

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B1)

(11) 特許番号

特許第5718533号
(P5718533)

(45) 発行日 平成27年5月13日 (2015. 5. 13)

(24) 登録日 平成27年3月27日 (2015. 3. 27)

| | | | | | | |
|-------------------|------------------|------|-------|------|--|--|
| (51) Int. Cl. | F I | | | | | |
| G06F 3/06 | (2006.01) | G06F | 3/06 | 304F | | |
| G06F 13/14 | (2006.01) | G06F | 3/06 | 301Z | | |
| G06F 12/00 | (2006.01) | G06F | 13/14 | 310H | | |
| | | G06F | 12/00 | 501A | | |
| | | G06F | 12/00 | 531D | | |

請求項の数 21 (全 48 頁)

| | | | |
|---------------|------------------------------|-----------|---------------------------------|
| (21) 出願番号 | 特願2014-534683 (P2014-534683) | (73) 特許権者 | 000005108 |
| (86) (22) 出願日 | 平成26年4月22日 (2014. 4. 22) | | 株式会社日立製作所 |
| (86) 国際出願番号 | PCT/JP2014/061245 | | 東京都千代田区丸の内一丁目6番6号 |
| 審査請求日 | 平成26年7月17日 (2014. 7. 17) | (74) 代理人 | 110000062 |
| 早期審査対象出願 | | | 特許業務法人第一国際特許事務所 |
| | | (72) 発明者 | 鈴木 秀典 |
| | | | 日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内 |
| | | (72) 発明者 | 河野 太貴 |
| | | | 日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内 |
| | | (72) 発明者 | 神 梓 |
| | | | 日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内 |

最終頁に続く

(54) 【発明の名称】 ストレージシステムのデータ移行方法

(57) 【特許請求の範囲】

【請求項1】

ホストに接続され、移行元プライマリストレージ装置のプライマリボリュームと移行元セカンダリストレージ装置のセカンダリボリュームとでリモートコピーを構成している移行元ストレージシステムの移行元リモートコピー構成を、プライマリボリュームを有する移行先プライマリストレージ装置とセカンダリボリュームを有する移行先セカンダリストレージ装置とで構成される移行先ストレージシステムに、移行先リモートコピー構成として移行させるストレージシステムであって、

前記移行先プライマリストレージ装置は、

前記移行元プライマリストレージ装置のプライマリボリュームを記憶領域とする仮想ボリュームを定義し、

前記移行先のプライマリストレージ装置とセカンダリストレージ装置とは、

前記仮想ボリュームと前記移行先セカンダリストレージ装置の前記セカンダリボリュームとでリモートコピーペアを形成し、当該セカンダリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームのデータをコピーし、

前記移行先プライマリストレージ装置は、

前記移行先プライマリストレージ装置の前記プライマリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームからデータを移行させる

ことを特徴とするストレージシステム。

【請求項 2】

前記移行先プライマリストレージ装置は、

前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーと、前記移行先プライマリストレージ装置のプライマリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータの移行とが、完了すると、

前記仮想ボリュームに替えて、前記移行先セカンダリストレージ装置のセカンダリボリュームのコピー元を前記移行先プライマリストレージ装置のプライマリボリュームとする前記移行先リモートコピー構成を形成する

10

ことを特徴とする請求項 1 に記載のストレージシステム。

【請求項 3】

前記移行先プライマリストレージ装置は、

前記仮想ボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとで前記リモートコピーペアを形成するとき、前記移行元プライマリストレージ装置のプライマリボリュームの識別子である第 1 の識別子を有するボリュームとして、前記仮想ボリュームを前記ホストに提供し、

前記移行先プライマリストレージ装置のプライマリボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとで前記移行先リモートコピーペアを形成するとき、前記第 1 の識別子を有するボリュームとして、前記移行先プライマリストレージ装置のプライマリボリュームを前記ホストに提供する

20

ことを特徴とする請求項 2 に記載のストレージシステム。

【請求項 4】

前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーの完了前、前記移行先プライマリストレージ装置は、前記ホストから前記仮想ボリュームへのアクセス要求を受け付けると、前記移行元プライマリストレージ装置に転送し、

前記移行先プライマリストレージ装置から前記移行元プライマリストレージ装置にライト要求が転送されると、

前記移行元プライマリストレージ装置は、当該ライト要求にかかるデータを前記移行元プライマリストレージ装置のプライマリボリュームに格納し、前記移行元セカンダリストレージ装置のセカンダリボリュームに当該データをコピーすることで、前記移行元システムでリモートコピー構成を維持する、

30

ことを特徴とする請求項 1 乃至 3 いずれかに記載のストレージシステム。

【請求項 5】

前記移行先プライマリストレージ装置は、

リモートコピー機能により、前記仮想ボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとの間でリモートコピーを実行し、

データ移行機能により、前記仮想ボリュームと前記移行先プライマリストレージ装置のプライマリボリュームとの間でのデータ移行と、当該データ移行が完了すると当該仮想ボリュームから当該プライマリボリュームへの前記第 1 の識別子の引き継ぎと、を実行し、

40

前記移行先プライマリストレージ装置は、

前記データ移行機能によるデータ移行が完了しても、前記リモートコピー機能によるリモートコピーの完了を待ってから、前記仮想ボリュームから前記移行先プライマリストレージ装置のプライマリボリュームに前記第 1 の識別子を引き継ぐよう、前記データ移行機能を実行することを特徴とする請求項 3 に記載のストレージシステム。

【請求項 6】

前記移行元リモートコピー構成と前記移行先リモートコピー構成とは同期リモートコピー及び非同期リモートコピーのいずれか一方である、

ことを特徴とする請求項 1 乃至 5 のいずれかに記載のストレージシステム。

50

【請求項 7】

前記移行先プライマリストレージ装置のプライマリボリュームの記憶領域は前記移行先プライマリ装置が有する記憶デバイスで構成され、

前記移行先プライマリストレージ装置のプライマリボリュームへのライト処理として、前記記憶デバイスにデータを書き込むことを特徴とする請求項 1 乃至 6 のいずれかに記載のストレージシステム。

【請求項 8】

前記仮想ボリュームは、第 3 の識別子を有し、

前記移行元セカンダリストレージ装置のセカンダリボリュームは、第 2 の識別子を有し

、

前記移行先セカンダリストレージ装置のセカンダリボリュームは、第 5 の識別子を有し

、

前記移行先プライマリストレージ装置は、

前記第 3 の識別子と前記第 1 の識別子とを対応付け、

前記ホストからの Inquiry コマンドに対して、前記仮想ボリュームの識別子として、前記第 1 の識別子を返却し、

前記移行先セカンダリストレージ装置は、

前記第 5 の識別子と前記第 2 の識別子とを対応付け、

前記ホストからの Inquiry コマンドに対して、前記移行先セカンダリストレージ装置のセカンダリボリュームの識別子として、前記第 2 の識別子を返却することを特徴とする請求項 3 又は 5 のいずれかに記載のストレージシステム。

【請求項 9】

前記移行先プライマリストレージ装置のプライマリボリュームは、第 4 の識別子を有し、前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーと、前記移行先プライマリストレージ装置のプライマリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータの移行とが完了すると、

前記移行先プライマリストレージ装置は、

前記仮想ボリュームの識別子と前記移行先プライマリストレージ装置のプライマリボリュームの識別子とを入れ替え、前記移行先プライマリストレージ装置のプライマリボリュームを前記第 1 の識別子が対応付けられた前記第 3 の識別子を有するボリュームとし、前記仮想ボリュームを前記第 4 の識別子を有するボリュームとすることで、前記仮想ボリュームから前記移行先プライマリストレージ装置のプライマリボリュームに前記第 1 の識別子を引き継ぐ

ことを特徴とする請求項 3 又は 5 又は 8 のいずれかに記載のストレージシステム。

【請求項 10】

前記ストレージシステムは、管理端末をさらに有し、

前記移行先プライマリストレージ装置は、前記ホスト又は前記管理端末から前記第 1 の識別子と前記第 2 の識別子を指定したペア作成指示を受領すると、

前記第 1 の識別子と対応付けられた前記第 3 の識別子を有するボリュームと、前記第 2 の識別子と対応付けられた前記第 5 の識別子を有するボリュームとで、前記移行先ストレージシステムにおいてリモートコピーペアを形成する、ことを特徴とする請求項 8 に記載のストレージシステム。

【請求項 11】

ホストに接続され、移行元プライマリストレージ装置のプライマリボリュームと移行元セカンダリストレージ装置のセカンダリボリュームとでリモートコピーを構成している移行元ストレージシステムの移行元リモートコピー構成を、プライマリボリュームを有する移行先プライマリストレージ装置とセカンダリボリュームを有する移行先セカンダリストレージ装置とで構成される移行先ストレージシステムに、移行先リモートコピー構成として移行させるリモートコピー構成の移行方法であって、

10

20

30

40

50

前記移行元プライマリストレージ装置のプライマリボリュームを記憶領域とする仮想ボリュームを前記移行先プライマリストレージ装置に定義し、

前記仮想ボリュームと前記移行先セカンダリストレージ装置の前記セカンダリボリュームとでリモートコピーペアを形成し、

当該セカンダリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームのデータをコピーし、

前記移行先プライマリストレージ装置の前記プライマリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームからデータを移行する

ことを特徴とするリモートコピー構成の移行方法。

10

【請求項 1 2】

前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーと、前記移行先プライマリストレージ装置のプライマリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータの移行とが、完了すると、

前記仮想ボリュームに替えて、前記移行先セカンダリストレージ装置のセカンダリボリュームのコピー元を前記移行先プライマリストレージ装置のプライマリボリュームとする前記移行先リモートコピー構成を形成する

ことを特徴とする請求項 1 1 に記載のリモートコピー構成の移行方法。

【請求項 1 3】

20

前記仮想ボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとで前記リモートコピーペアを形成するとき、前記移行元プライマリストレージ装置のプライマリボリュームの識別子である第 1 の識別子を有するボリュームとして、前記仮想ボリュームを前記ホストに提供し、

前記移行先プライマリストレージ装置のプライマリボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとで前記移行先リモートコピーペアを形成するとき、前記第 1 の識別子を有するボリュームとして、前記移行先プライマリストレージ装置のプライマリボリュームを前記ホストに提供する

ことを特徴とする請求項 1 2 に記載のリモートコピー構成の移行方法。

【請求項 1 4】

30

前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーの完了前、

前記ホストから前記仮想ボリュームへのアクセス要求を受け付けると、前記移行元プライマリストレージ装置に転送し、

前記移行先プライマリストレージ装置から前記移行元プライマリストレージ装置にライト要求が転送されると、

当該ライト要求にかかるデータを前記移行元プライマリストレージ装置のプライマリボリュームに格納し、前記移行元セカンダリストレージ装置のセカンダリボリュームに当該データをコピーすることで、前記移行元システムでリモートコピー構成を維持する、

ことを特徴とする請求項 1 1 乃至 1 3 のいずれかに記載のリモートコピー構成の移行方法

40

【請求項 1 5】

リモートコピー機能により、前記仮想ボリュームと前記移行先セカンダリストレージ装置のセカンダリボリュームとの間でリモートコピーを実行し、

データ移行機能により、前記仮想ボリュームと前記移行先プライマリストレージ装置のプライマリボリュームとの間でのデータ移行と、当該データ移行が完了すると当該仮想ボリュームから当該プライマリボリュームへの前記第 1 の識別子の引き継ぎと、を実行し、

前記データ移行機能によるデータ移行が完了しても、前記リモートコピー機能によるリモートコピーの完了を待ってから、前記仮想ボリュームから前記移行先プライマリストレージ装置のプライマリボリュームに前記第 1 の識別子を引き継ぐよう、前記データ移行機

50

能を実行する、

ことを特徴とする請求項 1 3 に記載のリモートコピー構成の移行方法。

【請求項 1 6】

前記移行元リモートコピー構成と前記移行先リモートコピー構成とは同期リモートコピー又は非同期リモートコピーである、

ことを特徴とする請求項 1 1 乃至 1 5 のいずれかに記載のリモートコピー構成の移行方法。

【請求項 1 7】

前記移行先プライマリストレージ装置のプライマリボリュームの記憶領域は前記移行先プライマリ装置が有する記憶デバイスで構成され、

前記移行先プライマリストレージ装置のプライマリボリュームへのライト処理として、前記記憶デバイスにデータを書き込む、

ことを特徴とする請求項 1 1 乃至 1 6 のいずれかに記載のリモートコピー構成の移行方法。

【請求項 1 8】

前記仮想ボリュームは、第 3 の識別子を有し、

前記移行元セカンダリストレージ装置のセカンダリボリュームは、第 2 の識別子を有し、

前記移行先セカンダリストレージ装置のセカンダリボリュームは、第 5 の識別子を有し、

前記第 3 の識別子と前記第 1 の識別子とを対応付け、

前記ホストからの Inquiry コマンドに対して、前記仮想ボリュームの識別子として、前記第 1 の識別子を返却し、

前記第 5 の識別子と前記第 2 の識別子とを対応付け、

前記ホストからの Inquiry コマンドに対して、前記仮想ボリュームの識別子として、前記第 2 の識別子を返却し、

ことを特徴とする請求項 1 3 又は 1 5 のいずれかに記載のリモートコピー構成の移行方法。

【請求項 1 9】

前記移行先プライマリストレージ装置のプライマリボリュームは、第 4 の識別子を有し、前記移行先セカンダリストレージ装置のセカンダリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータのコピーと、前記移行先プライマリストレージ装置のプライマリボリュームへの前記移行元プライマリストレージ装置のプライマリボリュームのデータの移行とが完了すると、

前記仮想ボリュームの識別子と前記移行先プライマリストレージ装置のプライマリボリュームの識別子とを入れ替え、前記移行先プライマリストレージ装置のプライマリボリュームを前記第 1 の識別子が対応付けられた前記第 3 の識別子を有するボリュームとし、前記仮想ボリュームを前記第 4 の識別子を有するボリュームとすることで、前記仮想ボリュームから前記移行先プライマリストレージ装置のプライマリボリュームに前記第 1 の識別子を引き継ぐこと、

ことを特徴とする請求項 1 3 又は 1 5 又は 1 8 に記載のリモートコピー構成の移行方法。

【請求項 2 0】

前記移行先プライマリストレージ装置は、前記ホスト又は管理端末から前記第 1 の識別子と前記第 2 の識別子を指定したペア作成指示を受領すると、

前記第 1 の識別子と対応付けられた前記第 3 の識別子を有するボリュームと、前記第 2 の識別子と対応付けられた前記第 5 の識別子を有するボリュームとで、前記移行先ストレージシステムにおいてリモートコピーペアを形成する、

ことを特徴とする請求項 1 8 に記載のリモートコピー構成の移行方法。

【請求項 2 1】

移行元プライマリストレージ装置が有する移行元プライマリボリュームと移行元セカン

10

20

30

40

50

ダリストレージ装置が有する移行元セカンダリボリュームとで移行元リモートコピー構成を形成する移行元ストレージシステム及びホストに接続され、移行先プライマリストレージ装置が有する移行先プライマリボリュームと移行先セカンダリストレージ装置が有する移行先セカンダリボリュームとに前記移行元リモートコピー構成を移行先リモートコピー構成として移行する移行先ストレージシステムであって、

前記移行先プライマリストレージ装置は、

前記移行元プライマリストレージ装置のプライマリボリュームを記憶領域とする仮想ボリュームを定義し、

前記移行先のプライマリストレージ装置とセカンダリストレージ装置は、

前記仮想ボリュームと前記セカンダリボリュームとでリモートコピーペアを形成し、当該セカンダリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームのデータをコピーし、

前記移行先プライマリストレージ装置は、

当該プライマリボリュームに前記仮想ボリュームの記憶領域である前記移行元プライマリストレージ装置のプライマリボリュームからデータを移行させることを特徴とする移行先ストレージシステム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ装置間のボリューム移行技術に関する。

20

【背景技術】

【0002】

ITの進歩やインターネットの普及などにより、企業等における計算機システムが扱うデータ量は増加を続けており、データを格納するストレージ装置に対しても、より大容量、より高性能なストレージ装置が求められる。そのため、ストレージ装置は所定の年数を経過すると、新世代の装置に置き換えられる。

【0003】

企業等における計算機システムは、無停止で運用されることが原則であり、ストレージ装置も例外ではない。そのため、ストレージ装置の置き換えの際にも、ダウンタイムを最小限にとどめることが求められる。たとえば特許文献1では、移行元ストレージAから移行先ストレージBへのボリューム移行が開示されている。特許文献1には、ストレージ装置内またはストレージ装置間でボリュームのミラーリングが行われているボリュームペアを移行すること、及び移行元装置から移行先装置へのボリュームの移行中にI/O要求を受付可能であることが開示されている。

30

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2007-115221号公報

【特許文献2】米国特許出願公開第2005/0149577号明細書

【特許文献3】米国特許第6101497号明細書

【特許文献4】米国特許第5742792号明細書

【特許文献5】米国特許第7152079号明細書

40

【発明の概要】

【発明が解決しようとする課題】

【0005】

ボリュームペアを用いた運用を行っている計算機システムでは、必要に応じてボリュームペアの操作を行う。たとえばボリュームペアをバックアップ用途に用いる場合、ボリュームのバックアップを行う際に、コピー元ボリューム（正ボリューム）からコピー先ボリューム（副ボリューム）へのミラーリングを一時停止し、ミラーリングが停止された副ボ

50

リユームのデータをバックアップデバイスにバックアップし、バックアップが終了すると、ミラーリングを再開する、といった操作が行われる。この操作の際、ボリュームペアの所属するストレージ装置の識別番号（たとえば製造番号）やボリューム番号といった、ボリュームを一意に特定可能な識別子を指定したコマンドを発行することで、ボリュームペアの操作を可能にしている。また、バックアップなどは定型的作業であるため、バックアップ時に必要となるボリュームペアの一連の操作コマンドは、サーバ上のスクリプトファイルに記述されていることが多い。

【 0 0 0 6 】

ここでストレージ装置の置き換えが行われると、ストレージ装置の製造番号はもちろん、ボリューム番号も変更されることがある。そうすると、バックアップ用のスクリプトファイルについても、スクリプトファイルに記述されているボリューム番号などの情報をすべて書き替える必要がある。大規模な計算機システムで、多数（たとえば数百以上）のボリュームペアを管理しているような場合、その変更量は膨大で、スクリプトファイルの変更に長時間を要する。結果として計算機システムのダウンタイムが長期化することになる。特許文献 1 に開示の技術では、データの移行中であっても計算機システムからの I / O は可能であるが、ボリューム番号などの識別子の変更に係る考慮は無い。

10

【 0 0 0 7 】

本発明の目的は、ボリュームの移行に際し、計算機システムへの影響を最小限に抑えることにある。

【課題を解決するための手段】

20

【 0 0 0 8 】

本発明のボリューム移行方法では、移行元ストレージシステム内でボリュームペアを形成している少なくとも 2 つのボリューム（移行元 P - V O L、移行先 S - V O L）を、無停止で移行先ストレージシステムへと移行する。移行処理の最初に、移行先ストレージシステムに、移行元ストレージシステム内のボリュームペアの複製元ボリューム（移行元 P - V O L）の記憶領域をマッピングした仮想的な論理ボリューム（移行先 P - V O L）と、移行先 P - V O L の複製先ボリュームとなる移行先 S - V O L とを作成する。

【 0 0 0 9 】

移行先ストレージシステムは、移行先 P - V O L と移行先 S - V O L に対して、移行元 P - V O L 及び移行元 S - V O L の有する識別子と同一の識別子を付し、ホスト計算機に対して移行先 P - V O L と移行先 S - V O L が、移行元 P - V O L 及び移行元 S - V O L と同一ボリュームであると認識させる。続いてホスト計算機から移行元 P - V O L へのアクセスパスを、移行先 P - V O L へと切り替える。

30

【 0 0 1 0 】

その後、移行先 P - V O L と移行先 S - V O L との間で、ボリュームペアの作成を行い、移行先 S - V O L に移行先 P - V O L のデータの複製が格納される状態にする。最後に移行先 P - V O L の記憶領域を、移行先ストレージシステム内の記憶領域へと移動することで、移行処理を完了させる。

【発明の効果】**【 0 0 1 1 】**

40

本発明の移行方法では、ホスト I / O を停止することなく、ストレージ装置間でボリュームペアを移行できるとともに、ストレージ装置間でのボリュームペア移行を行っても、各ボリュームの識別子が変わらないため、ボリュームを使用するホスト計算機側での設定変更が不要となる。

【図面の簡単な説明】**【 0 0 1 2 】**

【図 1】本発明の実施例 1 に係る計算機システムの概略構成図を示す。

【図 2】ストレージ装置のハードウェア構成を示す。

【図 3】パリティグループ管理テーブルの構成例を示す。

【図 4】外部ボリュームグループ管理テーブルの構成例を示す。

50

【図 5】論理ボリューム管理テーブルの構成例を示す。

【図 6】LU管理テーブルの構成例を示す。

【図 7】ペア管理テーブルの構成例を示す。

【図 8】移行先プライマリストレージ装置に仮想ストレージを定義した場合の例を示した図である。

【図 9】V - B O X 管理テーブルの構成例を示す。

【図 10】仮想 L D E V 管理テーブルの構成例を示す。

【図 11】ストレージマネージャ 101 が使用する設定ファイルの例を示す。

【図 12】ストレージマネージャ 101 が使用する設定ファイルの例を示す。

【図 13】ストレージマネージャ 101 が使用する設定ファイルの例を示す。

【図 14】本発明の実施例 1 に係る計算機システムにおける、移行処理の概要を表した図である。

【図 15】本発明の実施例 1 に係る計算機システムにおける、移行処理後の状態を表した図である。

【図 16】本発明の実施例 1 に係る計算機システムにおける、ボリューム移行処理のフローチャート (1) である。

【図 17】本発明の実施例 1 に係る計算機システムにおける、ボリューム移行処理のフローチャート (2) である。

【図 18】本発明の実施例 1 に係る計算機システムにおける、ボリューム移行処理のフローチャート (3) である。

【図 19】本発明の変形例 2 に係る計算機システムの構成図である。

【図 20】本発明の変形例 2 に係る計算機システムにおける、移行処理中のデータの流れを表した図である。

【図 21】本発明の実施例 2 に係る計算機システムの概略構成図である。

【図 22】本発明の実施例 2 に係る計算機システムにおける、ボリューム移行処理のフローチャート (1) である。

【図 23】本発明の実施例 2 に係る計算機システムにおける、ボリューム移行処理のフローチャート (2) である。

【図 24】本発明の実施例 2 に係る計算機システムにおける、ボリューム移行処理のフローチャート (3) である。

【図 25】本発明の実施例 3 に係る計算機システムの構成図である。

【図 26】本発明の実施例 3 に係る計算機システムにおける、ボリューム移行処理のフローチャートである。

【図 27】ストレージシステムの拡張に適用した例を表した図である。

【発明を実施するための形態】

【0013】

以下、図面を参照して、本発明の一実施形態に係るストレージシステムを説明する。なお、本発明は、以下に説明する実施形態に限定されるものではない。

【実施例 1】

【0014】

[計算機システムの構成]

図 1 は、本発明の実施例 1 に係る計算機システムの概略構成図である。計算機システムは、移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b、移行先プライマリストレージ装置 20 a、移行先セカンダリストレージ装置 20 b、プライマリサーバ 30 a、セカンダリサーバ 30 b からなる。

【0015】

なお、図 1 は、以下で説明する移行処理の処理中の構成を表したものであり、移行処理前は、移行先プライマリストレージ装置 20 a と移行先セカンダリストレージ装置 20 b は存在しない。移行処理前の状態では、プライマリサーバ 30 a は S A N (S t o r a g e A r e a N e t w o r k) 50 を介して移行元プライマリストレージ装置 10 a に

10

20

30

40

50

接続されており、移行元プライマリストレージ装置 10 a の有する論理ボリューム 130 a へのアクセスを行っている。またセカンダリサーバ 30 b は移行処理を実施する前は、SAN 60 を介して移行元セカンダリストレージ装置 10 b に接続され、移行元セカンダリストレージ装置 10 b の論理ボリューム 130 b にアクセス可能な状態にある。また、移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b はポート (11 a、11 b) により相互接続されている。なお、本発明の実施例 1 に係る計算機システムにおいて、SAN 50 及び SAN 60 はファイバチャネルケーブル及びファイバチャネルスイッチで構成されたネットワークであるが、イーサネットを用いて構成されたネットワークであってもよい。

【0016】

10

移行処理を行う際に、移行先プライマリストレージ装置 20 a と移行先セカンダリストレージ装置 20 b が計算機システム内に設置される。そして図 1 に示されているように、移行先プライマリストレージ装置 20 a は SAN 50 を介してプライマリサーバ 30 a に接続されるとともに、移行元プライマリストレージ装置 10 a 及び移行先セカンダリストレージ装置 20 b にも接続される。また移行先セカンダリストレージ装置 20 b は、SAN 60 を介してセカンダリサーバ 30 b に接続される。

【0017】

なお、以下では、移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b をまとめて、「移行元ストレージシステム 10」と呼ぶ。また移行先プライマリストレージ装置 20 a、移行先セカンダリストレージ装置 20 b をまとめて、「移行先ストレージシステム 20」と呼ぶ。そして移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b、移行先プライマリストレージ装置 20 a、移行先セカンダリストレージ装置 20 b を総称して、「ストレージ装置」と呼ぶ。また、プライマリサーバ 30 a、セカンダリサーバ 30 b を総称して、「ホスト計算機」と呼ぶ。

20

【0018】

ここで、計算機システム内のストレージ装置内部のハードウェア構成について、図 2 を用いて説明する。ここでは主に、移行先プライマリストレージ装置 20 a のハードウェア構成について説明するが、他のストレージ装置 (移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b、移行先セカンダリストレージ装置 20 b) も同様のハードウェア構成を採る。ただし、必ずしもすべてのストレージ装置が同一のハードウェア構成を採用している必要はない。

30

【0019】

移行先プライマリストレージ装置 20 a は、フロントエンドパッケージ (FEPK) 201、バックエンドパッケージ (BEPK) 202、プロセッサパッケージ (MPPK) 203、キャッシュメモリパッケージ (CMPK) 204 が相互結合網 205 で互いに結合されて構成されたコントローラ 200 と、複数のドライブ 221 を搭載したディスクユニット (DKU) 210 から構成される。

【0020】

フロントエンドパッケージ (FEPK) 201 は、プライマリサーバ 30 a などのホスト計算機や、ストレージ装置 (10 a, 10 b, 20 b) と接続するためのポート 21 を複数有し、CMPK 204 や MPPK 203 との間で送受信される制御情報やデータの中継を行うコンポーネントである。FEPK 201 はポート 21 の他、図示しないバッファ、CPU、内部バス、及び内部ポートを有する。バッファは、FEPK 201 が中継する制御情報及びデータを一時的に記憶するための記憶領域であり、CMPK 204 と同様に、各種の揮発性メモリや不揮発性メモリを用いて構成される。内部バスは、FEPK 201 内の各種コンポーネントを相互接続する。FEPK 201 は、内部ポートを介して相互結合網 205 に接続されている。また、ポート 21 は、図で示された数に限定されるものではない。

40

【0021】

バックエンドパッケージ (BEPK) 202 は、ドライブ 221 と接続するためのイン

50

タフェース（I/F）2021を複数有し、ドライブ221と、CMPK204またはMPPK203との間で送受信される制御情報やデータの中継を行うコンポーネントである。BEPK202はインタフェース（I/F）2021の他、図示しないバッファ、CPU、内部バス、及び内部ポートを有する。バッファは、BEPK202が中継する制御情報やデータを一時的に記憶するための記憶領域であり、CMPK204と同様に、各種の揮発性メモリや不揮発性メモリを用いて構成される。内部バスは、FEPK201内の各種コンポーネントを相互接続する。

【0022】

MPPK203は、CPU2031、メモリ（LM）2032、そして非図示の内部バス及び内部ポートとから構成される。メモリ（LM）2032は、キャッシュメモリ2041と同様に、各種の揮発性メモリや不揮発性メモリを用いて構成することができる。CPU2031、メモリ（LM）2032及び内部ポートは、内部バスを介して相互接続される。MPPK203は、MPPK203の内部ポートを介して相互結合網205に接続されている。

10

【0023】

またMPPK203は、ストレージ装置内の各種データ処理を行うためのコンポーネントである。CPU2031がメモリ（LM）2032上のプログラムを実行することにより、以下で説明する、ストレージ装置が有するボリュームコピー機能等の各種機能が実現される。

【0024】

CMPK204は、キャッシュメモリ2041と共有メモリ2042を備えたコンポーネントである。キャッシュメモリ2041は、プライマリサーバ30aあるいは別のストレージ装置から受信したデータを一時的に記憶し、及びドライブ221から読み出したデータを一時的に記憶するため記憶領域である。キャッシュメモリ2041は例えば、DRAMやSRAMのような揮発性メモリ、あるいは、フラッシュメモリ等の揮発性メモリを用いて構成される。

20

【0025】

共有メモリ2042は、ストレージ装置内の各種データ処理に関連する管理情報を記憶するための記憶領域である。共有メモリ2042は、キャッシュメモリ2041と同様に、各種の揮発性メモリや不揮発性メモリを用いて構成することができる。なお、共有メモリ2042のハードウェアとしては、キャッシュメモリ2041と共通のハードウェアを用いることもできるし、共通でないハードウェアを用いることもできる。

30

【0026】

相互結合網205は、コンポーネント間を相互に接続し、相互に接続されたコンポーネント間で制御情報やデータを転送するためのコンポーネントである。相互結合網は、例えば、スイッチやバスを用いて構成することができる。

【0027】

ドライブ221は、プライマリサーバ30a上の各種プログラムが使用するデータ（ユーザデータ）、冗長データ（パリティデータ）を格納するための記憶デバイスである。ドライブ221の記憶メディアとしては、HDDで用いられるような磁気記憶媒体のほか、NANDフラッシュメモリ、MRAM、ReRAM、及びPRAMのような不揮発性半導体記憶媒体を用いることができる。

40

【0028】

図1の説明に戻る。本発明の実施例1に係る各ストレージ装置は、1以上の論理ボリューム（130a、130b、230a、230b）とコマンドデバイス（135a、135b、235a、235b）を有する。論理ボリューム（130a、130b、230a、230b）は、プライマリサーバ30a等のホスト計算機からのライトデータを格納する記憶デバイスであり、詳細は後述する。コマンドデバイス（135a、135b、235a、235b）は、論理ボリュームの一種であるが、ホスト計算機からのライトデータを格納するために用いられるわけではない。プライマリサーバ30aまたはセカンダリサ

50

サーバ30b上のストレージマネージャ101から、論理ボリュームのコピー指示などの各種制御命令を受け付けるために用いられる記憶デバイスである。本発明の実施例1に係る計算機システムでは、各ストレージ装置が1つのコマンドデバイスを有する。コマンドデバイスは、管理者が管理端末を用いて、各ストレージ装置にコマンドデバイスの作成指示の命令を発行することによって作成される。

【0029】

本発明の実施例1に係る各ストレージ装置は、多数のポートを有するため、以下では各ポートを区別するため、図1に示されている通り、各ポートにA～Rの識別子を付して説明する。移行元プライマリストレージ装置10aは、ポートA～Eを有し、ポートA、Bはプライマリサーバ30aと接続するために用いられ、またポートD、Eは、移行元セカンダリストレージ装置10bと接続するために用いられる。またポートCは移行処理時に移行先プライマリストレージ装置20aと接続するために用いられる。

10

【0030】

移行元セカンダリストレージ装置10bは、ポートF～Iを有し、ポートF、Gは移行元プライマリストレージ装置10aと接続するために用いられる。ポートH、Iはセカンダリサーバ30bと接続するために用いられる。

【0031】

移行先プライマリストレージ装置20aはポートJ～Nを有し、ポートJ、Kはプライマリサーバ30aと接続するために用いられ、またポートM、Nは、移行先セカンダリストレージ装置20bと接続するために用いられる。またポートLは移行元プライマリストレージ装置10aと接続するために用いられる。

20

【0032】

移行先セカンダリストレージ装置20bは、ポートO～Rを有し、ポートO、Pは移行先プライマリストレージ装置20aと接続するために用いられる。ポートQ、Rはセカンダリサーバ30bと接続するために用いられる。

【0033】

なお、各ストレージ装置は、上で説明した数のポートのみを有する構成に限定されるわけではなく、上で説明した数より多くのポートを有していてもよい。

【0034】

プライマリサーバ30a、セカンダリサーバ30bは、非図示のCPU、メモリを有した計算機である。また、プライマリサーバ30aとセカンダリサーバ30bは、非図示の通信ネットワークで接続されている。通信ネットワークには、LAN(Local Area Network)やWAN(Wide Area Network)が用いられる。CPUはメモリ上にロードされた各種のプログラムを実行する。プライマリサーバ30a上及びセカンダリサーバ30b上で動作するプログラムには、アプリケーションプログラム(AP)100、ストレージマネージャ101、交替パスソフト102、クラスタソフトウェア103がある。

30

【0035】

アプリケーションプログラム(AP)100は一例として、データベース管理システムなどのプログラムで、論理ボリューム130aに対するアクセス(リード、ライト)を行う。また移行元ストレージシステムでは論理ボリューム130aにライトされたデータを、論理ボリューム130bに複製するリモートコピー機能が稼働しており、論理ボリューム130a上のデータは、常に論理ボリューム130bにも複製される。

40

【0036】

クラスタソフトウェア103は、いわゆるディザスタリカバリ運用を行うためのプログラムで、プライマリサーバ30aとセカンダリサーバ30bとで稼働する。例えばプライマリサーバ30a、移行元プライマリストレージ装置10aが障害などでダウンした場合、セカンダリサーバ30bで稼働しているクラスタソフトウェア103は、プライマリサーバ30aのAP100で行っていた業務を再開させる処理を行う。

【0037】

50

交替パスソフト102は、プライマリサーバ30a等のホスト計算機から、ストレージ装置のボリューム（論理ボリューム）へのアクセス経路（パス）を管理しており、1つのパスが障害で切断された場合でもボリュームへのアクセスを可能にするため、代替パスを用いてボリュームへのアクセスを継続させるプログラムである。

【0038】

ストレージマネージャ101は、ストレージ装置の設定、制御を行うためのプログラムである。詳細な説明は後述する。

【0039】

[移行処理で用いられるストレージ装置の機能]

続いて、本発明の実施例における移行方法の説明に必要な、ストレージ装置の機能について説明する。ここでは主に、移行先プライマリストレージ装置20aの持つ機能を中心に説明する。なお、先に述べたとおり、以下で説明するストレージ装置の各機能は、ストレージ装置のCPU2031がメモリ(LM)2032上のプログラムを実行することにより実現される。

【0040】

(1) ボリューム作成機能

本発明の実施例における移行先プライマリストレージ装置20aは、1または複数のドライブ221を1つのグループとして管理しており、このグループをパリティグループ（図1の要素22a、22b等がパリティグループである）と呼ぶ。また移行先プライマリストレージ装置20aは、いわゆるRAID技術を用いて、パリティグループ22を構成するドライブ群にデータを格納する前に、データから冗長データ（パリティ）を作成する。そしてデータとパリティとをパリティグループ22を構成するドライブ群に格納する。

【0041】

移行先プライマリストレージ装置20aで管理されるパリティグループ22の情報（たとえばパリティグループを構成するドライブの情報等）は、図3に示されるような、パリティグループ管理テーブルT100で管理される。この情報は、共有メモリ2042に格納される。パリティグループ管理テーブルT100は、パリティグループ名T101、ドライブ名T102、RAIDレベルT103、サイズT104、残サイズT105の欄から構成される。パリティグループ名T101は、パリティグループの識別子が格納され、ドライブ名T102には、パリティグループを構成するドライブの識別子が格納される。RAIDレベルT103には、当該パリティグループの冗長化方式であるRAIDレベルの情報が格納され、サイズT104にはパリティグループの容量の情報が格納される。

【0042】

後述するが、論理ボリュームを作成する際にパリティグループの記憶領域が使用される。残サイズT105には、パリティグループの容量のうち、まだ論理ボリュームの作成のために使用されていないサイズの情報が格納される。

【0043】

さらに移行先プライマリストレージ装置20aは、パリティグループの提供する記憶領域を用いて、1つまたは複数の論理ボリューム(LDEVとも呼ばれる)を形成する機能を有する。この機能のことをボリューム作成機能と呼ぶ。

【0044】

移行先プライマリストレージ装置20aは各論理ボリュームにストレージ装置内で一意な番号を付与して管理しており、それを論理ボリューム番号(LDEV#)と呼ぶ。論理ボリュームの情報は、図5に示されている論理ボリューム管理テーブルT200で管理される。論理ボリューム管理テーブルT200の各行が、論理ボリュームの情報を表しており、LDEV#(T201)が論理ボリュームの識別子である、論理ボリューム番号(LDEV#)であり、サイズT202には、論理ボリュームのサイズが格納される。グループ名T203、先頭アドレスT204は、論理ボリュームの記憶領域が対応付けられているパリティグループ及び当該パリティグループの記憶領域の先頭アドレスの情報が格納される。

10

20

30

40

50

【 0 0 4 5 】

図5の例では、LDEV#(T201)が44である論理ボリュームの情報(行906の情報)を参照すると、グループ名T203が“RGA”で、先頭アドレスT204が0であり、また論理ボリュームのサイズ(T202)が100GBであるので、パリティグループ“RGA”の先頭(アドレス0)から100GBの領域が、論理ボリューム“44”に割り当てられていることを表している。

【 0 0 4 6 】

なお、ここでは移行先プライマリストレージ装置20aがボリューム作成機能を有していることを説明したが、その他のストレージ装置も、上で説明したボリューム作成機能を有している。

10

【 0 0 4 7 】

(2) 外部ストレージ接続機能

移行先プライマリストレージ装置20aは、他のストレージ装置(移行元プライマリストレージ装置10a等)の有するボリュームの記憶領域を、移行先プライマリストレージ装置20aの有する記憶領域として扱い、当該記憶領域をプライマリサーバ30a等のホスト計算機に提供する機能を有している。以下、この機能のことを「外部ストレージ接続機能」と呼ぶ。また、外部ストレージ接続機能により管理される記憶領域のことを、外部ボリュームグループ(24a)と呼ぶ。

【 0 0 4 8 】

移行先プライマリストレージ装置20aで管理される外部ボリュームグループの情報は、図4に示されるような、外部ボリュームグループ管理テーブルT150で管理される。各行が外部ボリュームグループの情報を表し、グループ名T151には外部ボリュームグループの識別子が格納され、WWN(T152)、LUN(T153)、サイズ(T154)には、外部ボリュームグループに対応付けられた、外部のストレージ装置の論理ボリュームの情報が格納される。図4の例では、WWN(T152)が“50060e8005111112b”、LUN(T153)が“0”である論理ボリュームが、外部ボリュームグループに対応付けられている。以下では、論理ボリュームを外部ボリュームグループに対応付ける操作のことを、「マッピング」と呼ぶ。なお、移行先プライマリストレージ装置20aの外部ボリュームグループから外部のストレージの論理ボリュームへのアクセス経路(パス)を複数設け、パス障害時にアクセスパスを交替したりすることもできる。

20

30

【 0 0 4 9 】

外部ボリュームグループ24aは、パリティグループ22aと同様に扱われる。そのため、ボリューム作成機能では、外部ボリュームグループの記憶領域を用いて、1または複数の論理ボリュームを作成することもできる。図5の論理ボリューム管理テーブルT200の行907に、外部ボリュームグループ“EG1”の記憶領域から作成された論理ボリューム(LDEV#(T201)が“33”のボリューム)の例が示されている。

【 0 0 5 0 】

なお、外部ストレージ接続機能は、本発明の実施例1に係る計算機システムにおいて、移行先プライマリストレージ装置20aのみが有している機能である。ただし、移行先プライマリストレージ装置20a以外のストレージ装置も、外部ストレージ接続機能を有していてもよい。

40

【 0 0 5 1 】

(3) LUパス設定機能

上で説明した論理ボリュームを、プライマリサーバ30a等のホスト計算機が認識できるようにするために、LUN(論理ユニット番号)やポートの識別子を付す機能を「論理パス作成機能」あるいは「LUパス設定機能」と呼ぶ。本発明の実施例では、各ポート21aが有する識別子であるWWN(World Wide Name)をポート識別子として用いる。ただし、その他の識別子を用いてもよい。図6に示すLU管理テーブルT300は、ストレージ装置内の各論理ボリュームに対応付けられている論理ユニット番号(

50

LUN)とポート名の情報を管理するためのテーブルである。このテーブルは、ストレージ装置の共有メモリ2042に格納される。

【0052】

ポート名T301とLUN(T302)には、LDEV#(T303)で特定される論理ボリュームに対応付けられているポート名及びLUNの情報が格納される。たとえば図6のLU管理テーブルT300の先頭のエン트리(行)300-5には、ポート名T301が「3e243174 a a a a a a a a」、LUN(T302)が「0」、LDEV#(T303)が「11」の情報が格納されている。この場合、SAN50に接続されているプライマリサーバ30a等のホスト計算機は、ポートの配下に、LUNが0番のボリュームが存在すると認識する。以下、ホスト計算機(プライマリサーバ30a等)が認識するボリュームのことを、「論理ユニット」あるいは「LU」と呼ぶ。

10

【0053】

また、本発明の実施例におけるストレージ装置では、1つの論理ボリュームに対して(ポート名、LUN)の組を複数対応付けることができる。たとえば図6の例では、LDEV#(T303)が11番の論理ボリュームの情報が、エン트리(行)300-5に格納されているとともに、下段のエン트리(行)300-6にも格納されている。この場合、ホスト計算機は、ポート(3e243174 a a a a a a a a)の下に、LUNが0番のLUが存在することを認識するとともに、ポート(3e243174 b b b b b b b b)の下にも、LUNが0番のLUが存在すると認識する。

【0054】

20

またストレージ装置のLUパス設定機能では、LUパス設定とは逆に、論理ボリュームに対応付けられたLUNとポート名を削除する機能も有する。論理ボリュームに対応付けられたLUNとポート名を削除する処理のことを、LUパス削除処理と呼ぶ。ある論理ボリュームに対してLUパス削除が行われると、ホスト計算機から当該論理ボリュームにはアクセスできない。

【0055】

(4) ボリュームコピー機能

ストレージ装置は、論理デバイス内のデータを別のボリュームに複製する機能を有し、この機能をボリュームコピー機能と呼ぶ。ボリュームコピー機能には、コピー元ボリュームの属するストレージ装置と同一のストレージ装置内のボリュームにデータをコピーする機能、コピー元ボリュームの属するストレージ装置とは別のストレージ装置内のボリュームにデータをコピーする機能(たとえば移行元プライマリストレージ装置10a内のボリュームのデータを、移行元セカンダリストレージ装置10b内のボリュームへとコピー)とがある。本発明の実施例ではそれぞれを、「ローカルコピー機能」、「リモートコピー機能」と呼ぶ。

30

【0056】

ローカルコピー機能としてはたとえば、先に挙げた特許文献2や特許文献3に開示されている技術を用いて実現できる。

【0057】

またリモートコピー機能は、たとえば特許文献4や特許文献5に開示されている機能である。リモートコピー機能の種類として、同期型リモートコピー、非同期型リモートコピーという2種類が存在するが、ストレージ装置がいずれの種類のリモートコピー機能を採用している場合であっても、本発明は有効である。またリモートコピー機能の中には、たとえば特許文献5に開示されているように、コピーデータをジャーナルとしてボリューム(ジャーナルボリューム)に蓄積し、ジャーナルボリュームに格納されたデータをコピー先のストレージ装置へとコピーする、いわゆるジャーナルを用いたリモートコピーを行う方法も存在するが、その場合でも本発明は有効である。

40

【0058】

以下、本明細書で用いられる、ボリュームコピー機能に関する各種用語の定義を説明しておく。

50

【 0 0 5 9 】

(a) P - V O L : ボリュームコピー機能によりコピーされるボリューム (コピー元ボリューム) は、プライマリボリュームと呼ばれる。また、P - V O L と表記される。

【 0 0 6 0 】

(b) S - V O L : P - V O L 内データのコピー先のボリュームは、セカンダリボリュームと呼ばれる。またS - V O L と表記される。

【 0 0 6 1 】

(c) ペア : P - V O L 内のデータ複製が格納されるS - V O L のことを、P - V O L と「ペア関係にあるボリューム」と呼ぶ。同様に、S - V O L の複製元データが格納されているP - V O L も、当該S - V O L とペア関係にあるボリューム、と呼ばれることもある。そして、P - V O L 及び当該P - V O L のデータの複製が格納されるS - V O L のセットのことは、「ペア」または「ボリュームペア」と呼ばれる。

10

【 0 0 6 2 】

(d) ペア形成、ペア削除 : ボリュームコピー機能を用いてストレージ装置 (たとえば移行先プライマリストレージ装置 2 0 a) のL D E V (P - V O L) とストレージ装置 (たとえば移行先セカンダリストレージ装置 2 0 b) のL D E V (S - V O L) とをペア関係にあるボリュームにするよう、ストレージ装置に指示することを、「ペア形成する」 (または「ペア作成する」) と呼ぶ。逆に、ペア関係にあるボリュームの関係を解除する (P - V O L のデータ複製がS - V O L に作成されないようにする) オペレーションのことを「ペア解除する」 (または「ペア削除する」) と呼ぶ。ストレージ装置は、ペア形成またはペア削除などの指示を、ストレージマネージャ 1 0 1 あるいは管理端末から受け付ける。なお、ストレージ装置が受信する、ペア作成などを指示する制御用命令には、P - V O L の属するストレージ装置のシリアル番号、P - V O L のL D E V # 、S - V O L の属するストレージ装置のシリアル番号、S - V O L のL D E V # が含まれている。これによりストレージ装置は、ペア操作対象になるP - V O L 、S - V O L を一意に特定する。

20

【 0 0 6 3 】

(e) ペア状態 : ボリュームペアは、ボリューム内のデータのコピー状態によって、いくつかの状態をとる。ストレージ装置がペア形成指示を受信すると、P - V O L 内データを全てS - V O L にコピーする作業を開始する。データのコピーを開始した時点では、まだP - V O L 内データがすべてS - V O L に反映されていないが、この状態のことを、“ c o p y p e n d i n g ” と呼ぶ。そしてP - V O L 内全データのコピーが完了し、P - V O L とS - V O L の内容が同一になった (同期した) 状態のことを、“ p a i r ” 状態と呼ぶ。“ p a i r ” 状態にあるボリュームペアのP - V O L に対して、ホスト計算機から書き込みがあると、ライトデータはP - V O L に書き込まれるとともに、S - V O L にもライトデータの複製が書き込まれ、P - V O L とS - V O L は同一データが格納されている状態が維持されるように制御される。また、ペア解除された後の状態のことを“ s i m p l e x ” 状態と呼ぶ。

30

【 0 0 6 4 】

さらに、“ p a i r ” 状態にあるボリュームペアで行われる複製処理を一旦停止することも可能である。複製処理が一時停止した状態のことを“ s u s p e n d e d ” 状態と呼ぶ。また“ s u s p e n d e d ” 状態のボリュームペアに対して、複製処理を再開し、再び“ p a i r ” 状態にすることもできる。この処理のことを「再同期」と呼ぶ。

40

【 0 0 6 5 】

ストレージ装置は、ペア管理テーブルT 2 0 0 0 を用いて各ボリュームペアを管理している。図 7 にペア管理テーブルT 2 0 0 0 の構成例を示す。ストレージ装置は各ペアに識別子を付して管理しており、この識別子をペア番号 (T 2 0 0 1) と呼ぶ。そしてペア管理テーブルT 2 0 0 0 には、各ペアのP - V O L 、S - V O L の情報 (P - V O L の属するストレージ装置を特定可能な識別番号であるP D K C # (T 2 0 0 3) 、P - V O L のL D E V # であるP - V O L # (T 2 0 0 4) 、S - V O L の属するストレージ装置を特定可能な識別番号であるS D K C # (T 2 0 0 5) 、S - V O L のL D E V # であるS -

50

VOL#(T2006))を管理する。PDKC#(T2003)、SDKC#(T2005)には、それぞれストレージ装置のシリアル番号、ストレージ装置20bのシリアル番号が格納される。なお、ペア管理テーブルT2000はP-VOLの属するストレージ装置とS-VOLの属するストレージ装置の両方に格納される。そのため、移行先プライマリストレージ装置20aと移行先セカンダリストレージ装置20bで、ペアボリュームが作成される場合、ペア管理テーブルT2000は移行先プライマリストレージ装置20aと移行先セカンダリストレージ装置20bの両方に格納される。

【0066】

(5) ボリュームマイグレーション機能

ボリュームマイグレーション機能は、移行先プライマリストレージ装置20a内の、1つの論理ボリュームに格納されているデータを、移行先プライマリストレージ装置20a内のその他の論理ボリュームに移動する機能である。また、論理ボリューム内のデータの移動に加えて、LDEV#等の識別子も移動する。

【0067】

一例として、移行先プライマリストレージ装置20a内に、LDEV#44、LDEV#33の論理ボリュームが作成されていた場合、そして、論理ボリューム管理テーブルT200の状態が図5に示されている状態であった場合について説明する。移行先プライマリストレージ装置20aが、LDEV#33のデータをLDEV#44に移動するよう指示された場合、移行先プライマリストレージ装置20aは、LDEV#33のデータをLDEV#44に複製する処理を行う。以下では、LDEV#33のデータの移行先となる論理ボリューム(LDEV#44)のことを、「ターゲットボリューム」または「マイグレーションターゲットボリューム」と呼ぶ。また、複製中にプライマリサーバ30a等のホスト計算機からLDEV#33に対するデータ書き込みを受け付けることが可能である。

【0068】

LDEV#33からLDEV#44への複製が完了した時点で、移行先プライマリストレージ装置20aは、LDEV#33とLDEV#44の役割を入れ替える。具体的には、論理ボリューム管理テーブルT200の、LDEV#(T201)の内容を入れ替える。つまり、図5の例では、行906のLDEV#(T201)に44が、行907のLDEV#(T201)に33が格納されているが、ボリュームマイグレーション機能によるボリューム移動後は、行906のLDEV#(T201)に33が、行907のLDEV#(T201)に44が格納される。すなわち、LDEV#と、データ格納先であるグループ名及び先頭アドレスと、の対応関係をマイグレーション元とマイグレーション先とでスワップしている。これにより、データの格納されている論理ボリュームのLDEV#(及び、論理ボリュームに対応付けられているLUNやポート名)は変化させずに、ホスト計算機などの上位装置に透過的にデータの格納位置を移動することができる。いいかえると、LDEV#33へのアクセスの際、複製完了前は外部ボリュームEG1からデータが読出し/書き込みされていたのに対して、複製完了後はパリティグループRGAからデータが読出し/書き込みされるようになる。

【0069】

(6) 仮想ストレージ

続いて、仮想ストレージについて説明する。本発明の実施例に係る移行先プライマリストレージ装置20aは、移行先プライマリストレージ装置20a等の物理的なストレージ装置とは異なる、1または複数の仮想的なストレージ装置25a(以下、これを仮想ストレージと呼ぶ)を定義し、プライマリサーバ30aに対して、物理的なストレージ装置に加えて仮想ストレージがSAN50上に存在するように見せかける機能を有する。なお、以下では移行先プライマリストレージ装置20aを例にとって説明するが、仮想ストレージを定義する機能は、移行先セカンダリストレージ装置20bも備えているので、移行先セカンダリストレージ装置20bも以下で説明する機能を備えている。また、移行元プライマリストレージ装置10a及び/または移行元セカンダリストレージ装置10bが仮想

10

20

30

40

50

ストレージを定義する機能を備えていても、備えていなくてもよい。移行元プライマリストレージ装置 10 a 及び / または移行元セカンダリストレージ装置 10 b が仮想ストレージを定義する機能を備えており、移行元プライマリストレージ装置 10 a 及び / または移行元セカンダリストレージ装置 10 b に定義された仮想ストレージに属するボリュームペアを移行する場合、以下の説明では、移行元のストレージ装置は移行元の仮想ストレージ装置に、移行元のストレージ装置の装置シリアル番号は仮想製番に、L D E V # は V L D E V # になる。

【 0 0 7 0 】

移行先プライマリストレージ装置 20 a に仮想ストレージ 25 a を定義した場合の例を、図 8 を用いて説明する。仮想ストレージは、装置シリアル番号 (S / N) を有し (以下、仮想ストレージの有する装置シリアル番号のことを、「仮想製番」と呼ぶ)、また仮想ストレージ内のリソースとして論理ボリュームを有することができる。図 8 は、移行元プライマリストレージ装置 10 a のシリアル番号が 1、移行先プライマリストレージ装置 20 a のシリアル番号が 2、そして定義された仮想ストレージ 25 a のシリアル番号が 1 (つまり移行元プライマリストレージ装置 10 a のシリアル番号と同じ) である構成例を示している。また図 8 において、移行元プライマリストレージ装置 10 a は L D E V # が 11 番の論理ボリューム 130 a を有し、仮想ストレージ 25 a は L D E V # が 33 番の論理ボリューム 230 a を有している。

10

【 0 0 7 1 】

仮想ストレージの有する論理ボリューム 230 a は、L D E V # とは別の、仮想的な識別番号を有しており、これを仮想 L D E V # あるいは V L D E V # と呼ぶ。図 8 の仮想ストレージ 25 a の有する論理ボリューム 230 a は、V L D E V # が 11 番 (つまり移行元プライマリストレージ装置 10 a の有する論理ボリューム 130 a の L D E V # と等しい番号が付されている) である。

20

【 0 0 7 2 】

この場合、プライマリサーバ 30 a が S C S I の I N Q U I R Y コマンド等を移行先プライマリストレージ装置 20 a に発行することにより、論理ボリューム 230 a の情報を取得すると、ボリューム番号として 11 番が、装置シリアル番号として 1 番が返却される。つまりプライマリサーバ 30 a に対し、仮想製番と V L D E V # が返却される。一方プライマリサーバ 30 a が移行元プライマリストレージ装置 10 a から論理ボリューム 130 a の情報を取得すると、論理ボリューム 230 a と同一のボリューム番号 (11 番) 及び同一の装置シリアル番号 (1 番) が返却される。

30

【 0 0 7 3 】

そのため、プライマリサーバ 30 a (の交替パスソフト 102) はこの 2 つのボリュームを同一ボリュームとして認識するので、プライマリサーバ 30 a から論理ボリューム 130 a へのパス (図中の実線矢印。以下このパスを「パス 1」と呼ぶ) の交替パスが、プライマリサーバ 30 a から論理ボリューム 230 a へのパス (図中の点線矢印。以下このパスを「パス 2」と呼ぶ) であると判断する。

【 0 0 7 4 】

そしてパス 1 が削除された場合、交替パスソフト 102 がアプリケーションプログラムなどから論理ボリューム 130 a へのアクセス要求を受け付けると、交替パスソフト 102 はパス 2 を介してアクセス要求を発行する (つまり論理ボリューム 230 a にアクセス要求を発行する) 。

40

【 0 0 7 5 】

仮想ストレージの管理で用いられる管理情報について、図 9 を用いて説明する。図 9 は、移行先プライマリストレージ装置 20 a が有する V - B O X 管理テーブル T 1000 の構成を示している。V - B O X 管理テーブル T 1000 は移行先プライマリストレージ装置 20 a の共有メモリ 2042 に格納される。また、ここでは移行先プライマリストレージ装置 20 a が有する V - B O X 管理テーブル T 1000 の構成についてのみ説明するが、本発明の実施例 1 において、データ移行作業時には移行先セカンダリストレージ装置 2

50

0 bにも仮想ストレージが定義されるため、移行先セカンダリストレージ装置20bもV-BOX管理テーブルT1000を有する。

【0076】

移行先プライマリストレージ装置20aは仮想ストレージを定義する際、ストレージ装置20内にV-BOXと呼ばれる管理情報を生成する。V-BOXとは、ある仮想ストレージに所属させるべき、論理ボリューム等のリソースの情報を管理するための管理情報である。図9のV-BOX管理テーブルT1000内の1つの行(たとえばR1011, R1012)が、1つのV-BOXを表す情報に相当する。

【0077】

V-BOXは、移行先プライマリストレージ装置20aで定義されるV-BOXの識別子であるID(T1001)、V-BOXの機種名情報(T1002)、V-BOXのシリアル番号(S/N)つまり仮想製番(T1003)、V-BOXに属する論理ボリュームに付されているVLDEV#(T1004)、LDEV#(T1005)から成る。VLDEV#(T1004)、LDEV#(T1005)はそれぞれ、仮想ストレージに所属する論理ボリュームの仮想LDEV#とLDEV#の情報が格納される。

10

【0078】

移行先プライマリストレージ装置20aは初期状態では、移行先プライマリストレージ装置20a内に、仮想製番として移行先プライマリストレージ装置20aの装置シリアル番号が設定されたV-BOXが1つ定義されている。以下、このV-BOXを「仮想ストレージv0」と呼ぶ。初期状態ではV-BOX管理テーブルT1000に、行R1011の情報のみが格納された状態にある。また論理ボリュームが作成されると、作成された論理ボリュームはすべて仮想ストレージv0に所属した状態になる。

20

【0079】

仮想ストレージは、移行先プライマリストレージ装置20aがストレージマネージャ101あるいは管理端末から、仮想ストレージの機種名、仮想製番の指定とともに、仮想ストレージの定義を指示する命令を受信することに応じて作成される。ただし仮想ストレージが定義された直後は、仮想ストレージ内に論理ボリュームが1つも所属していない仮想ストレージが定義される(つまりT1004, T1005の欄に論理ボリュームの情報が登録されていない)。仮想ストレージの定義後に、仮想ストレージに論理ボリュームを所属させる旨の指示を移行先プライマリストレージ装置20aが受け付けると、V-BOX管理テーブル1000のT1004, T1005の欄に論理ボリュームの情報(仮想LDEV#及びLDEV#)が登録される。

30

【0080】

V-BOX管理テーブル1000に論理ボリュームの情報を登録する際に用いられる、仮想LDEV管理テーブルT1500について、図10を用いて説明する。仮想LDEV管理テーブルT1500は、論理ボリュームのLDEV#とVLDEV#との対応関係を管理するテーブルで、LDEV#(T1501)、VLDEV#(T1502)、ID(T1503)、属性(T1504)の欄からなる。

【0081】

VLDEV#(T1502)には、LDEV#(T1501)で特定される論理ボリュームに付された仮想LDEV#が格納される。初期状態では、全ての論理ボリュームについて、LDEV#(T1501)とVLDEV#(T1502)は同一の値が格納される。ただし外部(ストレージマネージャ101など)からの指示によって、VLDEV#(T1502)は変更可能である。また外部からの指示に応じて、VLDEV#(T1502)の値を無効値(NULL)に設定することもできる。

40

【0082】

ID(T1503)は、論理ボリュームが所属しているV-BOXの識別子であるIDが格納される。

【0083】

属性(T1504)には、当該論理ボリュームが仮想ストレージに登録(予約)されて

50

いるか否かを表す情報が格納される。論理ボリュームが仮想ストレージに登録（予約）された場合、属性（T1504）には“reserved”が格納される。

【0084】

仮想LDEV管理テーブルT1500またはV-BOX管理テーブル1000に登録されるLDEV#は、論理ボリュームが作成された状態にあるか否か（論理ボリューム管理テーブルT200に論理ボリュームの情報が登録されているか否か）とは独立に登録されることができる。そのため、論理ボリュームが作成されていない場合であっても、移行先プライマリストレージ装置20aは仮想ストレージに論理ボリュームを登録する処理を実施することができる。

【0085】

一例として、図5の論理ボリューム管理テーブルT200には、LDEV#（T201）が33、44、55の3つの論理ボリュームの情報が登録されている。つまり、移行先プライマリストレージ装置20aの論理ボリューム管理テーブルT200の状態が図5の状態である場合、移行先プライマリストレージ装置20aは3つの論理ボリュームが作成された状態にある。この状態でも、論理ボリューム管理テーブルT200に登録されていない論理ボリューム、たとえばLDEV#が2番の論理ボリュームを仮想ストレージに所属させる処理を行うことが許されている。

【0086】

先に述べたとおり、LDEV#に対応づけられたVLDEV#を、後から変更することは可能で、本発明の実施例1ではこの作業を「仮想化」と呼ぶ。たとえば移行先プライマリストレージ装置20aが、LDEV#が44番の論理ボリュームに仮想LDEV#として11番を付す指示を受け付けると、移行先プライマリストレージ装置20aは、仮想LDEV管理テーブルT1500のLDEV#（T1502）が44番の行に対応するVLDEV#（T1501）に11を、そしてV-BOX管理テーブル1000のLDEV#（T1005）が44番の行に対応するVLDEV#（T1004）に11を格納する。

【0087】

本発明の実施例1に係る移行先プライマリストレージ装置20aで仮想化を行う場合、以下の手順を踏むことになっている。LDEV#が2番の論理ボリュームを仮想ストレージv1（V-BOX管理テーブルT1000において、ID（T1001）がv1である仮想ストレージ）に所属させ、仮想化により4番のVLDEV#を付す時の処理の流れを概説する。なお、以下では、LDEV#がn番の論理ボリュームのことを、「LDEV#n」と略記する。また、LDEV#2は仮想ストレージv0に所属している状態で、LDEV#2に対応するVLDEV#が2である状態を想定する。

【0088】

まず、移行先プライマリストレージ装置20aは、LDEV#2に対応付けられているVLDEV#をNULLにする。これにより仮想LDEV管理テーブルT1500の、LDEV#（T1501）が2番の行において、VLDEV#（T1502）をNULLにする。また仮想LDEV管理テーブルT1500の属性T1504に“reserved”を格納する。続いて同じ行のID（T1503）を参照し、LDEV#の所属するV-BOXが「v0」であることを認識すると、V-BOX管理テーブルT1000の行R1011において、LDEV#（T1005）が2の行のVLDEV#（T1004）をNULLにする。

【0089】

続いて、LDEV#2を仮想ストレージv1に所属させる処理が行われる。移行先プライマリストレージ装置20aは、V-BOX管理テーブルT1000の行R1012のLDEV#（T1005）の欄に2を格納する。また対応するVLDEV#（T1004）にはNULLが格納される。仮想LDEV管理テーブルT1500のLDEV#2に対応する行のID（T1503）には「v1」が格納される。

【0090】

その後、LDEV#2を仮想化する。移行先プライマリストレージ装置20aは、V-

10

20

30

40

50

BOX管理テーブルT1000に登録されているLDEV#T1005が2番の行に対応するVLDEV#(T1004)に4を格納する。仮想LDEV管理テーブルT1500のLDEV#2に対応する行のVLDEV#(T1502)にも、4を格納する。

【0091】

詳細は後述するが、本発明の実施例1に係る計算機システムでは、仮想ストレージに所属しているボリュームペアに対してペア操作を行う場合、P-VOLやS-VOLを特定する識別情報として仮想製番と仮想LDEV#の情報、及び命令中に仮想の識別子が使用されている旨の情報が含まれた、ペア操作の制御用命令がストレージ装置に発行される。ストレージ装置(移行先プライマリストレージ装置20a等)が、P-VOLやS-VOLの識別子として仮想製番と仮想LDEV#の情報を受信すると、V-BOX管理テーブルT1000を参照することによって、仮想LDEV#をLDEV#に変換し、処理対象の論理ボリュームを特定する。また、命令中に仮想の識別子が使用されている旨の情報が含まれていない場合、命令に含まれている仮想製番と、命令を受信したストレージ装置の装置シリアル番号が一致していないとき、ストレージ装置は命令を拒絶する。

10

【0092】

(7)ストレージマネージャの機能

ストレージマネージャ101は、プライマリサーバ30aまたはセカンダリサーバ30bからストレージ装置の設定、制御を行うためのプログラムである。本発明の実施例1に係る計算機システムの各ストレージ装置はいずれも、ストレージマネージャ101から制御される。以下では、ストレージマネージャ101の機能及びストレージマネージャ101が使用する設定情報について説明する。

20

【0093】

ストレージマネージャ101は、管理者またはクラスタソフトウェア103等のプログラムがストレージ装置の管理操作を行う場合に用いられるプログラムである。ここでの「管理操作」とは、LUパス設定などの設定操作、あるいはリモートコピー機能を用いたペア操作などである。

【0094】

ストレージマネージャ101は、管理操作のコマンドをいくつかサポートしており、管理者またはクラスタソフトウェア103等のプログラムはストレージマネージャ101にコマンドを発行することで、ストレージ装置の管理操作を行う。ストレージマネージャ101は、受信したコマンドを加工し、ストレージ装置に対して発行するための制御用命令を作成する。そして作成された制御用命令を、ストレージ装置に定義されているコマンドデバイスに対して発行する。制御用命令を受信したストレージ装置は、受信した制御用命令の内容に応じて、所定の処理(LUパス設定等)を行う。そのため、本発明の実施例1に係る計算機システムに存在する各ストレージ装置には、コマンドデバイスが1つ定義される。

30

【0095】

ストレージマネージャ101でサポートされる管理操作のコマンドは、大別すると二種類のコマンドがある。1つ目の種類のコマンドは、ボリュームコピー機能(ローカルコピー機能、リモートコピー機能)及びボリュームマイグレーション機能を用いた、ボリュームペアの操作のコマンド(以下これを、「ペア操作コマンド」と呼ぶ)である。

40

【0096】

2つ目の種類のコマンドは、ペア操作コマンド以外のコマンドで、これにはLUパス設定など、ストレージ装置の設定を行うコマンドが含まれる(以下これを、「設定用コマンド」と呼ぶ)。ストレージマネージャ101は、以下で説明する移行処理において必要となるストレージ装置の各種設定処理のうち、コマンドデバイスの作成処理以外の全ての設定処理を行うことができる。

【0097】

なお、コマンド発行時に操作対象の論理ボリュームの識別子を指定する必要があるコマンドが存在する。たとえば、ボリュームペアを作成するコマンドや、LUパス設定コマン

50

ドなどである。操作対象の論理ボリュームに仮想LDEV#が付されている場合、指定すべき論理ボリュームの識別子として、仮想LDEV#を指定するべきか、LDEV#を指定するべきかという問題がある。本発明の実施例に係るストレージマネージャ101では、設定用コマンドでは、論理ボリュームを特定する識別子に、仮想LDEV#ではなくLDEV#が用いられる。またペア操作のコマンド(あるいは後述する設定ファイル)で指定される、論理ボリュームを特定する識別子には、仮想LDEV#が用いられる。

【0098】

ストレージマネージャ101は、ホスト計算機(プライマリサーバ30aまたはセカンダリサーバ30b)上で動作するオペレーティングシステム(OS)の常駐プログラム(サービス)として動作する。本発明の実施例ではこの常駐プログラムを「インスタンス」と呼ぶ。インスタンスは、1つのホスト計算機で複数稼働させることが可能である。またインスタンスの稼働を開始させることを、「インスタンスを起動する」と呼ぶ。

【0099】

各インスタンスが起動される時、インスタンス番号と呼ばれる番号が指定される。各インスタンスは起動時に、指定されたインスタンス番号に従って、1つの設定ファイルを読み込む(設定ファイルにはインスタンス番号を含むファイル名が付されており、インスタンスは、指定されたインスタンス番号をファイル名に含むファイルを読み込む)。

【0100】

設定ファイルには後述する通り、様々な情報が格納されているが、少なくとも使用するコマンドデバイスの情報が格納されている。各インスタンスは設定ファイルに格納されているコマンドデバイスの情報に基づいて、制御用命令の発行先となるコマンドデバイスを特定する。たとえば2つのインスタンスを起動する場合(以下、この2つのインスタンスをそれぞれ、「インスタンス0」、「インスタンス1」と呼ぶ)、インスタンス0用の設定ファイルには移行元プライマリストレージ装置10aのコマンドデバイスの情報を格納し、インスタンス1用の設定ファイルに移行先プライマリストレージ装置20aのコマンドデバイスの情報を格納しておけば、インスタンス0は移行元プライマリストレージ装置10aの管理操作を行うために使用可能で、インスタンス1は移行先プライマリストレージ装置20aの管理操作を行うために使用可能である。

【0101】

またインスタンス0用の設定ファイルとインスタンス1用の設定ファイルの両方に、同一コマンドデバイスの情報を格納しておけば、いずれのインスタンスも、同じストレージ装置の管理操作を行うために用いることができる。これは、一方のインスタンスを、LUPAS設定などの設定用用途で用い、もう一方のインスタンスをペア操作の用途で用いる場合に好適である。

【0102】

また、ストレージマネージャ101を用いてストレージ装置のリモートコピー機能の制御(ペア操作)を行う場合、P-VOLにアクセスするホスト計算機(プライマリサーバ30a)と、S-VOLにアクセスするホスト計算機(セカンダリサーバ30b)のそれぞれで、少なくとも1つのインスタンスを起動する。

【0103】

リモートコピー機能の制御を行う場合の設定ファイルの例を図11に示す。設定ファイル3000-1はプライマリサーバ30aに格納される設定ファイルで、設定ファイル3000-2はセカンダリサーバ30bに格納される設定ファイルである。また、設定ファイル3000-1は移行元プライマリストレージ装置10aのペア操作の設定ファイルで、設定ファイル3000-2は移行元セカンダリストレージ装置10bのペア操作の設定ファイルである。そしてP-VOLのLDEV#が11番、P-VOLの属するストレージ装置(移行元プライマリストレージ装置10a)のシリアル番号が1番、そしてS-VOLのLDEV#が22番、S-VOLの属するストレージ装置(移行元セカンダリストレージ装置10b)のシリアル製番が11番である場合の例である。

【0104】

各設定ファイル3000-1, 3000-2には、3種類の情報が格納される。1つ目の情報が、先に述べたコマンドデバイスの情報であり、コマンドデバイス名3001のフィールドには、インスタンスが制御用命令を発行するコマンドデバイスの識別子が記録される。コマンドデバイスの識別子の名称は、ストレージ装置毎に一意に定まっており、設定用ファイルを作成する際にはその名称を記録する。

【0105】

2つ目の情報はボリュームペアの情報であり、ペア情報フィールド3002には、リモートコピー機能による操作対象となるボリュームペアの情報が格納される。記述形式については後述する。

【0106】

3つ目の情報は、相手先ホスト計算機の情報3003である。ストレージマネージャ101を用いてストレージ装置のリモートコピー機能の制御を行う場合、プライマリサーバ30aのインスタンスはセカンダリサーバ30bのインスタンスからS-VOLの情報を取得する。相手先ホスト計算機の情報3003はそのため用いられる情報で、相手先ホスト計算機のIPアドレスが記述されている。

【0107】

2つ目の情報である、ペア情報フィールド3002の記述形式について説明する。ペア情報フィールド3002には、グループ名3002-1、装置シリアル番号3002-2、LDEV#(3002-3)が格納される。LDEV#3002-3は、ボリュームペアを形成するうちの1つの論理ボリュームのLDEV#が記述される。装置シリアル番号3002-2には、LDEVの存在する装置のシリアル番号(S/N)が格納される。グループ名3002-1にはボリュームペアに付されるグループ名が記述される。グループ名3002-1の詳細については後述する。

【0108】

なお、プライマリサーバ30aの設定ファイルには、移行元(または移行先)プライマリストレージ装置(10aまたは20a)にあるP-VOLのLDEV#が格納され、セカンダリサーバ30bの設定ファイルには、移行元(または移行先)セカンダリストレージ装置(10bまたは20b)にあるS-VOLのLDEV#が格納される。また、少なくともペア関係にあるLDEVについては、同じグループ名が付与される必要がある。図11の例では、LDEV#(3002-3)が11の論理ボリュームと、LDEV#(3002-3)が22の論理ボリュームがペア関係にあり、そのグループ名(3002-1)が“devg01”であることが記述されている。

【0109】

なお、管理者またはクラスタソフトウェア103等のプログラムが、ストレージマネージャ101(のインスタンス)を用いてボリュームペア操作を行う場合、操作対象となるボリュームペアを特定する情報としてグループ名を用いる。たとえばペア作成を指示するコマンドの書式は、以下の通りである。

paircreate <グループ名>

【0110】

プライマリサーバ30aで稼働するインスタンスに対して、このコマンドが発行されると、プライマリサーバ30aで稼働するインスタンスは、設定ファイルの内容に基づいて、コマンドで指定されたグループ名に対応するLDEV#(3002-3)及び装置シリアル番号(3002-1)を特定する。またセカンダリサーバ30bで稼働するインスタンスと通信することにより、セカンダリサーバ30bで稼働するインスタンス用の設定ファイルに記述されている、グループ名に対応するLDEV#(3002-3)及び装置シリアル番号(3002-1)を取得する。これによりプライマリサーバ30aで稼働するインスタンスは、P-VOL及びS-VOLのLDEV#及び装置シリアル番号を特定できる。この情報に基づいて、コマンドデバイスに対して発行する制御用命令を作成して、コマンドデバイスに対してペア操作の制御用命令を発行する。

【0111】

10

20

30

40

50

また、複数のボリュームペアに同一のグループ名を付与してもよく、その場合、上で説明したペア操作コマンドを発行すると、複数のボリュームペアのペア操作が行われることになる。

【0112】

この設定ファイルは、計算機システムの管理者によって作成されていた。移行元ストレージシステム10でボリュームペアの制御を行う場合には、LDEV#(3002-3)及び装置シリアル番号3002-1には、移行元ストレージシステム10におけるP-VOLまたはS-VOLのLDEV#及び移行元ストレージシステム10の装置シリアル番号が格納されている。ボリューム移行により、ボリュームペアが移行先ストレージシステム20に移動した場合、装置シリアル番号やLDEV#も変更されることになるが、その場合、管理者はこの設定ファイルを書き換える必要がある。つまり装置シリアル番号あるいはLDEV#を、移行先ストレージシステムの装置シリアル番号やLDEV#に書き換える必要がある。この作業は、特にボリューム数が増大すると時間のかかる作業となる。

【0113】

次に、ストレージ装置に仮想ストレージが定義されており、ペア操作対象となる論理ボリュームが仮想ストレージ装置に所属している場合の設定ファイルの例について、図12を用いて説明する。図12の例は、移行先ストレージシステムにおいて仮想ストレージを定義し、P-VOLの仮想LDEV#が11番、P-VOLの属する仮想ストレージの仮想製番が1番、そしてS-VOLの仮想LDEV#が22番、S-VOLの属する仮想ストレージの仮想製番が11番にした場合の例である(つまり、移行先ストレージシステムにおけるP-VOL(これを移行先P-VOLと呼ぶ)及びS-VOL(これを移行先S-VOLと呼ぶ)に付される仮想製番はそれぞれ、移行元プライマリストレージ装置10a及び移行元セカンダリストレージ装置10bのシリアル番号と同じで、また移行先P-VOL及び移行先S-VOLに付される仮想LDEV#はそれぞれ、移行元プライマリストレージ装置10aのP-VOL(移行元P-VOL)のLDEV#及び移行元セカンダリストレージ装置10bのS-VOL(移行元S-VOL)のLDEV#と等しい関係にある)。

【0114】

先に述べたとおり、仮想ストレージが定義されているストレージ装置は、仮想ストレージに所属しているボリュームペアに対するペア操作の指示を受け付ける場合、ストレージマネージャ101から、P-VOLやS-VOLの識別子として、仮想製番と仮想LDEV#の情報を受信することによって動作する。そのため、設定ファイルに記述されるボリュームペアの情報としては、P-VOLやS-VOLを特定する情報として、仮想製番と仮想LDEV#が記述される。

【0115】

図12において、ファイル3000-1'が、移行先プライマリストレージ装置20a用の設定ファイルで、ファイル3000-2'が、移行先セカンダリストレージ装置20b用の設定ファイルである。以下、ファイル3000-1'を例にとって、図11の設定ファイルとの差異を説明する。

【0116】

図11に示されている設定ファイルの例と比較すると、ブロック3001の内容が、ブロック3001-1'、3001-2'に変更されている点のみが異なる。ブロック3001-1'は、ストレージマネージャ101がコマンドを発行する先であるコマンドデバイスの識別子である。またブロック3001-2'は、仮想ストレージに対してコマンドを発行することを表す情報(HORCM_VCMD)、そして仮想ストレージの仮想製番(1番)を表している。

【0117】

そしてブロック3001以外の点は、図11の設定ファイルと同じである。そのため、本発明の実施例に係る移行先ストレージシステムを使用した場合、移行元ストレージシステムから移行先ストレージシステムにボリュームを移行する際の設定ファイル変更の必要

10

20

30

40

50

がないため、移行に係る作業負担を大幅に削減することができる。

【0118】

設定ファイルの別の例を図13に示す。図13に示されている設定ファイル3000-1'には、コマンドデバイス名3001のフィールドだけが存在する。LUパス設定などの設定用途だけに用いられるインスタンスを起動する場合には、このようにコマンドデバイス名3001のフィールドだけを定義すればよい。

【0119】

[移行処理の概要]

図14は、本発明の実施例1に係る計算機システムにおける、移行処理の概要を表した図である。移行元プライマリストレージ装置10a、移行元セカンダリストレージ装置10bは、リモートコピー機能を有している。リモートコピー機能により、移行元プライマリストレージ装置10aの論理ボリューム130aと移行元セカンダリストレージ装置10bの論理ボリューム130bとはペア関係にあり、論理ボリューム130aがP-VOL、論理ボリューム130bがS-VOLという関係にある。つまりプライマリサーバ30aから移行元プライマリストレージ装置10aの論理ボリューム130aに書き込まれたデータは、移行元セカンダリストレージ装置10bの論理ボリューム130bへとコピーされ、論理ボリューム130bには常時、論理ボリューム130aのデータの複製が格納されている状態(ペア状態)にある。以下、論理ボリューム130aのことを「移行元P-VOL」と呼び、論理ボリューム130bのことを「移行元S-VOL」と呼ぶ。

【0120】

本発明の実施例1に係るボリューム移行方法では、このペア状態にある論理ボリューム130aと論理ボリューム130bを、ペア状態を維持したまま移行先ストレージシステム20へと移行する。まず、移行先プライマリストレージ装置20aにおいて、仮想ストレージ25aを作成する。そして仮想ストレージ25a内に、移行先プライマリボリュームとなる論理ボリューム230a、及び論理ボリューム231aを作成する。同様に移行先セカンダリストレージ装置20bにおいて、仮想ストレージ25bを作成し、仮想ストレージ25b内に移行先セカンダリボリュームとなる論理ボリューム230bを作成する。

【0121】

仮想ストレージ25aは、移行元プライマリストレージ装置10aのシリアル番号(S/N=1番)と同じ仮想製番(1番)が設定される。また仮想ストレージ25bは、移行元セカンダリストレージ装置10bのシリアル番号(S/N=11番)と同じ仮想製番(11番)が設定される。さらに論理ボリューム230aには、移行元P-VOLのLDEV#と同じ番号の仮想LDEV#が付され、論理ボリューム230bには、移行元S-VOLのLDEV#と同じ番号の仮想LDEV#が付される。

【0122】

なお、論理ボリューム230aを作成する際、外部ストレージ接続機能により、論理ボリューム130aが記憶領域として対応付けられる。そのため、論理ボリューム230aは「論理ボリューム130aにマッピングされている」と表現されることもある。一方、論理ボリューム231aは、移行先プライマリストレージ装置20a内のパリティグループ22a(図1に記載のパリティグループ22a)を記憶領域として作成された論理ボリュームである。論理ボリューム230a、論理ボリューム231aは、論理ボリューム130aと同サイズのボリュームとして作成される。論理ボリューム230aは、実体(データの記憶される記憶領域)が外部のストレージ装置(移行元プライマリストレージ装置10a)に存在する。

【0123】

続いてプライマリサーバ30aの論理ボリュームへのアクセス経路(パス)を切り替えることにより、プライマリサーバ30aが論理ボリューム130a(移行元P-VOL)ではなく、移行先P-VOLである論理ボリューム230aにアクセスするように設定を変更する。論理ボリューム230aは、論理ボリューム130aにマッピングされている

10

20

30

40

50

ため、プライマリサーバ30aが論理ボリューム230aにデータをライトすると、そのデータは移行先プライマリストレージ装置20aから移行元プライマリストレージ装置10aに転送され、論理ボリューム130aへと書き込まれる。またプライマリサーバ30aが論理ボリューム230aにリードリクエストを発行すると、論理ボリューム130aからデータが読み出され、読み出されたデータは移行元プライマリストレージ装置10aから移行先プライマリストレージ装置20aに転送され、移行先プライマリストレージ装置20aからプライマリサーバ30aへと返送される。さらに、論理ボリューム230aには、論理ボリューム130aと同一のボリューム番号(仮想ボリューム番号)が付されており、プライマリサーバ30aのアクセス経路は移行先プライマリストレージ装置20aへと変更されたものの、プライマリサーバ30aは、論理ボリューム230aが論理ボリューム130aと同一ボリュームであると認識する。

10

【0124】

次に、移行先ストレージシステム20のリモートコピー機能により、論理ボリューム230aと論理ボリューム230bの間でボリュームコピーを開始する。これにより、論理ボリューム230a(実際には論理ボリューム130a)から順次データが読み出されて、論理ボリューム230bへと複製される。なお、この複製処理中も、プライマリサーバ30aは論理ボリューム230aに対してI/O要求を発行することができる。プライマリサーバ30aが論理ボリューム230aに対してライト要求を発行した場合、論理ボリューム230a(実際には論理ボリューム130a)に対してライトデータが書き込まれるとともに、論理ボリューム230bにもライトデータの複製が格納される。移行元ストレージシステム10においても、論理ボリューム130aと論理ボリューム130bはペア状態が維持されているので、論理ボリューム230bに加えて論理ボリューム130bにも、論理ボリューム130aのデータの複製が格納される状態にある。

20

【0125】

論理ボリューム230aと論理ボリューム230bの間でボリュームコピーが完了し、2つのボリュームの内容が同一になった時点(つまり、ペア状態が“pair”になった時点)で、セカンダリサーバ30bの論理ボリュームへのアクセス経路を切り替えて、セカンダリサーバ30aが論理ボリューム130bではなく論理ボリューム230bにアクセスするように設定を変更する。これで、プライマリサーバ30aや移行先プライマリストレージ装置20aが障害で停止した場合にも、セカンダリサーバ30b及び移行先セカンダリストレージ装置20bを用いた業務継続が可能になる。ただしこの時点では、移行先プライマリストレージ装置21aには論理ボリューム230aのデータは存在しない。これは、論理ボリューム230aの実体が移行元プライマリストレージ装置10aの論理ボリューム130aだからであり、論理ボリューム230aの実体を移行先プライマリストレージ装置20aに移動しない限り、移行は完了したことになる。そのため、移行先プライマリストレージ装置20aはボリュームマイグレーション機能を用いることにより、論理ボリューム230aの内容を論理ボリューム231aへと移行する処理を行う。

30

【0126】

この移行処理では、移行先プライマリストレージ装置20aは論理ボリューム230aの内容を論理ボリューム231aへと複製する。このとき、論理ボリューム230aのデータの实体は、論理ボリューム130aにあるため、実際は、論理ボリューム130aから論理ボリューム231aに対応付けられた記憶領域にデータがコピーされる。そして複製が完了した時点で移行先プライマリストレージ装置20aは、論理ボリューム230aと論理ボリューム231aの役割を入れ替える。その結果、論理ボリューム231aが、論理ボリューム230bのボリュームペアになり、論理ボリューム231aのボリューム番号が、これまで論理ボリューム230aに付されていたボリューム番号へと変更される。

40

【0127】

図15に、ボリュームマイグレーション機能により、論理ボリューム230aと論理ボリューム231aの役割が入れ替わった後の状態を示す。これにより、プライマリサーバ

50

30 a からライト要求が到来しても、ライトデータが移行元プライマリストレージシステム 10 に転送されることはなくなるため、移行元ストレージシステム 10 を撤去しても良い状態（つまりボリュームペアの移行が完了した状態）になる。

【0128】

[移行処理のフロー]

続いて、本発明の実施例 1 に係る計算機システムにおける、ボリューム移行処理のフローを、図 16、17、18 のフローチャートを用いて説明する。なお、以下では説明の簡単化のため、移行元ストレージシステムの 1 ボリュームペア（P-VOL、S-VOL のセット）のみの移行処理について説明するが、同時に複数のボリュームペアの移行を行うことも可能である。

10

【0129】

最初にストレージマネージャ 101 は、移行元ストレージシステムの構成の調査を行う（S10）。移行元ストレージシステムの構成の調査では、主に以下の情報を取得する。

(ア) 移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b のシリアル番号（S/N）、機種名

(イ) 移行対象の P-VOL、S-VOL の LDEV # 及びサイズ

(ウ) 移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b で使用されていない LDEV #

【0130】

これらの情報を調査する際、移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b に対して、ストレージマネージャ 101 から構成情報取得用の制御用命令を発行する。ストレージ装置は構成情報取得用の制御用命令を受信すると、装置のシリアル番号、ストレージ装置内で定義されている論理ボリュームの LDEV # やサイズ等の情報を返却する。ストレージ装置内で定義されている論理ボリュームの情報として、一例では論理ボリューム管理テーブル T200 の内容が返却される。管理者は、論理ボリューム管理テーブル T200 の LDEV #（T201）を参照することで、LDEV #（T201）に記録されていない LDEV # が「使用されていない LDEV #」であると判断することができる。

20

【0131】

なお、以下で説明する移行処理では、各ストレージ装置の構成が以下の場合について説明する。

30

(a) 移行元プライマリストレージ装置 10 a の P-VOL（移行元 P-VOL と呼ぶ）の LDEV # が 11 番、サイズが 100 GB

(b) 移行元セカンダリストレージ装置 10 b の S-VOL（移行元 S-VOL と呼ぶ）の LDEV # が 22 番、サイズが 100 GB

(c) 移行元プライマリストレージ装置 10 a のシリアル番号が 1 番

(d) 移行元セカンダリストレージ装置 10 b のシリアル番号が 11 番

(e) 移行先プライマリストレージ装置 20 a のシリアル番号が 2 番

(f) 移行先セカンダリストレージ装置 20 b のシリアル番号が 22 番

【0132】

40

また、S10 において、移行元プライマリストレージ装置 10 a、移行元セカンダリストレージ装置 10 b で使用されていない LDEV # を調査した結果、移行先 P-VOL、移行先 S-VOL、ボリュームマイグレーション機能のターゲットボリュームとして用いる論理ボリュームとして、以下の LDEV # の論理ボリュームを用いると決定された場合を想定する。

(g) 移行先プライマリストレージ装置 20 a の P-VOL（移行先 P-VOL と呼ぶ）として選定される論理ボリュームの LDEV # が 33 番

(h) 移行先プライマリストレージ装置 20 a において、ボリュームマイグレーション機能のターゲットボリュームとして選定される論理ボリュームの LDEV # が 44 番、及び仮想 LDEV # が 99 番

50

(i) 移行元セカンダリストレージ装置 1 0 b の S - V O L (移行元 S - V O L と呼ぶ) の L D E V # が 5 5 番

【 0 1 3 3 】

また、プライマリサーバ 3 0 a、セカンダリサーバ 3 0 b で稼働している、ストレージマネージャ 1 0 1 のインスタンス番号、及び各インスタンスが使用している設定ファイルを特定しておく。以下の例では、プライマリサーバ 3 0 a 及びセカンダリサーバ 3 0 b で稼働しているインスタンスのインスタンス番号がそれぞれ、1 0 0 番及び 1 0 1 番であったとする。また以下では、設定ファイルの内容が図 1 1 に示されているものであった場合について説明する。

【 0 1 3 4 】

続いて、移行先ストレージシステムを設置し、移行先ストレージシステムをホスト計算機や他のストレージ装置とファイバチャネルケーブルで接続する (S 2 0)。なお、ストレージ装置とホスト計算機、あるいはストレージ装置同士を接続する、ファイバチャネルケーブルなどの伝送線のことを、「物理パス」と呼び、ストレージ装置とホスト計算機、あるいはストレージ装置同士を物理パスで接続する作業のことを、「物理パスの接続」と呼ぶ。

【 0 1 3 5 】

物理パスの接続作業では、以下の物理パスを接続する。

(ア) 移行先プライマリストレージ装置 2 0 a (ポート J、ポート K) とプライマリサーバ 3 0 a 間の物理パス

(イ) 移行先セカンダリストレージ装置 2 0 b (ポート Q、ポート R) とセカンダリサーバ 3 0 b 間の物理パス

(ウ) 移行元プライマリストレージ装置 1 0 a (ポート C) と移行先プライマリストレージ装置 1 0 b (ポート L) 間の物理パス

(エ) 移行先プライマリストレージ装置 2 0 a (ポート M、ポート N) と移行先セカンダリストレージ装置 2 0 b (ポート O、ポート P) 間の物理パス

【 0 1 3 6 】

(ア) は、プライマリサーバ 3 0 a と移行先プライマリストレージ装置 2 0 a の間の I / O 要求、移行先プライマリストレージ装置 2 0 a のコマンドデバイスへの指示をやり取りするための物理パスである。(イ) は、セカンダリサーバ 3 0 b と移行先セカンダリストレージ装置 2 0 b の間の I / O 要求、移行先セカンダリストレージ装置 2 0 b のコマンドデバイスへの指示をやり取りするための物理パスである。

【 0 1 3 7 】

(ウ) は、データ移行のために用いられる物理パスで、外部ストレージ接続機能により移行元プライマリストレージ装置 1 0 a の移行元ボリューム (P - V O L) を、移行先プライマリストレージ装置 2 0 a にマッピングするために用いられる。実施例 1 では、(ウ) の物理パスは 1 本の例を説明したが、複数の物理パスを接続する構成をとることもできる。また (エ) は、移行先ストレージシステムで、リモートコピー機能によるデータ伝送のために用いられる物理パスである。

【 0 1 3 8 】

S 3 0 では、ストレージマネージャ 1 0 1 は移行元プライマリストレージ装置 1 0 a の論理ボリュームに対して、L U パス設定を行う。ここで L U パス設定の行われる論理ボリュームは、移行対象となる P - V O L である。またここで行われる L U パス設定は、移行先プライマリストレージ装置 2 0 a が外部ストレージ接続機能により、移行元プライマリストレージ装置 1 0 a の移行対象ボリューム (P - V O L) を認識できるようにするために行われる。そのため、S 2 0 で移行先プライマリストレージ装置 2 0 a と接続した、移行元プライマリストレージ装置 1 0 a のポート C に対して L U パス設定を行う。なお、L U パス設定の際、L U N を指定する必要があるが、任意の L U N を指定すればよい。

【 0 1 3 9 】

プライマリサーバ 3 0 a のストレージマネージャ 1 0 1 が移行元プライマリストレージ

10

20

30

40

50

装置 10 a に L U パス設定の指示を発行すると、移行元プライマリストレージ装置 10 a は、その指示に従って P - V O L に対する L U パス設定を行う。また、別の実施態様として、管理端末から L U パス設定が行われてもよい。

【 0 1 4 0 】

続いて S 4 0 では、移行先プライマリストレージ装置 2 0 a の設定作業を行う。以下、図 1 7 を用いて、S 4 0 で行われる処理を説明する。まず、管理端末は、コマンドデバイスを作成する。そしてコマンドデバイスの作成後、プライマリサーバ 3 0 a でストレージマネージャ 1 0 1 のインスタンスを起動する。この時起動されるインスタンスが読み込む設定ファイルは、図 1 3 に示されているようなファイル、つまり少なくともコマンドデバイスの識別子の情報が記録された設定ファイルであればよい。また、ここで起動されるインスタンスのインスタンス番号は、すでにプライマリサーバ 3 0 a 及びセカンダリサーバ 3 0 b で稼働しているインスタンスのインスタンス番号（1 0 0 番及び 1 0 1 番）とは異なるインスタンス番号を用いる。これ以降の各作業は、ストレージマネージャ 1 0 1 から行われる。ただし別の実施形態として、管理端末が用いられてもよい。

【 0 1 4 1 】

続いて S 4 0 b では、起動されたインスタンスを用いて、移行先プライマリストレージ装置 2 0 a のポートのうち、移行元プライマリストレージ装置 1 0 a と接続されるポート（ポート L）の属性を、外部ストレージ接続機能用の属性に変更する。さらに移行先プライマリストレージ装置 2 0 a のポートのうち、移行先セカンダリストレージ装置 2 0 b と接続されるポート（ポート M、ポート N）の属性を、リモートコピー機能用の属性に変更する。具体的にはポート M、ポート N のうち、一方のポートの属性を、移行先プライマリストレージ装置 2 0 a から移行先セカンダリストレージ装置 2 0 b へのデータ送信用ポートの属性にし、残りのポートの属性を移行先セカンダリストレージ装置 2 0 b から移行先プライマリストレージ装置 2 0 a へのデータ受信用ポートの属性に変更する。同様に、移行先セカンダリストレージ装置 2 0 b のポートのうち、移行先プライマリストレージ装置 2 0 a と接続されるポート（ポート O、ポート P）の属性を、リモートコピー機能用の属性に変更する。

【 0 1 4 2 】

S 4 0 c では、ストレージマネージャ 1 0 1 は、移行先プライマリストレージ装置 2 0 a に仮想ストレージを作成する。仮想ストレージを作成する際に、ストレージマネージャ 1 0 1 が移行先プライマリストレージ装置 2 0 a に発行する仮想ストレージ作成用の制御用命令には、作成対象の仮想ストレージの識別子、仮想ストレージのシリアル番号、機種名の情報が含まれる。S 4 0 c で発行される制御用命令に含まれる仮想ストレージの識別子は V - B O X 管理テーブル T 1 0 0 0 の I D（T 1 0 0 1）に格納される識別子であり、定義済みである仮想ストレージの識別子（たとえば、先に述べたように、識別子 v 0 は初期状態で定義済みである）と異なる識別子が指定される。また S 4 0 c で発行される制御用命令に含まれる仮想ストレージのシリアル番号は移行元プライマリストレージ 1 0 a のシリアル番号、仮想ストレージ機種名は移行元プライマリストレージ 1 0 a の機種名が指定される。ストレージマネージャ 1 0 1 はこれらの情報とともに仮想ストレージ作成用の制御用命令を移行先プライマリストレージ装置 2 0 a に発行する。仮想ストレージ作成用の制御用命令を受信した移行先プライマリストレージ装置 2 0 a は、V - B O X 管理テーブル T 1 0 0 0 の、I D（T 1 0 0 1）、機種名（T 1 0 0 2）、S / N（T 1 0 0 3）に、受信した仮想ストレージの識別子、機種名、シリアル番号の情報を格納する。

【 0 1 4 3 】

S 4 0 d では、ストレージマネージャ 1 0 1 は移行先プライマリストレージ装置 2 0 a に対し、移行先ボリュームとして用いるボリュームに付されている仮想 L D E V # を削除する指示を発行する。先に述べたとおり、移行先プライマリストレージ装置 2 0 a で用意する移行先ボリューム（移行先 P - V O L、ターゲットボリューム）の L D E V # が 3 3 番と 4 4 番であるので、その場合について説明する。移行先プライマリストレージ装置 2 0 a は、仮想 L D E V # を削除する制御用命令を受信すると、仮想 L D E V 管理テーブル

10

20

30

40

50

に登録されているLDEVのうち、LDEV#(T1501)が44番と33番である行のVLDEV#(T1502)の内容を無効化する(NULL値を格納する)。また、ストレージマネージャ101は、仮想LDEV#を削除したLDEVのLDEV#を、仮想ストレージに登録する指示を発行する。移行先プライマリストレージ装置20aはこの指示を受けると、V-BOX管理テーブルT1000のLDEV#(T1005)にLDEV#を登録し、また仮想LDEV管理テーブルT1500の属性T1504に“reserved”を格納する。属性T1504に“reserved”が格納されていると、移行先プライマリストレージ装置20aは、当該LDEVが仮想ストレージに所属していると認識し、当該LDEVを他の用途で用いない。

【0144】

S40eでは、ストレージマネージャ101は、仮想ストレージに登録したLDEVを仮想化、つまり仮想ストレージに登録したLDEVに対して仮想LDEV#を付す指示を発行する。先に述べたとおり、ターゲットボリュームの仮想LDEV#には99番を使用する。そして移行先P-VOLの仮想LDEV#には、移行元P-VOLのLDEV#を付すので、LDEV#が33番の論理ボリュームに対して、11番の仮想LDEV#を付す制御用命令を、またLDEV#が44番の論理ボリュームに対して、99番の仮想LDEV#を付す制御用命令を発行する。

【0145】

移行先プライマリストレージ装置20aはこの指示を受け付けると、V-BOX管理テーブルT1000に登録されているLDEV#(T1005)が33番の行に対応するVLDEV#(T1004)に11を格納し、またLDEV#(T1005)が44番の行に対応するVLDEV#(T1004)に99を格納する。これにより、LDEV#が33番の論理ボリュームに仮想LDEV#として11番が付され、LDEV#が44番の論理ボリュームに、仮想LDEV#として99番が付されたことになる。

【0146】

S40fでは、外部ストレージ接続機能により、移行元プライマリストレージ装置10aの移行元ボリューム(P-VOL)を、移行先プライマリストレージ装置20aにマッピングする。ストレージマネージャ101は、移行元ボリュームを移行先プライマリストレージ装置20aの外部ボリュームグループに登録する制御用命令を発行する。移行先プライマリストレージ装置20aはこのコマンドを受け付けると、外部ボリュームグループ管理テーブルT150に、移行元ボリュームの情報を登録する。この時移行先プライマリストレージ装置20aは、移行元ボリュームのサイズの情報を移行元プライマリストレージ装置10aから取得し、サイズT154に格納する。図4が移行元P-VOLをマッピングした後の、外部ボリュームグループ管理テーブルT150の例である。ここでは外部ボリュームグループ名(T151)がEG1である外部ボリュームグループに、移行元P-VOLがマッピングされた場合の例が示されている。移行元P-VOLのサイズが100GBであるので、サイズT154には「100GB」が格納される。

【0147】

さらに、ストレージマネージャ101は外部ボリュームグループの記憶領域から論理ボリューム(移行先P-VOL)を作成する制御用命令を発行する。移行先プライマリストレージ装置20aはこの制御用命令を受信すると、論理ボリューム管理テーブルT200のLDEV#(T201)に33、サイズT202に100GB(つまり移行元P-VOLの全領域を使って移行先P-VOLを作成する)、グループ名T203に“EG1”、先頭アドレスT204に0を格納し、LDEV#が33番の論理ボリュームの記憶領域が外部ボリュームグループEG1の記憶領域である旨の情報を格納する。

【0148】

S40gでは、ボリュームマイグレーション機能で用いるターゲットボリュームを作成する。ここで作成するターゲットボリュームのサイズは、移行先P-VOLと同サイズであり、記憶領域として、移行先プライマリストレージ装置20a内のパリティグループの記憶領域を用いる。パリティグループが複数存在する場合、パリティグループの残サイズ

10

20

30

40

50

(パリティグループ管理テーブルT100の残サイズT105)が移行先P-VOLのサイズ以上である、任意のパリティグループを選定して良い。

【0149】

ストレージマネージャ101は、LDEV#が44番の論理ボリュームに、パリティグループの記憶領域を割り当てる制御用命令を発行する。移行先プライマリストレージ装置20aはこの制御用命令を受け付けると、論理ボリューム管理テーブルT200に、LDEV#及びLDEVに割り当てた記憶領域(パリティグループ)の情報を登録する。図5に、割り当て後の論理ボリューム管理テーブルT200の内容の一例を示す。図5では、LDEV#44に、パリティグループRGAの領域が100GB割り当てられていること(行906)、LDEV#33に、外部ボリュームグループEG1の領域が100GB

10

【0150】

S40hでは、移行先ボリュームにLUパス設定を行う。ストレージマネージャ101は、LDEV#33に、ポート名、LUNを割り当てる制御用命令を発行する。移行先プライマリストレージ装置20aはこの制御用命令を受け付けると、LU管理テーブルT300に情報を登録する。たとえば、LDEV#33に、ポート名3e243174aaaaaaa、LUN0を割り当てる旨の指示が到来した場合、移行先プライマリストレージ装置20aは、LU管理テーブルT300の、ポート名T301が“3e243174aaaaaaa”、LUN(T303)が0の行のLDEV#(T304)の欄に、“33”を登録する。

20

【0151】

以上がS40で行われる処理である。続いて図16に戻り、S50以降の処理について説明する。S50では、移行先セカンダリストレージ装置20bの設定作業が行われる。この処理は、S40で移行先プライマリストレージ装置20aに対して行った作業とほとんど同様である。コマンドデバイスの作成(S40a相当の作業)、仮想ストレージの作成(S40c相当の作業)、移行先ボリューム(S-VOL)として用いるボリュームに付されている仮想LDEV#の削除及び移行先ボリュームの仮想ストレージへの登録(S40d相当の作業)、移行先ボリュームの仮想化(S40e相当の作業)、移行先ボリュームの作成(S40f相当の作業)、移行先ボリュームへのLUパス設定(S40h相当の作業)を行う。また、コマンドデバイスの作成作業は、管理端末から実施し、コマンド

30

【0152】

S60では、プライマリサーバ30aから移行先プライマリストレージ装置20aへの交替パスを追加する。プライマリサーバ30aで、プライマリサーバ30aのオペレーティングシステム及び交替パスソフト102が提供するコマンドを実行することで、S40の作業で移行先プライマリストレージ装置20aに作成された移行先P-VOLを、プライマリサーバ30aのオペレーティングシステム及び交替パスソフト102が認識する。移行先P-VOLが交替パスソフト102から認識されると、移行元プライマリストレージ装置10aの移行元P-VOLと、移行先プライマリストレージ装置20aの移行先P-VOLは、プライマリサーバ30aに対して同じ属性情報(装置シリアル番号、LDEV#)を見せているため、交替パスソフト102は双方のボリュームが同一ボリュームであると認識し、交替パスが構築される。

40

【0153】

続いて、プライマリサーバ30aから移行元プライマリストレージ装置10aの移行元ボリュームへのパスが無効化されるよう、交替パスソフト102の設定が変更される。これは交替パスソフト102が提供するコマンドが実行されることによって行われる。交替パスソフト102は、移行元P-VOLと移行先P-VOLが同一ボリュームであると認識しているので、交替パスが削除された場合、交替パスソフト102は、プライマリサー

50

バ30aで稼働するアプリケーションプログラムなどが移行元P-VOLに対して発行していたI/O要求を、移行先P-VOLへと発行するようになる。

【0154】

その後移行元プライマリストレージ装置10aで、移行元P-VOLのLUパスの削除が行われる。この作業は、ストレージマネージャ101または管理端末から、移行元プライマリストレージ装置10aに対してLUパスの削除指示を発行することによって行われる。

【0155】

S70では、移行先ストレージシステム10で、移行先P-VOLと移行先S-VOL間のペア作成を行う。プライマリサーバ30a及びセカンダリサーバ30bで、移行元P-VOL、移行先S-VOLのボリュームペア操作用に用いられていた設定ファイル(設定ファイルの例は図11)のデータに基づいて、図12のように、仮想ストレージに対して制御用命令を発行するための設定ファイルが作成される。そして作成された設定ファイルを用いて、プライマリサーバ30a及びセカンダリサーバ30bのストレージマネージャ101のインスタンスが起動される。なお、ここで起動されるインスタンスのインスタンス番号は、すでにプライマリサーバ30a及びセカンダリサーバ30bで稼働しているインスタンスのインスタンス番号(100番及び101番)とは異なるインスタンス番号を用いる。その後プライマリサーバ30aのストレージマネージャ101がペア作成の制御用命令を発行すると、移行先ストレージシステム10は、移行先P-VOLと移行先S-VOL間でのデータ複製を開始する。

【0156】

ペア作成が完了する、つまり移行先P-VOLのデータが完全に移行先S-VOLに複製される間、プライマリサーバ30aが移行先P-VOL(LDEV#33、仮想LDEV#11)にデータをライトすると、そのデータは外部ストレージ接続機能により移行元プライマリストレージ装置10aの移行元P-VOL(LDEV#11)へと書き込まれ、書き込まれたデータは移行元S-VOL(LDEV#22)へと書き込まれる。よって、移行先P-VOLと移行先S-VOLとの間のペア形成中も移行元P-VOLと移行元S-VOLとの間で冗長構成が保たれている。また、ペア状態は、ストレージマネージャ101から移行先プライマリストレージ装置20aに対してペア状態を表示する制御用命令を発行することによって、確認できる。

【0157】

移行先P-VOLと移行先S-VOLのペア状態が“pair”状態に変更された後、プライマリサーバ30a及びセカンダリサーバ30b上のクラスタソフトウェア103等、ディザスタリカバリ用のソフトウェアの動作が停止される(S80)。

【0158】

続いて、移行元ストレージシステム用に起動していた、ストレージマネージャ101のインスタンスを停止し、ストレージマネージャ101の設定ファイルの書き換えを行う。設定ファイルの書き換えは、S70で行ったものと同様である。設定ファイルの書き換えでは、ボリュームペアの情報が記載されているブロック3002等の変更は不要であるため、設定ファイルの書き換えは短時間で実施できる。設定ファイルの書き換えが終了すると、プライマリサーバ30a及びセカンダリサーバ30bで、当該設定ファイルを用いたインスタンス起動が行われる(S90)。ここで起動されるインスタンスのインスタンス番号は、移行処理前にプライマリサーバ30a及びセカンダリサーバ30bで稼働しているインスタンスのインスタンス番号(100番及び101番)と同じインスタンス番号が用いられる。

【0159】

S100では、セカンダリサーバ30bから移行元セカンダリストレージ装置10bの移行元ボリュームへのパスの無効化、移行元セカンダリストレージ装置10bの移行元ボリュームのLUパス削除が行われる。この作業はS60で、移行元プライマリストレージ装置10aの移行元P-VOLに対して行われた作業と同様の作業を、移行元セカンダリ

ストレージ装置 10 b の移行元 S - V O L に対して行うものである。また、セカンダリサーバ 30 b から移行先セカンダリストレージ装置 20 b の移行先 S - V O L への L U パス設定を行い、セカンダリサーバ 30 b に移行先 S - V O L を認識させる。その後、S 80 で停止していたプライマリサーバ 30 a 及びセカンダリサーバ 30 b 上のクラスタソフトウェア 103 等のディザスタリカバリ用のソフトウェアの動作を再開させる。

【0160】

続いて S 110 では、移行先 P - V O L をターゲットボリュームに移行する作業が行われる。移行の際、S 90 で起動したストレージマネージャ 101 のインスタンスとは別のインスタンス（第 2 インスタンスと呼ぶ）がプライマリサーバ 30 a で起動され、移行先プライマリストレージ装置 20 a に対するボリュームの移行の指示は第 2 インスタンスを用いて行われる。指示を受け付けた移行先プライマリストレージ装置 10 a は、ボリュームマイグレーション機能を用いて、移行先 P - V O L をターゲットボリュームに移行する。

10

【0161】

移行の完了を待つ、つまり移行先 P - V O L のデータが完全にターゲットボリュームに複製される間、プライマリサーバ 30 a が移行先 P - V O L (L D E V # 33、仮想 L D E V # 11) にデータをライトすると、そのデータは外部ストレージ接続機能により移行元プライマリストレージ装置 10 a の移行元 P - V O L (L D E V # 11) へと書き込まれる。また、書き込まれたデータは移行先 S - V O L (L D E V # 55、仮想 L D E V # 22) へと書き込まれる。よって、移行先 P - V O L とターゲットボリュームとの間のデータ移行中も移行元 P - V O L と移行先 S - V O L との間で冗長構成が保たれている。また、ボリューム状態は、ストレージマネージャ 101 から移行先プライマリストレージ装置 20 a に対して、移行状態を取得する制御用命令を発行することによって、確認できる。

20

【0162】

移行が完了すると、移行元ストレージシステムのデータはすべて移行先ストレージシステムに格納された状態になる。また、ボリュームマイグレーション機能により、L D E V # とデータ格納先記憶領域との対応関係が、マイグレーション元とマイグレーション先とで入れ替わる。そのため、マイグレーションが完了すると、仮想 L D E V # 11 を指定したプライマリサーバ 30 a からのアクセス要求に対して、移行先プライマリストレージ装置 20 a 内のパリティグループ 22 a の記憶領域が対応付けられた論理ボリュームに対してアクセス処理が実行される。プライマリサーバ 30 a 及びセカンダリサーバ 30 b は移行元ストレージシステムにはアクセスしない状態になっている（パスが削除されているため）ので、管理者は移行元ストレージシステムを撤去する（S 120）。これにより移行処理が完了する。ここまで説明した手順により、プライマリサーバ 30 a からのアクセス（リード、ライト要求等）受け付けを停止することなく、冗長構成を維持したまま、P - V O L と S - V O L のボリュームペアを移行元ストレージシステム 10 から移行先ストレージシステム 20 に移行することができる。

30

【0163】

[変形例 1]

上で説明した移行処理では、S 70 でペア作成の指示を行った後、ペア状態が “ p a i r ” に変化したことが確認されてから、S 80 以降の処理が行われていた。ただし本発明の移行処理は、上で説明した処理手順に限定されるものではない。別の実施形態として、S 70 でペア作成の指示を行った直後に、それ以降の作業を実施するようにしてもよい。その場合の処理の流れについて、上で説明した移行処理との相違点を中心に説明する。

40

【0164】

S 10 から S 60 の処理は、上で説明した処理と同一である。S 70 において、ペア作成の指示を行うと、ペア状態が “ p a i r ” に変化するのを待たずに S 80 の処理（ディザスタリカバリ用のソフトウェアの動作の停止）に進む。

【0165】

50

続いて、上で説明した移行処理のうち、S 1 1 0のボリュームの移行指示を移行先プライマリストレージ装置2 0 aに対して発行する。ボリューム移行指示の後、ペア作成が完了するのを待つ。この場合、ペア作成が完了すると、同時にボリュームマイグレーション機能により、移行元P - V O Lとターゲットボリュームが入れ替わる。

【 0 1 6 6 】

ペア作成が完了した後、S 9 0、S 1 0 0の処理が実行される。その後S 1 2 0の処理（移行元ストレージシステムの撤去）を行って、移行処理が終了する。

【 0 1 6 7 】

この移行処理の場合、リモートコピー機能による移行元P - V O Lから移行先S - V O Lへの複製と、移行元P - V O Lからターゲットボリュームへのデータ移行（複製）とが並行して実施されるため、実施例1に係るデータ移行処理よりも、データ移行に要する時間を短縮することができる。

10

【 0 1 6 8 】

[変形例 2]

続いて、本発明の変形例2に係る計算機システムについて説明する。図19は、本発明の変形例2に係る計算機システムの構成図である。本発明の変形例2に係る計算機システムは実施例1に係る計算機システムと同じく、移行元プライマリストレージ装置1 0 a、移行元セカンダリストレージ装置1 0 b、移行先プライマリストレージ装置2 0 a、移行先セカンダリストレージ装置2 0 b、プライマリサーバ3 0 a、セカンダリサーバ3 0 bから構成され、各ストレージ装置のハードウェア構成も実施例1で説明したものと同じである。

20

【 0 1 6 9 】

実施例1に係る計算機システムとの相違は、リモートコピー機能に、ジャーナルを用いたリモートコピー機能が用いられる点である。なお、変形例2に係る計算機システムにおいて、ジャーナルとは、P - V O Lのあるストレージ装置からS - V O Lのあるストレージ装置に対して送信される、P - V O Lの複製データを一時蓄積するために用いられる記憶領域である。移行元プライマリストレージ装置1 0 aと移行元セカンダリストレージ装置1 0 bにはそれぞれ、ジャーナル1 3 a、ジャーナル1 3 bが設けられている。また、移行処理の際、移行先プライマリストレージ装置2 0 aと移行先セカンダリストレージ装置2 0 bにもそれぞれ、ジャーナル2 3 a、ジャーナル2 3 bが設けられる。

30

【 0 1 7 0 】

ジャーナル(1 3 a、1 3 b、2 3 a、2 3 b)には、ジャーナルIDと呼ばれる、ストレージ装置内で一意な識別子が付されている。また、ジャーナル1 3 a、1 3 b、2 3 a、2 3 bにはそれぞれ、ジャーナルボリューム1 3 3 a、1 3 3 b、2 3 3 a、2 3 3 bが登録されている。以下、ジャーナルのことを「JNL」と略記することがあり、またジャーナルボリュームのことを「JVOL」と略記することがある。

【 0 1 7 1 】

ジャーナルボリューム(1 3 3 a、1 3 3 b、2 3 3 a、2 3 3 b)は、実施例1で説明してきた論理ボリューム1 3 0 a等と同じ論理ボリュームである。なお、実施例1で説明してきた論理ボリューム1 3 0 aは、論理ボリューム作成時に、静的にストレージ装置内のパリティグループの記憶領域が対応付けられるボリュームであったが、いわゆるThin Provisioning技術を用いて形成されるボリュームをジャーナルボリュームとして用いてもよい。Thin Provisioning技術を用いて形成されるボリュームの場合、ボリュームに対するアクセスを受け付けた時点で、アクセスのあった領域に対して動的に記憶領域を割り当てるため、記憶領域の節約になる。

40

【 0 1 7 2 】

変形例2に係る計算機システムにおいて、移行元P - V O Lである論理ボリューム1 3 0 aのデータは、ジャーナルを用いたリモートコピー機能によって、常時移行元S - V O Lである論理ボリューム1 3 0 bに複製されている状態（ペア状態）にある。移行処理中のデータの流れを図20に示すが、移行元P - V O Lの複製データが移行元S - V O Lへ

50

と複製される前に、JNL13a、13bを経由する点以外は、実施例1と変わりはない。つまり、移行元P-VOLの複製データを移行元セカンダリストレージ装置10bに転送する際、複製データは一旦JNL13a(のJVOL133a)に格納される。なお、この時複製データには、シーケンス番号と呼ばれる、複製データの書き込み順を指示する番号が付される。シーケンス番号は、プライマリサーバ30aが移行元P-VOLにデータを書き込んだ順序と同じ順序で、移行元セカンダリストレージ装置10bが移行元S-VOLに複製データを書き込めるようにするために付されるものである。

【0173】

JNL13aに格納された複製データは、移行元セカンダリストレージ装置10bのJNL13b(のJVOL133b)へと転送される。その後移行元セカンダリストレージ装置10bは、JNL13bに格納された移行元P-VOLの複製データを取り出し、複製データに付されているシーケンス番号の順に、複製データを移行元S-VOLへと反映する。ここでは、移行元ストレージシステム10におけるボリューム複製の概要を説明したが、移行先ストレージシステム20でも、移行先P-VOLと移行先S-VOLのボリュームペア作成が行われると同様の処理が行われる。

10

【0174】

[移行処理のフロー]

続いて、ボリューム移行処理のフローを説明する。変形例2に係る計算機システムにおける、ボリューム移行処理の流れは、実施例1で説明したものとほとんど同じであるため、実施例1で用いられた図16~18を用いて説明する。

20

【0175】

変形例2に係る計算機システムでは、リモートコピー機能としてジャーナルを用いたりリモートコピー機能を用いるため、移行先ストレージシステム20にジャーナルを用意する処理が加わる点が、主な相違点である。移行処理のS10において、実施例1で説明した情報に加えて、移行元P-VOL、移行元S-VOLのボリュームペアで使用されているジャーナルのジャーナルIDの情報を取得する。

【0176】

後で説明するS40、S50において、移行先ストレージシステム20にジャーナルを作成するが、ここでは移行元ストレージシステム10のジャーナル13a、13bのジャーナルIDと同じIDのジャーナル(23a、23b)を作成する。以下では、ジャーナルIDの情報の取得の結果、ジャーナル13aのジャーナルIDは0、ジャーナル13bのジャーナルIDは1であった場合を例にとって説明する。また、ジャーナルボリューム233aのLDEV#は77番、ジャーナルボリューム233bのLDEV#は88番とする場合を例にとって説明する。

30

【0177】

S20、S30は、実施例1で説明した処理と同じである。

【0178】

S40では、実施例1で説明した処理(図17のS40a~S40h)に加えて、S40gとS40hの間に、ジャーナルボリューム233aの作成、そしてジャーナルボリューム233aのジャーナル23aへの登録を行う。それ以外の点は、実施例1で説明したS40の処理と同じである。ジャーナルボリューム233aの作成及びジャーナル23aへの登録は、ストレージマネージャ101から移行先プライマリストレージ装置20aに、ジャーナルID=0のジャーナルに、LDEV#77の論理ボリューム(233a)を登録する制御用命令を発行することで実施される。この制御用命令を受信した移行先プライマリストレージ装置20aは、ジャーナルID=0のジャーナルを作成し、さらにLDEV#77の論理ボリューム(233a)を作成されたジャーナルに登録する処理を行う。なお、この処理を行う順序は上で説明した順に限定されるわけではない。たとえばS40fの前に行うようにしてもよい。

40

【0179】

S50でも、実施例1で説明した処理に加えて、ジャーナルボリューム233bの作成

50

及び作成されたジャーナルボリューム 233b のジャーナル 23b への登録を行う。それ以外の点は実施例 1 で説明した S50 の処理と同じである。

【0180】

S60 は、実施例 1 で説明した処理と同じである。

【0181】

S70 は、ストレージマネージャ 101 が、ボリュームペア（つまり移行先 P-VOL と移行先 S-VOL）のペア作成の制御用命令を発行する際に、移行先 P-VOL と移行先 S-VOL の識別子を指定する他に、S40 で作成されたジャーナル 23a、ジャーナル 23b のジャーナル ID（0 と 1）を指定した制御用命令を発行する。それ以外の点は、実施例 1 で説明した処理と同じである。これにより、ジャーナル 23a、ジャーナル 23b を用いた、移行先 P-VOL から移行先 S-VOL へのコピーが行われる。

10

【0182】

S80 以降の処理は、実施例 1 で説明した処理と同じである。

【実施例 2】

【0183】

続いて、本発明の実施例 2 に係る計算機システムについて説明する。図 21 は、本発明の実施例 2 に係る計算機システムの概略構成図である。計算機システムは、移行元ストレージ装置 10a、移行先ストレージ装置 20a、プライマリサーバ 30a、セカンダリサーバ 30b からなる。移行元ストレージ装置 10a、移行先ストレージ装置 20a のハードウェア構成は、実施例 1 で説明した構成と同様である。

20

【0184】

本発明の実施例 2 に係る計算機システムでは、移行前には移行元ストレージ装置 10a 内の論理ボリューム 130a、130a' がボリュームペアとして運用されている。論理ボリューム 130a が P-VOL、論理ボリューム 130a' が S-VOL である。つまり、実施例 1 に係る計算機システムにおいて、移行元セカンダリストレージ 10b 内にあった論理ボリューム 130b が、実施例 2 に係る計算機システムにおいては、論理ボリューム 130a と同じストレージ装置（移行元ストレージ装置 10a）に存在する構成である。

【0185】

論理ボリューム 130a' は、論理ボリューム 130a のバックアップ取得用に用いられるボリュームで、セカンダリサーバ 30b がバックアップ取得時に論理ボリューム 130a' のデータを、非図示のバックアップデバイスにバックアップする。セカンダリサーバ 30b には、バックアップ運用のためのプログラムであるバックアップソフト 104 が稼働している。

30

【0186】

たとえばバックアップ取得時には、バックアップソフト 104 はストレージマネージャ 101 を用いてボリュームペアの状態を “suspended” にし（あるいは、すでにボリュームペアが “suspended” の状態であった場合には、ボリュームペアの再同期を行ってからボリュームペアの状態を “suspended” にする）、セカンダリサーバ 30b のバックアップソフト 104 は論理ボリューム 130a' のデータを、非図示のバックアップデバイスにコピーする処理を行う。バックアップ処理が終了すると、セカンダリサーバ 30b はボリュームペアの再同期を行い、ボリュームペアを再び “pair” 状態にする。

40

【0187】

以下では、このような構成の計算機システムにおいて、論理ボリューム 130a、130a' のペア状態を維持したまま移行する際の処理の流れを、図 22、23、24 を用いて説明する。この処理の流れは、実施例 1 の移行処理と類似しているため、以下では実施例 1 との相違点を中心に説明する。

【0188】

最初に、移行元ストレージ装置 10a の構成の調査を行う（S10'）。実施例 1 の S

50

10では、移行元セカンダリストレージ装置10bについての調査を行ったが、実施例2の計算機システムでは移行元セカンダリストレージ装置10bは存在しないため、S10'では移行元セカンダリストレージ装置10bについての調査は行わない。その他の点は、S10'と実施例1のS10とは変わりがない。

【0189】

続いて管理者は、移行先ストレージ装置20aを設置し、移行先ストレージシステムとホスト計算機や移行元ストレージ装置10aとの間の物理パス接続を行う(S20')。物理パスの接続作業では、以下の物理パスを接続する。

(ア) 移行先ストレージ装置20a(ポートJ、ポートK)とプライマリサーバ30a間の物理パス

(イ) 移行先ストレージ装置20a(ポートQ、ポートR)とセカンダリサーバ30b間の物理パス

(ウ) 移行元ストレージ装置10a(ポートC)と移行先ストレージ装置20a(ポートL)間の物理パス

【0190】

S30'では、ストレージマネージャ101は、移行元ストレージ装置10aの論理ボリュームに対して、LUパス設定を行う。ここで行われるLUパス設定は、実施例1のS30で行われた処理と同様である。つまり移行先ストレージ装置20aが外部ストレージ接続機能により、移行元プライマリストレージ装置10aの移行対象ボリューム(P-VOL)を認識できるようにするために行われる。そのため、S20で移行先ストレージ装置20aと接続した、移行元ストレージ装置10aのポートCに対してLUパス設定を行う。

【0191】

続いてS40'では、移行先ストレージ装置20aの設定作業が行われる。この作業も、移行先セカンダリストレージ装置に対しての処理を行わない点を除けば、実施例1のS40と同様の作業が行われる。以下、図23を用いて、S40'で実施される処理を説明する。まず、管理端末は、コマンドデバイスを作成する(S40a)。コマンドデバイスの作成後、プライマリサーバ30aでストレージマネージャ101のインスタンスが起動される。以降の各作業では、ストレージマネージャ101から各種設定が行われる。ただし別の実施形態として、管理端末が用いられてもよい。

【0192】

続いてS40bで、プライマリサーバ30aで稼働するストレージマネージャ101は、移行先ストレージ装置20aのポートのうち、移行元プライマリストレージ装置10aと接続されるポート(ポートL)の属性を、外部ストレージ接続機能用の属性に変更する。

【0193】

S40cでは、ストレージマネージャ101は、移行先ストレージ装置20aに仮想ストレージ25aを作成する。

【0194】

S40d'では、ストレージマネージャ101は移行先ストレージ装置20aに対し、移行先ボリュームとして用いるボリュームに付されている仮想LDEV#を削除する制御命令を発行する。続いて、ストレージマネージャ101は、仮想LDEV#を削除したLDEVのLDEV#を、仮想ストレージに登録する制御命令を発行する。ここで、移行先ストレージ装置20aで用いられる移行先ボリュームには、移行先P-VOL及び移行先S-VOL(実施例1では移行先セカンダリストレージ装置20b内にあったもの)が存在する。そのためS40d'では、2つのボリュームに付されている仮想LDEV#を削除し、この2つのボリュームを仮想ストレージに登録する。

【0195】

S40e'では、ストレージマネージャ101は、仮想ストレージに登録したLDEVの仮想化を実施する。S40d'と同じく、2つのボリュームに対して仮想化が実施され

10

20

30

40

50

る。

【0196】

S40fでは、ストレージマネージャ101は、移行元ストレージ装置10aの移行元P-VOLを、移行先プライマリストレージ装置20aにマッピングする。さらにストレージマネージャ101は、マッピングされた記憶領域から移行先ボリューム(移行先P-VOL)を作成する。

【0197】

S40g'では、ボリュームマイグレーション機能で用いるターゲットボリュームを作成する。ここで作成するターゲットボリュームのサイズは、移行先P-VOLと同サイズであり、記憶領域として、パリティグル プ22aの記憶領域を用いる。同時に、移行先S-VOLの作成を行う。移行先S-VOLも記憶領域として、パリティグル プ22aの記憶領域が用いられる。

10

【0198】

S40hでは、移行先P-VOLにLUパス設定を行う。ストレージマネージャ101は、LDEV#33に、ポート名、LUNを割り当てる制御用命令を発行する。

【0199】

以上がS40'で行われる処理である。続いて図22に戻り、S60'以降の処理について説明する。

【0200】

S60'では、プライマリサーバ30aから移行先ストレージ装置20aへの交替パスが追加され、またプライマリサーバ30aから移行元ストレージ装置10aへの交替パスが削除される。これはS60と同様の処理である。

20

【0201】

S70'では、移行先ストレージ装置20aで、移行先P-VOLと移行先S-VOL間のペア作成が行われる。プライマリサーバ30aのストレージマネージャ101がペア作成の制御用命令を発行すると、移行先ストレージ装置20aはローカルコピー機能により、移行先P-VOLと移行先S-VOL間でのデータ複製を開始する。

【0202】

移行先P-VOLと移行先S-VOLのペア状態が“pair”状態になった後、プライマリサーバ30a及びセカンダリサーバ30b上のバックアップソフト104が停止される(S80')。

30

【0203】

続いて、移行元ストレージ装置10a用に起動していた、ストレージマネージャ101のインスタンスを停止し、ストレージマネージャ101の設定ファイルの書き換えを行う。設定ファイルの書き換え内容は、実施例1と同様で、仮想ストレージに対する制御を行うように、設定ファイルの内容を変更すればよい。設定ファイルの書き換え終了後、プライマリサーバ30a及びセカンダリサーバ30bで、該設定ファイルを用いたインスタンス起動が行われる(S90')。

【0204】

続いてS100'で、セカンダリサーバ30bは移行元ストレージ装置10aの移行元ボリュームへのパスの無効化、移行元セカンダリストレージ装置10bの移行元ボリュームのLUパス削除を行う。また、セカンダリサーバ30bは移行先ストレージ装置20aの移行先S-VOLへのLUパス設定を行い、セカンダリサーバ30bに移行先S-VOLを認識させる。その後、S80'で停止していたバックアップソフト104の動作を再開させる。

40

【0205】

続いてS110'では、移行先P-VOLをターゲットボリュームに移行する作業が行われる。移行の際、S90'で起動したストレージマネージャ101のインスタンスとは別のインスタンス(第2インスタンスと呼ぶ)がプライマリサーバ30aで起動され、移行先ストレージ装置20aに対するボリュームの移行の指示は第2インスタンスから行わ

50

れる。指示を受け付けた移行先ストレージ装置 20 a は、ボリュームマイグレーション機能を用いて、移行先 P - VOL をターゲットボリュームに移行する。

【 0 2 0 6 】

移行が完了すると、管理者は移行元ストレージシステムを撤去し (S 1 2 0 ')、移行処理を終了する。実施例 2 では、プライマリサーバ 30 a からのアクセス受け付けを停止することなく、ストレージ装置内の論理ボリュームのペアを維持したまま、P - VOL と S - VOL とを移行元ストレージ装置 10 から移行先ストレージ装置 20 に移行することができる。

【実施例 3】

【 0 2 0 7 】

続いて、本発明の実施例 3 に係る計算機システムについて説明する。図 25 は、本発明の実施例 3 に係る計算機システムの構成図である。実施例 1 に係る計算機システムとの相違は、プライマリサーバ 30 a に、移行処理を実施するプログラムである移行マネージャ 105 が、そしてセカンダリサーバ 30 b には、移行マネージャ 105 から呼び出されることで移行先ストレージシステムやセカンダリサーバ 30 b の設定処理を行うプログラムである移行サブマネージャ 106 が存在する点である。その他の点は、実施例 1 に係る計算機システムと同様である。なお、図示を省略しているが、プライマリサーバ 30 a とセカンダリサーバ 30 b、そして管理端末 (16 a、16 b、26 a、26 b) とは、LAN または WAN で接続され、相互に通信可能になっている。

【 0 2 0 8 】

実施例 3 に係る計算機システムでは、移行マネージャ 105 が移行処理を一括実行する。そのため、移行マネージャ 105 は、ストレージマネージャ 101、交替パスソフト 102、クラスタソフトウェア 103 に対してコマンドを発行し、所定の処理を行わせることができる。また管理端末 16 a、26 a に対してコマンドを発行して、コマンドデバイス作成等の処理を行わせることができる。また、ストレージマネージャ 101 が使用する設定ファイルの書き換えを行う機能も有する。

【 0 2 0 9 】

移行サブマネージャ 106 も移行マネージャ 105 と同様に、ストレージマネージャ 101、交替パスソフト 102、クラスタソフトウェア 103 に対してコマンドを発行し、所定の処理を行わせることができる。また、ストレージマネージャ 101 が使用する設定ファイルの書き換えを行う機能も有する。ただし移行サブマネージャ 106 は、移行マネージャ 105 からの指示に従って、ストレージマネージャ 101、交替パスソフト 102、クラスタソフトウェア 103 に対してコマンドを発行する、あるいはストレージマネージャ 101 が使用する設定ファイルの書き換えを行う。

【 0 2 1 0 】

[移行処理の流れ]

続いて、本発明の実施例 3 に係る計算機システムにおける、ボリューム移行処理の流れを、図 26 のフローチャートを用いて説明する。なお、以下では説明の簡単化のため、移行元ストレージシステムの 1 ボリュームペア (P - VOL、S - VOL のセット) のみの移行処理について説明するが、同時に複数のボリュームペアの移行を行うことも可能である。

【 0 2 1 1 】

S 10、S 20 は、実施例 1 の移行処理と同じである。ただし、管理者は S 10 において、移行元ストレージシステムでボリュームペアとして運用されていたボリューム (移行元 P - VOL、移行元 S - VOL) の情報を調査する。そしてこのボリュームペアの調査結果と S 10 での移行元ストレージシステムの構成調査の結果に基づいて、管理者は移行マネージャ 105 に対して移行処理を指示する際に必要となる、以下の情報を準備する。

- (1) ストレージ装置 10 a、10 b、20 a、20 b の装置シリアル番号
- (2) ストレージ装置 10 a、10 b、20 a、20 b で未使用の L D E V #
- (3) ストレージ装置 20 a、20 b の、各パリティグループの残サイズ

10

20

30

40

50

- (4) 移行元 P - V O L の L D E V #、移行元 S - V O L の L D E V #
- (5) 移行元プライマリストレージ装置 1 0 a のポートのうち、プライマリサーバ 3 0 a、移行元セカンダリストレージ装置 1 0 b と接続されているポート名
- (6) 移行元セカンダリストレージ装置 1 0 b のポートのうち、セカンダリサーバ 3 0 b、移行元プライマリストレージ装置 1 0 a と接続されているポート名
- (7) 移行先プライマリストレージ装置 2 0 a と接続される、移行元プライマリストレージ装置 1 0 a のポート名 (仮にこのポート名を「ポート C」と呼ぶ)
- (8) ポート C に対して L U パス設定される論理ボリューム (移行元 P - V O L) の L D E V #

【 0 2 1 2 】

続いて管理者は移行マネージャ 1 0 5 に対し、データ移行の指示を発行する (S 2 0 ')。その際、以下の情報をパラメータとして指定する。

- (1) 移行元プライマリストレージ装置 1 0 a の装置シリアル番号
- (2) 移行元セカンダリストレージ装置 1 0 b の装置シリアル番号
- (3) 移行先プライマリストレージ装置 2 0 a の装置シリアル番号
- (4) 移行先セカンダリストレージ装置 2 0 b の装置シリアル番号
- (5) 移行元 P - V O L の L D E V #、移行元 S - V O L の L D E V #
- (6) 移行先プライマリストレージ装置 1 0 a 及び移行先セカンダリストレージ装置 2 0 b で使用可能な L D E V # (S 1 0 で調査する、未使用の L D E V # をもとに決定する)
- (7) 移行先プライマリストレージ装置 1 0 a 及び移行先セカンダリストレージ装置 2 0 b で使用可能なパリティグループのグループ名 (S 1 0 で調査する、パリティグループの残サイズをもとに、移行先 P - V O L (のマイグレーションターゲットボリューム)、移行先 S - V O L を作成可能なパリティグループを決定する)
- (8) 移行先プライマリストレージ装置 2 0 a と接続される、移行元プライマリストレージ装置 1 0 a のポート名 (仮にこのポート名を「ポート C」と呼ぶ)
- (9) プライマリサーバ 3 0 a と接続される移行先プライマリストレージ装置 2 0 a のポート名
- (10) セカンダリサーバ 3 0 b と接続される移行先セカンダリストレージ装置 2 0 b のポート名
- (11) 移行先セカンダリストレージ装置 2 0 b と接続される、移行先プライマリストレージ装置 2 0 a のポート名及び移行先セカンダリストレージ装置 2 0 b のポート名
- (12) プライマリサーバ 3 0 a 及びセカンダリサーバ 3 0 b で稼働しているインスタンスのインスタンス番号、及びインスタンスが読み込む設定ファイルのファイル名

【 0 2 1 3 】

移行マネージャ 1 0 5 は指示を受信すると、移行処理を開始する (S 3 0 ')。移行処理で行われる内容は、実施例 1 の S 3 0 以降の処理と同様である。以下、実施例 1 の移行処理と異なる点を中心に説明する。

【 0 2 1 4 】

S 3 0、S 4 0 の処理は、実施例 1 の S 3 0、S 4 0 と同様である。S 5 0 では、移行マネージャ 1 0 5 は移行サブマネージャ 1 0 6 に指示を出すことにより、移行サブマネージャ 1 0 6 に移行先セカンダリストレージ装置 2 0 b の設定処理を行わせる。移行サブマネージャ 1 0 6 が実施する処理の内容は、実施例 1 の S 5 0 と同様である。

【 0 2 1 5 】

S 5 0 の処理が終了すると、移行マネージャ 1 0 5 は S 6 0、S 7 0 の処理を行う。S 7 0 の処理が完了 (つまりペア作成完了) すると、移行マネージャ 1 0 5 は S 8 0 の処理を実施する。S 8 0 では、セカンダリサーバ 3 0 b 上のクラスタソフトウェア 1 0 3 等の停止を行う必要もあるため、移行マネージャ 1 0 5 はプライマリサーバ 3 0 a 上のクラス

10

20

30

40

50

タソフトウェアの動作を停止する処理に加えて、移行サブマネージャ106に対し、セカンダリサーバ30b上のクラスタソフトウェア103等の停止を指示する。

【0216】

S90では、移行マネージャ105はストレージマネージャ101の設定ファイルの書き換えを行う。セカンダリサーバ30bの設定ファイルの書き換えも必要であるので、移行マネージャ105は移行サブマネージャ106に、セカンダリサーバ30bの設定ファイルの書き換えを行わせる。続いてS100で、移行マネージャ105は移行サブマネージャ106に指示して、セカンダリサーバ30bから移行元セカンダリストレージ装置10bの移行元ボリュームへのパスの無効化、移行元セカンダリストレージ装置10bの移行元ボリュームのLUパス削除を行わせる。

10

【0217】

最後にS110で移行マネージャ105は、移行先P-VOLをターゲットボリュームに移行する処理を行う。移行が終了すると、移行マネージャ105は管理者に、ボリュームの移行処理が完了したことを通知する。管理者はこの通知を受信してから移行元ストレージシステムを撤去(S120)し、移行処理が完了する。

【0218】

なお、実施例3の移行処理は、上で説明した処理に限定されるものではなく、様々な変形が可能である。上で説明したすべての処理を移行マネージャ105あるいは移行サブマネージャ106で実行する以外に、一部の処理は管理者が手動で実施するようにしてもよい。たとえば設定ファイルの書き換えは管理者が実施するようにしてもよい。

20

【0219】

逆に移行先ストレージシステムで作成する論理ボリュームのLDEV#の選定、あるいは論理ボリュームの作成先となるパリティグループの選定を、移行マネージャ105が構成情報の調査結果に基づいて自動決定してもよい。

【0220】

また、実施例3では、実施例1に係る計算機システムと同様の構成下でのボリュームペア移行、つまりリモートコピー機能によるボリュームペアの移行について説明したが、実施例3で説明した移行方法を、実施例2に係る計算機システムでのボリュームペア移行(ローカルコピー機能によるボリュームペアの移行)に適用することもできる。

【0221】

以上が、本発明の実施例に係る移行処理の内容である。本発明の移行処理では、移行元ストレージ装置から移行先ストレージ装置にボリュームを移行しても、移行対象の論理ボリュームに付されている、ホスト計算機から認識される識別情報である、ボリューム番号や装置シリアル番号などの情報が変化しない。そのためホスト計算機からは透過的にボリューム移行が実行されるので、ホスト計算機ではI/O処理を停止する必要がない。

30

【0222】

また、リモートコピー機能などによるボリュームペアを形成しているボリューム群を移行する場合でも、移行中でもボリュームが二重化された状態が常時維持されるため、移行処理によってデータの可用性が低下しない。さらにボリューム番号等の識別情報が変化しないため、ホスト計算機上に存在するペアボリュームの制御用の設定ファイルをほとんど変更する必要がない。そのため、バックアップやディザスタリカバリ用のソフトウェア・サービスの停止時間を、極めて短時間に限定することができる。

40

【0223】

以上、本発明の実施形態に係るストレージシステムを説明してきたが、本発明は上で説明した実施形態に限定されるものではない。上で説明してきた移行処理は、ホスト計算機(プライマリサーバ)からの指示によって行われるものであったが、ストレージ装置が管理操作の命令を発行することにより、移行を制御する構成もあり得る。また、本発明の実施例に係る移行処理は、装置間のデータ移行、装置の置き換え用途にのみ用いられるものではなく、システムの拡張に用いることもできる。

【0224】

50

一例として、図27に示されているように、ストレージ装置10a、10a'から構成されているストレージシステム10'において、ストレージ装置10aの負荷が高くなった場合を想定する。その場合、ストレージシステム10'に新規ストレージ装置20a'を導入し、ストレージシステム10'内のボリュームペアを、新規ストレージ装置20a'に移行するという場合に、本発明の実施例に係る移行処理を適用することが可能である。またこの時、ストレージシステム10'内のストレージ装置10a'(または10a)内に、実施例1で説明したストレージマネージャ相当のプログラムであるストレージ制御プログラム150'を設け、ストレージ装置10a'が各装置の構成情報の取得、ボリュームの設定、ペア作成、ボリューム移行等の処理の流れを制御するようにしてもよい。

【符号の説明】

【0225】

10a: 移行元プライマリストレージ装置
 10b: 移行元セカンダリストレージ装置
 20a: 移行先プライマリストレージ装置
 20b: 移行先セカンダリストレージ装置
 30a: プライマリサーバ
 30b: セカンダリサーバ
 50: SAN
 60: SAN

【要約】

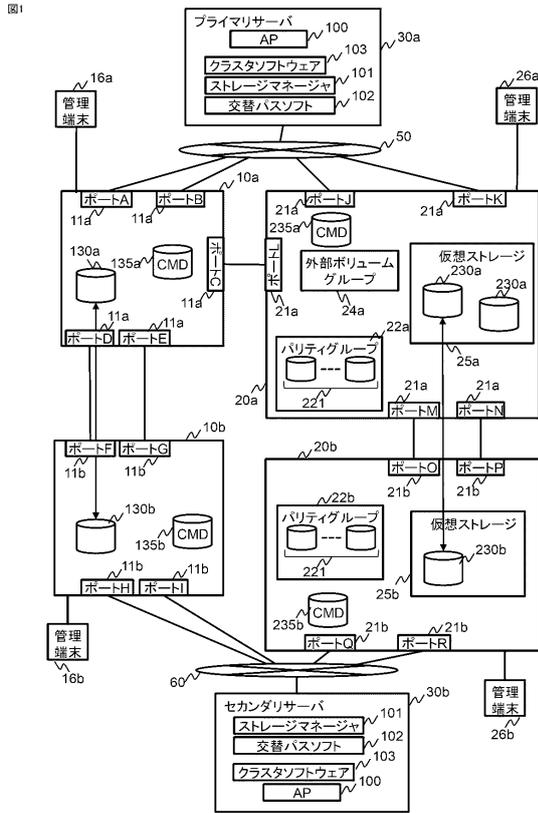
本発明のボリューム移行方法では、移行元ストレージシステム内でボリュームペアを形成している移行元P-VOL及び移行先S-VOLを、無停止で移行先ストレージシステムへと移行する。最初に移行先ストレージシステムに、移行元ストレージシステム内の移行元P-VOLの記憶領域をマッピングした論理ボリューム(移行先P-VOL)と、その複製先ボリュームとなる移行先S-VOLとを作成する。移行先ストレージシステムは、移行先P-VOLと移行先S-VOLに対して、移行元P-VOL及び移行元S-VOLの有する識別子と同一の識別子を付し、ホスト計算機に対し、移行先P-VOLと移行先S-VOLが移行元P-VOL及び移行元S-VOLと同一ボリュームであると認識させる。続いてホスト計算機から移行元P-VOLへのアクセスパスを、移行先P-VOLへと切り替える。その後、移行先P-VOLと移行先S-VOL間でボリュームペアの作成を行い、移行先S-VOLに移行先P-VOLのデータの複製が格納する。最後に移行先P-VOLの記憶領域を、移行先ストレージシステム内の記憶領域へと移動することで、移行処理を完了させる。

10

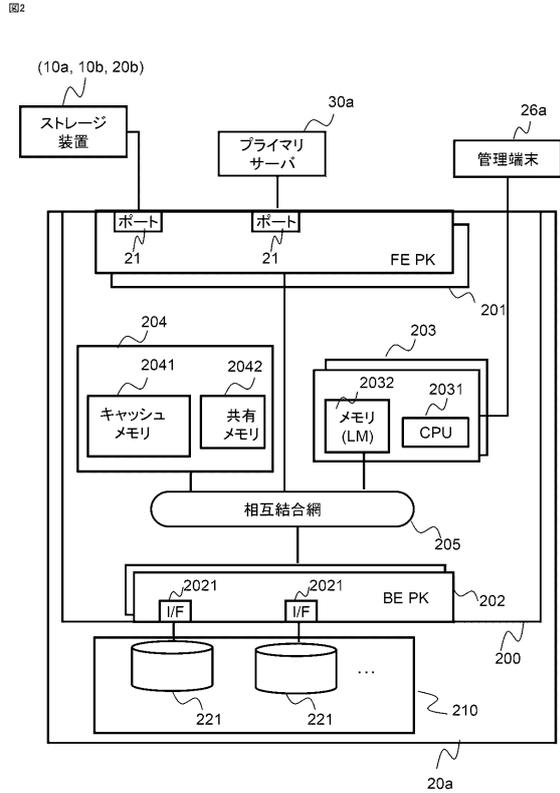
20

30

【図1】



【図2】



【図3】

図3

| パリティグループ名 | ドライブ名 | RAIDレベル | サイズ | 残サイズ |
|-----------|--------------------------------|---------|-------|-------|
| RG A | 0, 1, 2, 3 | RAID 5 | 3TB | 2TB |
| RG B | 10, 11, 12, 13 | RAID 5 | 1.5TB | 1.5TB |
| RG C | 20, 21, 22, 23, 24, 25, 26, 27 | RAID 5 | 6 TB | 6TB |

【図5】

図5

| LDEV# | サイズ | グループ名 | 先頭アドレス |
|-------|-------|-------|--------|
| 44 | 100GB | RG A | 0 |
| 33 | 100GB | EG1 | 0 |
| 55 | 300GB | RG C | 0 |

【図4】

図4

| グループ名 | WWN | LUN | サイズ |
|-------|-------------------|-----|-------|
| EG1 | 50060e8005111112b | 0 | 100GB |
| EG2 | ... | ... | ... |

【図6】

図6

| ポート名 | LUN | LDEV# |
|--------------------------|-----|-------|
| 3e243174 aaaaaaaa (ポートA) | 0 | 11 |
| | 1 | 12 |
| | 2 | 13 |
| 3e243174 bbbbbbbb | 0 | 11 |
| | : | : |
| | : | : |

【図7】

図7

| ベア番号 | 状態 | PDKC# | P-VOL# | SDKC# | S-VOL# |
|------|----|-------|--------|-------|--------|
| 1 | 1 | 1 | 1 | 11 | 0 |
| 2 | 2 | 1 | 2 | 11 | 1 |
| 3 | 2 | 1 | 3 | 11 | 3 |
| 4 | 1 | 1 | 0 | 11 | 2 |
| : | : | : | : | : | : |
| : | : | : | : | : | : |

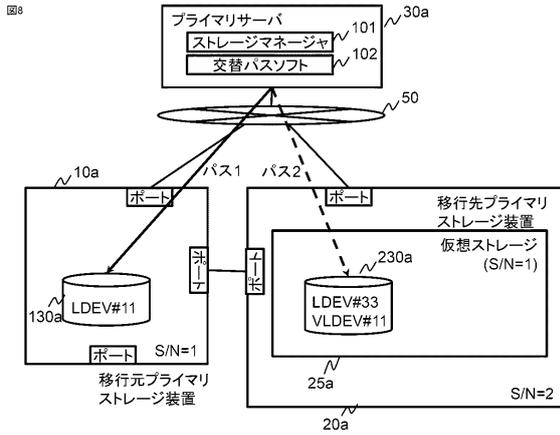
状態:
 1: copy pending
 2: pair
 3: suspended

【図9】

図9 V-BOX管理テーブル T1000

| ID | 機種名 | S/N | VLDEV# | LDEV# |
|-----|-----|-----|--------|-------|
| v0 | VSP | 2 | 0 | 0 |
| | | | 1 | 1 |
| | | | 2 | 2 |
| | | | ... | ... |
| v1 | USP | 1 | NULL | 44 |
| | | | 99 | 33 |
| | | | ... | ... |
| ... | ... | ... | ... | ... |

【図8】

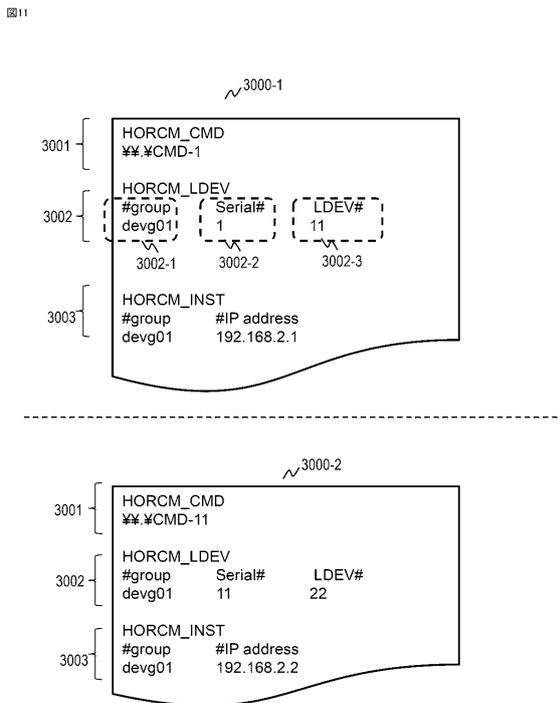


【図10】

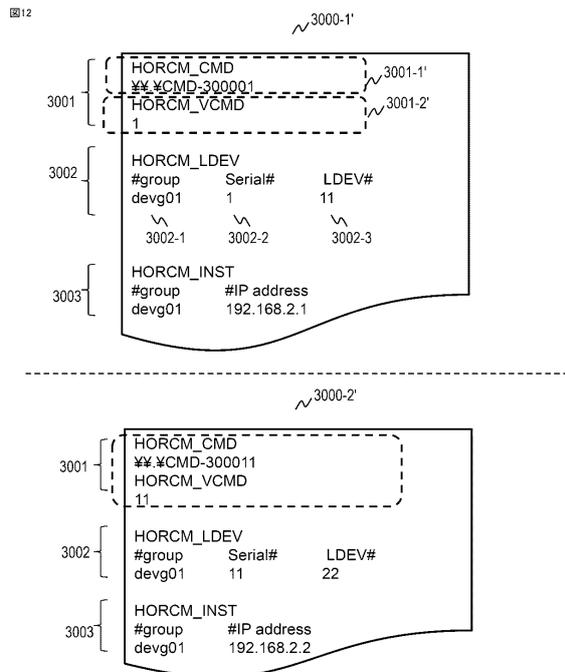
図10 仮想LDEV管理テーブル T1500

| LDEV# | VLDEV# | ID | 属性 |
|-------|--------|-----|----------|
| 0 | 0 | v0 | |
| 1 | 1 | v0 | |
| 2 | NULL | v0 | reserved |
| ... | ... | ... | ... |
| 44 | NULL | v1 | reserved |
| ... | ... | ... | ... |

【図11】



【図12】

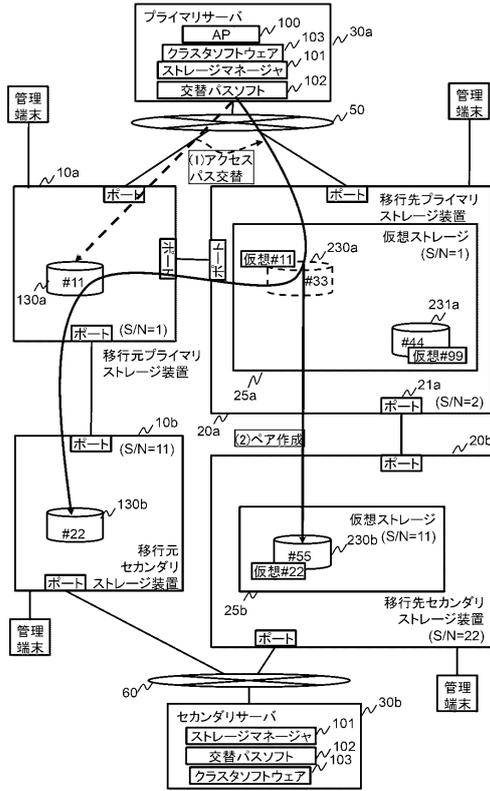


【図13】



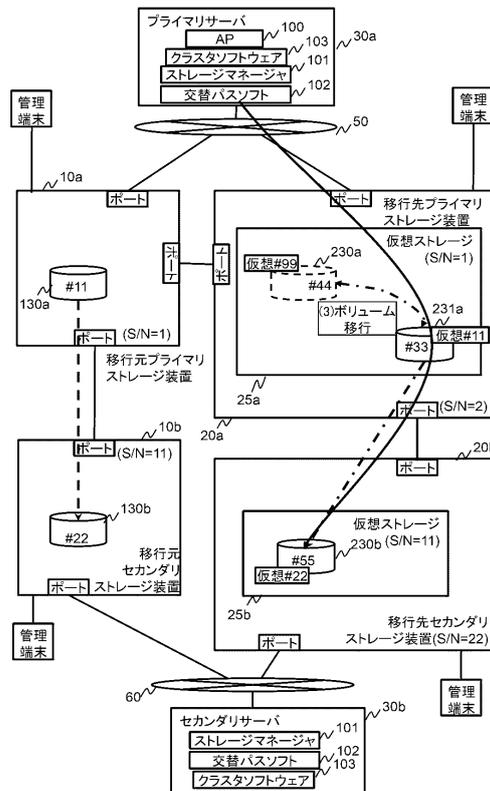
【図14】

図14



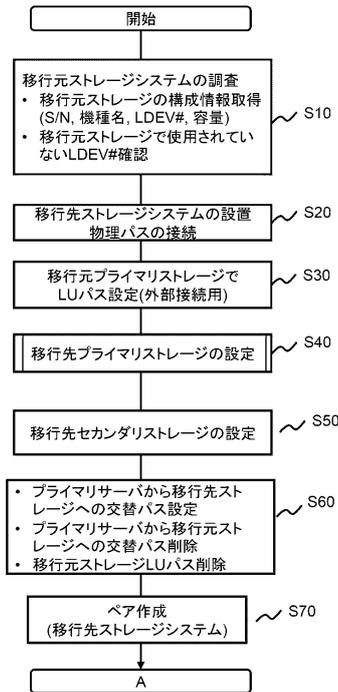
【図15】

図15



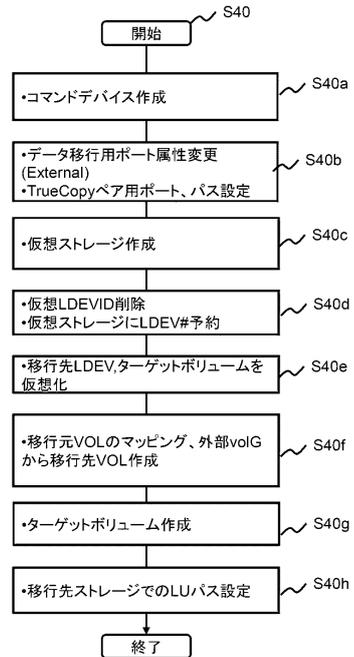
【図16】

図16



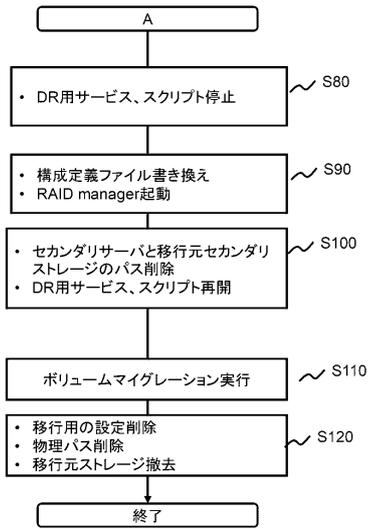
【図17】

図17



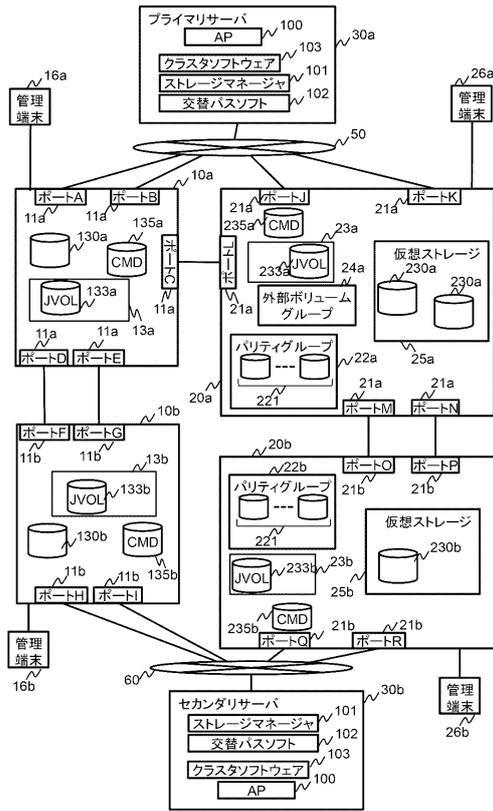
【図18】

図18



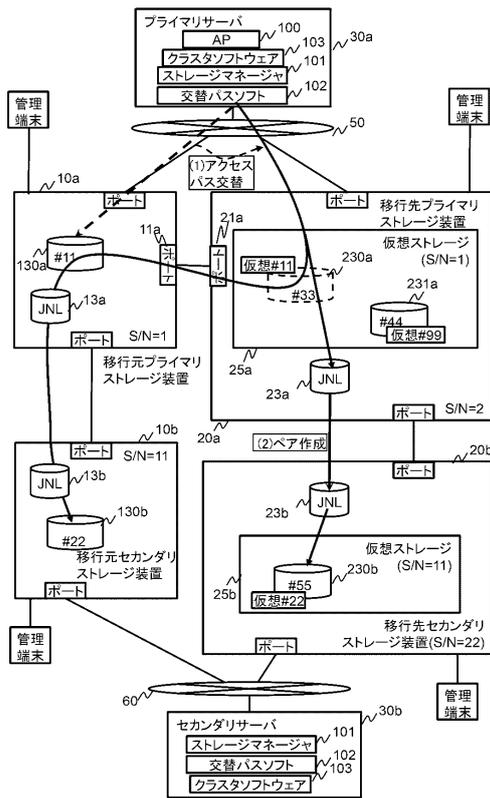
【図19】

図19



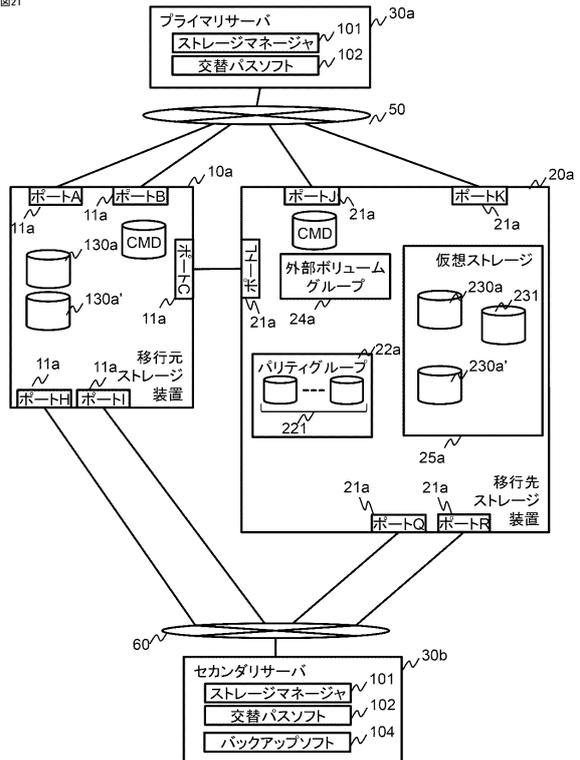
【図20】

図20



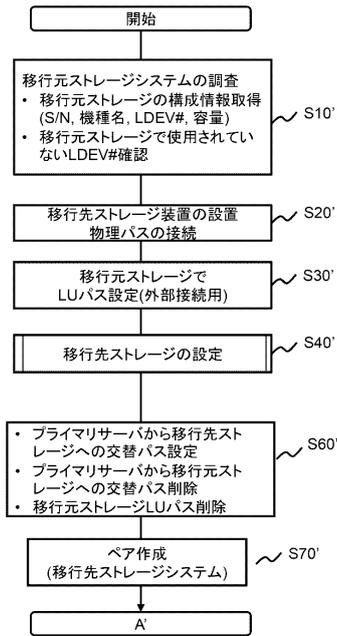
【図21】

図21



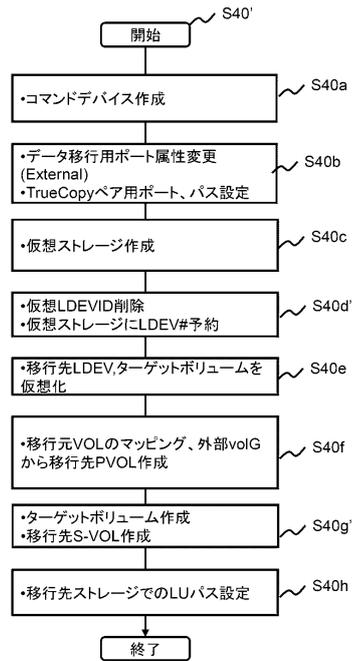
【 図 2 2 】

図22



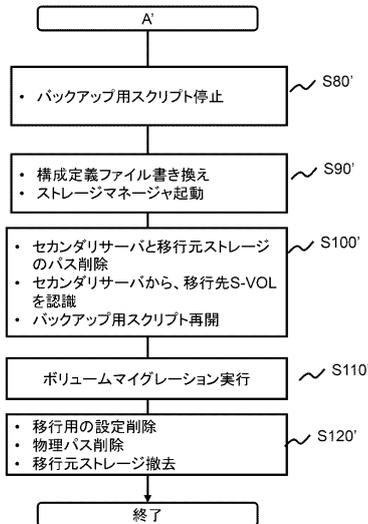
【 図 2 3 】

図23



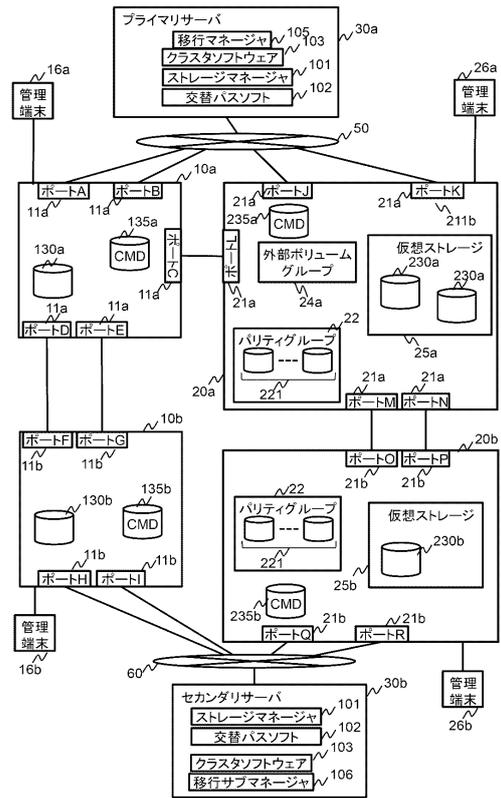
【 図 2 4 】

図24



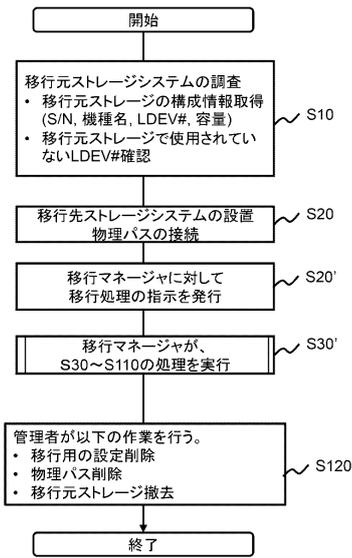
【 図 2 5 】

図25



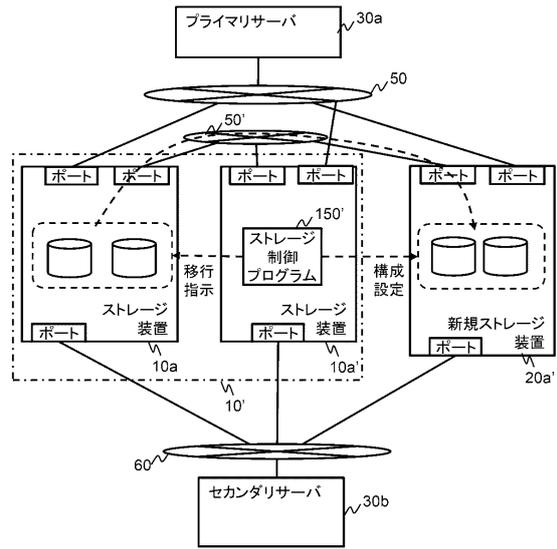
【図26】

図26



【図27】

図27



フロントページの続き

- (72)発明者 那須 弘志
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 杉野 昇史
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 田村 圭史
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 梨本 国彦
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 川口 智大
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 斎藤 秀雄
日本国東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内

審査官 木村 雅也

- (56)参考文献 特開2008-134986(JP,A)
特開2007-066154(JP,A)
特開2007-042008(JP,A)
特開2008-015984(JP,A)
特開2009-093316(JP,A)
特開2009-104421(JP,A)
特開2010-191958(JP,A)
特開2007-122531(JP,A)
特表2013-535037(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06
G06F 12/00
G06F 13/14