス 202 a - 1, 202 a - 2 を一つに集めた記憶領域であり、LUSE 構成を有している。拡張デバイス 206 a - 1 は、外部ボリューム 301 - 1 を仮想化したものであり、仮想デバイス 203 a - 1 にマッピングされている。仮想デバイス 203 a - 1 は、論理デバイス 202 a - 1 にマッピングされている。拡張デバイス 206 a - 2 は、外部ボリューム 301 - 2 を仮想化したものであり、仮想デバイス 203 a - 2 にマッピングされている。仮想デバイス 203 a - 2 は、論理デバイス 202 a - 2 にマッピングされている。

30

【0217】

ストレージシステム 130 a (移動元ストレージ) 内の論理ユニット 201 a (移動元論理ユニット) にあるデータをストレージシステム 130 b (移動先ストレージ) 内の論理ユニット 201 b (移動先論理ユニット) に移動する処理について概説する。論理ユニット移動処理は、下記の(1)~(9)の処理手順を含む。

【0218】

(1) ストレージシステム 130 b は、外部ボリューム 301 - 1, 301 - 2 と、ストレージシステム 130 b との間を外部接続するためのパス 405 - 1, 405 - 2 を定義し、外部ボリューム 301 - 1, 301 - 2 を仮想化する拡張デバイス 206 b - 1, 206 b - 2 をストレージシステム 130 b 内に作成する。

40

【0219】

(2) ストレージシステム 130 b は、外部ボリューム 301 - 1 の記憶容量と同一の記憶容量を有する仮想デバイス 203 b - 1 と、外部ボリューム 301 - 2 の記憶容量と同一の記憶容量を有する仮想デバイス 203 b - 2 をストレージシステム 130 b 内に作成する。

【0220】

(3) ストレージシステム 130 b は、論理ユニット 201 a とストレージシステム 130 b との間を外部接続するためのパス 404 を定義し、論理ユニット 201 a を仮想化するための拡張デバイス 206 c をストレージシステム 130 b 内に作成する。拡張デバイ

50

ス 2 0 6 c は、複数の記憶領域 2 0 6 c - 1 , 2 0 6 c - 2 に分割される。

【 0 2 2 1 】

(4) ストレージシステム 1 3 0 b は、拡張デバイス 2 0 6 c を仮想デバイス 2 0 3 b - 1 , 2 0 3 b - 2 にマッピングする(より詳細には、記憶領域 2 0 6 c - 1 を仮想デバイス 2 0 3 b - 1 にマッピングし、記憶領域 2 0 6 c - 2 を仮想デバイス 2 0 3 b - 2 にマッピングする。)。

【 0 2 2 2 】

(5) ストレージシステム 1 3 0 b は、論理デバイス 2 0 2 a - 1 の論理構成と同一の論理構成を有する論理デバイス 2 0 2 b - 1 と、論理デバイス 2 0 2 a - 2 の論理構成と同一の論理構成を有する論理デバイス 2 0 2 b - 2 とをそれぞれストレージシステム 1 3 0 b 内に作成する。仮想デバイス 2 0 3 b - 1 は、論理デバイス 2 0 2 b - 1 にマッピングされ、仮想デバイス 2 0 3 b - 2 は、論理デバイス 2 0 2 b - 2 にマッピングされる。

【 0 2 2 3 】

(6) ストレージシステム 1 3 0 b は、論理ユニット 2 0 1 a の論理構成と同一の論理構成を有する論理ユニット 2 0 1 b をストレージシステム 1 3 0 b 内に作成し、ホスト計算機 1 0 0 と論理ユニット 2 0 1 b とを接続するためのパス 4 0 3 を定義する。論理デバイス 2 0 2 b - 1 , 2 0 2 b - 2 は、論理ユニット 2 0 1 b にマッピングされる。

【 0 2 2 4 】

(7) ホスト計算機 1 0 0 は、外部ボリューム 3 0 1 - 1 , 3 0 1 - 2 にアクセスするための経路として、パス 4 0 1 からパス 4 0 3 に切り替える。このとき、パス 4 0 1 を削除する等して、ホスト計算機 1 0 0 から論理ユニット 2 0 1 a へのアクセスを禁止する。尚、パス 4 0 1 からパス 4 0 3 へのパス切り替えは、ストレージシステム 1 3 0 a , 1 3 0 b , 又は管理サーバ 1 1 0 等が行っても良い。また、必ずしもパス 4 0 1 を削除する必要はなく、例えば、パス 4 0 1 をパス 4 0 3 の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス 4 0 1 の状態を無効としたまま、パス 4 0 1 を残しても良い。

【 0 2 2 5 】

上記の処理により、論理ユニット 2 0 1 b は、論理ユニット 2 0 1 a を仮想化したものであるだけでなく、外部ボリューム 3 0 1 - 1 , 3 0 1 - 2 をも仮想化したものでもあるので、論理ユニット移動処理中における、ホスト計算機 1 0 0 から論理ユニット 2 0 1 a へのデータ入出力経路は、パス 4 0 3 ストレージシステム 1 3 0 b 内の記憶階層 (L U / L D E V / V D E V / E D E V) パス 4 0 4 ストレージシステム 1 3 0 a 内の記憶階層 (L U / L D E V / V D E V / E D E V) パス 4 0 2 - 1 , 4 0 2 - 2 外部ボリューム 3 0 1 - 1 , 3 0 1 - 2 となる。

【 0 2 2 6 】

(8) ストレージシステム 1 3 0 a は、ストレージシステム 1 3 0 a のディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム 3 0 1 - 1 , 3 0 1 - 2 にデステージするためにストレージシステム 1 3 0 a の動作モードをキャッシュスルーモードに設定する。尚、ストレージシステム 1 3 0 a は、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部(例えば、ストレージシステム 1 3 0 b 、ホスト計算機 1 0 0 、 又は管理サーバ 1 1 0 など)からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

【 0 2 2 7 】

(9) ストレージシステム 1 3 0 a のディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム 3 0 1 - 1 , 3 0 1 - 2 にデステージできたならば、ストレージシステム 1 3 0 b は、拡張デバイス 2 0 6 c と仮想デバイス 2 0 3 b - 1 , 2 0 3 b - 2 との間のマッピング関係を解除し、拡張デバイス 2 0 6 b - 1 を仮想デバイス 2 0 3 b - 1 にマッピングし、拡張デバイス 2 0 6 b - 2 を仮想デバイス 2 0 3 b - 2 にマッピングする。更に、ストレージシステム 1 3 0 b は、パス 4 0 4 を削除する。

【 0 2 2 8 】

10

20

30

40

50

次に、図37を参照しながら外部ボリュームRAID構成を有する論理ユニットの移動処理の概要について説明する。

【0229】

ストレージシステム130aは、論理ユニット201a、論理デバイス202a、仮想デバイス203a、及び拡張デバイス206a-1、206a-2を備える。論理ユニット201aは、パス401を介してホスト計算機100に接続されている。論理ユニット201aは、複数の外部ボリューム301-1、301-2にRAID構成を適用することにより、これらの外部ボリューム301-1、301-2を一つの記憶領域に仮想化したものである。複数の外部ボリューム301-1、301-2に適用されるRAID構成として、本実施例では、RAID1(ミラーリング)を例示する。RAID1構成においては、外部ボリューム301-1は、プライマリボリュームであるのに対し、外部ボリューム301-2は、セカンダリボリューム(ミラーボリューム)である。外部ボリューム301-1、301-2に適用されるRAID構成としては、RAID1に限らず、RAID0、RAID5、RAID6等でもよい。拡張デバイス206a-1は、外部ボリューム301-1を仮想化したものである。拡張デバイス206a-2は、外部ボリューム301-2を仮想化したものである。仮想デバイス203aは、複数の拡張デバイス206a-1、206a-2にRAID構成を適用することにより、一つの記憶領域に仮想化したものである。仮想デバイス203aの中でプライマリボリュームに対応する記憶領域は、論理デバイス202aにマッピングされている。論理デバイス202aは、論理ユニット201aにマッピングされている。

10

20

【0230】

ストレージシステム130a(移動元ストレージ)内の論理ユニット201a(移動元論理ユニット)にあるデータをストレージシステム130b(移動先ストレージ)内の論理ユニット201b(移動先論理ユニット)に移動する処理について概説する。論理ユニット移動処理は、下記の(1)~(9)の処理手順を含む。

【0231】

(1)ストレージシステム130bは、外部ボリューム301-1、301-2と、ストレージシステム130bとの間を外部接続するためのパス405-1、405-2を定義し、外部ボリューム301-1、301-2を仮想化する拡張デバイス206b-1、206b-2をストレージシステム130b内に作成する。

30

【0232】

(2)ストレージシステム130bは、複数の外部ボリューム301-1、301-2を仮想化するための仮想デバイス203bをストレージシステム130b内に作成する。複数の外部ボリューム301-1、301-2を仮想化するためのRAID構成として、ミラーリングを適用する場合には、仮想デバイス203bの記憶容量は、複数の外部ボリューム301-1、301-2のそれぞれの記憶容量の総和に等しい。

【0233】

(3)ストレージシステム130bは、論理ユニット201aとストレージシステム130bとの間を外部接続するためのパス404を定義し、論理ユニット201aを仮想化するための拡張デバイス206cをストレージシステム130b内に作成する。

40

【0234】

(4)ストレージシステム130bは、拡張デバイス206cを仮想デバイス203b(より詳細には、仮想デバイス203bの中でプライマリボリュームに対応する記憶領域)にマッピングする。

【0235】

(5)ストレージシステム130bは、論理デバイス202aの論理構成と同一の論理構成を有する論理デバイス202bをストレージシステム130b内に作成する。仮想デバイス203bの中でプライマリボリュームに対応する記憶領域は、論理デバイス202bにマッピングされる。

【0236】

50

(6) ストレージシステム 130b は、論理ユニット 201a の論理構成と同一の論理構成を有する論理ユニット 201b をストレージシステム 130b 内に作成し、ホスト計算機 100 と論理ユニット 201b とを接続するためのパス 403 を定義する。論理デバイス 202b は、論理ユニット 201b にマッピングされる。

【0237】

(7) ホスト計算機 100 は、外部ボリューム 301-1, 301-2 にアクセスするための経路として、パス 401 からパス 403 に切り替える。このとき、パス 401 を削除する等して、ホスト計算機 100 から論理ユニット 201a へのアクセスを禁止する。尚、パス 401 からパス 403 へのパス切り替えは、ストレージシステム 130a, 130b, 又は管理サーバ 110 等が行っても良い。また、必ずしもパス 401 を削除する必要はなく、例えば、パス 401 をパス 403 の交替パス（例えば、システム障害時に使用するためのパス）として設定し、パス 401 の状態を無効としたまま、パス 401 を残しても良い。

【0238】

上記の処理により、論理ユニット 201b は、論理ユニット 201a を仮想化したものであるだけでなく、外部ボリューム 301-1, 301-2 をも仮想化したものである。論理ユニット移動処理中における、ホスト計算機 100 から論理ユニット 201a へのデータ入出力経路は、パス 403 ストレージシステム 130b 内の記憶階層 (LU/LDEV/VDEV/EDDEV) パス 404 ストレージシステム 130a 内の記憶階層 (LU/LDEV/VDEV/EDDEV) パス 402-1, 402-2 外部ボリューム 301-1, 301-2 となる。

【0239】

(8) ストレージシステム 130a は、ストレージシステム 130a のディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム 301-1, 301-2 にデステージするためにストレージシステム 130a の動作モードをキャッシュスルーモードに設定する。尚、ストレージシステム 130a は、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部（例えば、ストレージシステム 130b、ホスト計算機 100、又は管理サーバ 110 など）からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

【0240】

(9) ストレージシステム 130a のディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム 301-1, 301-2 にデステージできたならば、ストレージシステム 130b は、拡張デバイス 206c と仮想デバイス 203b との間のマッピング関係を解除し、拡張デバイス 206b-1, 206b-2 を仮想デバイス 203b にマッピングする。更に、ストレージシステム 130b は、パス 404 を削除する。

【0241】

次に、図 38 を参照しながら、VMA 付きボリュームの移動処理の概要について説明する。

【0242】

ストレージシステム 130a は、論理ユニット 201a, 論理デバイス 202a, 仮想デバイス 203a, 及び拡張デバイス 206a を備える。論理ユニット 201a は、パス 401 を介してホスト計算機 100 に接続されている。論理デバイス 202a は、VMA 付きボリュームであり、ホスト計算機 100 からアクセス可能なユーザ領域 202a-1 と、管理情報を格納するための管理領域 202a-2 とを含む。管理情報には、例えば、VMA 付きボリュームのアクセス属性が含まれる。ユーザ領域 202a-1 は、論理ユニット 201a に割り当てられることにより、ホスト計算機 100 により認識可能である。管理領域 202a-2 は、論理ユニット 201a に割り当てられておらず、ホスト計算機 100 から認識することはできない。仮想デバイス 203a は、ホスト計算機 100 からアクセス可能なユーザ領域 203a-1 と、管理情報を格納するための管理領域 203a-2 とを含む。拡張デバイス 206a は、ホスト計算機 100 からアクセス可能なユーザ

10

20

30

40

50

領域 206 a - 1 と、管理情報を格納するための管理領域 206 a - 2 とを含む。拡張デバイス 206 a 内のユーザ領域 206 a - 1 は、外部ボリューム 301 内のユーザ領域 301 - 1 を仮想化したものである。拡張デバイス 206 a 内の管理領域 206 a - 2 は、外部ボリューム 301 内の管理領域 301 - 2 を仮想化したものである。ユーザ領域 206 a - 1 は、ユーザ領域 203 a - 1 にマッピングされている。管理領域 206 a - 2 は、管理領域 203 a - 2 にマッピングされている。ユーザ領域 203 a - 1 は、ユーザ領域 202 a - 1 にマッピングされている。管理領域 203 a - 2 は、管理領域 202 a - 2 にマッピングされている。

【0243】

尚、移動元のストレージシステム 130 a は、外部ボリューム 301 に管理情報を格納するのではなく、ストレージシステム 130 a 内のディスクキャッシュ又はその他のメモリに管理情報を格納している場合には、論理ユニット 201 a の移動処理を実行する前準備として、外部ボリューム 301 に管理情報をデステージしておく必要がある。

【0244】

ストレージシステム 130 a (移動元ストレージ) 内の論理ユニット 201 a (移動元論理ユニット) にあるデータをストレージシステム 130 b (移動先ストレージ) 内の論理ユニット 201 b (移動先論理ユニット) に移動する処理について概説する。論理ユニット移動処理は、下記の (1) ~ (9) の処理手順を含む。

【0245】

(1) ストレージシステム 130 b は、外部ボリューム 301 と、ストレージシステム 130 b との間を外部接続するためのパス 405 を定義し、外部ボリューム 301 を仮想化する拡張デバイス 206 b をストレージシステム 130 b 内に作成する。拡張デバイス 206 b は、ホスト計算機 100 からアクセス可能なユーザ領域 206 b - 1 と、管理情報を格納するための管理領域 206 b - 2 とを含む。ストレージシステム 130 b は、管理領域 301 - 2 を管理領域 206 b - 2 にマッピングする。

【0246】

(2) ストレージシステム 130 b は、外部ボリューム 301 を仮想化するための仮想デバイス 203 b をストレージシステム 130 b 内に作成する。仮想デバイス 203 b は、ホスト計算機 100 からアクセス可能なユーザ領域 203 b - 1 と、管理情報を格納するための管理領域 203 b - 2 とを含む。ストレージシステム 130 b は、管理領域 206 b - 2 を管理領域 203 b - 2 にマッピングする。

【0247】

(3) ストレージシステム 130 b は、論理ユニット 201 a とストレージシステム 130 b との間を外部接続するためのパス 404 を定義し、論理ユニット 201 a を仮想化するための拡張デバイス 206 c をストレージシステム 130 b 内に作成する。

【0248】

(4) ストレージシステム 130 b は、拡張デバイス 206 c を仮想デバイス 203 b 内のユーザ領域 203 b - 1 にマッピングする。

【0249】

(5) ストレージシステム 130 b は、論理デバイス 202 a の論理構成と同一の論理構成を有する論理デバイス 202 b をストレージシステム 130 b 内に作成する。論理デバイス 202 b は、ホスト計算機 100 からアクセス可能なユーザ領域 202 b - 1 と、管理情報を格納するための管理領域 202 b - 2 とを含む。ユーザ領域 203 b - 1 は、ユーザ領域 202 b - 1 にマッピングされる。管理領域 203 b - 2 は、管理領域 202 b - 2 にマッピングされる。

【0250】

(6) ストレージシステム 130 b は、論理ユニット 201 a の論理構成と同一の論理構成を有する論理ユニット 201 b をストレージシステム 130 b 内に作成し、ホスト計算機 100 と論理ユニット 201 b とを接続するためのパス 403 を定義する。ユーザ領域 202 b - 1 は、論理ユニット 201 b にマッピングされる。

10

20

30

40

50

【 0 2 5 1 】

(7) ホスト計算機 1 0 0 は、外部ボリューム 3 0 1 にアクセスするための経路として、パス 4 0 1 からパス 4 0 3 に切り替える。このとき、パス 4 0 1 を削除する等して、ホスト計算機 1 0 0 から論理ユニット 2 0 1 a へのアクセスを禁止する。尚、パス 4 0 1 からパス 4 0 3 へのパス切り替えは、ストレージシステム 1 3 0 a , 1 3 0 b , 又は管理サーバ 1 1 0 等が行っても良い。また、必ずしもパス 4 0 1 を削除する必要はなく、例えば、パス 4 0 1 をパス 4 0 3 の交替パス (例えば、システム障害時に使用するためのパス) として設定し、パス 4 0 1 の状態を無効としたまま、パス 4 0 1 を残しても良い。

【 0 2 5 2 】

上記の処理により、論理ユニット 2 0 1 b は、論理ユニット 2 0 1 a を仮想化したものであるだけでなく、外部ボリューム 3 0 1 内のユーザ領域 3 0 1 - 1 をも仮想化したものでもあるので、論理ユニット移動処理中における、ホスト計算機 1 0 0 から論理ユニット 2 0 1 a へのデータ入出力経路は、パス 4 0 3 ストレージシステム 1 3 0 b 内の記憶階層 (L U / L D E V / V D E V / E D E V) パス 4 0 4 ストレージシステム 1 3 0 a 内の記憶階層 (L U / L D E V / V D E V / E D E V) パス 4 0 2 外部ボリューム 3 0 1 となる。

10

【 0 2 5 3 】

尚、ストレージシステム 1 3 0 b は、外部ボリューム 1 3 0 内の管理領域 3 0 1 - 2 を仮想デバイス 2 0 3 b によって仮想化しているため、パス 4 0 5 を介して管理情報を参照することができる。論理ユニット 2 0 1 a の移動処理中において、ストレージシステム 1 3 0 a がパス 4 0 2 を介して外部ボリューム 1 3 0 内の管理情報を参照するためには、仮想デバイス 2 0 3 b 内の管理領域 2 0 3 b - 2 から外部ボリューム 3 0 1 内の管理領域 3 0 1 - 2 への書き込みをキャッシュスループードに設定しておき、管理領域 3 0 1 - 2 内の管理情報を常時最新にしておく必要がある。外部ボリューム 3 0 1 内の管理領域 3 0 1 - 2 に書き込まれている管理情報の更新は、移動先のストレージシステム 1 3 0 b によって行われる。

20

【 0 2 5 4 】

(8) ストレージシステム 1 3 0 a は、ストレージシステム 1 3 0 a のディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム 3 0 1 にデステージするためにストレージシステム 1 3 0 a の動作モードをキャッシュスループードに設定する。尚、ストレージシステム 1 3 0 a は、自発的に動作モードをキャッシュスループードに設定してもよく、或いは外部 (例えば、ストレージシステム 1 3 0 b 、ホスト計算機 1 0 0 、又は管理サーバ 1 1 0 など) からの指示に基づいて、動作モードをキャッシュスループードに設定してもよい。

30

【 0 2 5 5 】

(9) ストレージシステム 1 3 0 a のディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム 3 0 1 にデステージできたならば、ストレージシステム 1 3 0 b は、拡張デバイス 2 0 6 c と仮想デバイス 2 0 3 b との間のマッピング関係を解除し、拡張デバイス 2 0 6 b 内のユーザ領域 2 0 6 b - 1 を仮想デバイス 2 0 3 b 内のユーザ領域 2 0 3 b - 1 にマッピングする。更に、ストレージシステム 1 3 0 b は、パス 4 0 4 を削除する。

40

【 0 2 5 6 】

尚、図 3 8 に示す方法は、外部ボリューム 3 0 1 が V M A 付きボリュームの場合だけでなく、メインフレームボリュームの場合にも、上述の処理手順 (1) ~ (9) と同様の手順を経て論理ユニット 2 0 1 a を移動させることができる。

【 0 2 5 7 】

次に、図 3 9 を参照しながら、V M A 付きボリュームの移動処理の他の方法について説明する。

【 0 2 5 8 】

ストレージシステム 1 3 0 a は、論理ユニット 2 0 1 a - 1 , 2 0 2 a - 1 , 論理デバ

50

イス 202 a, 仮想デバイス 203 a, 及び拡張デバイス 206 a を備える。論理ユニット 201 a - 1 は、パス 401 を介してホスト計算機 100 に接続されており、ホスト計算機 100 から認識可能である。一方、論理ユニット 201 a - 2 は、ホスト計算機 100 に接続されておらず、ホスト計算機 100 からの認識は不可能である。論理ユニット 201 a - 2 は、論理ユニット 201 a - 1 をストレージシステム 130 a からストレージシステム 130 b に移動させる間のときにのみ一時的に作成される。論理デバイス 202 a は、VMA 付きボリュームであり、ホスト計算機 100 からアクセス可能なユーザ領域 202 a - 1 と、管理情報を格納するための管理領域 202 a - 2 とを含む。管理情報には、例えば、VMA 付きボリュームのアクセス属性が含まれる。ユーザ領域 202 a - 1 は、論理ユニット 201 a - 1 に割り当てられることにより、ホスト計算機 100 により認識可能である。管理領域 202 a - 2 は論理ユニット 201 a - 2 に割り当てられているが、ホスト計算機 100 から認識することはできない。仮想デバイス 203 a は、ホスト計算機 100 からアクセス可能なユーザ領域 203 a - 1 と、管理情報を格納するための管理領域 203 a - 2 とを含む。拡張デバイス 206 a は、ホスト計算機 100 からアクセス可能なユーザ領域 206 a - 1 と、管理情報を格納するための管理領域 206 a - 2 とを含む。拡張デバイス 206 a 内のユーザ領域 206 a - 1 は、外部ボリューム 301 内のユーザ領域 301 - 1 を仮想化したものである。拡張デバイス 206 a 内の管理領域 206 a - 2 は、外部ボリューム 301 内の管理領域 301 - 2 を仮想化したものである。ユーザ領域 206 a - 1 は、ユーザ領域 203 a - 1 にマッピングされている。管理領域 206 a - 2 は、管理領域 203 a - 2 にマッピングされている。ユーザ領域 203 a - 1 は、ユーザ領域 202 a - 1 にマッピングされている。管理領域 203 a - 2 は、管理領域 202 a - 2 にマッピングされている。

10

20

【0259】

ストレージシステム 130 a (移動元ストレージ) 内の論理ユニット 201 a (移動元論理ユニット) にあるデータをストレージシステム 130 b (移動先ストレージ) 内の論理ユニット 201 b (移動先論理ユニット) に移動する処理について概説する。論理ユニット移動処理は、下記の (1) ~ (9) の処理手順を含む。

【0260】

(1) ストレージシステム 130 b は、外部ボリューム 301 と、ストレージシステム 130 b との間を外部接続するためのパス 405 を定義し、外部ボリューム 301 を仮想化する拡張デバイス 206 b をストレージシステム 130 b 内に作成する。拡張デバイス 206 b は、ホスト計算機 100 からアクセス可能なユーザ領域 206 b - 1 と、管理情報を格納するための管理領域 206 b - 2 とを含む。

30

【0261】

(2) ストレージシステム 130 b は、外部ボリューム 301 を仮想化するための仮想デバイス 203 b をストレージシステム 130 b 内に作成する。仮想デバイス 203 b は、ホスト計算機 100 からアクセス可能なユーザ領域 203 b - 1 と、管理情報を格納するための管理領域 203 b - 2 とを含む。

【0262】

(3) ストレージシステム 130 b は、論理ユニット 201 a - 1 とストレージシステム 130 b との間を外部接続するためのパス 404 - 1, 論理ユニット 201 a - 2 とストレージシステム 130 b との間を外部接続するためのパス 404 - 2 とをそれぞれ定義し、論理ユニット 201 a - 1 を仮想化するための拡張デバイス 206 c - 1, 論理ユニット 201 a - 2 を仮想化するための拡張デバイス 206 c - 2 をそれぞれストレージシステム 130 b 内に作成する。

40

【0263】

(4) ストレージシステム 130 b は、拡張デバイス 206 c - 1 を仮想デバイス 203 b 内のユーザ領域 203 b - 1 にマッピングし、拡張デバイス 206 c - 2 を仮想デバイス 203 b 内の管理領域 203 b - 2 にマッピングする。

【0264】

50

(5) ストレージシステム 130b は、論理デバイス 202a の論理構成と同一の論理構成を有する論理デバイス 202b をストレージシステム 130b 内に作成する。論理デバイス 202b は、ホスト計算機 100 からアクセス可能なユーザ領域 202b-1 と、管理情報を格納するための管理領域 202b-2 とを含む。ユーザ領域 203b-1 は、ユーザ領域 202b-1 にマッピングされる。管理領域 203b-2 は、管理領域 202b-2 にマッピングされる。

【0265】

(6) ストレージシステム 130b は、論理ユニット 201a-1 の論理構成と同一の論理構成を有する論理ユニット 201b をストレージシステム 130b 内に作成し、ホスト計算機 100 と論理ユニット 201b とを接続するためのパス 403 を定義する。ユーザ領域 202b-1 は、論理ユニット 201b にマッピングされる。

10

【0266】

(7) ホスト計算機 100 は、外部ボリューム 301 にアクセスするための経路として、パス 401 からパス 403 に切り替える。このとき、パス 401 を削除する等して、ホスト計算機 100 から論理ユニット 201a-1 へのアクセスを禁止する。尚、パス 401 からパス 403 へのパス切り替えは、ストレージシステム 130a, 130b, 又は管理サーバ 110 等が行っても良い。また、必ずしもパス 401 を削除する必要はなく、例えば、パス 401 をパス 403 の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス 401 の状態を無効としたまま、パス 401 を残しても良い。

【0267】

20

上記の処理により、論理ユニット 201b は、論理ユニット 201a-1 を仮想化したものであるだけでなく、外部ボリューム 301 内のユーザ領域 301-1 をも仮想化したものである。論理ユニット移動処理中における、ホスト計算機 100 から論理ユニット 201a へのデータ入出力経路は、パス 403 ストレージシステム 130b 内の記憶階層(LU/LDEV/VDEV/EDEV) パス 404-1 ストレージシステム 130a 内の記憶階層(LU/LDEV/VDEV/EDEV) パス 402 外部ボリューム 301 となる。

【0268】

尚、ストレージシステム 130a は、外部ボリューム 130 内の管理領域 301-1 を仮想デバイス 203a-1 によって仮想化しているため、パス 402 を介して管理情報を参照することができる。論理ユニット 201a の移動処理中において、ストレージシステム 130b が外部ボリューム 130 内の管理情報を参照するためには、パス 404-2 及び論理ユニット 201a-2 を介して外部ボリューム 130 内の管理領域 301-2 にアクセスすればよい。外部ボリューム 301 内の管理領域 301-2 に書き込まれている管理情報の更新は、移動元のストレージシステム 130a によって行われる。

30

【0269】

(8) ストレージシステム 130a は、ストレージシステム 130a のディスクキャッシュに蓄積されている全てのダーティデータを外部ボリューム 301 にデステージするためにストレージシステム 130a の動作モードをキャッシュスルーモードに設定する。尚、ストレージシステム 130a は、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部(例えば、ストレージシステム 130b、ホスト計算機 100、又は管理サーバ 110 など)からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

40

【0270】

(9) ストレージシステム 130a のディスクキャッシュに蓄積されているダーティデータの全てを外部ボリューム 301 にデステージできたならば、ストレージシステム 130b は、拡張デバイス 206c-1 とユーザ領域 203b-1 との間のマッピング関係を解除するとともに、拡張デバイス 206c-2 と管理領域 203b-2 との間のマッピング関係を解除する。更に、ストレージシステム 130b は、拡張デバイス 206b 内のユーザ領域 206b-1 をユーザ領域 203b-1 にマッピングするとともに、管理領域 20

50

6 b - 2を管理領域 2 0 3 b - 2にマッピングする。更に、ストレージシステム 1 3 0 b は、パス 4 0 4を削除する。

【 0 2 7 1 】

尚、図 3 9に示す方法は、外部ボリューム 3 0 1が V M A 付きボリュームの場合だけでなく、メインフレームボリュームの場合にも、上述の処理手順 (1) ~ (9)と同様の手順を経て論理ユニット 2 0 1 a - 1を移動させることができる。

【 0 2 7 2 】

図 3 8に示す方法と図 3 9に示す方法との間の主な相違点は次の通りである。図 3 9に示す方法では、移動先のストレージシステム 1 3 0 bは、移動元のストレージシステム 1 3 0 aを介して管理領域 3 0 1 - 2にアクセスするが、図 3 8に示す方法では、移動先のストレージシステム 1 3 0 bは、管理領域 3 0 1 - 2に直接アクセスする。図 3 9に示す方法では、管理領域 3 0 1 - 2内の管理情報の更新は、移動元のストレージシステム 1 3 0 aが実行するが、図 3 8に示す方法では、管理領域 3 0 1 - 2内の管理情報の更新は、移動先のストレージシステム 1 3 0 bが実行する。

【 0 2 7 3 】

次に、図 4 0を参照しながら、V M A 付きボリュームの移動処理の他の方法について説明する。

【 0 2 7 4 】

ストレージシステム 1 3 0 aは、論理ユニット 2 0 1 a , 2 0 1 a' , 論理デバイス 2 0 2 a , 仮想デバイス 2 0 3 a , 及び拡張デバイス 2 0 6 aを備える。論理ユニット 2 0 1 aは、パス 4 0 1を介してホスト計算機 1 0 0に接続されており、ホスト計算機 1 0 0から認識可能である。一方、論理ユニット 2 0 1 a'は、ホスト計算機 1 0 0に接続されておらず、ホスト計算機 1 0 0からの認識は不可能である。論理ユニット 2 0 1 a'は、論理ユニット 2 0 1 aをストレージシステム 1 3 0 aからストレージシステム 1 3 0 bに移動させる間のみ一時的に作成される。論理ユニット 2 0 1 a'は、ユーザ領域 2 0 1 a - 1'と、管理領域 2 0 1 a - 2'を含む。論理デバイス 2 0 2 aは、V M A 付きボリュームであり、ホスト計算機 1 0 0からアクセス可能なユーザ領域 2 0 2 a - 1と、管理情報を格納するための管理領域 2 0 2 a - 2を含む。管理情報には、例えば、V M A 付きボリュームのアクセス属性が含まれる。ユーザ領域 2 0 2 a - 1は、論理ユニット 2 0 1 aに割り当てられることにより、ホスト計算機 1 0 0により認識可能である。ユーザ領域 2 0 2 a - 1は、論理ユニット 2 0 1 a'内のユーザ領域 2 0 1 a - 1'にも割り当てられる。管理領域 2 0 2 a - 2は、論理ユニット 2 0 1 a'内の管理領域 2 0 1 a - 2'に割り当てられる。仮想デバイス 2 0 3 aは、ホスト計算機 1 0 0からアクセス可能なユーザ領域 2 0 3 a - 1と、管理情報を格納するための管理領域 2 0 3 a - 2を含む。拡張デバイス 2 0 6 aは、ホスト計算機 1 0 0からアクセス可能なユーザ領域 2 0 6 a - 1と、管理情報を格納するための管理領域 2 0 6 a - 2を含む。拡張デバイス 2 0 6 a内のユーザ領域 2 0 6 a - 1は、外部ボリューム 3 0 1内のユーザ領域 3 0 1 - 1を仮想化したものである。拡張デバイス 2 0 6 a内の管理領域 2 0 6 a - 2は、外部ボリューム 3 0 1内の管理領域 3 0 1 - 2を仮想化したものである。ユーザ領域 2 0 6 a - 1は、ユーザ領域 2 0 3 a - 1にマッピングされている。管理領域 2 0 6 a - 2は、管理領域 2 0 3 a - 2にマッピングされている。ユーザ領域 2 0 3 a - 1は、ユーザ領域 2 0 2 a - 1にマッピングされている。管理領域 2 0 3 a - 2は、管理領域 2 0 2 a - 2にマッピングされている。

【 0 2 7 5 】

ストレージシステム 1 3 0 a (移動元ストレージ) 内の論理ユニット 2 0 1 a (移動元論理ユニット) にあるデータをストレージシステム 1 3 0 b (移動先ストレージ) 内の論理ユニット 2 0 1 b (移動先論理ユニット) に移動する処理について概説する。論理ユニット移動処理は、下記の (1) ~ (9) の処理手順を含む。

【 0 2 7 6 】

(1) ストレージシステム 1 3 0 bは、外部ボリューム 3 0 1と、ストレージシステム 1

10

20

30

40

50

30bとの間を外部接続するためのパス405を定義し、外部ボリューム301を仮想化する拡張デバイス206bをストレージシステム130b内に作成する。拡張デバイス206bは、ホスト計算機100からアクセス可能なユーザ領域206b-1と、管理情報を格納するための管理領域206b-2とを含む。

【0277】

(2) ストレージシステム130bは、外部ボリューム301を仮想化するための仮想デバイス203bをストレージシステム130b内に作成する。仮想デバイス203bは、ホスト計算機100からアクセス可能なユーザ領域203b-1と、管理情報を格納するための管理領域203b-2とを含む。

【0278】

(3) ストレージシステム130bは、論理ユニット201a'とストレージシステム130bとの間を外部接続するためのパス404を定義し、論理ユニット201a'を仮想化するための拡張デバイス206cをストレージシステム130b内に作成する。拡張デバイス206cは、ホスト計算機100からアクセス可能なユーザ領域206c-1と、管理情報を格納するための管理領域206c-2とを含む。

【0279】

(4) ストレージシステム130bは、拡張デバイス206c内のユーザ領域206c-1を仮想デバイス203b内のユーザ領域203b-1にマッピングし、拡張デバイス206c内の管理領域206c-2を仮想デバイス203b内の管理領域203b-2にマッピングする。

【0280】

(5) ストレージシステム130bは、論理デバイス202aの論理構成と同一の論理構成を有する論理デバイス202bをストレージシステム130b内に作成する。論理デバイス202bは、ホスト計算機100からアクセス可能なユーザ領域202b-1と、管理情報を格納するための管理領域202b-2とを含む。ユーザ領域203b-1は、ユーザ領域202b-1にマッピングされる。管理領域203b-2は、管理領域202b-2にマッピングされる。

【0281】

(6) ストレージシステム130bは、論理ユニット201aの論理構成と同一の論理構成を有する論理ユニット201bをストレージシステム130b内に作成し、ホスト計算機100と論理ユニット201bとを接続するためのパス403を定義する。ユーザ領域202b-1は、論理ユニット201bにマッピングされる。

【0282】

(7) ホスト計算機100は、外部ボリューム301にアクセスするための経路として、パス401からパス403に切り替える。このとき、パス401を削除する等して、ホスト計算機100から論理ユニット201a-1へのアクセスを禁止する。尚、パス401からパス403へのパス切り替えは、ストレージシステム130a, 130b, 又は管理サーバ110等が行っても良い。また、必ずしもパス401を削除する必要はなく、例えば、パス401をパス403の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス401の状態を無効としたまま、パス401を残しても良い。

【0283】

上記の処理により、論理ユニット201bは、論理ユニット201aを仮想化したものであるだけでなく、外部ボリューム301内のユーザ領域301-1をも仮想化したものであるので、論理ユニット移動処理中における、ホスト計算機100から論理ユニット201aへのデータ入出力経路は、パス403 ストレージシステム130b内の記憶階層(LU/LDEV/VDEV/EDEV) パス404 ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV) パス402 外部ボリューム301となる。

【0284】

尚、ストレージシステム130aは、外部ボリューム130内の管理領域301-1を

10

20

30

40

50

仮想デバイス 203 a - 1 によって仮想化しているため、パス 402 を介して管理情報を参照することができる。論理ユニット 201 a の移動処理中において、ストレージシステム 130 b が外部ボリューム 130 内の管理情報を参照するためには、パス 404 及び論理ユニット 201 a' を介して外部ボリューム 130 内の管理領域 301 - 2 にアクセスすればよい。外部ボリューム 301 内の管理領域 301 - 2 に書き込まれている管理情報の更新は、移動元のストレージシステム 130 a によって行われる。

【0285】

(8) ストレージシステム 130 a は、ストレージシステム 130 a のディスクキャッシュに蓄積されている全てのパーティデータを外部ボリューム 301 にデステージするためにストレージシステム 130 a の動作モードをキャッシュスルーモードに設定する。尚、ストレージシステム 130 a は、自発的に動作モードをキャッシュスルーモードに設定してもよく、或いは外部(例えば、ストレージシステム 130 b、ホスト計算機 100、又は管理サーバ 110 など)からの指示に基づいて、動作モードをキャッシュスルーモードに設定してもよい。

10

【0286】

(9) ストレージシステム 130 a のディスクキャッシュに蓄積されているパーティデータの全てを外部ボリューム 301 にデステージできたならば、ストレージシステム 130 b は、拡張デバイス 206 c と仮想デバイス 203 b との間のマッピング関係を解除するとともに、拡張デバイス 206 b を仮想デバイス 203 b にマッピングする。より詳細には、ストレージシステム 130 b は、拡張デバイス 206 b 内のユーザ領域 206 b - 1 を仮想デバイス 203 b 内のユーザ領域 203 b - 1 にマッピングするとともに、拡張デバイス 206 b 内の管理領域 206 b - 2 を仮想デバイス 203 b 内の管理領域 203 b - 2 にマッピングする。更に、ストレージシステム 130 b は、パス 404 を削除する。

20

【0287】

次に、図 41 を参照しながら、複数の論理ユニットを移動させる処理の概要について説明する。

【0288】

ストレージシステム 130 a は、複数の論理ユニット 201 a - 1, 201 a - 2, 複数の論理デバイス 202 a - 1, 202 a - 2, 複数の仮想デバイス 203 a - 1, 203 a - 2, 及び複数の拡張デバイス 206 a - 1, 206 a - 2 を備えている。拡張デバイス 206 a - 1 は、外部ボリューム 301 - 1 を仮想化したものであり、パス 402 - 1 を介して外部ストレージシステム 150 に接続している。拡張デバイス 206 a - 1 は、仮想デバイス 203 a - 1 にマッピングされている。仮想デバイス 203 a - 1 は、論理デバイス 202 a - 1 にマッピングされている。論理デバイス 202 a - 1 は論理ユニット 201 a - 1 にマッピングされている。論理ユニット 201 a - 1 は、外部ボリューム 301 - 1 を仮想化したものであり、パス 401 - 1 を介してホスト計算機 100 に接続されている。

30

【0289】

一方、拡張デバイス 206 a - 2 は、外部ボリューム 301 - 2 を仮想化したものであり、パス 402 - 2 を介して外部ストレージシステム 150 に接続している。拡張デバイス 206 a - 2 は、仮想デバイス 203 a - 2 にマッピングされている。仮想デバイス 203 a - 2 は、論理デバイス 202 a - 2 にマッピングされている。論理デバイス 202 a - 2 は、論理ユニット 201 a - 2 にマッピングされている。論理ユニット 201 a - 2 は、外部ボリューム 301 - 2 を仮想化したものであり、パス 401 - 2 を介してホスト計算機 100 に接続されている。

40

【0290】

さて、上記の構成において、複数の論理ユニット 201 a - 1, 201 a - 2 をストレージシステム 130 a からストレージシステム 130 b へ移動するには、上述の処理手順と同様に、ストレージシステム 130 b は、移動先となる複数の論理ユニット 201 b - 1, 201 b - 2 を作成する。論理ユニット 201 b - 1 は、パス 403 - 1 を介してホ

50

スト計算機 100 に接続する。論理ユニット 201b-2 は、パス 403-2 を介してホスト計算機 100 に接続する。更に、ストレージシステム 130b は、移動元の複数の論理ユニット 201a-1, 201a-2 を仮想化するため複数の拡張デバイス 206c-1, 206c-2, 複数の仮想デバイス 203b-1, 203b-2, 複数の論理デバイス 202b-1, 202b-2 をそれぞれ作成する。拡張デバイス 206c-1 は、論理ユニット 201a-1 の移動処理中においては、論理ユニット 201a-1 を仮想化する一方で、論理ユニット 201a-1 の移動処理後（より詳細には、仮想デバイス 203a-1 に書き込まれたダーティデータのデステージ完了後）においては、パス 405-1 を介して外部ボリューム 301-1 に接続し、外部ボリューム 301-1 を仮想化する。拡張デバイス 206c-2 は、論理ユニット 201a-2 の移動処理中においては、論理ユニット 201a-2 を仮想化する一方で、論理ユニット 201a-2 の移動処理後（より詳細には、仮想デバイス 203a-2 に書き込まれたダーティデータのデステージ完了後）においては、パス 405-2 を介して外部ボリューム 301-2 に接続し、外部ボリューム 301-2 を仮想化する。

10

【0291】

尚、複数の論理ユニットを移動させるための指示は、管理サーバ 110（図 2 参照）から各ストレージシステム 130a, 130b に与えてもよく、或いはホスト計算機 100 から各ストレージシステム 130a, 130b に与えても良い。

【0292】

次に、図 4 2 乃至図 4 6 を参照しながら論理ユニット移動処理の詳細について説明を加える。

20

【0293】

図 4 2 は論理ユニット移動指示処理ルーチンを示す。論理ユニット移動指示処理ルーチンは、管理サーバ 110 内の論理ユニット移動指示処理プログラム 241 によって実行される。

【0294】

管理サーバ 110 がストレージシステム 130a（移動元ストレージ）、及びストレージシステム 130b（移動先ストレージ）に対して、一つ以上の論理ユニットを移動させる指示を与える（ステップ 4201）。このとき、管理サーバ 110 を操作するストレージ管理者によって、各ストレージシステム 130a, 130b に論理ユニット移動指示が与えられてもよく、或いは管理サーバ 110 内の管理ソフトウェアが各ストレージシステム 130a, 130b に論理ユニット移動指示を与えても良い。尚、論理ユニットの移動指示は、ホスト計算機 100 から各ストレージシステム 130a, 130b に与えても良い。

30

【0295】

各ストレージシステム 130a, 130b は、論理ユニット移動指示を受けると、指定された論理ユニットを移動させるための処理を実行する（ステップ 4202）。

【0296】

指定された全ての論理ユニットのうち一部の論理ユニットの移動が未だ完了していない場合には（ステップ 4203；NO）、各ストレージシステム 130a, 130b は、論理ユニット移動処理を実行する（ステップ 4202）。

40

【0297】

指定された全ての論理ユニットの移動が完了すると（ステップ 4203；YES）、各ストレージシステム 130a, 130b は、管理サーバ 110 に移動完了報告をする（ステップ 4204）。尚、論理ユニット移動指示がホスト計算機 100 から各ストレージシステム 130a, 130b に与えられる場合には、各ストレージシステム 130a, 130b は、ホスト計算機 100 に移動完了報告をする。

【0298】

図 4 3 は論理ユニット移動処理サブルーチンを示す。論理ユニット移動処理サブルーチンは、ストレージシステム 130 内の論理ユニット移動処理プログラム 222 によって実

50

行される。

【0299】

論理ユニット移動処理のサブルーチンが起動すると、各ストレージシステム130a, 130bは、移動体対象選出処理(ステップ4301)、移動準備処理(ステップ4302)、及び移動処理(ステップ4303)を実行する。

【0300】

図44は移動対象選出処理サブルーチンを示す。移動対象選出処理サブルーチンが起動すると、ストレージシステム130aは、移動対象となる論理ユニットに対応する論理デバイス(移動対象となる論理デバイス)を選出する(ステップ4401)。例えば、LUS E構成では、一つの論理ユニットにつき複数の論理デバイスが選出される。

10

【0301】

次に、ストレージシステム130aは、移動対象となる論理デバイスを使用する全てのホスト計算機100を選出する(ステップ4402)。例えば、同一のホストグループに属するホスト計算機100のみならず、他のホストグループに属するホスト計算機100も選出する。

【0302】

次に、ストレージシステム130aは、移動対象となる論理デバイスに対応するVDEV領域(移動対象となるVDEV領域)を選出する(ステップ4403)。VDEV領域は、一つの仮想デバイスの一部に対応する場合もあれば、複数の仮想デバイスに対応する場合もある。例えば、CVS構成では、移動対象となる論理デバイスに対応するVDEV領域として、仮想デバイスの一部が選出される。一方、VDEV連結構成では、移動対象となる論理デバイスに対応するVDEV領域として、複数の仮想デバイスが選出される。

20

【0303】

次に、ストレージシステム130aは、移動対象となるVDEV領域に対応するデバイスグループ(移動対象となるデバイスグループ)を選出する(ステップ4404)。デバイスグループは、一つのVDEV領域の一部に対応する場合もあれば、複数のVDEV領域に対応する場合もある。例えば、VDEV離散構成では、移動対象となるVDEV領域に対応するデバイスグループとして、複数のデバイスグループが選出される。

【0304】

次に、ストレージシステム130aは、移動対象となるデバイスグループに対応する拡張デバイス(移動対象となる拡張デバイス)を選出する(ステップ4405)。拡張デバイスは、一つのデバイスグループの一部に対応する場合もあれば、複数のデバイスグループに対応する場合もある。例えば、一つの拡張デバイスの一部の記憶領域はあるデバイスグループに割り当てられており、他の記憶領域は他のデバイスグループに割り当てられているような場合には、移動対象となるデバイスグループに対応する拡張デバイスとして、拡張デバイスの一部が選出される。一方、外部ボリュームRAID構成では、移動対象となるデバイスグループに対応する拡張デバイスとして、複数の拡張デバイスが選出される。

30

【0305】

次に、ストレージシステム130aは、移動対象となる拡張デバイスに対応する外部ボリュームを選出する(ステップ4406)。

40

【0306】

上記の処理により、移動対象となる論理ユニットに対応する論理デバイス、その論理デバイスに対応する仮想デバイス、その仮想デバイスに対応するデバイスグループ、そのデバイスグループに対応する拡張デバイス、その拡張デバイスに対応する外部ボリュームをそれぞれ選出することができる。

【0307】

図45は移動準備処理サブルーチンを示す。移動準備処理サブルーチンが起動すると、ストレージシステム130aは、ストレージシステム130bから外部ボリュームへのアクセス許可を設定する(ステップ4501)。

50

【 0 3 0 8 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b から外部ボリュームに接続するためのパスを定義し、外部ボリュームを仮想化するための第一の拡張デバイス（図 1 の拡張デバイス 2 0 6 b に相当）をストレージシステム 1 3 0 b 内に作成する（ステップ 4 5 0 2）。

【 0 3 0 9 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b 内の第一の拡張デバイスをストレージシステム 1 3 0 b 内のデバイスグループに対応付ける（ステップ 4 5 0 3）。例えば、外部ボリューム R A I D 構成では、複数の第一の拡張デバイスが一つのデバイスグループに対応付けられる。

10

【 0 3 1 0 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b 内の第一の拡張デバイスに対応付けられているデバイスグループをストレージシステム 1 3 0 b 内の仮想デバイスに対応付ける（ステップ 4 5 0 4）。

【 0 3 1 1 】

ストレージシステム 1 3 0 a は、ストレージシステム 1 3 0 b からストレージシステム 1 3 0 a 内の移動元論理ユニットへのアクセス許可を設定する（ステップ 4 5 0 5）。

【 0 3 1 2 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b 内の第二の拡張デバイス（図 1 の拡張デバイス 2 0 6 c に相当）をストレージシステム 1 3 0 a 内の移動対象となる論理ユニットに対応付ける（ステップ 4 5 0 6）。

20

【 0 3 1 3 】

例えば、移動対象となる論理ユニットが外部ボリューム R A I D 構成を有している場合には、ストレージシステム 1 3 0 b 内の第二の拡張デバイスは、ストレージシステム 1 3 0 a 内の仮想デバイスの一部（パリティデータ及びミラーデータを格納している記憶領域を除く。）に対応付けられる。

【 0 3 1 4 】

また例えば、移動対象となる論理ユニットが C V S 構成を有している場合には、ストレージシステム 1 3 0 b 内の第二の拡張デバイスは、ストレージシステム 1 3 0 a 内の仮想デバイスの一部に対応付けられる。

30

【 0 3 1 5 】

また例えば、移動対象となる論理ユニットが L U S E 構成を有している場合には、ストレージシステム 1 3 0 b 内の第二の拡張デバイスは、ストレージシステム 1 3 0 a 内の複数の仮想デバイスに対応付けられる。

【 0 3 1 6 】

また例えば、移動対象となる論理ユニットが V M A 付きボリュームである場合には、ストレージシステム 1 3 0 b 内の第二の拡張デバイスは、ストレージシステム 1 3 0 a 内の仮想デバイスの一部（ユーザ領域）に対応付けられる。

【 0 3 1 7 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b 内の第二の拡張デバイスをストレージシステム 1 3 0 b 内のデバイスグループに対応付ける（ステップ 4 5 0 7）。

40

【 0 3 1 8 】

ストレージシステム 1 3 0 b は、ストレージシステム 1 3 0 b 内の第二の拡張デバイスに対応付けられているデバイスグループをストレージシステム 1 3 0 b 内の仮想デバイスに対応付ける（ステップ 4 5 0 8）。

【 0 3 1 9 】

例えば、移動対象となる論理ユニットが外部ボリューム R A I D 構成を有している場合には、ストレージシステム 1 3 0 b 内の第二の拡張デバイスは、ストレージシステム 1 3 0 b 内の仮想デバイスの一部（パリティデータ及びミラーデータを格納している記憶領域

50

を除く。)に対応付けられる。

【0320】

また例えば、移動対象となる論理ユニットがCVS構成を有している場合には、ストレージシステム130b内の第二の拡張デバイスは、ストレージシステム130b内の仮想デバイスの一部に対応付けられる。

【0321】

また例えば、移動対象となる論理ユニットがLUSE構成を有している場合には、ストレージシステム130b内の第二の拡張デバイスは、ストレージシステム130b内の複数の仮想デバイスに対応付けられる。

【0322】

また例えば、移動対象となる論理ユニットがVMA付きボリュームである場合には、ストレージシステム130b内の第二の拡張デバイスは、ストレージシステム130b内の仮想デバイスの一部(ユーザ領域)に対応付けられる。

【0323】

ストレージシステム130bは、第一の拡張デバイスに対応付けられているストレージシステム130b内のデバイスグループとストレージシステム130b内の仮想デバイスとの間のマッピング関係を無効とし、第二の拡張デバイスに対応付けられているストレージシステム130b内のデバイスグループとストレージシステム130b内の仮想デバイスとの間のマッピング関係を有効とする(ステップ4509)。

【0324】

ストレージシステム130bは、ストレージシステム130bの動作モードをキャッシュスルーモードに設定し、ストレージシステム130aの動作モードをライトアフタモードに設定する(ステップ4510)。

【0325】

ストレージシステム130bは、ストレージシステム130b内に移動先の論理ユニットを作成する(ステップ4511)。

【0326】

ストレージシステム130bは、移動先の論理ユニットにアクセス属性を設定する(ステップ4512)。アクセス属性として、例えば、ホストグループに関する設定や交替パスに関する設定などがある。

【0327】

ストレージシステム130bは、ストレージシステム130bから外部ボリュームに書き込まれるデータを暗号化する機能を解除する(ステップ4513)。論理ユニットの移動処理中において、外部ボリュームへ書き込まれるデータは、移動元のストレージシステム130aによって暗号化される。

【0328】

ストレージシステム130bは、移動先の論理ユニットとホスト計算機100とを接続するためのパスを定義する(ステップ4514)。移動先の論理ユニットが複数存在する場合には、それぞれの移動先の論理ユニットとホスト計算機100とを接続するためのパスを定義する。

【0329】

ストレージシステム130bは、移動先の論理ユニットをホスト計算機100に認識させる(ステップ4515)。移動対象となる論理ユニットを使用するホスト計算機100が複数存在する場合には、それぞれのホスト計算機100に移動先の論理ユニットを認識させる。

【0330】

ストレージシステム130bは、移動先の論理ユニットとホスト計算機100とを接続するパスを、移動元の論理ユニットとホスト計算機100とを接続するパスの交替パスとして設定する(ステップ4516)。

【0331】

10

20

30

40

50

図46は移動処理サブルーチンを示す。移動処理サブルーチンが起動すると、ホスト計算機100は、移動元の論理ユニットへのI/O要求を停止する(ステップ4601)。ストレージシステム130aがホスト計算機100からのI/O要求を禁止してもよい。移動元の論理ユニットを使用するホスト計算機100が複数存在する場合には、それぞれのホスト計算機100から移動元の論理ユニットへのI/O要求を停止又は禁止する。

【0332】

ホスト計算機100は、移動元の論理ユニットに接続するパスから移動先の論理ユニットに接続するパスに切り替えるとともに、移動元の論理ユニットに接続するパスを削除する(ステップ4602)。移動元の論理ユニットを使用するホスト計算機100が複数存在する場合には、それぞれのホスト計算機100は、パス切り替えを同時に実行する。

10

【0333】

移動元のストレージシステム130aは、ホスト計算機100からのアクセスを禁止する(ステップ4603)。

【0334】

移動先のストレージシステム130bは、ストレージシステム130bの動作モードとして、キャッシュスルーモードを解除し、ライトアフタモードに設定する(ステップ4604)。

【0335】

移動元のストレージシステム130aは、ストレージシステム130aの動作モードとして、ライトアフタモードを解除し、キャッシュスルーモードに設定する(ステップ4605)。

20

【0336】

移動元のストレージシステム130aのディスクキャッシュに蓄積されているデータデータを全てを外部ボリュームにデステージできたならば(ステップ4606;YES)、移動先のストレージシステム130bは、第二の拡張デバイスに対応付けられているストレージシステム130b内のデバイスグループとストレージシステム130b内の仮想デバイスとの間のマッピング関係を無効とし、第一の拡張デバイスに対応付けられているストレージシステム130b内のデバイスグループとストレージシステム130b内の仮想デバイスとの間のマッピング関係を有効とする(ステップ4607)。

【0337】

30

ストレージシステム130bは、ストレージシステム130bから外部ボリュームへのアクセスを許可する一方で、ストレージシステム130aから外部ボリュームへのアクセスを禁止する設定(アクセス排他設定)を行う(ステップ4608)。

【0338】

ストレージシステム130bは、外部ボリュームに書き込まれるデータを暗号化する機能をアクティブに設定する(ステップ4609)。

【0339】

ストレージシステム130bは、ストレージシステム130b内の第二の拡張デバイスとストレージシステム130a内の移動元の論理ユニットとを接続するパスを削除する(ステップ4610)。

40

【0340】

移動元のストレージシステム130aと、移動先のストレージシステム130bとによって、外部ボリュームを共有する場合(例えば、移動元の論理ユニットがCVS構成を有している場合には(ステップ4611;YES)、移動元のストレージシステム130aは、移動処理後の論理ユニットとストレージシステム130a内の仮想デバイスとの間のマッピング関係を無効にする(ステップ4612)。

【0341】

移動元のストレージシステム130aと、移動先のストレージシステム130bとによって、外部ボリュームを共有しない場合には(ステップ4611;NO)、ストレージシステム130aは、ストレージシステム130aと外部ボリュームとを接続するパスを削

50

除する（ステップ4613）。

【0342】

移動元のストレージシステム130aは、外部ボリュームを仮想化するためにストレージシステム130a内に作成された拡張デバイス、その拡張デバイスに対応付けられているデバイスグループ、そのデバイスグループに対応付けられている仮想デバイス、その仮想デバイスに対応付けられている論理デバイス、その論理デバイスに対応付けられている論理ユニット、その論理ユニットとホスト計算機100とを接続するためのパスを何れも全て削除する（ステップ4614）。

【0343】

外部ストレージシステム150は、ストレージシステム130aから外部ボリュームへのアクセスを禁止する（ステップ4615）。

10

【0344】

図47はストレージI/O処理ルーチンを示す。ストレージI/O処理ルーチンは、ストレージI/O処理プログラム221によって実行される。

【0345】

ホスト計算機100からのI/O要求を受領したストレージシステム130は、論理デバイス管理情報603を参照し、ホスト計算機100のアクセス権を確認する（ステップ4701）。

【0346】

ホスト計算機100がアクセス権を有しない場合には（ステップ4701；NO）、ストレージシステム130は、I/O要求を拒否する（ステップ4708）。

20

【0347】

ホスト計算機100がアクセス権を有する場合には（ステップ4701；YES）、ストレージシステム130は、論理デバイス管理情報603を参照し、キャッシュスルードに設定されているか否かを判断する（ステップ4702）。

【0348】

ストレージシステム130の動作モードがキャッシュスルードであり（ステップ4702；YES）、且つホスト計算機100からのI/O要求がライト要求である場合には（ステップ4703；YES）、ストレージシステム130は、ホスト計算機100から受信したデータをディスクキャッシュ134に書き込む（ステップ4704）。

30

【0349】

その後、ストレージシステム130は、データを外部ボリュームに書き込み（ステップ4705）、キャッシュデータの管理情報をデステージ完了済みに更新し（ステップ4706）、ホスト計算機100にライト完了報告をする（ステップ4707）。

【0350】

ストレージシステム130の動作モードがキャッシュスルードでない場合は（ステップ4702；NO）、ストレージシステム130は、アクセス先の論理ユニットが移動中であるか否かを判定する（ステップ4710）。

【0351】

アクセス先の論理ユニットが移動中でない場合には（ステップ4710；NO）、ストレージシステム130は、移行元の論理ユニットに対してI/O処理を実行し（ステップ4709）、ホスト計算機100にライト完了報告をする（ステップ4707）。

40

【0352】

一方、アクセス先の論理ユニットが移動中である場合には（ステップ4710；YES）、ストレージシステム130は、アクセスアドレスに対応する記憶領域の移動が完了しているか否かを判定する（ステップ4711）。

【0353】

アクセスアドレスに対応する記憶領域の移動が完了してない場合には（ステップ4711；NO）、ストレージシステム130は、移行元の論理ユニットに対してI/O処理を実行し（ステップ4709）、ホスト計算機100にライト完了報告をする（ステップ4

50

707)。

【0354】

アクセスアドレスに対応する記憶領域の移動が完了している場合には(ステップ4711; YES)、ストレージシステム130は、移行先の論理ユニットに対してI/O処理を実行し(ステップ4712)、ホスト計算機100にライト完了報告をする(ステップ4707)。

【0355】

図48はホストI/O処理ルーチンを示す。ホストI/O処理ルーチンは、ホスト計算機100内のホストI/O処理プログラム261によって実行される。

【0356】

ホスト計算機100は、ホスト計算機100上で実行されるアプリケーションプログラムからストレージシステム130の論理ユニットに対応するデバイスファイルへのI/O要求を受信する(ステップ4801)。

【0357】

ホスト計算機100は、デバイスパス管理情報251を参照し、アクセス先のデバイスファイルに複数のパス(マルチパス)が設定されているか否かを判定する(ステップ4802)。

【0358】

複数のパスが設定されていない場合は(ステップ4802; NO)、ホスト計算機100は、指定デバイスファイルをアクセス対象として、ステップ4804の処理に進む。

【0359】

複数のパスが設定されている場合には(ステップ4802; YES)、ホスト計算機100は、各パスの状態とI/O振り分けの優先順位とに基づいて、アクセス対象となるデバイスファイルを決定する(ステップ4803)。

【0360】

次に、ホスト計算機100は、アクセス対象として決定したデバイスファイルを変換して、ストレージシステム130及びポート131の識別子とLUNとを算出する(ステップ4804)。

【0361】

ホスト計算機100は、アクセス先のストレージシステム130へI/O要求を送信する(ステップ4805)。

【0362】

I/O処理が正常に終了すると(ステップ4806; YES)、ホスト計算機100はアプリケーションプログラムに対して、I/O処理完了報告をする(ステップ4809)。

【0363】

I/O処理が正常に終了しない場合には(ステップ4806; NO)、ホスト計算機100は、交替パスの有無をチェックする(ステップ4807)。

【0364】

交替パスが存在すれば(ステップ4807; YES)、ホスト計算機100は、交替パス経由でストレージシステム130にI/O処理要求を再発行し(ステップ4808)、アプリケーションプログラムに対して、I/O処理完了報告をする(ステップ4809)。

【0365】

交替パスが存在しないならば(ステップ4807; NO)、ホスト計算機100は、アプリケーションプログラムに対して、I/O処理完了報告をする(ステップ4809)。

【0366】

図49はストレージ負荷監視処理ルーチンを示す。ストレージ負荷監視処理ルーチンはそれぞれのストレージシステム130の負荷をチェックし、高負荷状態又は性能低下状態にあるストレージシステム130内の論理ユニットを他のストレージシステム130に移

10

20

30

40

50

動させる目的で実行される。ストレージ負荷監視処理ルーチンは、管理サーバ110内のストレージ負荷監視プログラム242によって実行される。

【0367】

管理サーバ110は、IPネットワーク175を經由して、ストレージシステム130の構成情報を周期的に取得し(ステップ4901)、各ストレージシステム130内で構成異常が発生しているか否かをチェックする(ステップ4902)。

【0368】

ディスクキャッシュ134又は制御プロセッサ132などの障害によって、あるストレージシステム130のI/O処理能力が低下すると(ステップ4902)、管理サーバ110は、性能低下が見られるストレージシステム130を移動元ストレージシステム130aとして決定する(ステップ4908)。

10

【0369】

性能低下が見られるストレージシステム130を検出できない場合には(ステップ4902; NO)、管理サーバ110は、各ストレージシステム130から各論理デバイスへのI/O頻度、各制御プロセッサ132又はディスクキャッシュ134の利用率などの稼働情報を周期的に取得する(ステップ4903)。

【0370】

各ストレージシステム130間のI/O負荷の差が規定レベル以上である場合(ステップ4904; YES)、管理サーバ110は、高負荷状態にあるストレージシステム130を移動元ストレージ130aとして決定する(ステップ4905)。

20

【0371】

管理サーバ110は、ステップ4905又はステップ4908によって決定した移動元ストレージ130aに対して移動元となる論理ユニットを選定し、更に移動先のストレージ130bを選定するとともに移動先の論理ユニットを選定して(ステップ4906)、ストレージシステム130a、130bに論理ユニットの移動を指示する(ステップ4907)。

【0372】

計算機システム500を構成する各ストレージシステム130の負荷をストレージ負荷監視処理ルーチンによって監視し、ディスクキャッシュ134又は制御プロセッサ132の故障等により、あるストレージシステム130に障害が発生した場合、又は何等か原因でストレージシステム130のI/O処理性能が低下した場合には、外部ボリュームの制御担当を正常状態又は低負荷状態にあるストレージシステム130に切り替えることで、計算機システム500全体のI/O処理性能を最大限に発揮できる。

30

【0373】

次に、図50乃至図54を参照しながら、論理ユニット移動処理の他の方法について説明を加える。

【0374】

図50は計算機システム501のシステム構成を示す。計算機システム501は、ホスト計算機100、ストレージシステム130a、130b、及び外部ストレージシステム150を備える。図1に示す符号と同一符号のシステムは、同一のシステムを示すものとして、その詳細な説明を省略する。

40

【0375】

ストレージシステム130aの動作モードは、キャッシュスルーモードに設定されており、外部ボリューム301に格納されるデータは、常時最新の状態にある。ホスト計算機100は、アクセスパスをバス401からバス403に切り替えるだけで、デバイス同士のマッピングの切り替え等を行うことなく、論理ユニット201aを論理ユニット201bに移動させることができる。

【0376】

尚、ストレージシステム130aの動作モードがライトアフタモードに設定されている場合には、ホスト計算機100からストレージシステム130aへのI/O要求を一時的

50

に停止し、ストレージシステム 130 a 内のディスクキャッシュに蓄積されているデータデータを外部ボリューム 301 にデステージする。その後、ホスト計算機 100 は、アクセスパスをパス 401 からパス 403 に切り替えるだけで、デバイス同士のマッピングの切り替え等を行うことなく、論理ユニット 201 a を論理ユニット 201 b に移動させることができる。

【0377】

図 5 1 は計算機システム 502 のシステム構成を示す。計算機システム 502 は、ホスト計算機 100、及びストレージシステム 130 a、130 b を備える。ストレージシステム 130 a は、ディスクドライブ等の実記憶領域を有する記憶資源に対応付けられた物理デバイス 205 a を有しており、それぞれの記憶階層 (PDEV/VDEV/LDEV/LU) を介して物理デバイス 205 a を論理ユニット 201 a として仮想化している。一方、ストレージシステム 130 b は、ディスクドライブ等の実記憶領域を有する記憶資源に対応付けられた物理デバイス 205 b を有しており、それぞれの記憶階層 (PDEV/VDEV/LDEV/LU) を介して物理デバイス 205 b を論理ユニット 201 b として仮想化している。論理ユニット 201 b は、論理ユニット 201 a の移動先となることが予定されている。

10

【0378】

計算機システム 502 では、移動元となる論理ユニット 201 a と、移動先となる論理ユニット 201 b との間で同期コピーを行うためのコピーペアが設定されている。このようにコピーペアを設定することで、ホスト計算機 100 から論理ユニット 201 a (実際には、ディスクキャッシュ) に書き込まれたデータは、ストレージネットワーク上のパス (図示せず) を経由して、論理ユニット 201 b (実際には、ディスクキャッシュ) に書き込まれた後に、ホスト計算機 100 にライト完了が報告される。

20

【0379】

ストレージシステム 130 a に障害が発生すると、ホスト計算機 100 は、アクセスパスをパス 401 からパス 403 に切り替えるだけで、デバイス同士のマッピングの変更等をしなくても、論理ユニット 201 a を論理ユニット 201 b に移動させることが可能となる。

【0380】

図 5 2 は計算機システム 503 のシステム構成を示す。計算機システム 503 は、ホスト計算機 100、及びストレージシステム 130 a、130 b を備える。図 5 1 に示す符号と同一符号のシステム等は同一のシステム等を示すものとして、その詳細な説明を省略する。

30

【0381】

ストレージシステム 130 a (移動元ストレージ) 内の論理ユニット 201 a (移動元論理ユニット) にあるデータをストレージシステム 130 b (移動先ストレージ) 内の論理ユニット 201 b (移動先論理ユニット) に移動する処理について概説する。論理ユニット移動処理は、下記の (1) ~ (9) の処理手順を含む。

【0382】

(1) ストレージシステム 130 b は、移動元の論理ユニット 201 a と、移動先のストレージシステム 130 b との間を外部接続するためのパス 404 を定義し、論理ユニット 201 a を仮想化するための拡張デバイス 206 b をストレージシステム 130 b 内に作成する。

40

【0383】

(2) ストレージシステム 130 b は、仮想デバイス 203 a と同一の論理構成を有する仮想デバイス 203 b - 1、203 b - 2 をストレージシステム 130 b 内に作成する。

【0384】

(3) ストレージシステム 130 b は、論理デバイス 202 a の論理構成と同一の論理構成を有する論理デバイス 202 b - 1、202 b - 2 をストレージシステム 130 b 内に作成する。

50

【 0 3 8 5 】

(4) ストレージシステム 1 3 0 b は、拡張デバイス 2 0 6 b を仮想デバイス 2 0 3 b - 1 にマッピングし、更に物理デバイス 2 0 5 b を仮想デバイス 2 0 3 b - 2 にマッピングする。

【 0 3 8 6 】

(5) ストレージシステム 1 3 0 b は、仮想デバイス 2 0 3 b - 1 を論理デバイス 2 0 2 b - 1 にマッピングし、更に仮想デバイス 2 0 3 b - 2 を論理デバイス 2 0 2 b - 2 にマッピングする。

【 0 3 8 7 】

(6) ストレージシステム 1 3 0 b は、論理ユニット 2 0 1 a の論理構成と同一の論理構成を有する論理ユニット 2 0 1 b をストレージシステム 1 3 0 b 内に作成し、ホスト計算機 1 0 0 と論理ユニット 2 0 1 b とを接続するためのパス 4 0 3 を定義する。論理デバイス 2 0 2 b - 1 は、論理ユニット 2 0 1 b にマッピングされる。

10

【 0 3 8 8 】

(7) ホスト計算機 1 0 0 は、アクセスパスをパス 4 0 1 からパス 4 0 3 に切り替える。このとき、パス 4 0 1 を削除する等して、ホスト計算機 1 0 0 から論理ユニット 2 0 1 a へのアクセスを禁止する。尚、パス 4 0 1 からパス 4 0 3 へのパス切り替えは、ストレージシステム 1 3 0 a , 1 3 0 b , 又は管理サーバ 1 1 0 等が行っても良い。また、必ずしもパス 4 0 1 を削除する必要はなく、例えば、パス 4 0 1 をパス 4 0 3 の交替パス(例えば、システム障害時に使用するためのパス)として設定し、パス 4 0 1 の状態を無効としたまま、パス 4 0 1 を残しても良い。

20

(8) ストレージシステム 1 3 0 b は、論理デバイス 2 0 2 b - 1 を論理デバイス 2 0 2 b - 2 にコピーする。かかる処理により、物理デバイス 2 0 5 a と物理デバイス 2 0 5 b とを同期化させることができる。

【 0 3 8 9 】

尚、論理ユニット移動処理中における、ホスト計算機 1 0 0 からストレージシステム 1 3 0 b へのデータ入出力経路は、パス 4 0 3 ストレージシステム 1 3 0 b 内の記憶階層 (L U / L D E V / V D E V / E D E V) パス 4 0 4 ストレージシステム 1 3 0 a 内の記憶階層 (L U / L D E V / V D E V / P D E V) となる。

【 0 3 9 0 】

(9) ストレージシステム 1 3 0 b は、物理デバイス 2 0 5 a と物理デバイス 2 0 5 b とを同期化させることができたならば、仮想デバイス 2 0 3 b - 2 と論理デバイス 2 0 2 b - 2 との間のマッピング関係を解除し、仮想デバイス 2 0 3 b - 2 を論理デバイス 2 0 2 b - 1 にマッピングさせる。更に、ストレージシステム 1 3 0 b は、仮想デバイス 2 0 3 b - 1 と論理デバイス 2 0 2 b - 1 との間のマッピング関係を解除し、仮想デバイス 2 0 3 b - 1 を論理デバイス 2 0 2 b - 2 にマッピングさせる。更に、ストレージシステム 1 3 0 b は、パス 4 0 4 を削除する

30

【 0 3 9 1 】

図 5 3 は計算機システム 5 0 4 のシステム構成を示す。計算機システム 5 0 4 は、ホスト計算機 1 0 0、ストレージシステム 1 3 0 a , 1 3 0 b、及び外部ストレージシステム 1 5 0 を備える。図 5 2 に示す符号と同一符号のシステム等は同一のシステム等を示すものとして、その詳細な説明を省略する。

40

【 0 3 9 2 】

計算機システム 5 0 4 と、上述した計算機システム 5 0 3 との主な相違点は、次の点のみである。即ち、ストレージシステム 1 3 0 a は、物理デバイス 2 0 5 a に替えて、拡張デバイス 2 0 6 a を有している。拡張デバイス 2 0 6 a は、外部ボリューム 3 0 1 を仮想化するための仮想的な記憶領域である。計算機システム 5 0 4 が実行する論理ユニット移動処理の手順は、計算機システム 5 0 3 が実行する論理ユニット移動処理の手順と同じである。

【 0 3 9 3 】

50

図54は計算機システム505のシステム構成を示す。計算機システム505は、ホスト計算機100、ストレージシステム130a、130b、及び外部ストレージシステム150を備える。図52に示す符号と同一符号のシステム等は同一のシステム等を示すものとして、その詳細な説明を省略する。

【0394】

計算機システム505と、上述した計算機システム503との主な相違点は、次の点のみである。即ち、ストレージシステム130bは、物理デバイス205bに替えて、拡張デバイス206cを有している。拡張デバイス206cは、外部ボリューム301を仮想化するための仮想的な記憶領域である。計算機システム505が実行する論理ユニット移動処理の手順は、計算機システム503が実行する論理ユニット移動処理の手順と同じである。

10

【0395】

次に、図55乃至図56を参照しながら、論理ユニットの一部を移動元ストレージから移動先ストレージに移動させる機能を有する計算機システム506、507について説明を加える。

【0396】

図55は計算機システム506のシステム構成を示す。計算機システム506は、ホスト計算機100、ストレージシステム130a、130b、及び外部ストレージシステム150を備える。図1に示す符号と同一符号のシステムは、同一のシステムを示すものとして、その詳細な説明を省略する。

20

【0397】

外部ボリューム301は、複数の記憶領域301-1、301-2に分割されている。それぞれの記憶領域301-1、301-2は、外部ボリューム301を分割するパーティションである。それぞれの記憶領域301-1、301-2には、独立したファイルシステムを構築できる。例えば、記憶領域301-1はCドライブに対応し、記憶領域301-2はDドライブに対応する。

【0398】

ストレージシステム130aがパス401を介してホスト計算機100に提供する論理ユニット201aは、外部ボリューム301を仮想化したものであり、複数の記憶領域201a-1、201a-2に分割されている。記憶領域201a-1は、記憶領域301-1を仮想化したものである。記憶領域201a-2は、記憶領域301-2を仮想化したものである。

30

【0399】

さて、論理ユニット201aの一部（例えば、Dドライブに対応する記憶領域201a-2）をストレージシステム130aからストレージシステム130bに移動させるには、上述した処理手順(1)～(9)と同様の手順を経て、論理ユニット201aをストレージシステム130aからストレージシステム130bに移動させる。ストレージシステム130b内に作成される論理ユニット201bの論理構成は、論理ユニット201aの論理構成と同じである。つまり、ストレージシステム130bがパス403を介してホスト計算機100に提供する論理ユニット201bは、外部ボリューム301を仮想化したものであり、複数の記憶領域201b-1、201b-2に分割されている。記憶領域201b-1は、記憶領域301-1を仮想化したものである。記憶領域201b-2は、記憶領域301-2を仮想化したものである。

40

【0400】

ここで、ホスト計算機100から記憶領域201a-2へのアクセスを禁止するとともに、ホスト計算機100から記憶領域201b-1へのアクセスを禁止するように計算機システム506のシステム構成を設定すれば、ストレージシステム130aからストレージシステム130bへ論理ユニット201aの一部を移動させることが可能になる。

【0401】

尚、ストレージシステム130aから記憶領域301-2へのアクセスを禁止するとと

50

もに、ストレージシステム 130b から記憶領域 301 - 1 へのアクセスを禁止する必要がある。

【0402】

ホスト計算機 100 内のリンクマネージャ 108 は、アクセス先に応じてパスを切り替える。例えば、ホスト計算機 100 が C ドライブにアクセスするときには、パス 401 を選択し、ホスト計算機 100 が D ドライブにアクセスするときには、パス 403 を選択する。

【0403】

図 56 は計算機システム 507 のシステム構成を示す。計算機システム 507 は、ホスト計算機 100、ストレージシステム 130a、130b、及び外部ストレージシステム 150 を備える。図 55 に示す符号と同一符号のシステムは同一のシステムを示すものとして、その詳細な説明を省略する。

10

【0404】

計算機システム 507 の基本的なシステム構成は、計算機システム 506 のシステム構成とほぼ同じなので、両者の相違点を中心に説明する。外部ボリューム 301 は、複数の記憶領域 301 - 1、301 - 2、301 - 3 に分割される。記憶領域 301 - 3 は、それぞれの論理ユニット 201a、201b によって仮想化されている共用領域である。論理ユニット 201a の記憶領域 201a - 3 は、外部ボリューム 301 の記憶領域 301 - 3 を仮想化したものである。更に論理ユニット 201b の記憶領域 201b - 3 は、外部ボリューム 301 の記憶領域 301 - 3 を仮想化したものである。

20

【0405】

ホスト計算機 100 は、外部ボリューム 301 の共用領域である記憶領域 301 - 3 へのアクセス経路として、論理ユニット 201a を介してアクセスする経路と、論理ユニット 201b を介してアクセスする経路の二つの経路を有する。記憶領域 301 - 3 には、例えば、ホスト計算機 100 が業務を遂行する上で、重要なデータを格納する。ホスト計算機 100 は、記憶領域 301 - 3 にアクセスするための複数の経路を有するので、何れか一方の経路に障害が発生しても、他方の経路を経由してアクセスできる。

【0406】

次に、外部ボリューム 301 の共用領域である記憶領域 301 - 3 を複数の論理ユニット 201a、201b によって仮想化するための方式として、三つの方式を説明する。

30

【0407】

(1) 第一の方式

それぞれのストレージシステム 130a、130b の動作モードをキャッシュスルーモードに設定する。ホスト計算機 100 が論理ユニット 201a 又は論理ユニット 201b を介して外部ボリューム 301 の共用領域である記憶領域 301 - 3 にデータを書き込むと、記憶領域 301 - 3 上のデータが更新された後にホスト計算機 100 にライト完了が報告される。記憶領域 301 - 3 に格納されるデータは、常時最新の状態にあるので、ホスト計算機 100 が論理ユニット 201a を介して外部ボリューム 301 の共用領域にアクセスしても、或いはホスト計算機 100 が論理ユニット 201b を介して外部ボリューム 301 の共用領域にアクセスしても、共用領域上のデータに矛盾が生じることはない。

40

【0408】

(2) 第二の方式

論理ユニット 201a の記憶領域 201a - 3 と、論理ユニット 201b の記憶領域 201b - 3 との間で双方向に同期コピーを行うためのコピーペアを設定する。かかる設定により、ホスト計算機 100 が外部ボリューム 301 の共用領域である記憶領域 301 - 3 のデータを更新するために、論理ユニット 201a の記憶領域 201a - 3 上のデータを更新すると、論理ユニット 201b の記憶領域 201b - 3 上のデータも同じように更新される。また、ホスト計算機 100 が論理ユニット 201b の記憶領域 201b - 3 上のデータを更新すると、論理ユニット 201a の記憶領域 201a - 3 上のデータも同じように更新される。

50

【 0 4 0 9 】

(3) 第三の方式

外部ボリューム 3 0 1 を仮想化するための拡張デバイス 2 0 6 a の記憶領域のうち共用領域に対応する記憶領域 2 0 6 a - 3 を論理ユニット 2 0 1 a の記憶領域 2 0 1 a - 3 に対応付けるとともに、論理ユニット 2 0 1 b の記憶領域 2 0 1 b - 3 にも対応付ける。同様に、外部ボリューム 3 0 1 を仮想化するための拡張デバイス 2 0 6 b の記憶領域のうち共用領域に対応する記憶領域 2 0 6 b - 3 を論理ユニット 2 0 1 b の記憶領域 2 0 1 b - 3 に対応付けるとともに、論理ユニット 2 0 1 a の記憶領域 2 0 1 a - 3 にも対応付ける。このように、双方向に外部接続技術を適用することで、ホスト計算機 1 0 0 が論理ユニット 2 0 1 a の記憶領域 2 0 1 a - 3 上のデータを更新すると、論理ユニット 2 0 1 b の記憶領域 2 0 1 b - 3 上のデータも同じように更新される。また、ホスト計算機 1 0 0 が論理ユニット 2 0 1 b の記憶領域 2 0 1 b - 3 上のデータを更新すると、論理ユニット 2 0 1 a の記憶領域 2 0 1 a - 3 上のデータも同じように更新される。

10

【 0 4 1 0 】

尚、第三の方式においては、無限ループ処理を回避するため、ホストアクセスと、外部接続経路のアクセスとを区別する必要がある。

【 0 4 1 1 】

このように、外部ボリューム 3 0 1 が複数の論理ユニットによって仮想化されている場合、外部ボリューム 3 0 1 を構成するある記憶領域については、何れか一つの論理ユニットを介してアクセス可能であり、他の記憶領域（共用領域）については、複数の論理ユニットを介してアクセス可能である、という構成を適用できる。複数の論理ユニットのそれぞれの記憶領域のうち共用領域に対応する記憶領域については、同期コピーを適用してもよく、又は非同期コピーを適用してもよい。

20

【 0 4 1 2 】

次に、図 5 7 乃至図 5 8 を参照しながら、高可用性の観点から上述の計算機システム 5 0 0 のシステム構成に若干の変更を加えた計算機システム 5 0 8 , 5 0 9 について説明を加える。

【 0 4 1 3 】

図 5 7 は計算機システム 5 0 8 のシステム構成を示す。計算機システム 5 0 8 は、ホスト計算機 1 0 0、ストレージシステム 1 3 0 a , 1 3 0 b , 及び外部ストレージシステム 1 5 0 を備える。ホスト計算機 1 0 0、ストレージシステム 1 3 0 a , 1 3 0 b , 及び外部ストレージシステム 1 5 0 は、FC - SAN 等のストレージネットワーク（図示せず）を介して相互に接続されている。ストレージシステム 1 3 0 a は、常用系システム（稼働系システム）として機能し、ストレージシステム 1 3 0 b は、交替系システム（待機系システム）として機能する。ストレージシステム 1 3 0 a , 1 3 0 b は、外部ストレージシステム 1 5 0 内の外部ボリューム 3 0 1 を共有している。

30

【 0 4 1 4 】

ストレージシステム 1 3 0 a は、外部ボリューム 3 0 1 を仮想化するための複数の記憶階層（論理ユニット 2 0 1 a、論理デバイス 2 0 2 a、仮想デバイス 2 0 3 a、及び拡張デバイス 2 0 6 a）を有している。拡張デバイス 2 0 6 a は、外部ボリューム 3 0 1 を仮想化したものであり、パス 4 0 2 を介して外部ストレージシステム 1 5 0 に接続されている。拡張デバイス 2 0 6 a は、仮想デバイス 2 0 3 a にマッピングされている。仮想デバイス 2 0 3 a は、論理デバイス 2 0 2 a にマッピングされている。論理デバイス 2 0 2 a は、論理ユニット 2 0 1 a にマッピングされている。論理ユニット 2 0 1 a は、パス 4 0 1 を介してホスト計算機 1 0 0 に接続されている。ストレージシステム 1 3 0 a の動作モードは、キャッシュスルーモードに設定されており、外部ボリューム 3 0 1 には、ホストアクセスに伴う最新のデータが常時格納されている。

40

【 0 4 1 5 】

一方、ストレージシステム 1 3 0 b は、外部ボリューム 3 0 1 を仮想化するための複数の記憶階層（論理ユニット 2 0 1 b、論理デバイス 2 0 2 b、仮想デバイス 2 0 3 b、及

50

び拡張デバイス206b)を有している。拡張デバイス206bは、外部ボリューム301を仮想化したものであり、パス405を介して外部ストレージシステム150に接続されている。拡張デバイス206bは、仮想デバイス203bにマッピングされている。仮想デバイス203bは、論理デバイス202bにマッピングされている。論理デバイス202bは、論理ユニット201bにマッピングされている。論理ユニット201bは、パス403を介してホスト計算機100に接続されている。パス403は、パス401に対して交替パスの関係にある。ストレージシステム130bの動作モードは、キャッシュスルーモードでもよく、ライトアフタモードでもよい。

【0416】

さて、上記のシステム構成において、常用系のストレージシステム130aに障害が発生すると、ホスト計算機100は、パス401からパス403に切り替えて、交替系のストレージシステム130bにI/O要求を発行する。上述の如く、ストレージシステム130aの動作モードは、キャッシュスルーモードに設定されているので、ストレージシステム130a内のディスクキャッシュには、外部ボリューム301にデステージされていないデータは存在しない。それ故、ホスト計算機100は、常用系のストレージシステム130aに障害が発生したことを検出したときに、I/O要求の発行先を交替系のストレージシステム130bに直ちに切り替えることが可能となる。

【0417】

図58は計算機システム509のシステム構成を示す。計算機システム509は、ホスト計算機100、ストレージシステム130a、130b、及び外部ストレージシステム150を備える。図57に示す符号と同一符号のシステムは同一のシステムを示すものとして、その詳細な説明を省略する。

【0418】

ストレージシステム130aは、外部ボリューム301を仮想化するための複数の記憶階層(論理ユニット201a、論理デバイス202a、仮想デバイス203a、及び拡張デバイス206a、206d)を有している。拡張デバイス206aは外部ボリューム301を仮想化したものであり、パス402を介して外部ストレージシステム150に接続されている。拡張デバイス206dは、論理ユニット201bを仮想化したものであり、パス410を介してストレージシステム130bに接続されている。それぞれの拡張デバイス206a、206dは、仮想デバイス203aにマッピングされている。仮想デバイス203aは、論理デバイス202aにマッピングされている。論理デバイス202aは論理ユニット201aにマッピングされている。論理ユニット201aは、パス401を介してホスト計算機100に接続されている。

【0419】

一方、ストレージシステム130bは、外部ボリューム301を仮想化するための複数の記憶階層(論理ユニット201b、論理デバイス202b、仮想デバイス203b、及び拡張デバイス206b)を有している。拡張デバイス206bは、外部ボリューム301を仮想化したものであり、パス405を介して外部ストレージシステム150に接続されている。拡張デバイス206bは、仮想デバイス203bにマッピングされている。仮想デバイス203bは、論理デバイス202bにマッピングされている。論理デバイス202bは、論理ユニット201bにマッピングされている。論理ユニット201bは、パス403を介してホスト計算機100に接続されている。パス403は、パス401に対して交替パスの関係にある。

【0420】

次に、計算機システム509の可用性を高めるための方式として、三つの方式を説明する。

【0421】

(1) 第一の方式

常用系のストレージシステム130a及び交替系のストレージシステム130bのそれぞれの動モードをキャッシュスルーモードに設定する。すると、ホスト計算機100から

10

20

30

40

50

論理ユニット201aへのデータ入出力経路は、パス401 ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV) パス410 ストレージシステム130b内の記憶階層(LU/LDEV/VDEV/EDEV) パス405 外部ボリューム301となる。但し、ホスト計算機100から論理ユニット201aへリードアクセスするための経路として、例えば、パス401 ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV) パス402 外部ボリューム301となる経路を適用してもよい。

【0422】

第一の方式によれば、外部ボリューム301には、ホストアクセスに伴う最新のデータが常時格納されているので、常用系のストレージシステム130aに障害が発生したとしても、ホスト計算機100は、パス401からパス403に切り替えることで、交替系のストレージシステム130bを直ちに使用することができる。

10

【0423】

(2) 第二の方式

常用系のストレージシステム130a内にある論理ユニット201aと、交替系のストレージシステム130b内にある論理ユニット201bとの間のデータ入出力動作をキャッシュモードに設定する。つまり、ホスト計算機100から論理ユニット201a(実際には、ディスクキャッシュ)に書き込まれたデータは、ストレージシステム130a内の記憶階層(LU/LDEV/VDEV/EDEV)からパス410を経由して、論理ユニット201b(実際には、ディスクキャッシュ)に書き込まれた後に、ホスト計算機100にライト完了が報告される。一方、常用系のストレージシステム130a内にある論理ユニット201aと、外部ボリューム301との間のデータ入出力動作をライトアフタモードに設定する。つまり、ホスト計算機100から論理ユニット201a(実際には、ディスクキャッシュ)に書き込まれたデータは、ホスト計算機100にライト完了が報告された後に、外部ボリューム301にデステージされる。

20

【0424】

常用系のストレージシステム130aが正常に稼動している間は、ストレージシステム130aのみが外部ボリューム301へのアクセス権を有し、交替系のストレージシステム130bは、外部ボリューム301へのアクセス権を有しない。ストレージシステム130aが外部ボリューム301にデータを書き込んだことをストレージシステム103bがストレージシステム130aから通知を受けると、ストレージシステム130bは、ストレージシステム130aから論理ユニット201b(実際には、ディスクキャッシュ)に書き込まれたデータを削除する。

30

【0425】

第二の方式によれば、常用系のストレージシステム130aだけでなく、交替系のストレージシステム130bも、外部ボリューム301にデステージされていないダーティデータを保持することができる。それ故、常用系のストレージシステム130aに障害が発生した場合には、交替系のストレージシステム130bが外部ボリューム301へのアクセス権を取得し、ダーティデータを外部ボリューム301にデステージすることで、外部ボリューム301に格納されているデータを最新の状態に更新できる。外部ボリューム301に格納されているデータを最新の状態に更新できたならば、ホスト計算機100は、パス401からパス403に切り替えることで、交替系のストレージシステム130bにI/O要求を発行することができる。

40

【0426】

(3) 第三の方式

常用系のストレージシステム130a内にある論理ユニット201aと、交替系のストレージシステム130b内にある論理ユニット201bの間には、同期コピーのためのコピーペアが形成されている。このようにコピーペアを設定することで、ホスト計算機100から論理ユニット201a(実際には、ディスクキャッシュ)に書き込まれたデータは、ストレージネットワーク上のパス(図示せず)を経由して、論理ユニット201b(

50

実際には、ディスクキャッシュ)に書き込まれた後に、ホスト計算機100にライト完了が報告される。一方、常用系のストレージシステム130a内にある論理ユニット201aと、外部ボリューム301との間のデータ入出力動作をライトアフタモードに設定する。つまり、ホスト計算機100から論理ユニット201a(実際には、ディスクキャッシュ)に書き込まれたデータは、ホスト計算機100にライト完了が報告された後に、外部ボリューム301にデステージされる。

【0427】

常用系のストレージシステム130aが正常に稼動している間は、ストレージシステム130aのみが外部ボリューム301へのアクセス権を有し、交替系のストレージシステム130bは、外部ボリューム301へのアクセス権を有しない。ストレージシステム130aが外部ボリューム301にデータを書き込んだことをストレージシステム103bがストレージシステム130aから通知を受けると、ストレージシステム130bは、ストレージシステム130aから論理ユニット201b(実際には、ディスクキャッシュ)に書き込まれたデータを削除する。この点は、第二の方式と同じである。

【0428】

第三の方式によれば、常用系のストレージシステム130aだけでなく、交替系のストレージシステム130bも、外部ボリューム301にデステージされていないダーティデータを保持することができる。それ故、常用系のストレージシステム130aに障害が発生した場合には、交替系のストレージシステム130bが外部ボリューム301へのアクセス権を取得し、ダーティデータを外部ボリューム301にデステージすることで、外部ボリューム301に格納されているデータを最新の状態に更新できる。外部ボリューム301に格納されているデータを最新の状態に更新できたならば、ホスト計算機100は、パス401からパス403に切り替えることで、交替系のストレージシステム130bにI/O要求を発行することができる。

【0429】

図59は本実施例の変形例としての計算機システム580のシステム構成を示す。計算機システム580は、計算機システム500のシステム構成(図2参照)において、ストレージシステム130a, 130bを仮想化装置180a, 180b, 180cに置換してなるシステム構成を備えている。説明の便宜上、仮想化装置180a, 180b, 180cを仮想化装置180と総称する。仮想化装置180は、ストレージネットワークを介してファイバチャネルスイッチ120に接続するためのポート181と、外部ボリューム301を仮想化する処理等を制御する制御プロセッサ182と、制御プロセッサ182が実行する制御プログラム等を格納するメモリ183と、IPネットワーク175を介して管理サーバ110に接続するためのポート184とを備える。仮想化装置180は、例えば、記憶資源を仮想化する機能を備える専用の仮想化装置又は仮想化スイッチである。仮想化装置180は、外部ボリューム301を仮想化してなる論理ユニットをホスト計算機100に提供する機能を有する。

【0430】

計算機システム580は、例えば、仮想化装置180aがホスト計算機100に提供する論理ユニットを仮想化装置180bに移動させることができる。仮想化装置180aから仮想化装置180bへの論理ユニット移動処理の概要は、上述の処理内容とほぼ同じである。但し、仮想化装置180は、外部ボリューム301へ読み書きされるデータを一時的に格納するためのディスクキャッシュを有しないので、仮想化装置180の動作モードをキャッシュスルーモードへ設定する必要がないだけでなく、仮想化装置180から外部ボリューム301へのデステージ完了を待つ必要もない。移動元の仮想化装置180aがホスト計算機100に提供する論理ユニットに設定されている属性を、移動先の仮想化装置180bがホスト計算機100に提供する論理ユニットに引き継ぐだけで移動処理を完了させることができる。

【図面の簡単な説明】

【0431】

- 【図 1】本実施形態に関わる論理ユニット移動処理の原理図である。
- 【図 2】実施例に関わる計算機システムのネットワーク構成図である。
- 【図 3】ストレージシステム内に構築される記憶階層の説明図である。
- 【図 4】CVS 構成のマッピング関係を示す説明図である。
- 【図 5】LUSE 構成のマッピング関係を示す説明図である。
- 【図 6】外部ボリューム RAID 構成のマッピング関係を示す説明図である。
- 【図 7】VMA 付きボリュームの説明図である。
- 【図 8】メインフレームボリュームの説明図である。
- 【図 9】VDEV 連結のマッピング関係を示す説明図である。
- 【図 10】VDEV 離散のマッピング関係を示す説明図である。 10
- 【図 11】VDEV 離散のマッピング関係を示す説明図である。
- 【図 12】VDEV 離散のマッピング関係を示す説明図である。
- 【図 13】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 14】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 15】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 16】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 17】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 18】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 19】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。
- 【図 20】DEVGr - EDEV / PDEV マッピング関係を示す説明図である。 20
- 【図 21】AOU 構成の記憶階層を示す説明図である。
- 【図 22】スナップショット機能を有する記憶階層を示す説明図である。
- 【図 23】ホストグループに関する説明図である。
- 【図 24】交替パスに関する説明図である。
- 【図 25】交替パスに関する説明図である。
- 【図 26】ホストグループ管理情報の説明図である。
- 【図 27】LU パス管理情報の説明図である。
- 【図 28】論理デバイス管理情報の説明図である。
- 【図 29】仮想デバイス管理情報の説明図である。
- 【図 30】デバイスグループ管理情報の説明図である。 30
- 【図 31】VDEV - DEVGr マッピング情報の説明図である。
- 【図 32】物理デバイス管理情報の説明図である。
- 【図 33】拡張デバイス管理情報の説明図である。
- 【図 34】計算機システムのソフトウェア構成図である。
- 【図 35】CVS 構成を有する論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 36】LUSE 構成を有する論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 37】外部ボリューム RAID 構成を有する論理ユニットを移動させる処理の概要を示す説明図である。 40
- 【図 38】VMA 付きボリューム構成を有する論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 39】VMA 付きボリューム構成を有する論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 40】メインフレームボリューム構成を有する論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 41】複数の論理ユニットを移動させる処理の概要を示す説明図である。
- 【図 42】論理ユニット移動指示処理ルーチンを示す説明図である。
- 【図 43】論理ユニット移動処理サブルーチンを示す説明図である。
- 【図 44】移動対象選出処理サブルーチンを示す説明図である。 50

- 【図45】移動準備処理サブルーチンを示す説明図である。
- 【図46】移動処理サブルーチンを示す説明図である。
- 【図47】ストレージI/O処理ルーチンを示す説明図である。
- 【図48】ホストI/O処理ルーチンを示す説明図である。
- 【図49】ストレージ負荷監視処理ルーチンを示す説明図である。
- 【図50】論理ユニット移動処理の他の方法を示す説明図である。
- 【図51】論理ユニット移動処理の他の方法を示す説明図である。
- 【図52】論理ユニット移動処理の他の方法を示す説明図である。
- 【図53】論理ユニット移動処理の他の方法を示す説明図である。
- 【図54】論理ユニット移動処理の他の方法を示す説明図である。
- 【図55】論理ユニットの一部を移動させる処理の概要を示す説明図である。
- 【図56】論理ユニットの一部を移動させる処理の概要を示す説明図である。
- 【図57】高可用性用の計算機システムのシステム構成図である。
- 【図58】高可用性用の計算機システムのシステム構成図である。
- 【図59】高可用性用の計算機システムのシステム構成図である。
- 【符号の説明】

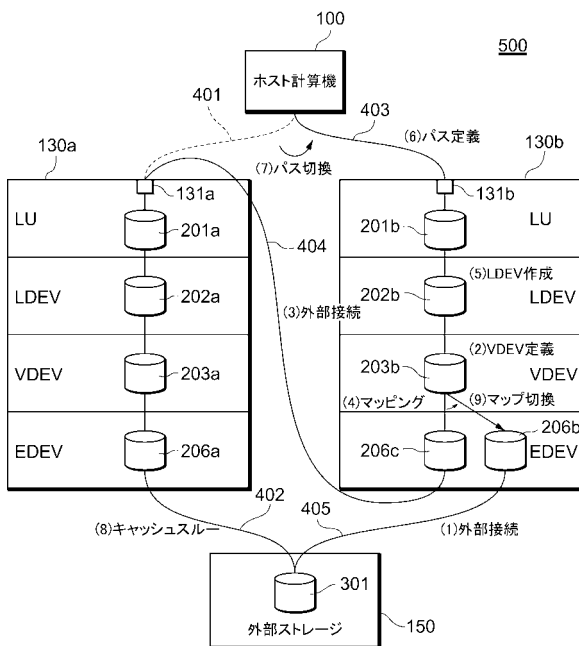
10

【0432】

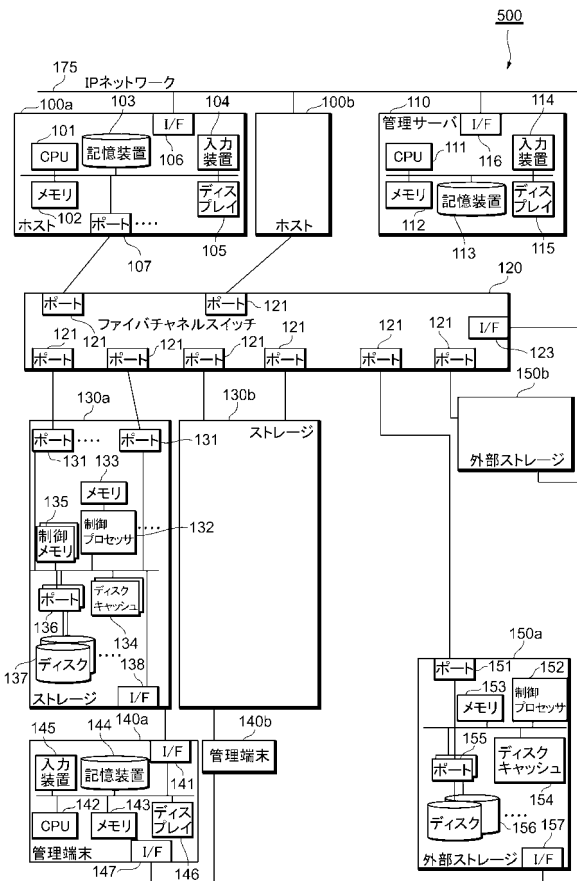
- 100 ... 120 ... 130 ... 140 ... 150 ... 500 ... 200 ... 201 ... 202 ... 203 ... 204 ... 205 ... 206 ...

20

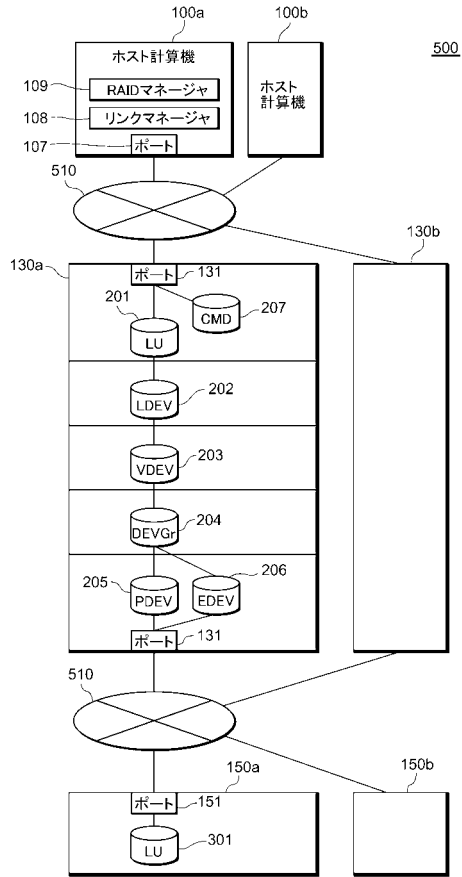
【図1】



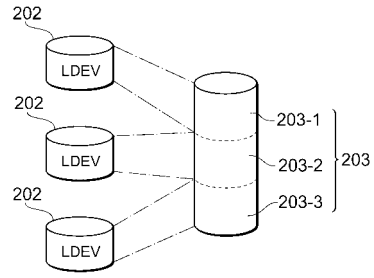
【図2】



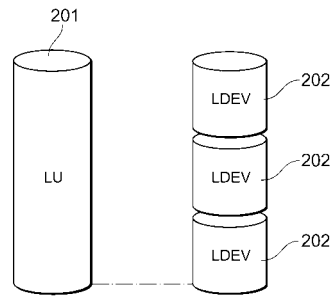
【図3】



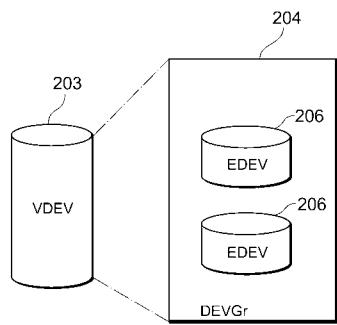
【図4】



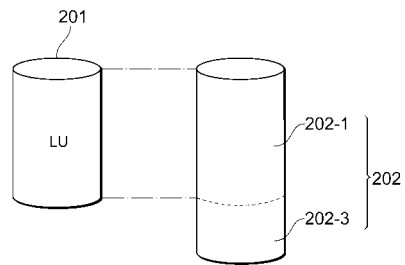
【図5】



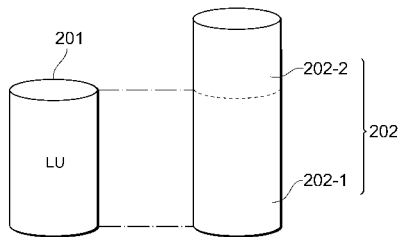
【図6】



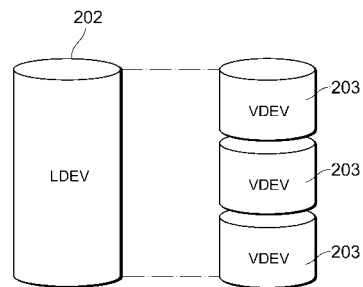
【図8】



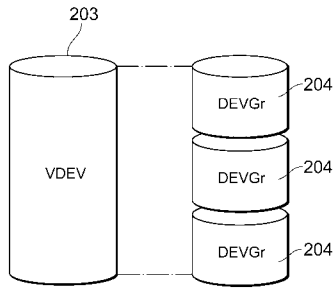
【図7】



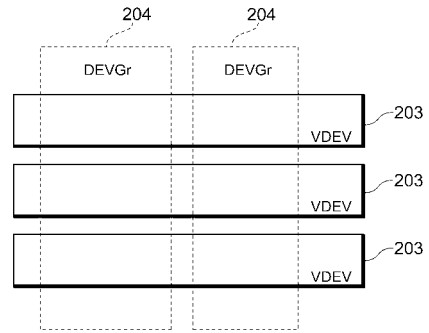
【図9】



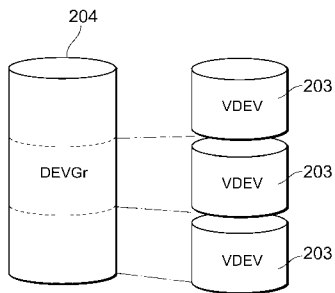
【図10】



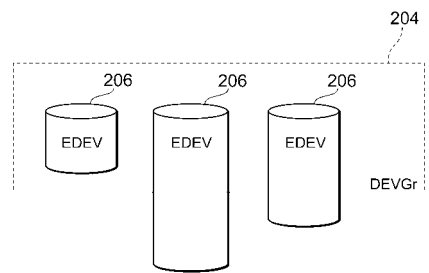
【図12】



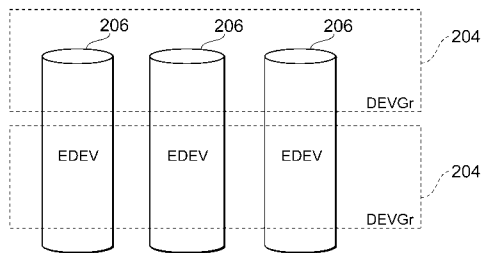
【図11】



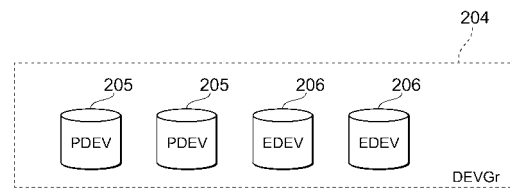
【図13】



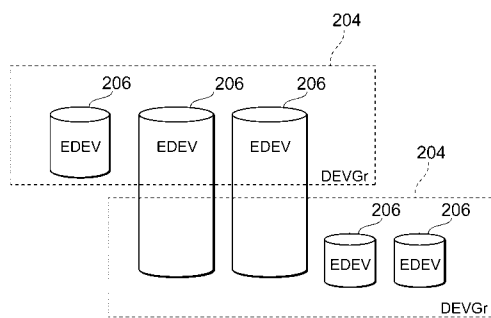
【図14】



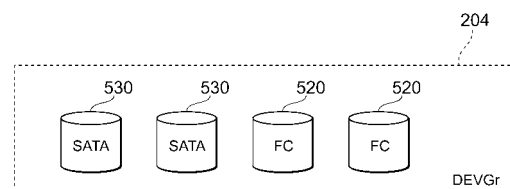
【図16】



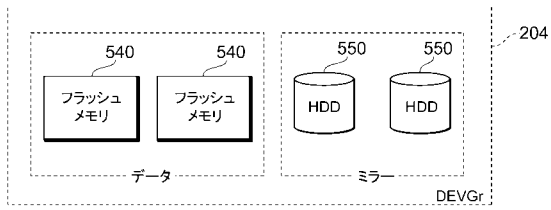
【図15】



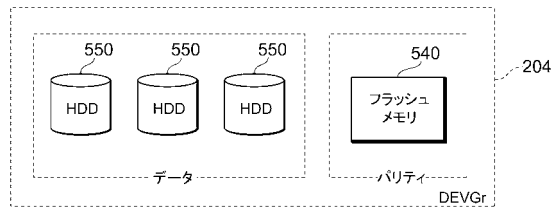
【図17】



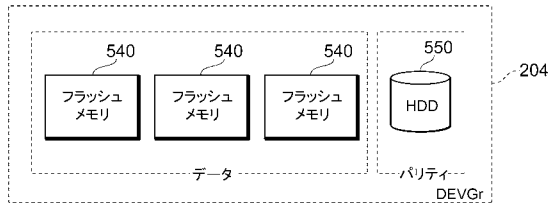
【図18】



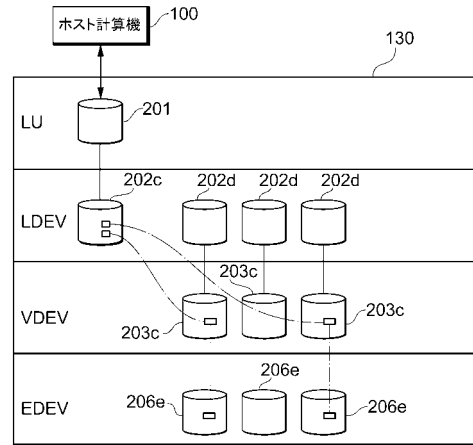
【図20】



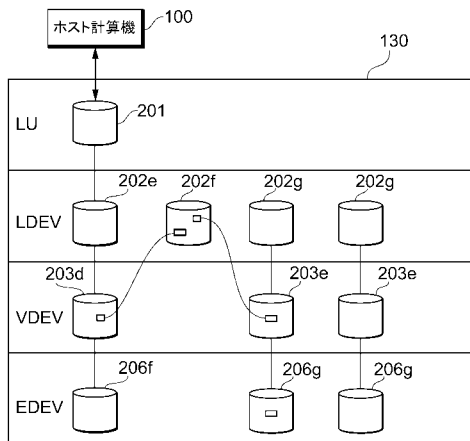
【図19】



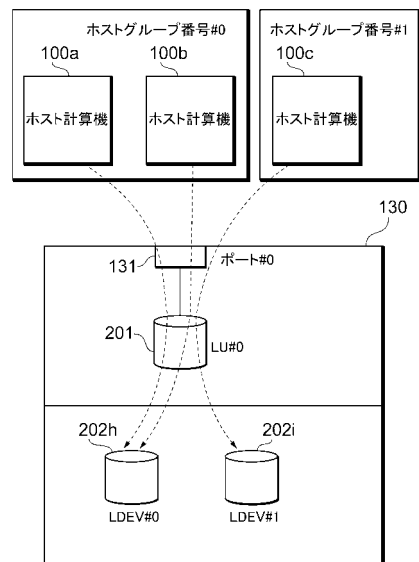
【図21】



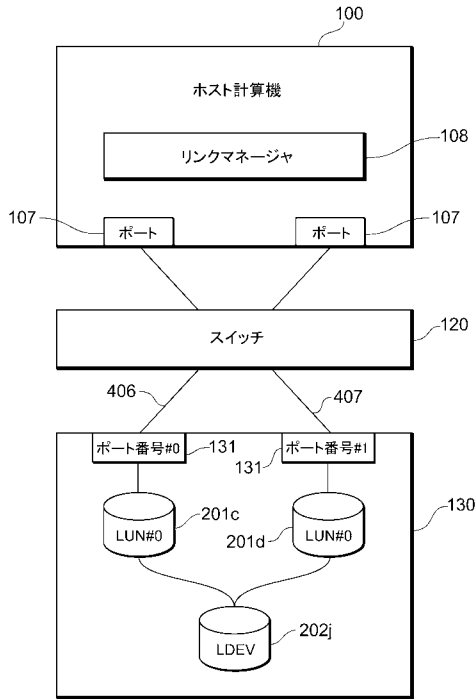
【図22】



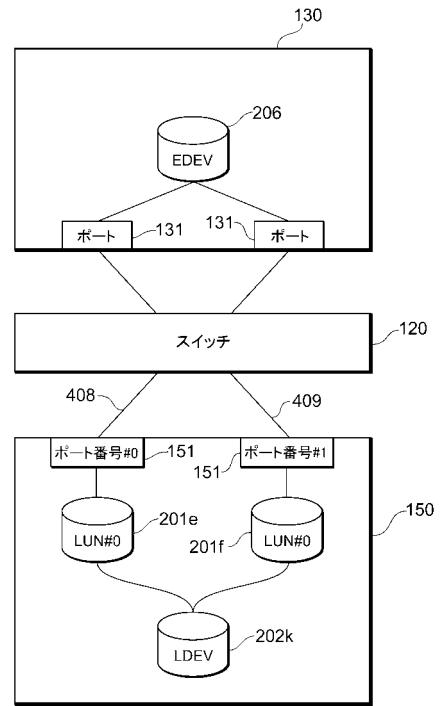
【図23】



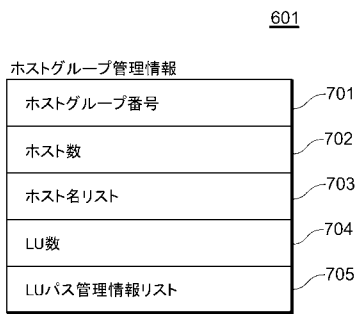
【図24】



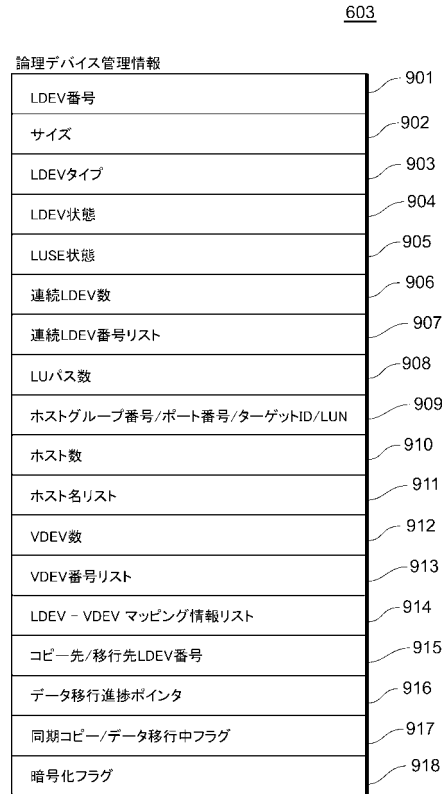
【図25】



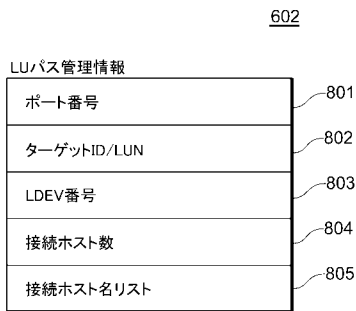
【図26】



【図28】



【図27】



【 図 2 9 】

604

仮想デバイス管理情報

VDEV番号	1001
サイズ	1002
VDEV状態	1003
RAID構成	1004
LDEV数	1005
LDEV番号リスト	1006
VDEV - LDEVマッピング情報リスト	1007
デバイスグループ数	1008
デバイスグループ番号リスト	1009
VDEV - DEVGrマッピング情報リスト	1010

【 図 3 0 】

605

デバイスグループ管理情報

デバイスグループ番号	1101
サイズ	1102
デバイスグループ状態	1103
RAID構成	1104
VDEV数	1105
VDEV番号リスト	1106
DEVGr - VDEVマッピング情報リスト	1107
デバイス数	1108
PDEV/EDEV番号リスト	1109
DEVGr - PDEV/EDEVマッピング情報リスト	1110

【 図 3 1 】

606

VDEV - DEVGrマッピング情報

VDEV領域番号	1201
デバイスグループ数	1202
デバイスグループ番号リスト	1203
有効フラグリスト	1204
デバイスグループ内オフセットリスト	1205
キャッシュスルーリード情報リスト	1206
キャッシュスルーライト情報リスト	1207

【 図 3 3 】

608

拡張デバイス管理情報

EDEV番号	1401
サイズ	1402
EDEV状態	1403
デバイスグループ数	1404
デバイスグループ番号リスト	1405
EDEV - DEVGrマッピング情報	1406
ストレージ識別情報	1407
外部ストレージ内LDEV番号	1408
イニシエータポート番号リスト	1409
ターゲットポートID/LUN	1410
アクセス制御情報	1411

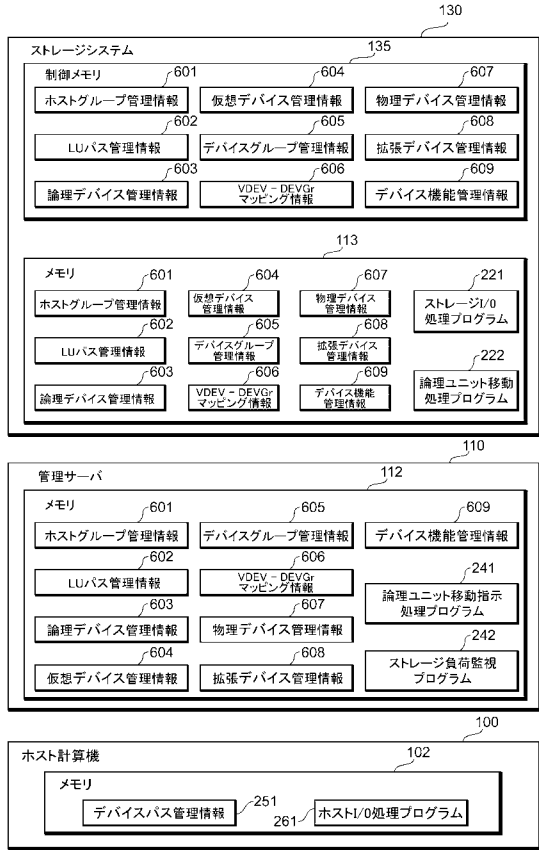
【 図 3 2 】

607

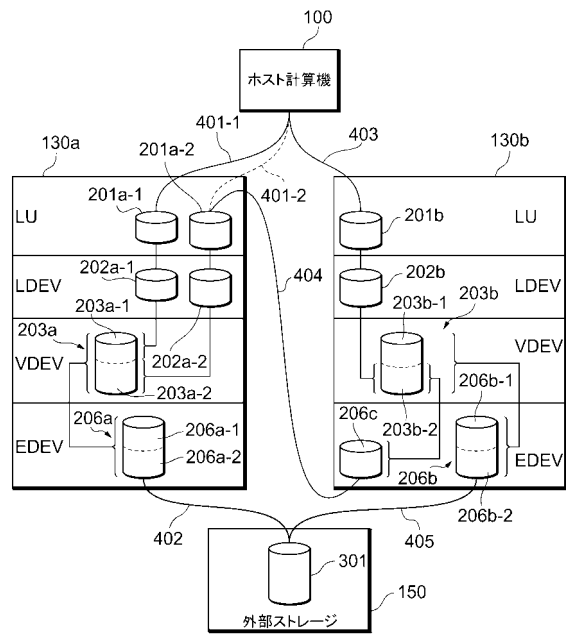
物理デバイス管理情報

PDEV番号	1301
サイズ	1302
PDEV状態	1303
デバイスグループ数	1304
デバイスグループ番号リスト	1305
PDEV - DEVGrマッピング情報	1306

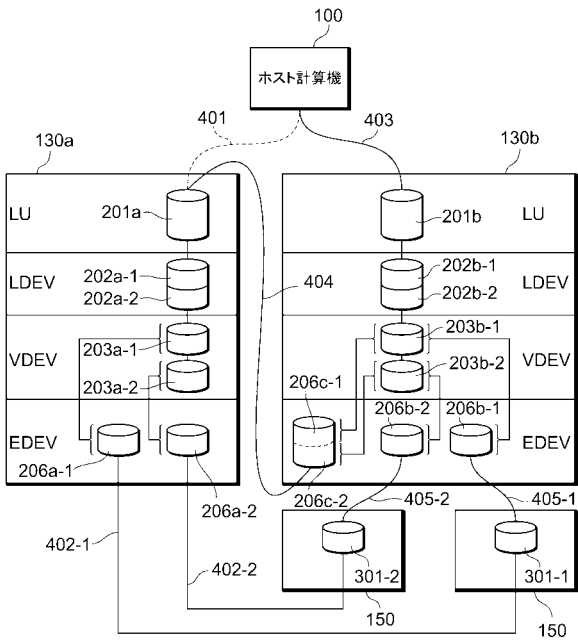
【図34】



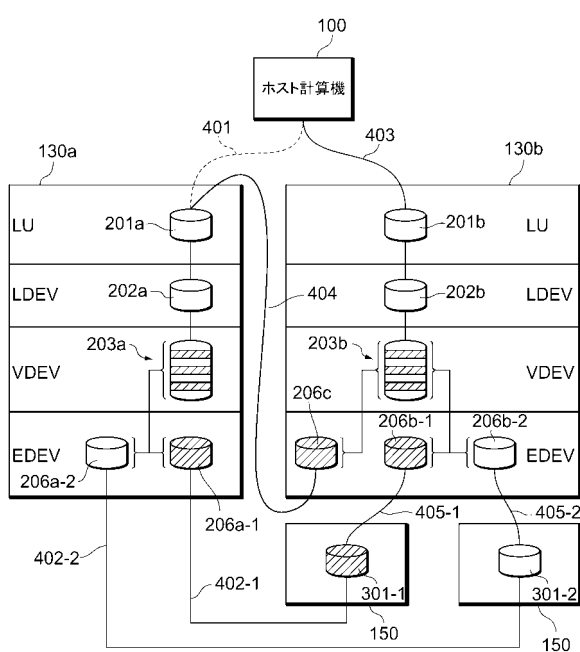
【図35】



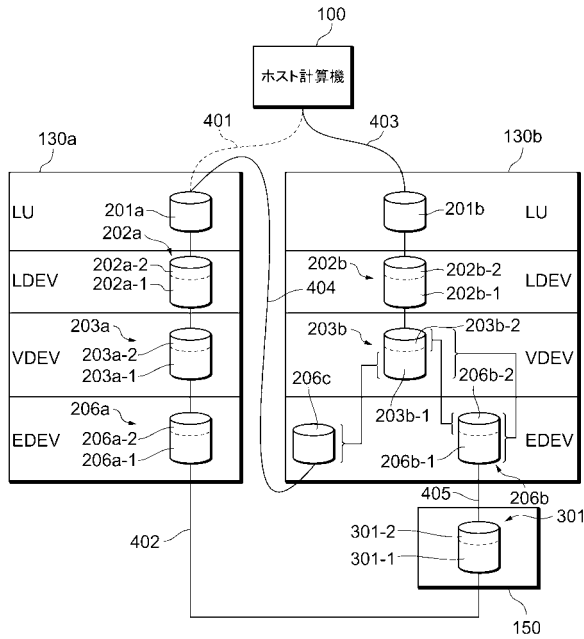
【図36】



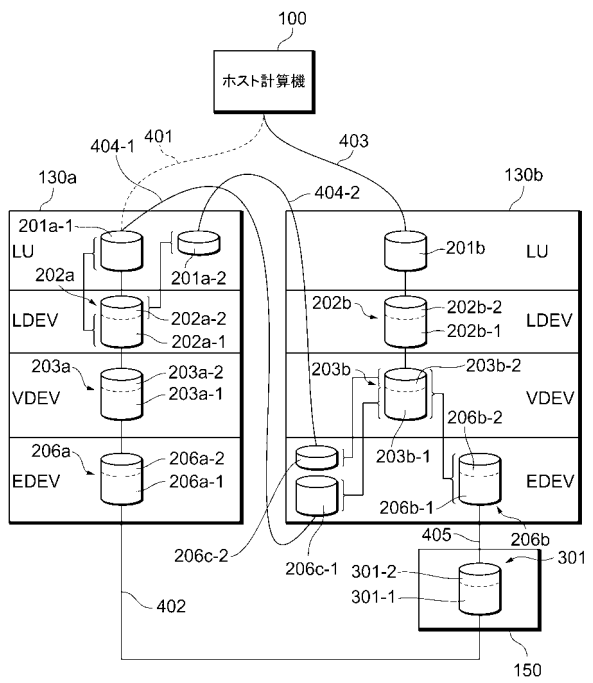
【図37】



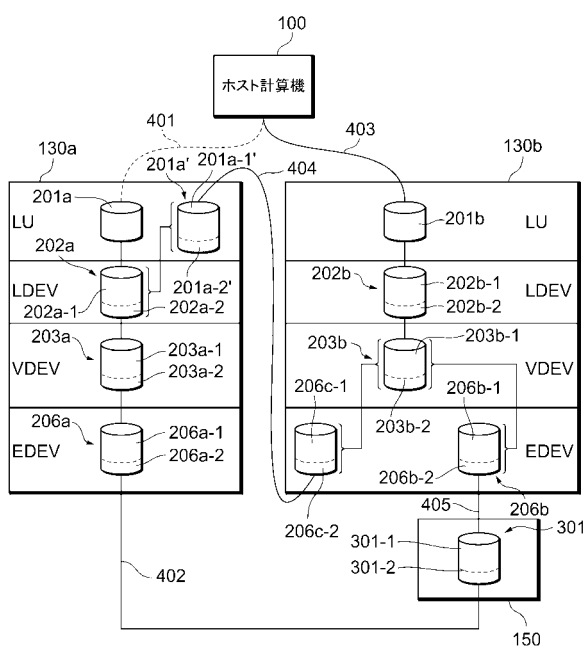
【図38】



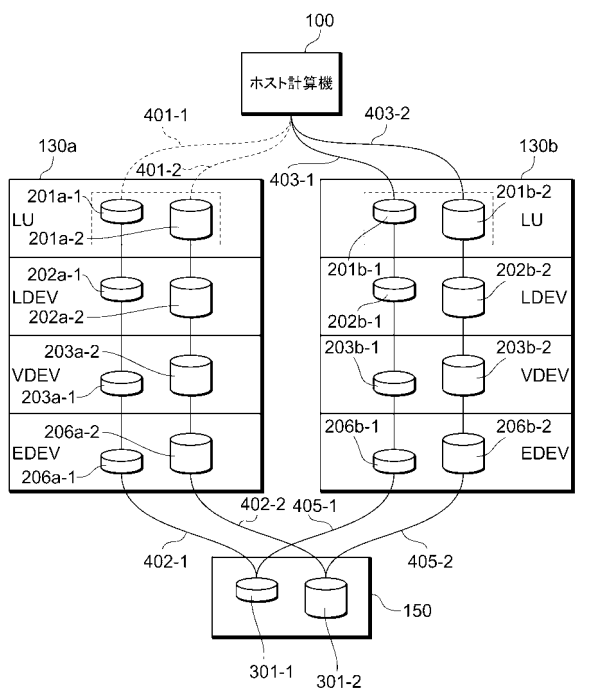
【図39】



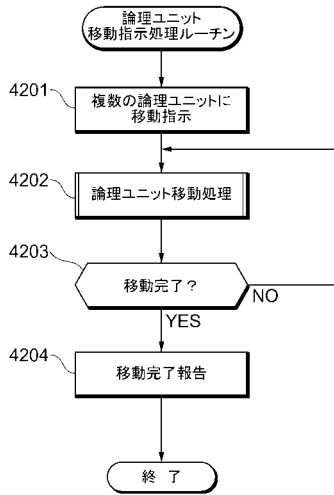
【図40】



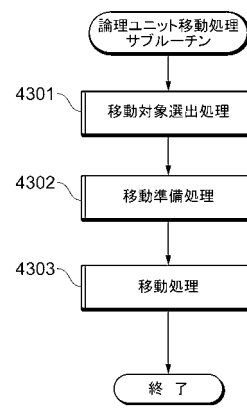
【図41】



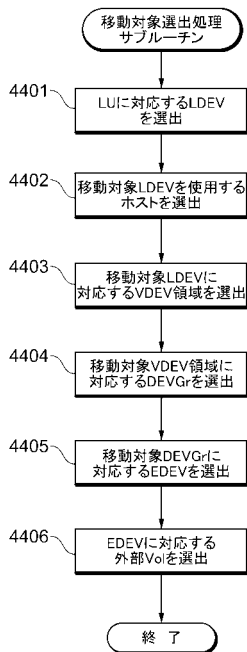
【図42】



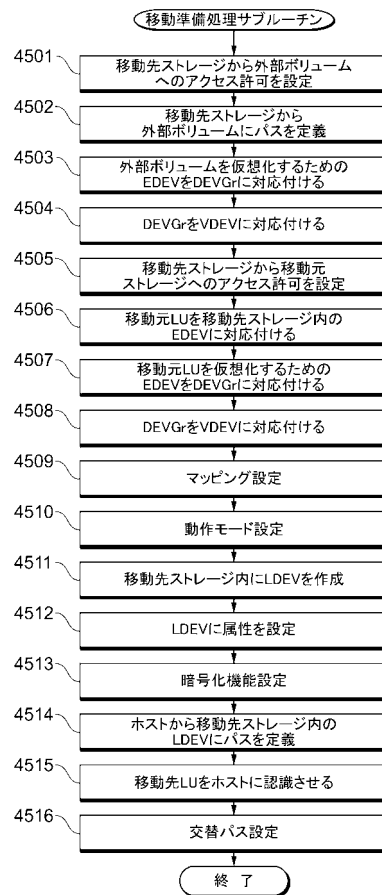
【図43】



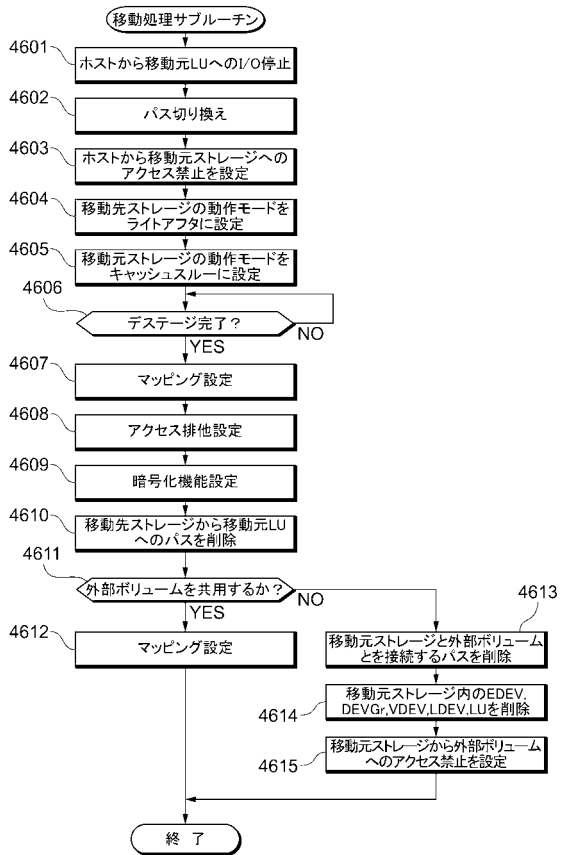
【図44】



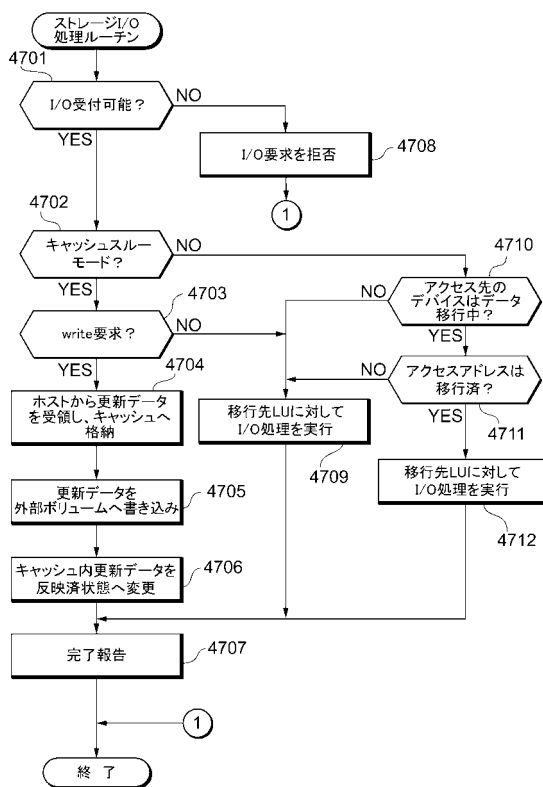
【図45】



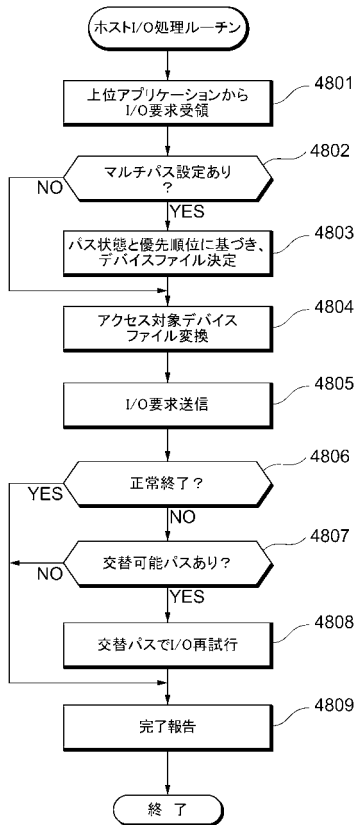
【図46】



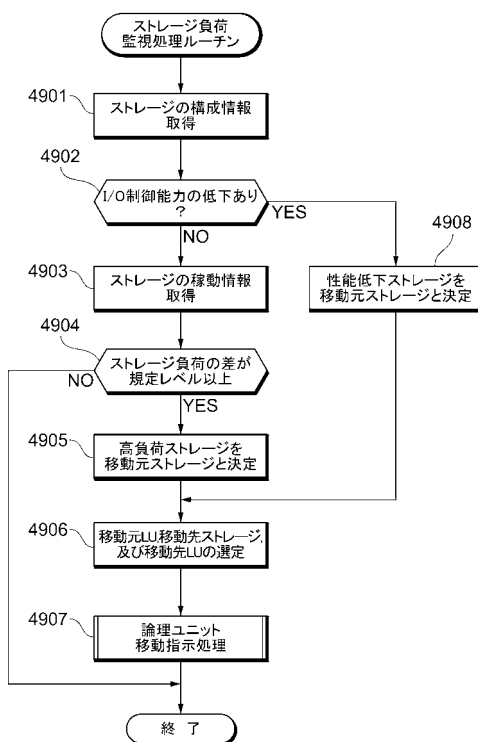
【図47】



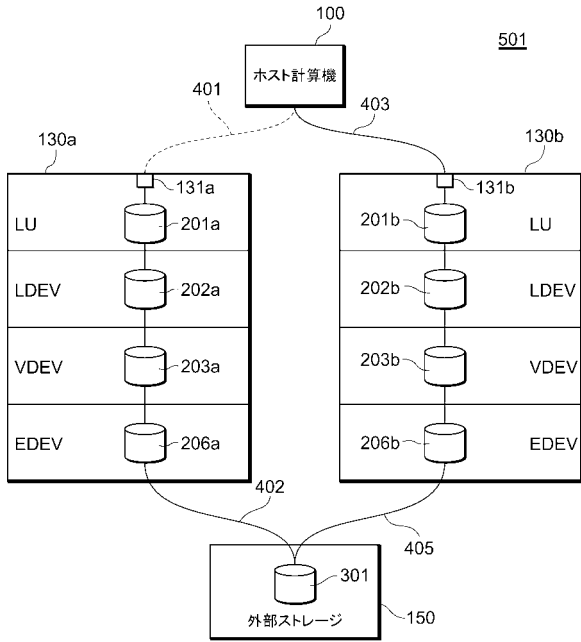
【図48】



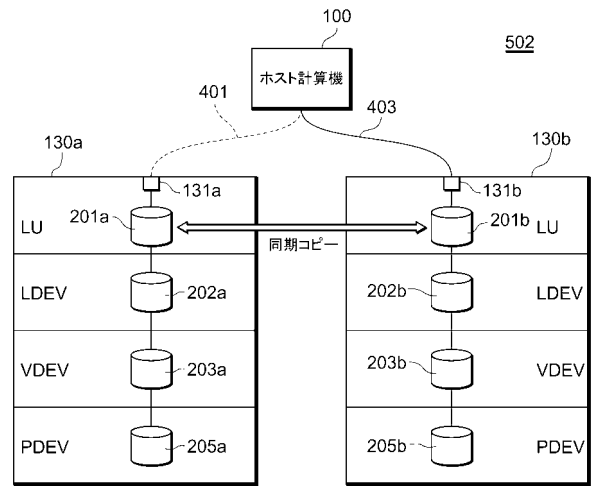
【図49】



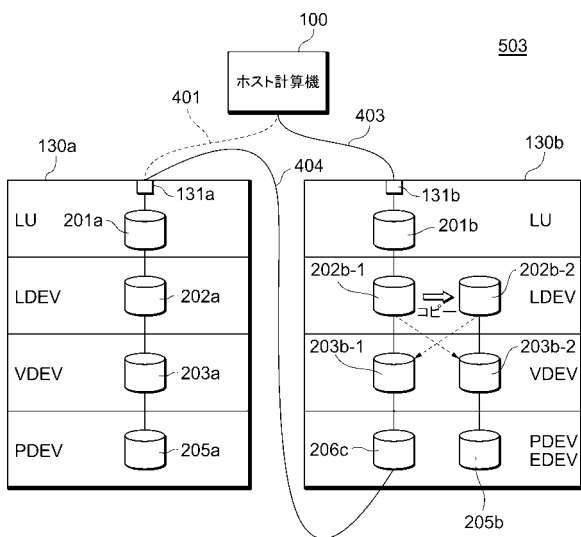
【図50】



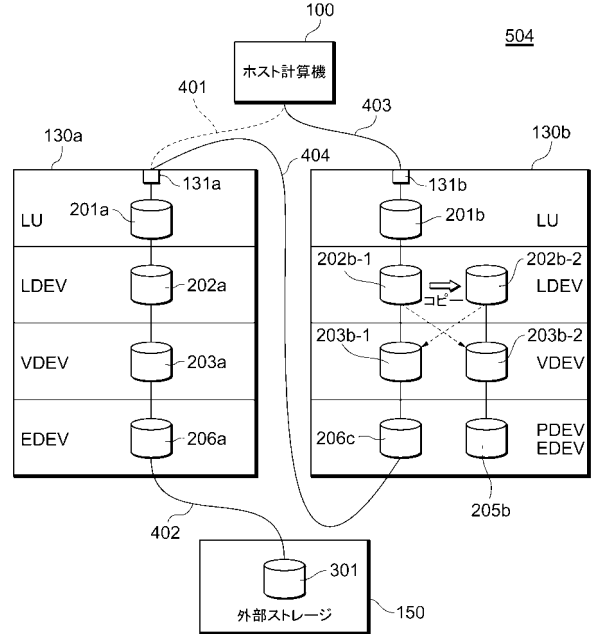
【図51】



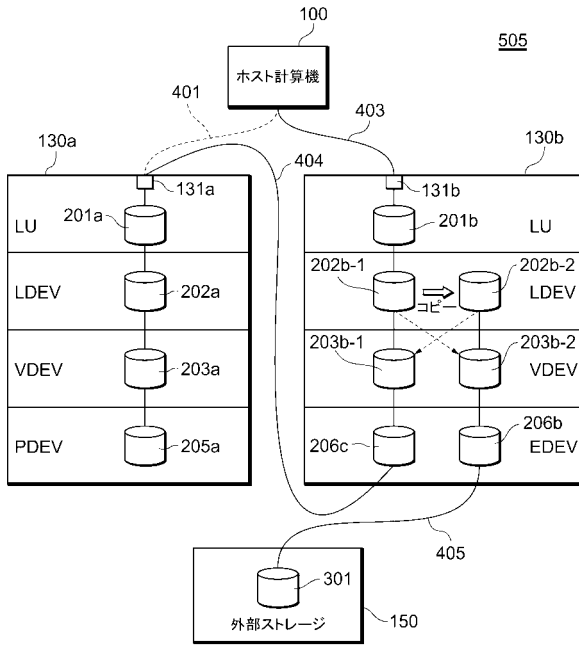
【図52】



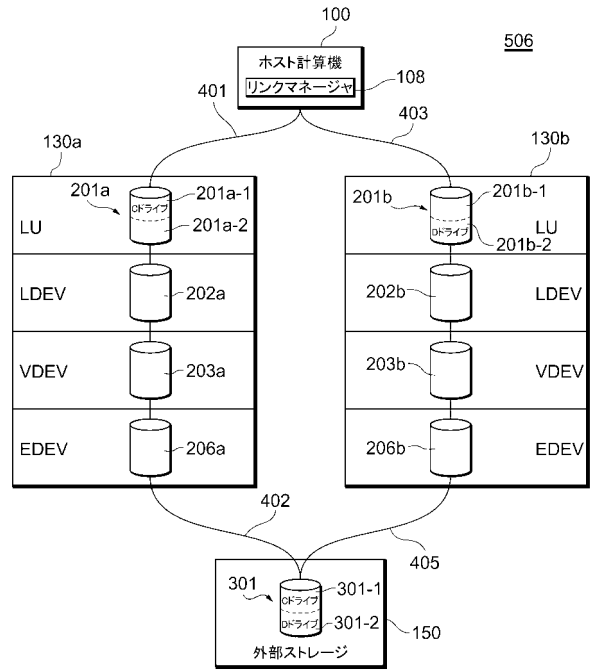
【図53】



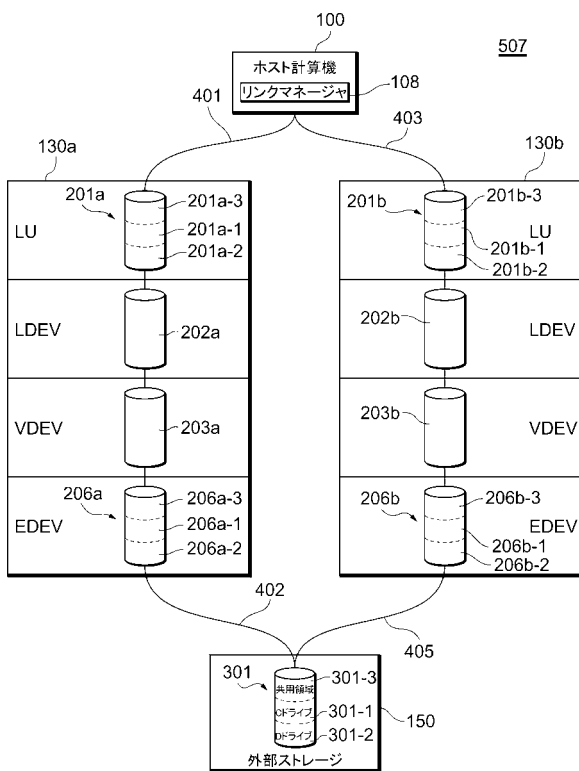
【図54】



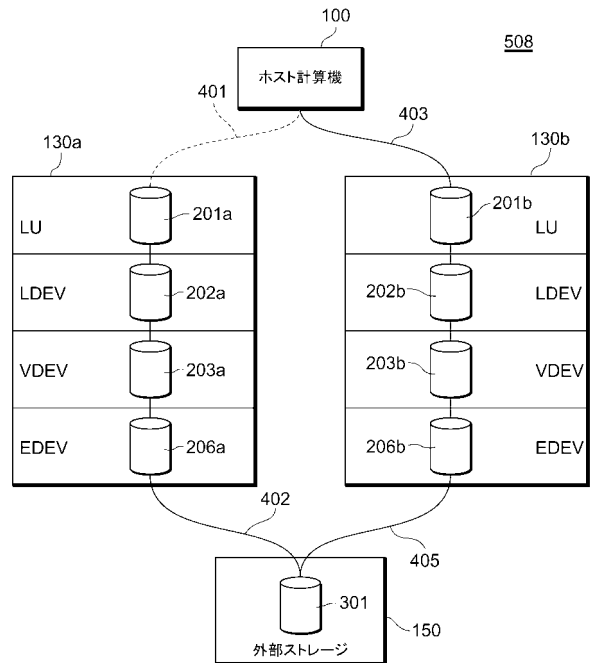
【図55】



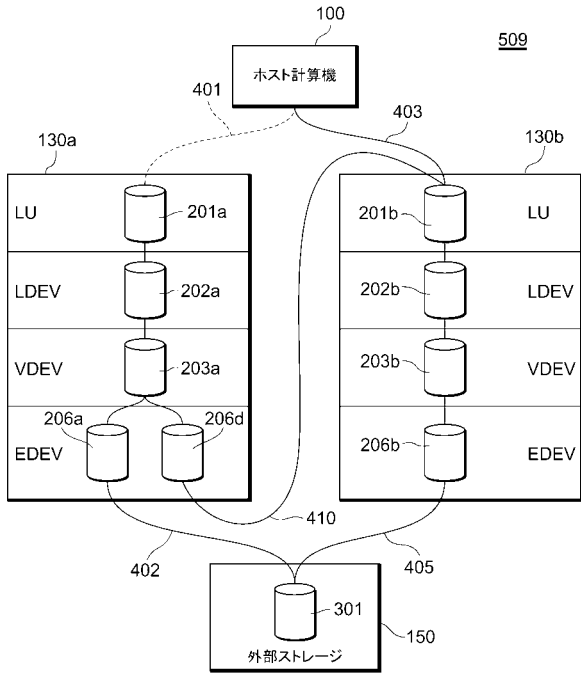
【図56】



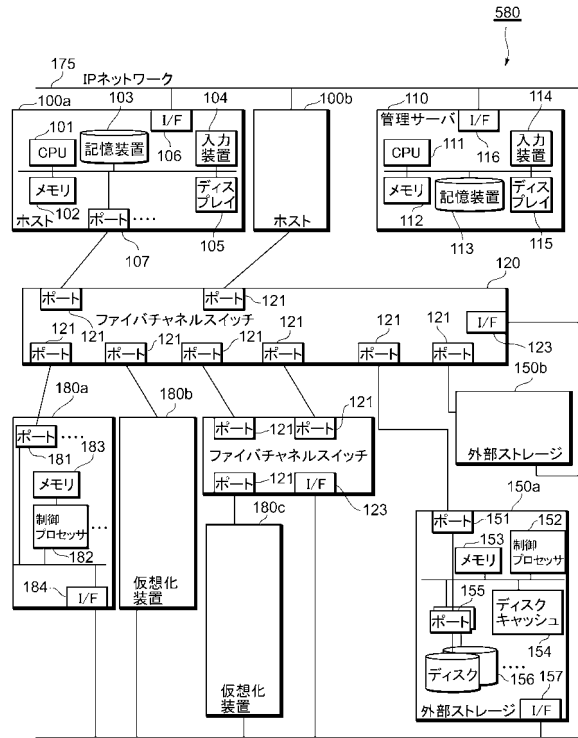
【図57】



【図58】



【図59】



フロントページの続き

(72)発明者 江口 賢哲

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

審査官 横山 佳弘

(56)参考文献 特開平11-007359(JP,A)
特開2002-333956(JP,A)
特開2001-022530(JP,A)
特開平04-365152(JP,A)
特開2004-013454(JP,A)
特開2005-326935(JP,A)
特開2006-031367(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 13/10

G06F 3/06

G06F 12/08