

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5512833号
(P5512833)

(45) 発行日 平成26年6月4日(2014.6.4)

(24) 登録日 平成26年4月4日(2014.4.4)

| | | | |
|-----------------------------|------|-------|------|
| (51) Int.Cl. | F I | | |
| G06F 3/06 (2006.01) | G06F | 3/06 | 301Z |
| G06F 12/00 (2006.01) | G06F | 3/06 | 301K |
| | G06F | 12/00 | 501M |
| | G06F | 12/00 | 537A |

請求項の数 13 (全 34 頁)

| | | | |
|---------------|-------------------------------|-----------|--|
| (21) 出願番号 | 特願2012-557097 (P2012-557097) | (73) 特許権者 | 000005108 |
| (86) (22) 出願日 | 平成22年12月22日 (2010.12.22) | | 株式会社日立製作所 |
| (65) 公表番号 | 特表2013-531283 (P2013-531283A) | | 東京都千代田区丸の内一丁目6番6号 |
| (43) 公表日 | 平成25年8月1日 (2013.8.1) | (74) 代理人 | 110000279 |
| (86) 国際出願番号 | PCT/JP2010/007457 | | 特許業務法人ウィルフォート国際特許事務所 |
| (87) 国際公開番号 | W02012/085975 | (72) 発明者 | 山本 彰 |
| (87) 国際公開日 | 平成24年6月28日 (2012.6.28) | | 東京都千代田区丸の内一丁目6番1号 株式会社日立製作所 研究開発本部内 |
| 審査請求日 | 平成24年12月20日 (2012.12.20) | (72) 発明者 | 森下 昇 |
| | | | 神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 横浜研究所内 |
| | | (72) 発明者 | 斎藤 秀雄 |
| | | | 神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 横浜研究所内 |

最終頁に続く

(54) 【発明の名称】 ストレージの仮想化機能と容量の仮想化機能との両方を有する複数のストレージ装置を含んだストレージシステム

(57) 【特許請求の範囲】

【請求項1】

外部のストレージ装置を共有する第1及び第2のストレージ装置を有し、
 前記第1のストレージ装置が、前記外部のストレージ装置が提供する論理ボリュームである外部ボリュームに基づく記憶領域でありそれぞれが複数のサブ記憶領域である外部ページに分割された外部容量プールと、複数の第1の仮想領域を含んだ仮想的な論理ボリュームである第1の仮想ボリュームとを有し、
 前記第2のストレージ装置が、複数の第2の仮想領域を含んだ仮想的な論理ボリュームである第2の仮想ボリュームを有し、
 前記第1及び第2のストレージ装置のうちのいずれか一方が、
ホストからライト要求を受け、
前記ライト要求で指定されているアドレスに属する前記仮想領域であるライト先仮想領域にいずれの前記外部ページが割り当てられておらず、且つ、その仮想領域にいずれか未割当ての前記外部ページを割り当てる場合、外部ページ割当て権限が、自分にあるか否かの第1の判断を行い、
前記第1の判断の結果が肯定的の場合、未割当ての前記外部ページを前記ライト先仮想領域に割り当て、
前記第1の判断の結果が否定的の場合、前記外部ページ割当て権限を有するストレージ装置に外部ページ割当て要求を送信し、
前記第1及び第2のストレージ装置のうち前記外部ページ割当て要求を受けたストレ

10

20

ジ装置は、未割当ての前記外部ページを前記ライト先仮想領域に割り当てる、ストレージシステム。

【請求項 2】

前記第 1 の仮想ボリュームから前記第 2 の仮想ボリュームへのデータ移行において、前記外部ページが割り当てられている第 1 の仮想領域から第 2 の仮想領域へのデータ移行は、下記により完了する、

前記第 1 のストレージ装置が、前記第 1 の仮想領域のアドレスである移行元アドレスと、前記第 1 の仮想領域に割り当てられている外部ページのアドレスである外部アドレスとを、前記第 2 のストレージ装置に通知し、且つ、

前記第 2 のストレージ装置が、前記通知された移行元アドレスと前記通知された外部アドレスとの対応関係を記憶する、
請求項 1 記載のストレージシステム。

10

【請求項 3】

前記第 1 及び第 2 のストレージ装置のうちのいずれか一方が、

前記第 1 及び第 2 の仮想ボリュームにおける仮想領域に割当て済の前記外部ページを未割当ての前記外部ページとすることである外部ページ解放を決定し、

外部ページ解放権限が、自分にあるか否かの第 2 の判断を行い、

前記第 2 の判断の結果が肯定的の場合、前記外部ページ解放を行い、

前記第 2 の判断の結果が否定的の場合、前記外部ページ解放権限を有するストレージ装置に外部ページ解放要求を送信し、

20

前記第 1 及び第 2 のストレージ装置のうち前記外部ページ解放要求を受けたストレージ装置は、前記外部ページ解放を行う、
請求項 1 記載のストレージシステム。

【請求項 4】

前記第 1 及び第 2 のストレージ装置のうちのいずれか一方が、

前記外部ページ割当て権限、又は、前記第 1 及び第 2 の仮想ボリュームにおける仮想領域に割当て済の前記外部ページを未割当ての前記外部ページとすることである外部ページ解放の権限を他のストレージ装置に移行するか否かの第 3 の判断を行い、

前記第 3 の判断の結果が肯定的の場合、前記外部ページ割当て権限又は前記外部ページ解放権限を前記他のストレージ装置に移行する、
請求項 1 記載のストレージシステム。

30

【請求項 5】

前記第 1 の仮想ボリュームから前記第 2 の仮想ボリュームへのデータ移行において、前記第 1 及び第 2 のストレージ装置のうちのいずれか一方が、

前記ホストからリード要求を受信し、

前記リード要求で指定されているアドレスに属する仮想領域に割り当てられている前記外部ページからデータを読み出し、読み出した前記データを前記ホストに送信する、
請求項 1 記載のストレージシステム。

【請求項 6】

前記第 1 及び第 2 のストレージ装置の各々が、そのストレージ装置が有する物理記憶デバイスに基づく記憶領域でありそれぞれが複数のサブ記憶領域である内部ページに分割された内部容量プールを有しており、

前記ライト要求は、前記第 1 の仮想ボリュームから前記第 2 の仮想ボリュームへのデータ移行において受けた要求であり、

前記第 1 及び第 2 のストレージ装置のうちのいずれか一方が、

前記ライト先仮想領域にいずれの前記外部ページも前記内部ページも割り当てられていない場合、前記外部ページを割り当てるか前記内部ページを割り当てるかの第 4 の判断を行い、

40

前記第 4 の判断の結果が前記外部ページを割り当てることである場合、前記第 1 の判断を行い、

50

前記第 4 の判断の結果が内部ページを割り当てることである場合、自分が有する前記内部容量プールから未割当ての前記内部ページを前記ライト先仮想領域に割り当てる、請求項 1 記載のストレージシステム。

【請求項 7】

前記第 1 のストレージ装置が、前記第 1 の仮想ボリュームから前記第 2 の仮想ボリュームへのデータ移行において、

前記ライト先仮想領域が前記第 1 の仮想ボリュームのうちのデータ移行が完了した仮想領域か否かの第 5 の判断を行い、

前記第 5 の判断の結果が肯定的であり、且つ、前記ライト先仮想領域に前記外部ページ又は前記内部ページが割り当てられている場合、前記ライト要求に従うライト対象データを、前記ライト先仮想領域に割り当てられている前記外部ページ又は前記内部ページに書込み、且つ、前記ライト対象データを、前記第 2 のストレージ装置に送信する、請求項 6 記載のストレージシステム。

10

【請求項 8】

前記第 1 の仮想ボリュームから前記第 2 の仮想ボリュームへのデータ移行において、前記第 2 のストレージ装置が、

前記ホストからリード要求を受信し、

前記リード要求で指定されたアドレスに従う仮想領域であるリード元仮想領域が、データ移行が完了した仮想領域であるか否かの第 6 の判断を行い、

前記第 6 の判断の結果が肯定的の場合、前記リード元仮想領域に割り当てられているページからデータを読み出し、読み出したデータを前記ホストに送信する、請求項 1 記載のストレージシステム。

20

【請求項 9】

前記第 1 のストレージ装置は、既存のストレージ装置であり、

前記第 2 のストレージ装置は、新たに用意されたストレージ装置である、請求項 1 記載のストレージシステム。

【請求項 10】

前記第 2 のストレージ装置は、前記第 2 の仮想ボリュームにおける第 2 の仮想領域のアドレスであるアクセス先アドレスを指定したアクセス要求を受信した場合、

前記アクセス先アドレスが、前記移行元アドレスに適合するアドレスであれば、前記第 2 のストレージ装置が、その移行元アドレスに対応する外部アドレスに従うアドレスを指定したアクセス要求を前記外部のストレージ装置に送信し、

30

前記アクセス先アドレスが、前記移行元アドレスに適合しないアドレスであれば、前記第 2 のストレージ装置が、前記外部ページの割当て要求を前記第 1 のストレージ装置に送信し、前記第 1 のストレージ装置が、その割当て要求に従い、前記外部容量プールにおける未割当ての前記外部ページを決定し、決定した前記外部ページのアドレスである外部ページアドレスを前記第 2 のストレージ装置に通知し、前記第 2 のストレージ装置が、前記アクセス先アドレスと通知された外部ページアドレスとの対応関係を記憶し、且つ、その外部ページアドレスに従うアドレスを指定したアクセス要求を前記外部のストレージ装置に送信する、

40

請求項 2 記載のストレージシステム。

【請求項 11】

前記第 2 の仮想ボリュームに割り当てられている前記外部ページにアクセスする権限は、前記第 2 のストレージ装置が有し、

前記第 2 の仮想ボリュームに前記外部ページを割り当てる権限は、前記第 1 のストレージ装置が有する、

請求項 1 記載のストレージシステム。

【請求項 12】

前記第 2 の仮想ボリュームに前記外部ページを割り当てる権限は、前記第 1 のストレージ装置の負荷が所定負荷以上に高い場合に、前記第 2 のストレージ装置に移行する、

50

請求項 1 1 記載のストレージシステム。

【請求項 1 3】

外部のストレージ装置を共有する第 1 及び第 2 のストレージ装置を有するシステムでの記憶制御方法であって、

前記第 1 のストレージ装置が、前記外部のストレージ装置が提供する論理ボリュームである外部ボリュームに基づく記憶領域でありそれぞれが複数のサブ記憶領域である外部ページに分割された外部容量プールと、複数の第 1 の仮想領域を含んだ仮想的な論理ボリュームである第 1 の仮想ボリュームとを有し、

前記第 2 のストレージ装置が、複数の第 2 の仮想領域を含んだ仮想的な論理ボリュームである第 2 の仮想ボリュームを有し、

前記記憶制御方法は、前記第 1 及び第 2 ストレージ装置のうちのいずれか一方により、
ホストからライト要求を受け、

前記ライト要求で指定されているアドレスに属する前記仮想領域であるライト先仮想領域にいずれの外部ページが割り当てられておらず、且つ、その仮想領域にいずれか未割当ての外部ページを割り当てる場合、外部ページ割当て権限が、自分にあるか否かの第 1 の判断を行い、

前記第 1 の判断の結果が肯定的の場合、未割当ての前記外部ページを前記ライト先仮想領域に割り当て、

前記第 1 の判断の結果が否定的の場合、前記外部ページ割当て権限を有するストレージ装置に外部ページ割当て要求を送信し、

前記第 1 及び第 2 のストレージ装置のうち前記外部ページ割当て要求を受けたストレージ装置が、未割当ての前記外部ページを前記ライト先仮想領域に割り当てる、
記憶制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージの仮想化機能と容量の仮想化機能との両方を有する複数のストレージ装置を含んだストレージシステムでの記憶制御に関する。

【背景技術】

【0002】

複数のストレージ装置を含んだストレージシステム（例えばデータセンタ）の運用コストを削減したいというニーズがある。この種のストレージシステムに関して、次のような技術が開示されている。

【0003】

特許文献 1 は、'ストレージの仮想化' と呼ばれる技術に関する。本明細書では、ストレージの仮想化機能を有するストレージ装置を「仮想ストレージ」と呼ぶ。また、本明細書では、仮想ストレージに接続されたストレージ装置であり、仮想ストレージによって仮想化される記憶資源を有するストレージ装置を「外部ストレージ」と呼ぶ。特許文献 1 では、1 台の仮想ストレージに、1 台以上の外部ストレージが接続され、見かけ上、外部ストレージは、仮想ストレージによって隠蔽される。本明細書では、サーバ等のホスト装置からのアクセス単位となる論理的なボリュームを「LU」（LU: Logical Unit）と呼ぶ。外部ストレージ上には、仮想ストレージ上の LU に 1 対 1 にマッピングされる LU が定義される。以降、サーバからアクセスされる LU を「サーバ LU」と呼び、外部ストレージ上の LU を「外部 LU」と呼ぶ。仮想ストレージは、外部 LU にマッピングされているサーバ LU を指定したアクセス要求をホスト装置から受信した場合、その外部 LU にアクセスすることで、そのアクセス要求を処理する。この技術によって、管理者は仮想ストレージの管理をすればよいので、管理コストを低減することができる。

【0004】

特許文献 2 は、'容量の仮想化' と呼ばれる技術に関する。容量の仮想化は、シン・プロビジョニング（或いは、ダイナミック・プロビジョニング）とも呼ばれる。本明細書

10

20

30

40

50

では、仮想ストレージは、容量の仮想化機能も有している。容量の仮想化機能によれば、「容量プール」と呼ばれる記憶領域が、「ページ」と呼ばれる区画に分割される。一般に、LUを定義する際、あわせて、そのLUの容量も指定し、ストレージ装置では、その容量に相当する記憶領域が確保される。一方、容量の仮想化機能をもつ仮想ストレージでは、仮想サーバLU（シン・プロビジョニングに従うサーバLU）の定義の際に、その仮想サーバLUの容量分の記憶領域を確保する必要は無い。実際に仮想サーバLUに対する書込みが発生した場合に、その書込み先を含んだ領域（仮想サーバLUにおける領域）に、ページが割り当てられる。これにより、消費される記憶領域を削減することができる。また、記憶領域を割り当てる契機が、書込みが発生した契機であるため、管理者は、仮想サーバLUの容量として適当な容量を定義すればよく、それ故、管理コストも低減することができる。

10

【0005】

特許文献3では、外部ストレージの記憶領域に、容量の仮想化機能を適用した技術が適用されている。この技術は、「ストレージの仮想化」と「容量の仮想化」の二つ技術が揃って始めて適用可能なため、付加価値の高い技術として重要であると考えられる。この場合、外部ストレージ上には、例えば、大容量の外部LUが定義される。この外部LUが、仮想ストレージ上では一つの容量プールとなる。仮想ストレージが、その容量プール（外部LUにマッピングされた容量プール）を複数のページに分割する。仮想サーバLUに書き込みが発生した場合、仮想ストレージが、その書込み先を含んだ領域にページを割り当てる。この場合、一つの仮想サーバLUに割り当てられている複数のページには、内部のページ（仮想ストレージ内部の物理記憶デバイスに基づく容量プール内のページ）と外部のページ（外部LUにマッピングされている容量プール内のページ）とが混在してもよい。

20

【0006】

特許文献1～特許文献3で開示されている技術では、仮想ストレージの台数が1台以上の場合に適用可能であった。しかし、さらなる運用管理コストの低減には、複数の仮想ストレージを連携させることも有効である。例えば、特許文献4では、2台の仮想ストレージを連携し、お互いをバックアップする技術が開示されている。また、特許文献4では、複数の仮想ストレージが外部ストレージを共有した環境で、負荷分散のため、外部ストレージ上の外部LUに対するアクセス権限を、当該LUのデータをコピーすることなく、ある仮想ストレージから別の仮想ストレージに移行する技術も開示されている。

30

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2004-005370号公報

【特許文献2】特開2003-015915号公報

【特許文献3】特開2007-102455号公報

【特許文献4】特開2007-072538号公報

【発明の概要】

【発明が解決しようとする課題】

40

【0008】

複数の仮想ストレージと1台以上の外部ストレージを含み、且つ、複数の仮想ストレージ（例えば、仮想ストレージAと仮想ストレージB）が外部ストレージを共有するストレージシステムが考えられる。そして、仮想ストレージA上に定義された外部LUに相当する容量プール内のページが割り当てられた仮想サーバLUのアクセス権限を、仮想ストレージAから仮想ストレージBに移すことが考えられる。

【0009】

本発明は、特許文献3で開示されたような技術（外部ストレージの記憶領域に、容量の仮想化機能を適用した技術）を、複数の仮想ストレージ間の連携技術に応用した場合に生ずる課題の解決を狙っており、重要度は高い。

50

【 0 0 1 0 】

仮想サーバLUのアクセス権限の移行の主な契機は、例えば、下記である。

(1) 仮想ストレージAと仮想ストレージBの間の負荷分散。

(2) 既存の古くなった仮想ストレージAから新しい仮想ストレージBへのデータ移行。

【 0 0 1 1 】

上記(2)で、仮想ストレージAを廃棄する場合、仮想ストレージAのすべての仮想サーバLUのアクセス権限が、仮想ストレージBを含む他の仮想ストレージに移されることになる。

【 0 0 1 2 】

外部ストレージが、仮想ストレージAと仮想ストレージB間で共有されている場合、仮想ストレージAがアクセス権限をもつ外部LUのアクセス権限は、仮想ストレージBが外部ストレージにアクセス可能であるため、特許文献4に開示された技術に従って、仮想ストレージAから仮想ストレージBに移行することが可能である。

10

【 0 0 1 3 】

しかし、仮想サーバLUのアクセス権限に加えて、容量プールに相当する外部LUのアクセス権限も、仮想ボリュームAから仮想ボリュームBに移行されると、その容量プールに関連付けられる複数の仮想サーバLUとして、仮想ストレージAがアクセス権限をもつ仮想サーバLUと仮想ストレージBがアクセス権限をもつ仮想サーバLUが混在してしまう可能性がある。従来は、このような課題の解決することは考慮されていない。

【課題を解決するための手段】

20

【 0 0 1 4 】

この課題を解決するために、本発明においては、外部LUに対応する容量プールの割当て・解放権限と、仮想サーバLU(以下、単に「LU」と言う)に対するアクセス権限が、独立に設けられる。

【 0 0 1 5 】

本発明では、仮想ストレージA及びBとそれら仮想ストレージA及びBが共有する外部ストレージ(外部ストレージが提供する外部LU)とを含んだストレージシステムでの、仮想ストレージAから仮想ストレージBへ、LU(LUのアクセス権限)が移行される際に、そのLUを構成する領域(以下、LU領域)毎に、下記(a)及び(b)の判断、(a)そのLU領域にページが割り当てられているか否か、(b)そのLU領域にページが割り当てられている場合、そのページは、仮想ストレージAが有する物理記憶デバイスに基づく容量プール内のページであるか、或いは、外部ストレージが有する外部LUに相当する容量プール内のページであるか、が行われ、その判断の結果に従う処理が実行される。具体的には、下記の通りである。以下、1つのLU領域を例に取り(以下、対象LU領域)、LU移行処理において行われる処理を説明する。

30

【 0 0 1 6 】

(1) 対象LU領域にページが割り当てられていないケース。

(1 a) 仮想ストレージAが、対象LU領域のアドレス及びその領域にページが割り当てられていないことを表す情報を仮想ストレージBに送信する。

40

(1 b) 仮想ストレージBが、仮想ストレージAから受信した情報(対象LU領域のアドレス及びその領域にページが割り当てられていないことを表す情報)を記憶する。

【 0 0 1 7 】

(2) 対象LU領域に、外部LUに相当する容量プール内のページ(以下、外部ページ)が割り当てられているケース。

(2 a) 仮想ストレージAが、対象LU領域のアドレスとその領域に割り当てられている外部ページのアドレスとを表す情報を仮想ストレージBに送信する。

(2 b) 仮想ストレージBが、仮想ストレージAから受信した情報(対象LU領域のアドレスとその領域に割り当てられている外部ページのアドレスとを表す情報)を記憶する。

【 0 0 1 8 】

50

これが、本発明の1つの特徴である。これにより、対象LU領域に割り当てられているページ内のデータをコピーする代わりに、対象LU領域に割り当てられている外部ページのアドレス（実質的に、外部LU内の領域のアドレス）を通知するだけでよい。このため、効率のよい、アクセス権限の移行が可能となる。なお、このケースでは、仮想ストレージBは、移行先のLUを指定したリード/ライト要求をホストから受けた場合、そのリード/ライト要求で指定されているアドレス（移行先LUにおけるLU領域のアドレス）が、仮想ストレージAから通知された対象LU領域アドレスに適合するアドレスであれば、仮想ストレージAから通知された外部ページアドレスに従う外部LU領域（外部LU内のLU領域）にアクセスする。

【0019】

(3) 対象LU領域に、仮想ストレージAが有する物理記憶デバイスに基づく容量プール内のページ（以下、内部ページ）が割り当てられているケース。

(3a) 仮想ストレージAが、対象LU領域のアドレスとその領域に割り当てられている内部ページのアドレスとを表す情報を仮想ストレージBに送信する。

(3b) 仮想ストレージAが、対象LU領域に割り当てられている内部ページ内のデータを仮想ストレージBに送信する。

(3c) 仮想ストレージBが、仮想ストレージBの物理記憶デバイスに基づく容量プール内のページを確保する。

(3d) 仮想ストレージBが、仮想ストレージAから受信したデータを、上記(3c)で確保したページに格納する。

【0020】

LUのアクセス権限を移行する契機は、上記(1) - (3)の処理の実行前と実行後のいずれでもよい。

【0021】

本発明では、原則として、LUのアクセス権限を仮想ストレージAから仮想ストレージBに移行しても、仮想ストレージAにおいてそのLUに関連付けられている容量プール（以下、容量プールA）の割当て・解放権限を、仮想ストレージAが保有していることになる。このため、仮想ストレージBで、仮想ストレージBに移行されたLU内の領域に、容量プールAのページの割当て・解放が必要になった場合、仮想ストレージBは、仮想ストレージAに、ページの割当て・解放処理を要求することになる。

【0022】

なお、例えば、以下の(X)又は(Y)のケースでは、容量プールAのページの割当て・解放権限も、仮想ストレージBに移行されても良い。

(X) 容量プールAの所定割合以上のページが割り当てられた1以上のLUのアクセス権限が仮想ストレージBに移行されたケース。

(Y) 仮想ストレージAの負荷（例えばCPU使用率）が所定負荷以上であるケース。例えば、仮想ストレージAにとって、容量プールAのページの割当て・解放の負荷が所定負荷以上であるケース。

【発明の効果】**【0023】**

本発明によれば、仮想ストレージA及びBとそれら仮想ストレージA及びBが共有する外部ストレージとを含んだストレージシステムにおいて、外部LUに相当する容量プール（仮想ストレージA内の容量プール）内のページが割り当てられたLUのアクセス権限が、仮想ストレージAから仮想ストレージBに移行される場合、そのLUに割り当てられている外部ページに関しては、その外部ページ内のデータをコピーすることに代えて、その外部ページのアドレスが仮想ストレージBに通知されるだけで、アクセス権限の移行が済む。これにより、高効率なアクセス権限の移行が可能となる。

【図面の簡単な説明】**【0024】**

【図1】実施例1に係るIT（InformationTechnology）システムの構成例を示す。

10

20

30

40

50

- 【図 2】実施例 1 に係る仮想ストレージの構成例を示す。
- 【図 3】実施例 1 に係る外部ストレージの構成例を示す。
- 【図 4】実施例 1 に係る内部容量プール情報のフォーマット例を示す。
- 【図 5】実施例 1 に係る外部 LU 情報のフォーマット例を示す。
- 【図 6】実施例 1 に係る外部容量プール情報のフォーマット例を示す。
- 【図 7】実施例 1 に係るホスト LU 情報のフォーマット例を示す。
- 【図 8】実施例 1 に係る外部 LU 情報のフォーマット例を示す。
- 【図 9】実施例 1 に係る仮想ストレージ 110 が発揮する複数の機能の一例を示す。
- 【図 10】実施例 1 に係る移行元の仮想ストレージにおけるデータ移行制御部 (a) の処理フローの一例を示す。 10
- 【図 11】実施例 1 に係る移行先の仮想ストレージにおけるデータ移行実行部 (a) の処理フローの一例を示す。
- 【図 12】実施例 1 に係る移行元の仮想ストレージにおける読み書き実行部 (a) の処理フローの一例を示す。
- 【図 13】実施例 1 に係るページ割り当て処理部の処理フローの一例を示す。
- 【図 14】実施例 1 に係る移行先の仮想ストレージにおけるライトデータ受領部の処理フローの一例を示す。
- 【図 15】実施例 1 に係るライトアフタ処理部の処理フローの一例を示す。
- 【図 16】実施例 1 に係る外部ページ割り当て処理部の処理フローの一例を示す。
- 【図 17】実施例 1 に係る外部ページ解放処理部の処理フローの一例を示す。 20
- 【図 18】実施例 1 に係る制御権移行部の処理フローの一例を示す。
- 【図 19】実施例 2 に係る仮想ストレージ 110 が発揮する複数の機能の一例を示す。
- 【図 20】実施例 2 に係る移行先の仮想ストレージにおけるデータ移行制御部 (b) の処理フローの一例を示す。
- 【図 21】実施例 2 に係る移行元の仮想ストレージにおけるデータ移行実行部 (b) の処理フローの一例を示す。
- 【図 22】実施例 2 に係る移行先の仮想ストレージにおける読み書き実行部 (b) の処理フローの一例を示す。
- 【図 23】実施例 2 に係る移行元の仮想ストレージにおけるリード要求受領部の処理フローの一例を示す。 30
- 【図 24】実施例 1 に係る制御権移行決定部の処理フローの一例を示す。
- 【発明を実施するための形態】
- 【0025】
- 以下、図面を参照して、本発明の幾つかの実施例を説明する。
- 【実施例 1】
- 【0026】
- 図 1 は、本発明の実施例 1 に係る IT (Information Technology) システムの構成例を示す。
- 【0027】
- ホスト 100 は、アプリケーションが実行される計算機であり、1 つ以上のホスト 100 が存在する。 40
- 【0028】
- 仮想ストレージ 110 は、ストレージの仮想化機能と容量の仮想化機能との両方を有するストレージ装置である。仮想ストレージ 110 は、ホスト 100 に接続され、ホスト 100 からリード/ライト要求 (リード要求又はライト要求) を受け付ける。また、仮想ストレージ 110 は、2 つ以上存在する。IT システムが、ストレージの仮想化機能及び容量の仮想化機能の少なくとも 1 つを有しないストレージ装置を含んでいても良い。例えば、そのようなストレージ装置を、複数の仮想ストレージ 110 に共有される外部ストレージ 130 として採用することが可能である。
- 【0029】 50

図1によれば、仮想ストレージ110は、SAN(a)120経由で、ホスト100に接続される。同様に、仮想ストレージ110同士も、SAN(a)120経由で接続されている。ただし、仮想ストレージ110はホスト100と直接接続されていても良いし(SAN(a)を介さずに接続されていてもよい)、仮想ストレージ110同士も、直接接続されていても良い(SAN(a)を介さずに接続されていてもよい)。また、1台の仮想ストレージ110は、少なくとも一つのホスト100と、少なくとも一つの仮想ストレージ110と接続されている。ホスト100、仮想ストレージ110、SAN(a)120それぞれの接続口を「ポート」140と呼ぶ。ホスト100、仮想ストレージ110は、少なくとも一つのポート140を持ち、SAN(a)120は、少なくとも二つのポート140をもつ。ホスト100は、仮想ストレージ110に発行するリード/ライト要求において、LU識別子(例えばLUN(Logical Unit Number))と、LU内の相対アドレス(例えばLBA(Logical Block Address))と、データ長を指定する。ホスト100が指定するLUを、以下、「ホストLU」と呼ぶ。

10

【0030】

外部ストレージ130は、SAN(b)150経由で、仮想ストレージ110に接続される。外部ストレージ130は、ホスト100からのリード/ライト要求は、直接受け付けず、仮想ストレージ110経由でこの要求を受け付けるのが基本である。例外があってもよいが、そのケースについては、後述する。外部ストレージ130は、少なくとも2台の仮想ストレージ110に接続される。外部ストレージ130上に定義されたLUを、「外部LU」と呼ぶ。外部LUは、基本的には、仮想ストレージ110からアクセスされる。

20

【0031】

管理サーバ160は、ホスト100と仮想ストレージ110を管理するサーバであり、ホスト100と仮想ストレージ110に接続されている。実施例1及び2では、管理サーバ160が、仮想ストレージ110のリード/ライト要求の頻度などの情報を収集し、仮想ストレージ110間で、ホストLUの移行することを決定する。また、実施例3及び4では、管理サーバ160が、新しい仮想ストレージ110をSAN(a)120に接続し、古い仮想ストレージ110のすべてのデータを新しい仮想ストレージ110に移し、古い仮想ストレージ110を廃棄するような制御を統括する。なお、いずれの実施例においても、管理サーバ160の機能が、管理サーバ160以外の装置(例えば、ホスト100、あるいは、仮想ストレージ110(例えばホストLUの移行元となる仮想ストレージ110、あるいは、移行先の仮想ストレージ110))が有していても良い。

30

【0032】

図2は、仮想ストレージ110の構成例を示す。

【0033】

仮想ストレージ110は、複数の物理記憶デバイス(例えばHDD(Hard Disk Drive)又はSSD(Solid State Drive))と、複数の物理記憶デバイスに接続されたコントローラとを有する。

【0034】

複数の物理記憶デバイスを基に、複数の論理ボリューム210が構築される。以下、仮想ストレージ110内の物理記憶デバイスに基づく論理ボリューム210を、「内部ボリューム210」と言う。内部ボリューム210は、2以上の物理記憶デバイスで構成されたRAIDグループ(Redundant Array of Independent Disks)に基づいていても良い。RAID構成の管理は、ストレージプロセッサ200が行う。

40

【0035】

コントローラは、仮想ストレージ110の動作を制御する装置である。コントローラは、複数のポート140と、キャッシュメモリ270を含んだ記憶資源と、それらの要素に接続されたストレージプロセッサ(例えばCPU(Central Processing Unit))200とを有する。記憶資源が、内部容量プール情報230、外部LU情報240、外部容量プール情報250及びホストLU情報260を記憶する。また、記憶資源は、ストレージプ

50

ロセッサ 200 が実行する 1 以上のコンピュータプログラムを記憶する。

【0036】

1 以上のストレージプロセッサ 200 は、仮想ストレージ 110 内に含まれ、これらのプロセッサ 200 は、ホスト 100 から受け付けたリード/ライト要求を処理する。

【0037】

内部容量プール情報 230 は、内部容量プールに関する情報である。「内部容量プール」とは、仮想ストレージ 110 内の物理記憶デバイスに基づく容量プールであり、1 以上の内部ボリューム 210 で構成されている。内部容量プールは、複数のページ（以下、内部ページ）に分割されている。

【0038】

外部 LU 情報 240 は、仮想ストレージ 110 に接続された外部ストレージ 130 に定義されている外部 LU に関する情報である。仮想ストレージ 110 に接続されたすべての外部ストレージ 130 上のすべての外部 LU について外部 LU 情報 240 が存在しても良いし、一部の外部 LU についてのみ外部 LU 情報 240 が存在しても良い。

【0039】

外部容量プール情報 250 は、外部容量プールに関する情報である。「外部容量プール」とは、1 以上の外部 LU がマッピングされた仮想的な論理ボリューム（ストレージ仮想化機能に従う論理ボリューム）で構成されている。外部容量プールは、複数のページ（以下、外部ページ）に分割されている。外部ページを書込み先としたデータは、実際には、その外部ページの基になっている外部 LU 内の LU 領域に書き込まれる。

【0040】

ホスト LU 情報 260 は、ホスト LU に関する情報である。

【0041】

キャッシュメモリ 270 は、ホスト 100 がホスト LU に対して読み書きするデータを格納するメモリである。メモリ 270 は、通常は、半導体メモリであるが、バッテリーなどで不揮発化されていてもよいし、高信頼化のため、2重化されていてもよい。

【0042】

図 3 は、外部ストレージ 130 の構成例を示す。

【0043】

外部ストレージ 130 は、複数の物理記憶デバイス（例えば HDD（Hard Disk Drive）又は SSD（Solid State Drive））と、複数の物理記憶デバイスに接続されたコントローラとを有する。

【0044】

複数の物理記憶デバイスを基に、複数の論理ボリューム 310 が構築される。以下、外部ストレージ 130 内の物理記憶デバイスに基づく論理ボリューム 310 を、「外部ボリューム 310」と言う。外部ボリューム 310 は、2 以上の物理記憶デバイスで構成された RAID グループに基づいていても良い。RAID 構成の管理は、外部ストレージプロセッサ 300 が行う。1 以上の外部ボリューム 310 で容量プールが構成され、その容量プールに、シン・プロビジョニングに従う仮想的な LU が設けられていても良い。外部ボリューム 310、或いは、1 以上の外部ボリューム 310 で構成された容量プールに関連付けられた仮想的な LU が、外部 LU であって良い。

【0045】

コントローラは、外部ストレージ 130 の動作を制御する装置である。コントローラは、複数のポート 140 と、外部キャッシュメモリ 330 を含んだ記憶資源と、それらの要素に接続された外部ストレージプロセッサ（例えば CPU）300 とを有する。記憶資源が、外部 LU 情報 320 を記憶する。また、記憶資源は、外部ストレージプロセッサ 300 が実行する 1 以上のコンピュータプログラムを記憶する。

【0046】

1 以上の外部ストレージプロセッサ 300 は、外部ストレージ 130 内に含まれ、これらのプロセッサ 300 は、ホスト 100 から受け付けたリード/ライト要求を処理する。

10

20

30

40

50

【 0 0 4 7 】

外部 L U 情報 3 2 0 は、外部 L U に関する情報である。

【 0 0 4 8 】

外部ストレージキャッシュメモリ 3 3 0 は、仮想ストレージが外部 L U に対して読み書きするデータを格納するメモリである。メモリ 3 3 0 は、通常は、半導体メモリであるが、バッテリーなどで不揮発化されていてもよいし、高信頼化のため、2重化されていてもよい。

【 0 0 4 9 】

以下、各種情報を説明する。その際、内部容量プールと外部容量プールを特に区別しない場合には、単に「容量プール」と言い、内部ページと外部ページを特に区別しない場合には、単に「ページ」と言うことがある。

10

【 0 0 5 0 】

図 4 は、内部容量プール情報 2 3 0 のフォーマット例を示す。

【 0 0 5 1 】

内部容量プール情報 2 3 0 は、内部容量プール毎に存在する。内部容量プール情報 2 3 0 は、ページサイズ 4 0 0、ページ数 4 1 0、割り当てページ数 4 2 0 及び L U アドレス 4 3 0 を含む。L U アドレス 4 3 0 は、内部ページ毎に存在する。以下、1つの内部容量プール（図 4 の説明において「対象プール」と言う）を例に取り、この情報 2 3 0 内の情報要素を説明する。

【 0 0 5 2 】

ページサイズ 4 0 0 は、内部ページのサイズを表す情報である。

20

【 0 0 5 3 】

ページ数 4 1 0 は、対象プールが有する内部ページの数（対象プールの容量をページサイズ 4 0 0 で割った値）を表す。

【 0 0 5 4 】

割り当てページ数 4 2 0 は、対象プールの中で、割り当て済みの内部ページの数を表す。

【 0 0 5 5 】

L U アドレス 4 3 0 は、内部ページの割り当て先の領域を表す情報（例えば、ホスト L U の L U N と、そのホスト L U の相対アドレス（例えば L B A ））である。ホスト L U に割り当てていない内部ページに対応する L U アドレス 4 3 0 として、ヌル値が格納されているとする。

30

【 0 0 5 6 】

図 5 は、外部 L U 情報 2 4 0 のフォーマット例を示す。

【 0 0 5 7 】

外部 L U 情報 2 4 0 は、外部 L U 毎に存在する情報である。外部 L U 情報 2 4 0 は、外部 L U 識別子 5 0 0、外部 L U 容量 5 1 0、容量仮想化フラグ 5 2 0、割り当て情報 5 3 0、制御権情報 5 4 0 及び制御権仮想ストレージ 5 5 0 を含む。以下、1つの外部 L U（図 5 の説明において「対象外部 L U」と言う）を例に取り、この情報 2 4 0 内の情報要素を説明する。

40

【 0 0 5 8 】

外部 L U 識別子 5 0 0 は、対象外部 L U の L U N と、対象外部 L U を有する外部ストレージ 1 3 0 の識別子とを含んだ情報である。

【 0 0 5 9 】

外部 L U 容量 5 1 0 は、対象外部 L U の容量を表す情報である。

【 0 0 6 0 】

容量仮想化フラグ 5 2 0 は、対象外部 L U に容量仮想化機能が適用されているかどうかを表す情報である。容量仮想化機能が適用されていない場合、基本的には、対象外部 L U と仮想ストレージ 1 1 0 内のホスト L U とが 1 対 1 に対応する。ただし、対象外部 L U にホスト L U がマッピングされていない場合もある。容量仮想化機能が適用されている場合、

50

対象外部LUは、外部容量プールに1対1に対応する。

【0061】

割り当て情報530は、対象外部LUに容量仮想化機能が適用されている場合、対象外部LUに対応する外部容量プールの識別子を表す。対象外部LUに容量仮想化機能が適用されていない場合、対象外部LUにマッピングされているホストLUの識別子が格納されている。対象外部LUにホストLUがマッピングされていない場合、割り当て情報530としてヌル値が格納される。

【0062】

制御権情報540は、対象外部LUに対応する外部容量プールの制御権（外部ページの割り当て・解放権限）をこの外部LU情報240を有する仮想ストレージ110が持っているか否かを表す。対象外部LUの制御権をこの仮想ストレージ110もっている場合、制御権情報540としてオンを表す値が格納される。具体的には、例えば、対象外部LUに容量仮想化機能が適用されている場合、この仮想ストレージ110が、対象外部LUに対応する外部容量プールの外部ページの割り当て・解放権限をもっている。また、対象外部LUに容量仮想化機能が適用されていない場合、対象外部LUにマッピングされているホストLUに対する読み書き処理の権限を、この仮想ストレージ110もっていることになる。

10

【0063】

制御権仮想ストレージ550は、対象外部LUの制御権をもっている仮想ストレージ110の識別子である。

20

【0064】

図6は、外部容量プール情報250のフォーマット例を示す。

【0065】

外部容量プール情報250は、外部容量プール毎に存在する。外部容量プール情報250は、外部LUポインタ600、ページサイズ400、ページ数410、割り当てページ数420及びLUアドレス430を含む。LUアドレス430は、外部ページ毎に存在する。以下、1つの外部容量プール（図6の説明において「対象プール」と言う）を例に取り、この情報250内の情報要素を説明する。

【0066】

外部LUポインタ600は、対象プールに対応する外部LUの外部LU情報240へのポインタである。

30

【0067】

それ以外の情報は、内部容量プール情報230に含まれている情報要素と実質的に同一である。すなわち、ページサイズ400は、外部ページのサイズを表す情報である。ページ数410は、対象プールが有する外部ページの数を表す。割り当てページ数420は、対象プールの中で、割り当て済みの外部ページの数を表す。LUアドレス430は、外部ページの割当て先の領域を表す情報（例えば、ホストLUのLUNと、そのホストLUの相対アドレス（例えばLBA））である。

【0068】

図7は、ホストLU情報260のフォーマット例を示す。なお、実施例1では、仮想ストレージ110内の全てのホストLUに容量仮想化機能が適用されているものとする。しかし、仮想ストレージ110において、一部のホストLUしか、容量仮想化機能が適用されていないとしても、本発明は有効である。

40

【0069】

ホストLU情報260は、ホストLU毎に存在する。ホストLU情報260は、ホストLU識別子700、仮想容量710、使用ページ数720、移行中フラグ730、移行中仮想ストレージ識別子740、移行中LU識別子750、コピーポインタ760、容量プール識別子770及び相対アドレス780を含む。容量プール識別子770及び相対アドレス780は、ホストLUを構成するLU領域毎に存在する。LU領域に1つの内部ページ又は外部ページが割り当てられる。従って、ホストLUを構成するLU領域の数は、実

50

施例 1 では、後述の仮想容量 7 0 0 をページサイズ 4 0 0 で割った値と同じである。なお、1 つの LU 領域に複数のページ（内部ページ及び / 又は外部ページ）が割り当てられても良い。以下、1 つのホスト LU（図 7 の説明において「対象ホスト LU」と言う）を例に取り、この情報 2 6 0 内の情報要素を説明する。

【 0 0 7 0 】

ホスト LU 識別子 7 0 0 は、対象ホスト LU の識別子（例えば LU N）である。

【 0 0 7 1 】

仮想容量 7 1 0 は、対象ホスト LU の仮想的な容量を表す情報である。

【 0 0 7 2 】

使用ページ数 7 2 0 は、対象ホスト LU に割り当てられているページの数を示す情報である。 10

【 0 0 7 3 】

移行中フラグ 7 3 0 は、対象ホスト LU を他の仮想ストレージに移行中であるか否かを示すフラグである。

【 0 0 7 4 】

移行中仮想ストレージ識別子 7 4 0 は、対象ホスト LU を移行する仮想ストレージ 1 1 0（移行元の仮想ストレージ 1 1 0）の識別子である。

【 0 0 7 5 】

移行中 LU 識別子 7 5 0 は、移行元の仮想ストレージ 1 1 0 内でのホスト LU の識別子（例えば LU N）である。 20

【 0 0 7 6 】

コピーポインタ 7 6 0 は、移行が完了した LU 領域のアドレスを示す。実施例 1 では、例えば、対象ホスト LU の先頭の LU 領域から末尾の LU 領域にかけて順次に移行処理が行われる。コピーポインタ 7 6 0 は、どの LU 領域まで移行処理が終了したかを示す。

【 0 0 7 7 】

容量プール識別子 7 7 0 は、LU 領域に割り当てられたページを含んだ容量プールの識別子である。

【 0 0 7 8 】

相対アドレス 7 8 0 は、LU 領域に割り当てられたページの容量プールでの相対アドレスを示す。ページが割り当てられていない場合、容量プール識別子 7 7 0 と相対アドレス 7 8 0 は、ともに、ヌル値である。 30

【 0 0 7 9 】

図 8 は、外部 LU 情報 3 2 0 のフォーマット例を示す。

【 0 0 8 0 】

外部 LU 情報 3 2 0 は、外部 LU 毎に存在する。外部 LU 情報 3 2 0 は、外部 LU 識別子 8 0 0 及び外部 LU 容量 8 1 0 を含む。以下、1 つの外部 LU（図 8 の説明において「対象外部 LU」と言う）を例に取り、この情報 3 2 0 内の情報要素を説明する。

【 0 0 8 1 】

外部 LU 識別子 8 0 0 は、外部ストレージ 1 3 0 での対象外部 LU の識別子（例えば LU N）である。 40

【 0 0 8 2 】

外部 LU 容量 8 1 0 は、対象外部 LU の容量を表す情報である。

【 0 0 8 3 】

以上が、各種情報についての説明である。なお、図 1 ~ 図 8 を参照して行った説明は、実施例 1 ~ 4 で共通である。

【 0 0 8 4 】

図 9 は、実施例 1 に係る仮想ストレージ 1 1 0 が発揮する複数の機能の一例を示す。

【 0 0 8 5 】

実施例 1 では、仮想ストレージ 1 1 0 においてストレージプロセッサ 2 0 0 が 1 以上のコンピュータプログラムを実行することにより、以下の複数の機能、すなわち、データ移 50

行制御部(a)900、データ移行実行部(a)910、読み書き実行部(a)920、ページ割り当て処理部930、ライトデータ受領部940、ライトアプタ処理部950、外部ページ割り当て処理部960、外部ページ解放処理部970、制御権移行部980及び制御権移行決定部990が発揮される。これらの機能によって実行される処理の少なくとも一部が、ハードウェア回路で実行されても良い。

【0086】

実施例1では、移行元の仮想ストレージ110から移行先の仮想ストレージ110にホストLUを移行するとき、移行処理の制御を移行元の仮想ストレージ110が実行する。また、移行処理中、ホスト100は、移行元の仮想ストレージ110にリード/ライト要求を発行することができる。

10

【0087】

本実施例では、管理サーバ160が、各仮想ストレージ110の負荷を調べ、移行すべき仮想ストレージ110のホストLUと移行先となる仮想ストレージ110を決定する。本実施例では、移行元となるホストLUの識別子は、移行先のホストLUに引き継がれないものとするが、本発明は、ファイバチャネルプロトコルで定義されたワールドワイドネームのような識別子を引き継ぐ場合も有効である。管理サーバ160は、移行元の仮想ストレージ110に、移行元のホストLUの識別子と、移行先の仮想ストレージ110の識別子と、移行先のホストLUの識別子とを送信する。また、管理サーバ160は、移行先の仮想ストレージ110に、移行元の仮想ストレージ110の識別子と、移行元のLUの識別子と、移行先のホストLUの識別子と、そのホストLUの仮想容量を表す情報とを送信する。

20

【0088】

実施例1では、移行元の仮想ストレージ110のデータ移行制御部(a)900が、管理サーバ160からLU移行要求を受け、その要求を受けたことを管理サーバ160に応答した後、図10に示す処理フローを実行する。

【0089】

図10は、移行元の仮想ストレージ110におけるデータ移行制御部(a)900の処理フローの一例を示す。

【0090】

ステップ1000では、データ移行制御部(a)900は、移行元のホストLUの識別子に対応するホストLU情報260を見つける。

30

【0091】

ステップ1001では、データ移行制御部(a)900は、ステップ1000で見つけたホストLU情報260(図10の説明において「対象情報260」)における移行中フラグ730の値を「オン」を表す値に変更する。さらに、データ移行制御部(a)900は、管理サーバ160から受け取った移行先の仮想ストレージ110の識別子と移行先のLUの識別子とを、対象情報260における移行中仮想ストレージ識別子740及び移行中ホストLU識別子750として設定する。さらに、データ移行制御部(a)900は、対象情報260におけるコピーポインタ760として、移行元のLUにおける先頭のLU領域を表す情報を設定する。

40

【0092】

ステップ1002では、データ移行制御部(a)900は、対象情報260におけるコピーポインタ750が示すLU領域(図10の説明において「対象移行元領域」)に対応する容量プール識別子770及び相対アドレス780を対象情報260から特定する。

【0093】

ステップ1003では、データ移行制御部(a)900は、ステップ1002で特定した識別子770及びアドレス780の値がヌルであるか否かを判断する。

【0094】

ステップ1003の判断の結果が肯定的の場合(ステップ1003:Y)、ステップ1004で、データ移行制御部(a)900は、対象移行元領域にページが割り当てられて

50

いないことを、移行先の仮想ストレージ 1 1 0 に通知する。この後、処理は、ステップ 1 0 1 1 へジャンプする。

【 0 0 9 5 】

ステップ 1 0 0 3 の判断の結果が否定的の場合（ステップ 1 0 0 3 : N）、ステップ 1 0 0 5 で、データ移行制御部 (a) 9 0 0 は、対象移行元領域に割り当てられているページが書き込み先であるが未だそのページに書き込まれていないデータがキャッシュメモリ 2 7 0 に有るか否かを判断する。

【 0 0 9 6 】

ステップ 1 0 0 5 の判断の結果が肯定的の場合（ステップ 1 0 0 5 : Y）、ステップ 1 0 0 6 で、データ移行制御部 (a) 9 0 0 は、そのデータを、対象移行元領域に割り当てられているページ（内部ページ又は外部ページ）に書き込む。外部ページが書き込み先の場合、データ移行制御部 (a) 9 0 0 は、その外部ページに対応した外部 L U 領域（外部ストレージ 1 3 0 内の外部 L U における L U 領域）に、データを書き込む。

10

【 0 0 9 7 】

ステップ 1 0 0 7 では、データ移行制御部 (a) 9 0 0 は、対象情報 2 6 0 における、対象移行元領域に対応した容量プール識別子 7 7 0 が、外部容量プールの識別子か否かを判断する。

【 0 0 9 8 】

ステップ 1 0 0 7 の判断の結果が肯定的の場合（ステップ 1 0 0 7 : Y）、ステップ 1 0 0 8 で、データ移行制御部 (a) 9 0 0 は、対象移行元領域に対応した容量プール識別子 7 7 0 （外部容量プールの識別子）に対応する外部容量プール情報 2 5 0 （図 1 0 において「対象外部容量プール情報 2 5 0 」）を特定する。データ移行制御部 (a) 9 0 0 は、対象外部容量プール情報 2 5 0 を基に、対象移行元領域に割り当てられている外部ページに対応する外部 L U 領域の相対アドレス（外部 L U での相対アドレス）と、その外部 L U 領域を含んだ外部 L U の識別子とを特定し、それらの情報を、移行先の仮想ストレージ 1 1 0 に送信する。この後、処理は、ステップ 1 0 1 1 へジャンプする。

20

【 0 0 9 9 】

ステップ 1 0 0 7 の判断の結果が否定的の場合（ステップ 1 0 0 7 : N）、データ移行制御部 (a) 9 0 0 は、ステップ 1 0 0 9 で、対象移行元領域に割り当てられている内部ページからデータを読み出し、そのデータを、キャッシュメモリ 2 7 0 に格納する。

30

【 0 1 0 0 】

ステップ 1 0 1 0 では、データ移行制御部 (a) 9 0 0 は、移行先の仮想ストレージ 1 1 0 に、対象移行元領域の相対アドレスと、対象移行元領域に移行先の仮想ストレージ 1 1 0 における内部ページを割り当てる必要があるということと、ステップ 1 0 0 9 でキャッシュメモリ 2 7 0 に格納したデータとを送信する。

【 0 1 0 1 】

ステップ 1 0 1 1 では、データ移行制御部 (a) 9 0 0 は、移行先の仮想ストレージ 1 1 0 から完了報告を待つ。

【 0 1 0 2 】

ステップ 1 0 1 2 では、データ移行制御部 (a) 9 0 0 は、移行元の L U の末尾の L U 領域について処理が完了したか否かを判断する。

40

【 0 1 0 3 】

ステップ 1 0 1 2 の判断の結果が否定的の場合（ステップ 1 0 1 2 : N）、ステップ 1 0 1 3 で、データ移行制御部 (a) 9 0 0 は、次の L U 領域を表す情報を、移行元ホスト L U に対応するホスト L U 情報 2 6 0 におけるコピーポインタ 7 5 0 として設定し、ステップ 1 0 0 2 へジャンプする。

【 0 1 0 4 】

ステップ 1 0 1 2 の判断の結果が肯定的の場合（ステップ 1 0 1 2 : Y）、ステップ 1 0 1 4 で、データ移行制御部 (a) 9 0 0 は、対象ホスト L U 情報 2 6 0 における移行中フラグ 7 3 0 の値を、オフを表す値に変更する。そして、データ移行制御部 (a) 9 0 0

50

は、移行処理が完了したことを管理サーバ160に報告し、次の指示をまつ。

【0105】

ステップ1015で、データ移行制御部(a)900は、管理サーバ160から、移行元のホストLUに関する情報を消去して良いという指示を受け、以下の処理を実行する。すなわち、データ移行制御部(a)900は、移行元のホストLUに割り当てられている全てのページを解放することにより、それぞれのページを空きページとする。また、この際、他の仮想ストレージ110が制御権をもっている外部容量プールの外部ページを解放する場合、データ移行制御部(a)900は、その外部容量プールに対応する外部LU情報230を特定する。データ移行制御部(a)900は、その情報230における制御権仮想ストレージ550から識別される他の仮想ストレージ110に、その外部ページの解放要求を送る。データ移行制御部(a)900は、その解放要求に回答した他の仮想ストレージ110からの要求により、外部ページを解放する。さらに、データ移行制御部(a)900は、移行元ホストLUに対応するホストLU情報260を消去する。この後、データ移行制御部(a)900は、この完了を、管理サーバ160に通知する。

10

【0106】

ステップ1016では、データ移行制御部(a)900は、移行元のホストLUに割り当てられていた外部ページが、どの外部容量プールに幾つ存在するかを特定する。

【0107】

ステップ1017では、データ移行制御部(a)900は、ステップ1016で特定された外部容量プール毎に、制御権をどの仮想ストレージ110がもっているかを判断する。

20

【0108】

ステップ1018では、データ移行制御部(a)900は、移行元のホストLUに割り当てられている外部ページを含んだ外部容量プールとして、移行元の仮想ストレージ110(自分)以外の他の仮想ストレージ110が制御権をもっている外部容量プールがあるか否かを判断する。外部容量プールについての制御権とは、外部ページの割り当て・解放権限である。

【0109】

ステップ1018の判断の結果が否定的の場合(ステップ1018:N)、処理は、ステップ1020へジャンプする。

30

【0110】

ステップ1018の判断の結果が肯定的の場合(ステップ1018:Y)、ステップ1019で、データ移行制御部(a)900は、外部容量プールの制御権をもつ他の仮想ストレージ110に、その外部容量プールについて、ステップ1016で特定された外部ページ数(移行元LUに割り当てられていた外部ページの数)を表す情報を送信する。データ移行制御部(a)900は、外部容量プールの制御権を移行するか否かの決定を、上記他の仮想ストレージ110から待つ。

【0111】

ステップ1020では、データ移行制御部(a)900は、移行元の仮想ストレージ110(自分)が制御権をもっている外部容量プールとして、移行元のホストLUに割り当てられている外部ページを含んだ外部容量プールがあるか否かを判断する。

40

【0112】

ステップ1020の判断の結果が否定的の場合(ステップ1020:N)、処理が終了する。

【0113】

ステップ1020の判断の結果が肯定的の場合(ステップ1020:Y)、ステップ1021で、データ移行制御部(a)900は、外部容量プールの制御権を移行するか否かを判断する。

【0114】

ステップ1021の判断の結果が否定的の場合(ステップ1021:N)、処理が終了

50

する。

【 0 1 1 5 】

ステップ 1 0 2 1 の判断の結果が肯定的の場合（ステップ 1 0 2 1 : Y）、ステップ 1 0 2 2 で、データ移行制御部（ a ） 9 0 0 は、どの仮想ストレージ 1 1 0 に外部容量プールの制御権を移行するかを決定する。ここでは、例えば、LU の移行先の仮想ストレージ 1 1 0 が決定される。その制御権の移行先の仮想ストレージ 1 1 0 を決定した場合、その制御権の移行先の仮想ストレージ 1 1 0 に、データ移行制御部（ a ） 9 0 0 は、その外部容量プールの制御権を移行することと、その外部容量プールに対応する外部容量プール情報 2 5 0 が表す全ての情報とを、移行先の仮想ストレージ 1 1 0 に送信する。また、それらの情報を送信した後、データ移行制御部（ a ） 9 0 0 は、その外部容量プール情報 2 5 0 内の情報を所定値にリセットし、且つ、その外部容量プールに対応する外部 LU 情報 2 4 0 における制御権仮想ストレージ 5 5 0 として移行先の仮想ストレージ 1 1 0 の識別子を設定する。

10

【 0 1 1 6 】

ステップ 1 0 2 3 では、データ移行制御部（ a ） 9 0 0 は、その外部容量プールに相当する外部 LU を有する外部ストレージ 1 3 0 に接続された他の全ての仮想ストレージ 1 1 0（移行元の仮想ストレージと制御権の移行先の仮想ストレージ以外の仮想ストレージ）に、制御権の移行先の仮想ストレージ 1 1 0 の識別子を通知する。さらに、データ移行制御部（ a ） 9 0 0 は、この外部 LU に対応する外部 LU 情報 2 4 0 における制御権情報 5 4 0 の値をオフを表す値に変更する。

20

【 0 1 1 7 】

以上で、データ移行制御部（ a ） 9 0 0 の処理が完了する。

【 0 1 1 8 】

実施例 1 では、移行先の仮想ストレージ 1 1 0 において、データ移行実行部（ a ） 9 1 0 が、管理サーバ 1 6 0 から、移行元の仮想ストレージ 1 1 0 の識別子と、移行元のホスト LU の識別子と、移行先のホスト LU の識別子と、そのホスト LU の仮想容量を表す情報とを受信し、図 1 1 に示す処理フローを実行する。

【 0 1 1 9 】

図 1 1 は、移行先の仮想ストレージ 1 1 0 におけるデータ移行実行部（ a ） 9 1 0 の処理フローの一例を示す。

30

【 0 1 2 0 】

ステップ 1 1 0 0 では、データ移行実行部（ a ） 9 1 0 は、受領したホスト LU の識別子に対応するホスト LU 情報 2 6 0 を生成する。その情報 2 6 0 において、ホスト LU 識別子 7 0 0 及び仮想容量 7 1 0 として、受領したホスト LU の識別子と仮想容量とが設定される。

【 0 1 2 1 】

ステップ 1 1 0 1 では、データ移行実行部（ a ） 9 1 0 は、移行元の仮想ストレージ 1 1 0 から情報を待つ。

【 0 1 2 2 】

ステップ 1 1 0 2 では、データ移行実行部（ a ） 9 1 0 は、移行元のホスト LU の LU 領域（対象移行元領域）に対応する、移行先のホスト LU における LU 領域（図 1 1 において「対象移行先領域」）に、ページが割り当てられていないという情報が送られてきたかを判断する。

40

【 0 1 2 3 】

ステップ 1 1 0 2 の判断の結果が肯定的の場合（ステップ 1 1 0 2 : Y）、データ移行実行部（ a ） 9 1 0 は、ステップ 1 1 0 3 で、対象移行先領域に対応する容量プール識別子 7 7 0 と相対アドレス 7 8 0 としてそれぞれヌル値を設定する。この後、処理は、ステップ 1 1 0 8 へジャンプする。

【 0 1 2 4 】

ステップ 1 1 0 4 では、データ移行実行部（ a ） 9 1 0 は、対象移行先領域が外部ペー

50

ジに割り当てられているという情報を受けたか否かを判断する。

【0125】

ステップ1104の判断の結果が肯定的の場合（ステップ1104：Y）、データ移行実行部（a）910は、ステップ1105では、受けた外部容量プール識別子及び相対アドレスを、対象移行先領域に対応する外部容量プール識別子770と相対アドレス780として設定する。そのあと、処理は、ステップ1108へジャンプする。

【0126】

ページを確保せよという要求を移行元の仮想ストレージ110から受けた場合、データ移行実行部（a）910は、ステップ1106で、移行先の仮想ストレージ110の内部容量プール情報230を基に、移行先の仮想ストレージ110における内部容量プールから空きの内部ページを特定し、特定した内部ページを対象移行先領域に割り当てる。具体的には、例えば、データ移行実行部（a）910は、対象移行先領域に対応した容量プール識別子770として、割り当てられた内部ページを含んだ内部容量プールの識別子を設定し、且つ、対象移行先領域に対応した相対アドレス780として、内部ページのアドレスを設定する。

10

【0127】

ステップ1107では、データ移行実行部（a）910は、移行元の仮想ストレージ110から受けたデータを、ステップ1106で割り当てた内部ページに書き込む。

【0128】

ステップ1108では、データ移行実行部（a）910は、対象移行先領域が移行先の

20

ホストLUにおける末尾のLU領域であるか否かを判断する（ステップ1108）。

【0129】

ステップ1108の判断の結果が否定的の場合（ステップ1108：N）、データ移行実行部（a）910は、ステップ1101へジャンプして、次の要求をまつ。

【0130】

ステップ1108の判断の結果が肯定的の場合（ステップ1108：Y）、ステップ1109で、データ移行実行部（a）910は、管理サーバ160へ移行が完了したことを報告する。

【0131】

以上で、データ移行実行部（a）910が処理を終了する。

30

【0132】

実施例1で、管理サーバ160は、移行元の仮想ストレージ110と移行先の仮想ストレージ110からホストLUの移行処理が完了した情報を受け付けると、ホスト100に、移行元のホストLUのLU識別子を指定したリード/ライト要求の送信先を移行先の仮想ストレージ110に切り替えるよう要求する。移行元のホストLUのLUNと移行先のホストLUのLUNが交換される或いは、ワールドワイドネームが移行元から移行先に送られ、SAN（a）120などで自動的に移行先の仮想ストレージ110にリード/ライト要求の送信先が切り替えられても良い。

【0133】

その後、管理サーバ160は、移行元の仮想ストレージ110に移行元のホストLUの情報を消去する要求を送信する。その要求に応答して、移行元の仮想ストレージ110が、移行元のホストLUに対応するホストLU情報260を所定値にリセットすることができる。

40

【0134】

図12は、移行元の仮想ストレージ110における読み書き実行部（a）920の処理フローの一例を示す。この処理フローは、移行元のホストLUの移行中に移行元のホストLUを指定したリード/ライト要求を移行元の仮想ストレージ110がホスト100から受けたときに実行される処理フローである。

【0135】

ステップ1200では、読み書き実行部（a）920は、受け取った要求が、リード要

50

求かライト要求かをチェックする。

【 0 1 3 6 】

要求がリード要求の場合、読み書き実行部 (a) 9 2 0 は、ステップ 1 2 0 1 で、リード対象データが、キャッシュメモリ 2 7 0 にあるかどうかを判断する。

【 0 1 3 7 】

ステップ 1 2 0 1 の判断の結果が肯定的の場合 (ステップ 1 2 0 1 : Y)、処理が、ステップ 1 2 0 3 へジャンプする。

【 0 1 3 8 】

ステップ 1 2 0 1 の判断の結果が否定的の場合 (ステップ 1 2 0 1 : N)、ステップ 1 2 0 2 で、読み書き実行部 (a) 9 2 0 は、リード元を含んだ L U 領域に割り当てられているページからリード対象データを読み出し、そのリード対象データを、キャッシュメモリ 2 7 0 に書き込む。

10

【 0 1 3 9 】

ステップ 1 2 0 3 では、読み書き実行部 (a) 9 2 0 は、キャッシュメモリ 2 7 0 からリード対象データを読み出し、そのデータをホスト 1 0 0 に転送する。以上で、処理が完了する。

【 0 1 4 0 】

要求がライト要求の場合、ステップ 1 2 0 4 では、読み書き実行部 (a) 9 2 0 は、ライト要求で指定されたライト先を含んだ L U 領域にページが割り当てられているかを判断する。

20

【 0 1 4 1 】

ステップ 1 2 0 4 の判断の結果が否定的の場合 (ステップ 1 2 0 4 : N)、読み書き実行部 (a) 9 2 0 は、ステップ 1 2 0 5 で、ページ割り当て処理部 9 3 0 をコールする。ステップ 1 2 0 6 で、読み書き実行部 (a) 9 2 0 は、ライト対象データをホスト 1 0 0 から受け、そのデータをキャッシュメモリ 2 7 0 に書き込む。

【 0 1 4 2 】

ステップ 1 2 0 7 で、読み書き実行部 (a) 9 2 0 は、ライト先のホスト L U に対応するホスト L U 情報 2 6 0 における移行中フラグ 7 3 0 がオンかを判断する。

【 0 1 4 3 】

ステップ 1 2 0 7 の判断の結果が否定的の場合 (ステップ 1 2 0 7 : N)、処理が完了する。

30

【 0 1 4 4 】

ステップ 1 2 0 7 の判断の結果が肯定的の場合 (ステップ 1 2 0 7 : Y)、ステップ 1 2 0 8 で、読み書き実行部 (a) 9 2 0 は、コピーポインタ 7 5 0 を基に、ライト先領域について移行処理を完了しているかをチェックする。コピーポインタ 7 5 0 が表す L U 領域の相対アドレスが、ライト先領域の相対アドレスよりも後のアドレスであれば、ライト先領域について移行処理が完了している。

【 0 1 4 5 】

ステップ 1 2 0 8 の判断の結果が否定的の場合 (ステップ 1 2 0 8 : N)、処理が完了する。

40

【 0 1 4 6 】

ステップ 1 2 0 8 の判断の結果が肯定的の場合 (ステップ 1 2 0 8 : Y)、ステップ 1 2 0 9 で、読み書き実行部 (a) 9 2 0 は、ライト先領域に割り当てられているページが、外部ページかを判断する。

【 0 1 4 7 】

ステップ 1 2 0 9 の判断の結果が否定的の場合 (ステップ 1 2 0 9 : N)、ステップ 1 2 1 0 で、読み書き実行部 (a) 9 2 0 は、キャッシュメモリ 2 7 0 に格納したライト対象データとそのライト要求 (ライト先 L U の L U 識別子とライト先 L U 領域の相対アドレスとを含んだ要求) を、移行先の仮想ストレージ 1 1 0 に送信する。その後、処理が完了する。なお、移行先の仮想ストレージ 1 1 0 は、そのライト要求を受けた場合、そのライ

50

ト要求で指定されているライト先LUに対応したホストLUと、そのライト要求で指定されているライト先LU領域に対応したLU領域とを特定し、特定したホストLU内の特定したLU領域に割り当てられている内部ページに、移行元の仮想ストレージ110からのライト対象データを書き込む。

【0148】

ステップ1209の判断の結果が肯定的の場合(ステップ1209:Y)、ステップ1211で、読み書き実行部(a)920は、キャッシュメモリ270に格納したライト対象データとそのライト要求(ライト先LU領域に割り当てられている外部ページを含んだ外部容量プールに対応する外部LUのLUNと、その外部ページに対応する相対アドレスとを含んだ要求)を、その外部LUを有する外部ストレージ130に送信する。以上で、
10

【0149】

図13は、ページ割り当て処理部930の処理フローの一例を示す。ページ割り当て処理部930は、ページが割り当てられていないLU領域を指定したライト要求を仮想ストレージ110が受けた場合に、そのLU領域にページを割り当てる処理を実行する。

【0150】

ステップ1300では、ページ割り当て処理部930は、内部ページを割り当てるべきか否かを判断する。内部ページを割り当てるべきか否かは、所定のルールに従い決定されて良い。ルールとしては、例えば、下記のうちの少なくとも1つが考えられる。
20

(*)ホストLUには内部ページと外部ページが交互に割り当てられる。

(*)ホストLUには内部ページが優先的に割り当てられ、ホストLUに関連付けられている内部容量プールの割当て済み内部ページの割合が所定割合以下になった場合に、ホストLUに関連付けられている外部容量プールから外部ページが割り当てられる。

【0151】

ステップ1300の判断の結果が肯定的の場合(ステップ1300:Y)、ステップ1301で、ページ割り当て処理部930は、ライト先LUに関連付けられている内部容量プールに対応した内部容量プール情報230を基に、割り当てるページを決定し、その内部容量プール情報230の中の情報で更新すべき情報を更新する。さらに、ページ割り当て処理部930は、ライト先LUに対応するホストLU情報260における容量プール識別子770と相対アドレス780として、決定した内部ページを含んだ内部容量プールの識別子と、決定した内部プールのアドレスとを設定する。その後、処理が完了する。
30

【0152】

ステップ1300の判断の結果が否定的の場合(ステップ1300:N)、ステップ1302で、ページ割り当て処理部930は、割当元の外部容量プールを決定する。また、ページ割り当て処理部930は、その処理部930を含んだ仮想ストレージ110が、その外部容量プールに対応する外部LUの制御権があるか否かを、その外部LUに対応する外部LU情報240内の制御権情報540を基に判断する。
40

【0153】

ステップ1302の判断の結果が肯定的の場合(ステップ1302:Y)、ステップ1303で、ページ割り当て処理部930は、その外部LUに対応する外部容量プール情報250を基に、割当てる外部ページを決定し、外部容量プール情報250の中の情報で更新すべき情報を更新する。さらに、ページ割り当て処理部930は、ライト先LUに対応するホストLU情報260内の容量プール識別子770と相対アドレス780として、ステップ1302で決定した外部容量プールの識別子と、ステップ1303で決定した外部ページのアドレスとを設定する。その後、処理が完了する。

【0154】

ステップ1302の判断の結果が否定的の場合(ステップ1302:N)、ステップ1
50

304で、ページ割り当て処理部930は、ステップ1302で決定した外部容量プールの制御権を有している仮想ストレージ110に、ページの割り当てを要求し、回答をまつ。

【0155】

ステップ1305で、ページ割り当て処理部930は、ステップ1304での要求先の仮想ストレージ110からページのアドレスを受け、ステップ1302で決定した外部容量プールの識別子と、受けたアドレスとを、ライト先LUに対応するホストLU情報260における容量プール識別子770及び相対アドレス780として設定する。その後、処理が完了する。

【0156】

図14は、移行先の仮想ストレージにおけるライトデータ受領部940の処理フローの一例を示す。ライトデータ受領部940は、移行元の仮想ストレージ110からライト対象データを受領する。

【0157】

ステップ1400で、ライトデータ受領部940は、移行元の仮想ストレージ110からライト対象データを受け、そのデータをキャッシュメモリ270に格納する。

【0158】

ステップ1401では、ライトデータ受領部940は、そのデータのライト先のLU領域にページが割り当てられているか否かを判断する。

【0159】

ステップ1401の判断の結果が肯定的の場合(ステップ1401:Y)、処理が完了する。

【0160】

ステップ1401の判断の結果が否定的の場合(ステップ1401:N)、ステップ1402で、ライトデータ受領部940は、ページ割り当て部930をコールして、ライト先のLU領域にページを割り当てる。その後、処理を完了する。

【0161】

図15は、ライトアフタ処理部950の処理フローの一例を示す。ライトアフタ処理部950は、キャッシュメモリ270内のライト対象データ(ライト先LU領域に割り当てられたページに書き込まれていないデータ)をページに書き込む。

【0162】

ステップ1500では、ライトアフタ処理部950は、ライト対象データのライト先LU領域に割り当てられているページが内部ページか外部ページかを判断する。

【0163】

ページが内部ページの場合、ステップ1501で、ライトアフタ処理部950は、データをキャッシュメモリ270から、内部ページに書き込む。このあと、処理が完了する。

【0164】

ページが外部ページの場合、ステップ1502で、ライトアフタ処理部950は、キャッシュメモリ270内のデータとそのライト要求(その外部ページを含んだ外部LUの識別子と、その外部ページの相対アドレスとを指定した要求)とを、外部ストレージ130に送信する。その後、処理が完了する。

【0165】

図16は、外部ページ割り当て処理部960の処理フローの一例を示す。外部ページ割り当て処理部960は、外部容量プールの制御権をもっている仮想ストレージ110が、他の仮想ストレージ110から、ページの割り当て要求を受けるときに、次の処理を実行する。

【0166】

ステップ1600では、外部ページ割り当て処理部960は、指定された外部容量プールに対応する外部容量プール情報250を基に、割り当てる外部ページを決定し、外部容量プール情報250の中の情報で更新すべき情報を更新する(外部ページをLU領域に割

10

20

30

40

50

り当てる)。

【0167】

ステップ1601では、外部ページ割り当て処理部960は、割り当てた外部ページのアドレスを、割り当てを要求した仮想ストレージ110に送る。以上で処理が完了する。

【0168】

図17は、外部ページ解放処理部970の処理フローの一例を示す。外部ページ解放処理部970は、外部容量プールの制御権をもっている仮想ストレージ110が、他の仮想ストレージ110から、ページの解放要求を受けたときに、次の処理を実行する。

【0169】

ステップ1700では、外部ページ解放処理部970は、受けた情報を基に、外部容量プール情報250から、解放する外部ページを特定し、外部容量プール情報250の中の情報で更新すべき情報を更新する(特定した外部ページを解放する)。

【0170】

図18は、制御権移行部980の処理フローの一例を示す。制御権移行部980は、外部LU(外部容量プール)の制御権の移行を、他の仮想ストレージ110から受けたときに、次の処理を実行する。

【0171】

ステップ1800で、制御権移行部980は、制御権の移行先が、この制御権移行部980を含んだ仮想ストレージ110(自仮想ストレージ)であるか否かを判断する。

【0172】

ステップ1800の判断の結果が否定的の場合(ステップ1800:N)、ステップ1801で、制御権移行部980は、他の仮想ストレージ110の識別子を、受けた情報で指定されている外部LU(外部容量プール)に対応する制御権仮想ストレージ550として設定する。この後、処理が完了する。

【0173】

ステップ1800の判断の結果が肯定的の場合(ステップ1800:Y)、ステップ1802で、制御権移行部980は、受けた情報を基に、この外部LU(外部容量プール)に対応する制御権情報540の値をオンとし、制御権仮想ストレージ550として、自仮想ストレージ110の識別子を設定する。さらに、制御権移行部980は、受けた情報に従う外部容量プールに関する情報を、外部容量プール情報250として生成し登録する。

【0174】

図24は、制御権移行決定部990の処理フローの一例を示す。制御権移行決定部990は、他の仮想ストレージ110が実行したホストLUの移行により、制御権を他の仮想ストレージ110に移行するかどうかの判断をするよう要求を受けたとき、次の処理を実行する。

【0175】

ステップ2400で、制御権移行決定部990は、受けた情報を基に、外部LU(外部容量プール)の制御権(割り当て・解放権限)を移行するか否かを決定する。ほとんどの場合、移行元の仮想ストレージ110から制御権は移行されないと決定されるが、例えば、移行元の仮想ストレージ110の負荷が所定負荷以上に高い場合に、移行元の仮想ストレージ110から制御権が移行されると決定される。

【0176】

ステップ2400の判断の結果が否定的の場合(ステップ2400:N)、処理が終了する。

【0177】

ステップ2401で、制御権移行決定部990は、どの仮想ストレージ110に権限を移行するかを決定する。制御権移行決定部990は、制御権の移行先の仮想ストレージ110に、制御権を移行することと、対応する外部容量プール情報250が表す全ての情報とを仮想ストレージ110に送る。また、この情報を送った後、制御権移行決定部990

10

20

30

40

50

は、対応する外部容量プール情報 250 の情報を所定値にリセットし、且つ、制御権仮想ストレージ 550 として、移行先の仮想ストレージ 110 の識別子を設定する。さらに、制御権移行決定部 990 は、この外部 LU に対応する外部 LU 情報 240 内の制御権情報 540 の値をオフとする。

【0178】

ステップ 2402 では、制御権移行決定部 990 は、その外部 LU を有する外部ストレージ 130 に接続された他の仮想ストレージ 110 に、制御権の移行先の仮想ストレージ 110 の識別子を通知する。以上で処理が完了する。

【実施例 2】

【0179】

以下、実施例 2 を説明する。その際、実施例 1 との相違点を主に説明し、実施例 1 との共通点については説明を省略或いは簡略する。

【0180】

図 19 は、実施例 2 に係る仮想ストレージ 110 が発揮する複数の機能の一例を示す。

【0181】

実施例 2 では、複数の機能として、データ移行制御部 (b) 1900、データ移行実行部 (b) 1910、読み書き実行部 (b) 1920、リード要求受領部 1930、ページ割り当て処理部 930、ライトデータ受領部 940、ライトアプタ処理部 950、外部ページ割り当て処理部 960、外部ページ解放処理部 970、制御権移行部 980、及び制御権移行決定部 990 がある。

【0182】

この中で、ページ割り当て処理部 930、ページ割り当て処理部 930、ライトデータ受領部 940、ライトアプタ処理部 950、外部ページ割り当て処理部 960、外部ページ解放処理部 970、制御権移行部 980、及び制御権移行決定部 990 は、実施例 1 と共通なので、説明を省略する。

【0183】

実施例 2 では、移行元の仮想ストレージ 110 から移行先の仮想ストレージ 110 にホスト LU を移行するとき、移行処理の制御を、移行先の仮想ストレージ 110 が実行する。また、移行処理中、ホスト 100 は、移行先の仮想ストレージにリード/ライト要求を発行する。

【0184】

実施例 2 でも、管理サーバ 160 が、各仮想ストレージ 110 の負荷を調べ、移行元のホスト LU と移行先の仮想ストレージ 110 を決定する。管理サーバ 160 は、移行元の仮想ストレージ 110 に、移行元のホスト LU の識別子と、移行先の仮想ストレージ 110 の識別子と、移行先のホスト LU の識別子を送る。また、管理サーバ 160 は、移行先の仮想ストレージ 110 に、移行元の仮想ストレージ 110 の識別子と、移行元の LU の識別子と、移行先のホスト LU の識別子と、移行元のホスト LU の仮想容量とを送る。移行元の仮想ストレージ 110 と移行先の仮想ストレージ 110 の双方から、受領報告が帰ってくると、管理サーバ 160 は、ホスト 100 に、移行元のホスト LU のリード/ライト要求のあて先を、移行先の仮想ストレージ 110 に切り替えるよう要求する。ワールドワイドネームが引き継がれ、SAN 120 などで自動的に、移行先の仮想ストレージに、リード/ライト要求のあて先が切り替えられても良い。

【0185】

実施例 2 では、移行先の仮想ストレージ 110 におけるデータ移行制御部 (b) 1900 が、管理サーバ 160 から、移行元の仮想ストレージ 110 の識別子、移行元のホスト LU の識別子、移行先のホスト LU の識別子と、移行元のホスト LU の仮想容量とを含んだ情報を受け、その情報を受けたことを、管理サーバ 160 に応答した後、以下の図 20 に示す処理フローを実行する。

【0186】

図 20 は、移行先の仮想ストレージ 110 におけるデータ移行制御部 (b) 1900 の

10

20

30

40

50

処理フローの一例を示す。

【0187】

ステップ2000では、データ移行制御部(b)1900は、移行先ホストLU(受領したホストLUの識別子)に対応するホストLU情報を生成する。データ移行制御部(b)1900は、ホストLU識別子700及び仮想容量710として、移行先ホストLUの識別子と仮想容量とを設定する。

【0188】

ステップ2001では、データ移行制御部(b)1900は、移行先ホストLUに対応する移行中フラグ730をオンにする。さらに、データ移行制御部(b)1900は、管理サーバ160から受けた移行先仮想ストレージ110の識別子と移行先LUの識別子とを、移行先ホストLUに対応する移行中仮想ストレージ識別子740及び移行中ホストLU識別子750として設定する。さらに、データ移行制御部(b)1900は、コピーポインタ760を最初のLU領域に設定する。

10

【0189】

ステップ2002では、データ移行制御部(b)1900は、コピーポインタ760が示すLU領域(以下、図20の説明で「対象移行先領域」)の処理要求を、移行元の仮想ストレージ110に要求する。

【0190】

ステップ2003では、データ移行制御部(b)1900は、移行先となる仮想ストレージ110から情報が送られてくるのを待つ。

20

【0191】

ステップ2004 - ステップ2009は、図11のステップ1102 - ステップ1107と同様のため、説明を省略する。

【0192】

ステップ2010では、データ移行制御部(b)1900は、対象移行先領域が移行先ホストLUにおける末尾のLU領域であるか否かを判断する。

【0193】

ステップ2010の判断の結果が否定的の場合(ステップ2010:N)、ステップ2011で、データ移行制御部(b)1900は、次のLU領域を表す情報を、移行先ホストLUに対応するホストLU情報260におけるコピーポインタ760として設定し、ステップ2002へジャンプする。

30

【0194】

ステップ2010の判断の結果が肯定的の場合(ステップ2010:Y)、ステップ2012以降が行われる。ただし、ステップ2012 - ステップ2014は、図11のステップ1109 - ステップ1111と同様のため、説明を省略する。

【0195】

実施例2では、移行元の仮想ストレージ110が、データ移行実行部(b)1910が、管理サーバ160から、移行先の仮想ストレージ110の識別子と、移行先のホストLUの識別子と、移行元のLUの識別子とを含んだ情報を受領し、その情報を受領したことを、管理サーバ160に回答した後、要求を受け、図21に示す処理フローを実行する。

40

【0196】

図21は、移行元の仮想ストレージ110におけるデータ移行実行部(b)1910の処理フローの一例を示す。

【0197】

ステップ2100では、データ移行実行部(b)1910は、移行元のホストLUの識別子に対応するホストLU情報を見つける。

【0198】

ステップ2101では、データ移行実行部(b)1910は、移行先の仮想ストレージ110から、処理要求が送られてくるのをまつ。

【0199】

50

移行先の仮想ストレージ 110 から、処理要求が来ると、その要求に従う対象移行先領域に対応する移行元の LU 領域（図 21 において「対象移行元領域」）について、ステップ 2102 以降が実行される。ただし、ステップ 2102 - ステップ 2110 は、図 10 のステップ 1002 - ステップ 1010 と同様であるため、説明を省略する。

【0200】

ステップ 2111 では、データ移行実行部 (b) 1910 は、対象移行元領域が移行元ホスト LU における末尾の LU 領域かどうかを判断する。

【0201】

ステップ 2111 の判断の結果が否定的の場合（ステップ 2111：N）、データ移行実行部 (b) 1910 は、ステップ 2101 へジャンプして、次の処理要求が、移行先の仮想ストレージ 110 から送られてくるのを待つ。

10

【0202】

ステップ 2111 の判断の結果が肯定的の場合（ステップ 2111：Y）、データ移行実行部 (b) 1910 は、ステップ 2112 で、移行処理が完了したことを管理サーバ 160 に報告する。この後、処理が完了する。

【0203】

ステップ 2113 - ステップ 2121 は、図 10 のステップ 1015 - ステップ 1023 と同様である。このため、説明を省略する。

【0204】

実施例 2 で、管理サーバ 160 は、移行元の仮想ストレージ 110 と移行先の仮想ストレージ 110 からホスト LU の移行処理が完了した情報を受け付けると、管理サーバ 160 は、移行元の仮想ストレージ 110 に、移行元のホスト LU の情報を消去するよう、指示する。

20

【0205】

図 22 は、移行先の仮想ストレージ 110 における読み書き実行部 (b) 1920 の処理フローの一例を示す。この処理は、ホスト LU の移行中に、移行先の仮想ストレージ 110 が、移行中のホスト LU を指定したリード/ライト要求をホスト 100 から受けた場合に実行される。

【0206】

ステップ 2200 では、読み書き実行部 (b) 1920 は、受けた要求が、リード要求かライト要求かを判断する。

30

【0207】

要求がリード要求の場合、ステップ 2201 で、読み書き実行部 (b) 1920 は、キャッシュメモリ 270 にリード対象データが存在するか否かを判断する。この判断の結果が肯定的であれば（ステップ 2210：Y）、処理がステップ 2206 へジャンプし、この判断の結果が否定的であれば（ステップ 2210：N）、ステップ 2202 が行われる。

【0208】

ステップ 2202 で、読み書き実行部 (b) 1920 は、読出し元のホスト LU に対応した移行中フラグ 730 がオンであり、且つ、読出し元の LU 領域について移行が完了しているか否かを判断する。読出し元のホスト LU に対応したコピーポインタ 760 が示すアドレスが、読出し元の LU 領域のアドレスより後のアドレスであれば、読出し元の LU 領域について移行が完了している。

40

【0209】

ステップ 2202 の判断の結果が否定的の場合（ステップ 2202：N）、処理がステップ 2205 へジャンプする。

【0210】

ステップ 2202 の判断の結果が否定的の場合（ステップ 2202：N）、そうでない場合、ステップ 2203 で、読み書き実行部 (b) 1920 は、移行元の仮想ストレージ 110 に、読出し元の LU 領域に対応する LU 領域（移行元ホスト LU 内の LU 領域）に

50

おけるデータを送るよう要求し、そのデータが送られてくるのを待つ。ステップ 2204 で、読み書き実行部 (b) 1920 は、送られてきたデータを、キャッシュメモリ 270 へ書き込む。

【0211】

ステップ 2205 - ステップ 2206 は、図 12 のステップ 1202 - ステップ 1203 と同様であるため、説明を省略する。

【0212】

ライト要求の場合、ステップ 2207 以降の処理を実行する。ただし、ステップ 2207 - ステップ 2211 の処理は、図 12 のステップ 1204 - ステップ 1208 と同様であるため、説明を省略する。

10

【0213】

図 22 において、ステップ 2211 で、読み書き実行部 (b) 1920 は、コピーポインタ 760 を基に、ライト先 LU 領域について移行処理が完了していないと判断すれば (ステップ 2211 : N)、ステップ 2212 で、キャッシュメモリ内のライト対象データとそのライト要求 (ライト先 LU の識別子とライト先 LU 領域のアドレスとを含んだ要求) を、移行元の仮想ストレージ 110 に送る。これにより、移行元の仮想ストレージにおいて、ライト先 LU 領域に対応する LU 領域 (移行元ホスト LU の LU 領域) に割り当てられているページに、ライト対象データが書き込まれる。読み書き実行部 (b) 1920 は、移行元の仮想ストレージ 110 から完了報告を受けると、キャッシュメモリ 270 からこのライト対象データを消去する。この後、処理が完了する。

20

【0214】

ライト先 LU 領域が移行処理が完了した領域の場合、ステップ 2213 - ステップ 2215 が実行される。ただし、ステップ 2213 - ステップ 2215 は、図 12 のステップ 1209 - ステップ 1211 (但し、ステップ 2214 を除く) と同様である。ステップ 2214 の処理は、図 12 のステップ 1210 に相当する。ステップ 2214 では、移行先の仮想ストレージ 110 のキャッシュメモリ 270 内のライト対象データが、移行元の仮想ストレージ 110 に送られる (図 12 のステップ 1210 では、逆に、移行元の仮想ストレージ 110 のキャッシュメモリ 270 内のライト対象データが、移行先の仮想ストレージに送られる)。

【0215】

30

図 23 は、移行元の仮想ストレージ 110 におけるリード要求受領部 1930 の処理フローの一例を示す。リード要求受領部 1930 は、移行元の仮想ストレージ 110 が、移行先の仮想ストレージ 110 から、データを送って欲しい要求を受け付けたとき、次の処理を実行する。

【0216】

ステップ 2300 では、リード要求受領部 1930 は、データを読み出し元が内部ページか外部ページかを判断する。

【0217】

ページが内部ページの場合、ステップ 2301 で、リード要求受領部 1930 は、データを内部ページから読み出し、そのデータをキャッシュメモリ 270 に格納する。このあと、処理がステップ 2303 にジャンプする。

40

【0218】

ページが外部ページの場合、ステップ 2302 で、リード要求受領部 1930 は、その外部ページに対応する外部アクセス先情報 (外部 LU の識別子と、外部 LU 領域の相対アドレスとを含んだ情報) を有するリード要求を外部ストレージ 130 に送信する。リード要求受領部 1930 は、そのリード要求に回答して外部ストレージ 130 から受けたデータを、キャッシュメモリ 270 に格納する。

【0219】

ステップ 2303 では、リード要求受領部 1930 は、キャッシュメモリ 270 から読み出したデータを、移行先の仮想ストレージ 110 に送る。その後、処理が完了する。

50

【実施例 3】**【0220】**

以下、実施例 3 を説明する。

【0221】

実施例 3 及び後述の実施例 4 では、既設の仮想ストレージ 110 が新たに用意される仮想ストレージ 110 に置き換えられる。まず、実施例 3 及び 4 に共通する事項を説明する。その際、既設の（置き換えられる）仮想ストレージ 110 を「旧仮想ストレージ 110」と言い、新たに用意される仮想ストレージ 110 を「新仮想ストレージ 110」と言う。

【0222】

まず、ホスト LU の移行に入る準備段階として、以下の手順が実行される。

(1) 新仮想ストレージ 110 が、SAN(a)120 経由で、ホスト 100 に接続され、ホスト 100 が、新仮想ストレージ 110 を認識する。ただし、新仮想ストレージ 110 はホスト 100 と直接接続されてもよい。

(2) 新仮想ストレージ 110 が、管理サーバ 160 に接続され、管理サーバ 160 が、新仮想ストレージ 110 を認識する。

(3) 旧仮想ストレージ 110 に接続されているすべての外部ストレージ 130 と、SAN(b)150 経由で、新仮想ストレージ 110 が接続される。新仮想ストレージ 110 は、外部ストレージ 130 内の外部 LU を認識し、その外部 LU に対応する外部 LU 情報 240 を記憶する。ただし、外部ストレージ 130 と新仮想ストレージ 110 は直接接続

されてもよい。

(4) 新仮想ストレージ 110 は、管理サーバ 160 から、認識したそれぞれの外部 LU について、下記の情報、

(a) 外部 LU は容量仮想化機能が適用されているか否か、

(b) 容量仮想化機能が適用されている外部 LU について、その外部 LU の制御権は、どの仮想ストレージ 110 が有しているか、

を受ける。これらの情報に基づき、新仮想ストレージ 110 は、認識している外部 LU に対応する外部 LU 情報 240 を更新する。

(5) 新仮想ストレージ 110 は、内部ボリューム 210 を有する。さらに、新仮想ストレージ 110 は、管理サーバ 160 からの情報を基に、内部容量プール情報 230 に必要な情報を設定する。

【0223】

以上の準備が整うと、旧仮想ストレージ 110 から新仮想ストレージ 110 へのホスト LU への移行が開始される。旧仮想ストレージ 110 に定義されている全てのホスト LU にそれぞれ容量仮想化機能が適用されていても良いし、一部のホスト LU に容量仮想化機能が適用されていても良い。

【0224】

管理サーバ 160 が、旧仮想ストレージ 110 上に定義されている移行元のホスト LU を認識し、新仮想ストレージ 110 に移行先のホスト LU を定義し、移行元のホスト LU から移行先のホスト LU へのデータの移行を移行元の仮想ストレージ 110 又は移行先の仮想ストレージ 110 に実行させる。

【0225】

実施例 3 は、ホスト LU の移行の仕方に、実施例 1 が適用されたケースである。

【0226】

実施例 3 では、旧仮想ストレージ（移行元の仮想ストレージ）110 から新仮想ストレージ（移行先の仮想ストレージ）110 にホスト LU を移行するとき、移行処理の制御を、旧仮想ストレージ 110 が実行する。また、移行処理中、ホスト 100 は、旧仮想ストレージ 110 にリード/ライト要求を発行する。

【0227】

本実施例 3 では、管理サーバ 160 が、旧仮想ストレージ 110 に定義されているホス

10

20

30

40

50

トLUを認識し、移行すべきホストLUを決定する。決定すると、管理サーバ160は、旧仮想ストレージ110に、移行元のホストLUの識別子と、新仮想ストレージ110の識別子と、移行先のホストLUの識別子を送る。また、管理サーバ160は、新仮想ストレージ110に、旧仮想ストレージ110の識別子と、移行元のホストLUの識別子と、移行先のホストLUの識別子と、移行元のホストLUの仮想容量とを送る。

【0228】

この後、実施例1で説明した処理が行われ、移行元ホストLUから移行先ホストLUへのデータ移行が完了する。また、この際、必要に応じて、外部LUの制御権も移行されて良い。

【0229】

この処理が完了すると、管理サーバ160は、旧仮想ストレージ110のホストLUのすべての移行が完了したかを認識する。まだ、移行が完了していないホストLUがある場合、管理サーバ160は、移行すべきホストLUを決定して、上記の処理を実行させる。すべてのホストLUの移行が完了すると、管理サーバ160は、処理を終了する。旧仮想ストレージ110が、全てのホストLUについてデータ移行が完了したか否かを判断し、その判断の結果が否定的の場合に、データ移行が済んでいないホストLUについてデータ移行を行って良い。

【実施例4】

【0230】

実施例4は、ホストLUの移行の仕方に、実施例2が適用されたケースである。

【0231】

実施例4では、旧仮想ストレージ（移行元の仮想ストレージ）110から新仮想ストレージ（移行先の仮想ストレージ）110にホストLUが移行されるとき、移行処理の制御を、新仮想ストレージ110が実行する。また、移行処理中、ホスト100は、新仮想ストレージにリード/ライト要求を発行する。

【0232】

実施例4でも、管理サーバ160が、旧仮想ストレージ110に定義されているホストLUを認識し、移行すべきホストLUを決定する。決定すると、管理サーバ160は、旧仮想ストレージ110に、移行元のホストLUの識別子と、新仮想ストレージ110の識別子と、移行先のホストLUの識別子を送る。また、管理サーバ160は、新しい仮想ストレージ110に、旧仮想ストレージ110の識別子と、移行元のLUの識別子と、移行先のホストLUの識別子と、移行先のホストLUの仮想容量とを送る。

【0233】

また、実施例2で示したように、管理サーバ160は、上記要求の受領報告を、旧仮想ストレージ110と新仮想ストレージの双方から受けると、ホスト100に、移行対象となるホストLUのリード/ライト要求のあて先を、新仮想ストレージ110に切り替えるよう要求して良い。

【0234】

実施例2と同様の処理が実行されることで、上記で指定されたホストLUの移行が完了する。また、この際、必要に応じて、外部LUの制御権も移行されて良い。

【0235】

この処理が完了すると、管理サーバ160は、旧仮想ストレージ110のホストLUのすべての移行が完了したかを認識する。まだ、移行が完了していないホストLUがある場合、管理サーバ160は、移行すべきホストLUを決定して、上記の処理を実行して良い。すべてのホストLUの移行が完了すると、管理サーバ160は、処理を終了する。

【0236】

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

【符号の説明】

10

20

30

40

50

【 0 2 3 7 】

1 0 0 : ホスト、 1 1 0 : 仮想ストレージ、 1 3 0 : 外部ストレージ

【 図 1 】

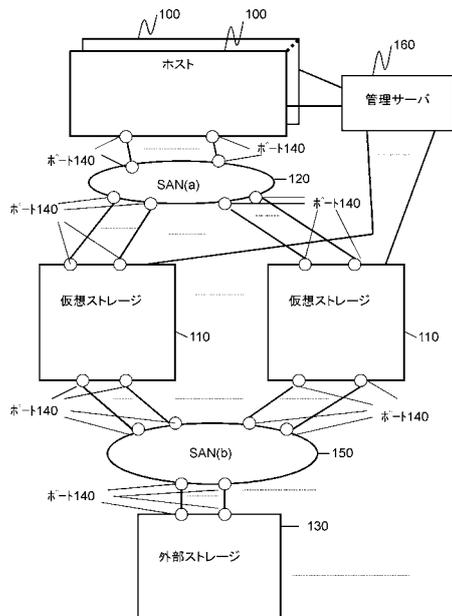


Fig.1

【 図 2 】

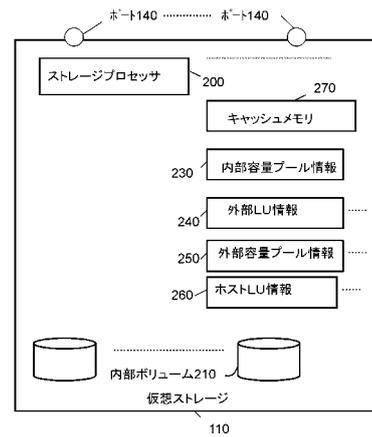


Fig.2

【 図 3 】

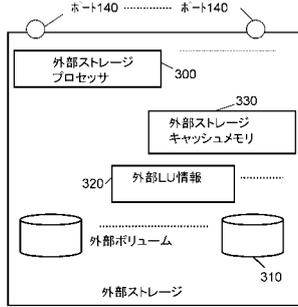


Fig.3

【 図 5 】



Fig.5

【 図 4 】

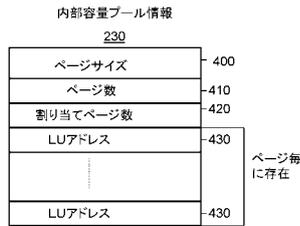


Fig.4

【 図 6 】

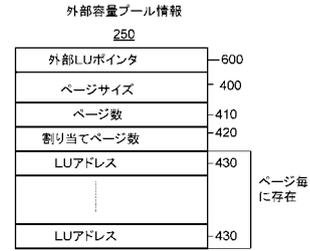


Fig.6

【 図 7 】

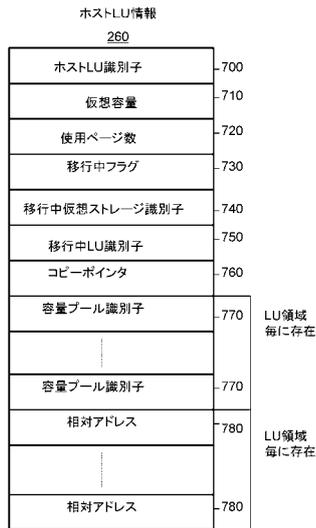


Fig.7

【 図 9 】

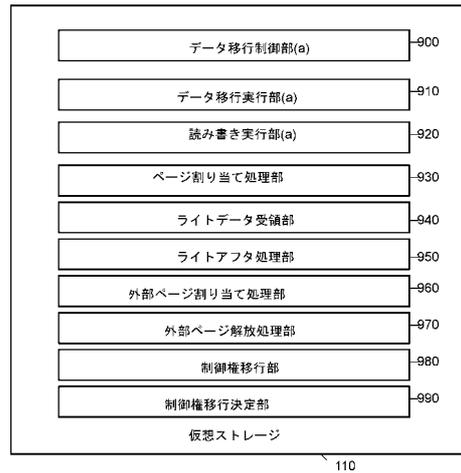


Fig.9

【 図 8 】

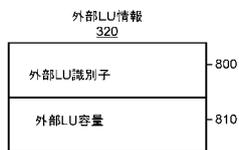


Fig.8

【図 10】

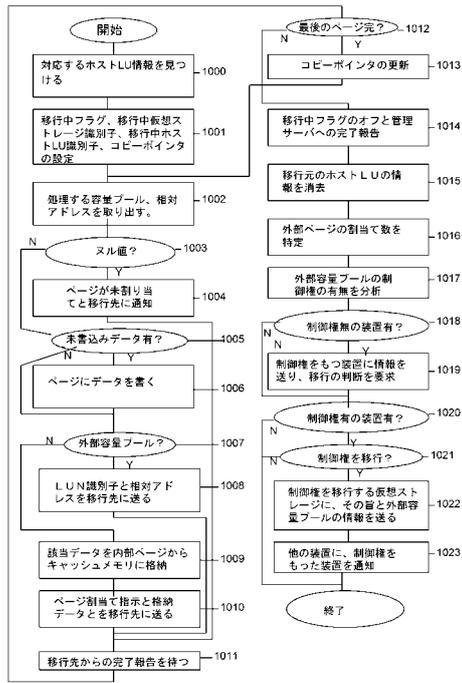


Fig.10

【図 11】

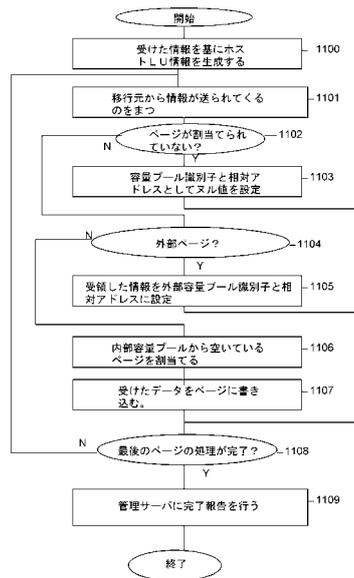


Fig.11

【図 12】

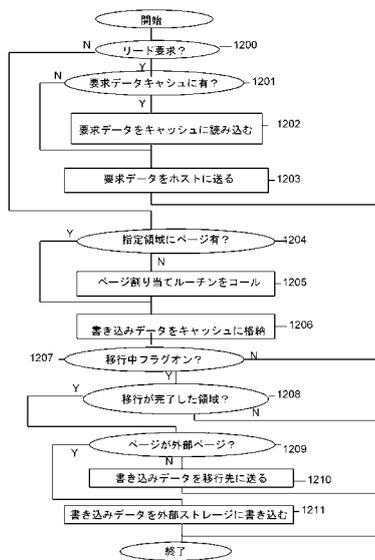


Fig.12

【図 13】

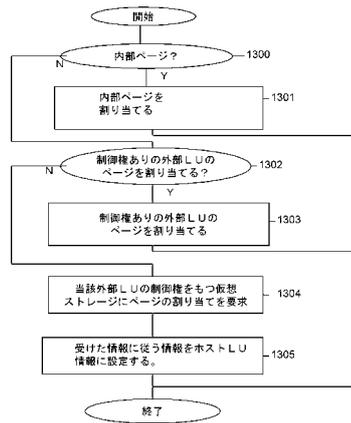


Fig.13

【図14】

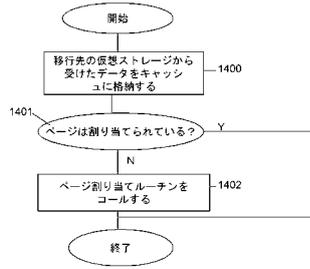


Fig.14

【図15】

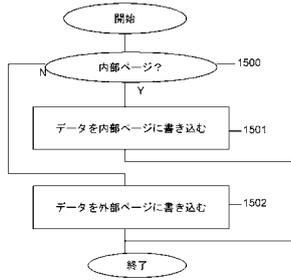


Fig.15

【図18】

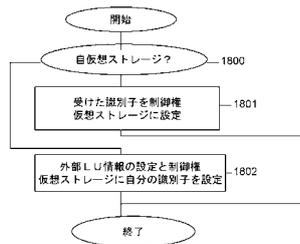


Fig.18

【図16】

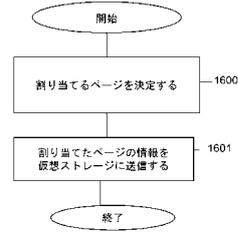


Fig.16

【図17】

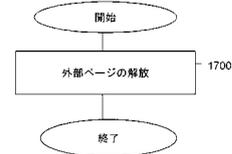


Fig.17

【図19】

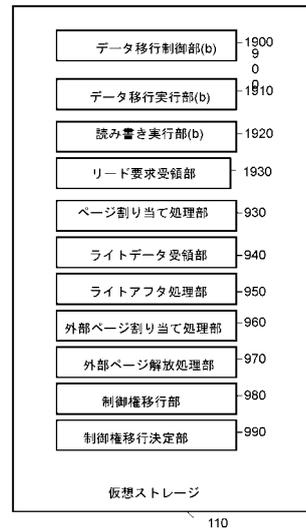


Fig.19

【図20】

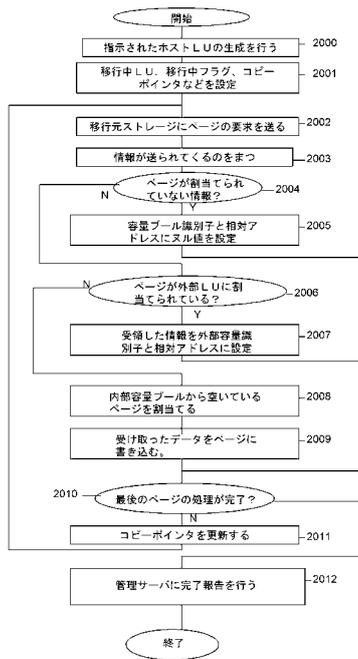


Fig.20

【図21】

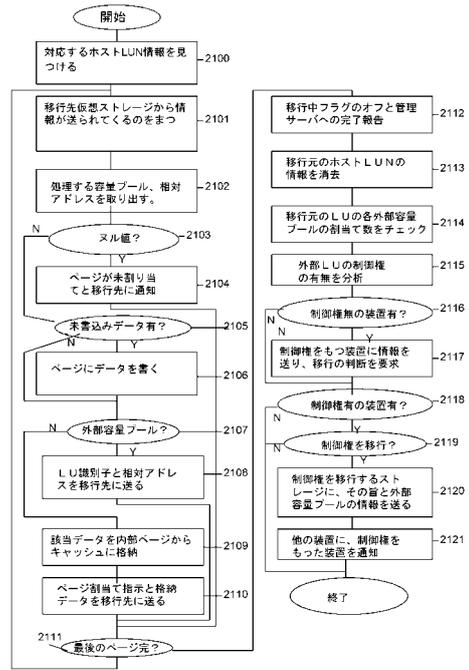


Fig.21

【図22】

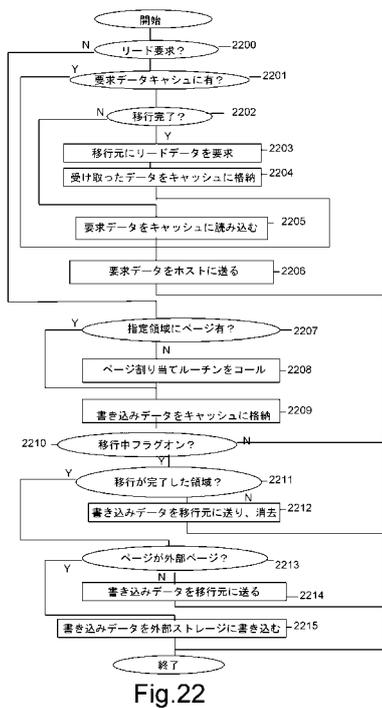


Fig.22

【図23】

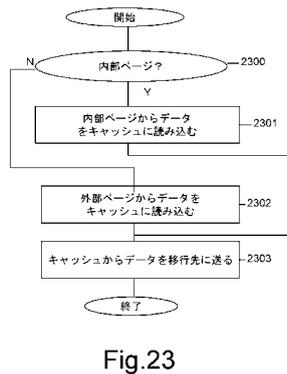


Fig.23

【図24】

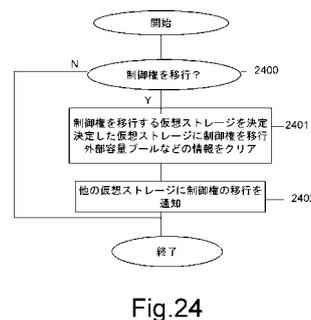


Fig.24

フロントページの続き

- (72)発明者 江口 賢哲
神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 横浜研究所内
- (72)発明者 山本 政行
神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 横浜研究所内

審査官 古河 雅輝

- (56)参考文献 特開2006-330895(JP,A)
特開2007-310495(JP,A)
特開2009-230742(JP,A)
特開2009-295049(JP,A)

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|---------------|
| G06F | 3/06 - 3/08 |
| G06F | 12/00 |
| G06F | 13/10 - 13/14 |