

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5323982号

(P5323982)

(45) 発行日 平成25年10月23日 (2013. 10. 23)

(24) 登録日 平成25年7月26日 (2013. 7. 26)

(51) Int. Cl.

F I

G O 6 F 12/00 (2006. 01)

G O 6 F 3/06 (2006. 01)

G O 6 F 13/10 (2006. 01)

G O 6 F 12/00 5 O 1 A

G O 6 F 12/00 5 3 1 D

G O 6 F 12/00 5 1 4 E

G O 6 F 12/00 5 4 5 A

G O 6 F 3/06 3 O 1 J

請求項の数 13 (全 42 頁) 最終頁に続く

(21) 出願番号 特願2012-504138 (P2012-504138)

(86) (22) 出願日 平成22年3月9日 (2010. 3. 9)

(86) 国際出願番号 PCT/JP2010/001667

(87) 国際公開番号 W02011/111093

(87) 国際公開日 平成23年9月15日 (2011. 9. 15)

審査請求日 平成24年3月30日 (2012. 3. 30)

(73) 特許権者 000005108

株式会社日立製作所

東京都千代田区丸の内一丁目6番6号

(74) 代理人 100093861

弁理士 大賀 真司

(74) 代理人 100129218

弁理士 百本 宏之

(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 里山 愛

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 田川 泰宏

最終頁に続く

(54) 【発明の名称】 ボリューム管理装置及びストレージシステム

(57) 【特許請求の範囲】

【請求項 1】

性能が異なる複数の物理デバイスの論理的記憶領域を形成する複数の論理デバイスを前記物理デバイスの性能に応じて階層化し、前記階層化された論理ボリュームを各階層に対応したプールボリュームとして格納する仮想化プールと、

コマンド発行元のアクセス対象となる正論理ボリュームと、

前記正論理ボリュームのペアの対象となる副論理ボリュームと、

前記正論理ボリュームの記憶領域を複数の正側ページに分割して管理するとともに、前記副論理ボリュームの記憶領域を複数の副側ページに分割して管理し、かつ前記仮想化プールに属する各階層のプールボリュームの記憶領域を複数のプール側ページに分割して管理し、前記コマンド発行元からのコマンドで指定された正側ページに前記いずれかの階層のプール側ページを割り当てるコントローラと、を備え、

前記コントローラは、

前記コマンド発行元から、前記正論理ボリュームと前記副論理ボリュームとをペアとするペア構成に関するコマンドを受けたときには、前記コマンドで指定された正側ページに前記いずれかのプール側ページを割り当て、前記副側ページに、前記正側ページに割り当てられたプール側ページと同じ階層に属するプール側ページを割り当て、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドを受けたときには、前記正側ページまたは前記副側ページのうち少なくとも一方に割り当てるプール側ページの階層を前記状態遷移コマンドに応じて選択する、ボリューム管理装置。

10

20

【請求項 2】

請求項 1 に記載のボリューム管理において、
前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーする更新コピーを停止し、

10

前記ペア分割のコマンドに、ページ移動指示が含まれているときには、前記副側ページに割り当てられているプール側ページを、前記正側ページに割り当てられたプール側ページとは異なる階層に属するプール側ページに移動し、

前記副側ページに、前記移動したプール側ページを割り当て、前記副側ページへのライトデータを、前記移動したプール側ページに格納する、ボリューム管理装置。

【請求項 3】

請求項 1 に記載のボリューム管理装置において、
前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーする、ボリューム管理装置。

20

【請求項 4】

請求項 1 に記載のボリューム管理装置において、
前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

30

前記ペア分割のコマンドに、ページ移動指示が含まれているときには、前記プール側ページの割り当てが解除された副側ページに、前記正側ページに割り当てられたプール側ページとは異なる階層に属するプール側ページを新たに割り当てる、ボリューム管理装置。

【請求項 5】

請求項 1 に記載のボリューム管理装置において、
前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

40

その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、前記コピーペアを再同期させるペア再同期のコマンドを受けたときには、前記正側ページと前記副側ページに、前記ペア再同期のコマンドを受ける前にそれぞれ割り当てられているプール側ページのうち、同じ階層に属するものを一組含むものを全て、前記ペア再同期におけ

50

るコピーペアとして、前記ペア再同期におけるコピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーする、ボリューム管理装置。

【請求項6】

請求項1に記載のボリューム管理装置において、

前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

10

その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、前記コピーペアをリバースリシンクさせるリバースリシンクのコマンドを受けたときには、前記正側ページと前記副側ページに、前記リバースリシンクのコマンドを受ける前にそれぞれ割り当てられているプール側ページのうち、同じ階層に属するものを一組含むものを全て、前記リバースリシンクにおけるコピーペアとして、前記リバースリシンクにおけるコピーペアに属する副側ページのデータを前記コピーペアに属する正側ページにコピーする、ボリューム管理装置。

【請求項7】

20

請求項1に記載のボリューム管理装置において、

前記コントローラは、

前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、その後、前記コマンド発行元から、前記ペアに関する状態遷移コマンドとして、前記コピーペアをクイックリストアさせるクイックリストアのコマンドを受けたときには、前記正論理ボリュームを新たな副論理ボリュームとし、前記副論理ボリュームを新たな正論理ボリュームとし、前記コマンド発行元から前記正論理ボリュームへのアクセスを前記新たな副論理ボリュームに切り替え、前記コマンド発行元から前記副論理ボリュームへのアクセスを前記新たな正論理ボリュームに切り替え、

30

前記新たな正論理ボリュームの正側ページに割り当てられている前記プール側ページと前記新たな副論理ボリュームの副側ページに割り当てられている前記プール側ページが異なる階層に属するときには、前記新たな正論理ボリュームの正側ページに、前記新たな副論理ボリュームの副側ページに割り当てられている前記プール側ページと同じ階層に属するプール側ページを新たに割り当てる、ボリューム管理装置。

【請求項8】

コマンド発行元のローカル側ストレージシステムにネットワークを介して接続されたりモート側ストレージシステムであって、

40

性能が異なる複数の物理デバイスと、

前記物理デバイスの論理的記憶領域を形成する複数の論理デバイスを前記物理デバイスの性能に応じて階層化し、前記階層化された論理ボリュームを各階層に対応したプールボリュームとして格納する仮想化プールと、

前記ローカル側ストレージシステムのアクセス対象となる正論理ボリュームと、

前記正論理ボリュームのペアの対象となる副論理ボリュームと、

前記正論理ボリュームの記憶領域を複数の正側ページに分割して管理するとともに、前記副論理ボリュームの記憶領域を複数の副側ページに分割して管理し、かつ前記仮想化プールに属する各階層のプールボリュームの記憶領域を複数のプール側ページに分割して管理し、前記コマンド発行元からのコマンドで指定された正側ページに前記いずれかの階層

50

のプール側ページを割り当てるコントローラと、を備え、

前記コントローラは、

前記ローカル側ストレージシステムから、前記正論理ボリュームと前記副論理ボリュームとをペアとするペア構成に関するコマンドを受けたときには、前記コマンドで指定された正側ページに前記いずれかのプール側ページを割り当て、前記副側ページに、前記正側ページに割り当てられたプール側ページと同じ階層に属するプール側ページを割り当て、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドを受けたときには、前記正側ページまたは前記副側ページのうち少なくとも一方に割り当てるプール側ページの階層を前記状態遷移コマンドに応じて選択する、ストレージシステム。

【請求項 9】

請求項 8 に記載のストレージシステムにおいて、

前記コントローラは、

前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーする、ストレージシステム。

【請求項 10】

請求項 8 に記載のストレージシステムにおいて、

前記コントローラは、

前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

前記ペア分割のコマンドに、ページ移動指示が含まれているときには、前記プール側ページの割り当てが解除された副側ページに、前記正側ページに割り当てられたプール側ページとは異なる階層に属するプール側ページを新たに割り当てる、ストレージシステム。

【請求項 11】

請求項 8 に記載のストレージシステムにおいて、

前記コントローラは、

前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組含むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、ペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、前記コピーペアを再同期させるペア再同期のコマンドを受けたときには、前記正側ページと前記副側ページに、前記ペア再同期のコマンドを受ける前にそれぞれ割り当てられているプール側ページのうち、同じ階層に属するものを一組含むものを全て、前記ペア再同期におけるコピーペアとして、前記ペア再同期におけるコピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーする、ストレージシステム。

【請求項 12】

請求項 8 に記載のストレージシステムにおいて、

前記コントローラは、

前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組合むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、前記コピーペアを分割させるペア分割のコマンドを受けたときには、前記副側ページに対して、前記プール側ページの割り当てを解除し、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、前記コピーペアをリバースリシンクさせるリバースリシンクのコマンドを受けたときには、前記正側ページと前記副側ページに、前記リバースリシンクのコマンドを受ける前にそれぞれ割り当てられているプール側ページのうち、同じ階層に属するものを一組合むものを全て、前記リバースリシンクにおけるコピーペアとして、前記リバースリシンクにおけるコピーペアに属する副側ページのデータを前記コピーペアに属する正側ページにコピーする、ストレージシステム。

【請求項 13】

請求項 8 に記載のストレージシステムにおいて、

前記コントローラは、

前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、コピーペア形成のコマンドを受けたときには、前記正側ページと前記副側ページのうち割り当てられているプール側ページが同じ階層に属するものを一組合むものをコピーペアとして、前記コピーペアに属する正側ページのデータを前記コピーペアに属する副側ページにコピーし、

その後、前記ローカル側ストレージシステムから、前記ペアに関する状態遷移コマンドとして、前記コピーペアをクイックリストアさせるクイックリストアのコマンドを受けたときには、前記正論理ボリュームを新たな副論理ボリュームとし、前記副論理ボリュームを新たな正論理ボリュームとし、前記ローカル側ストレージシステムから前記正論理ボリュームへのアクセスを前記新たな副論理ボリュームに切り替え、前記ローカル側ストレージシステムから前記副論理ボリュームへのアクセスを前記新たな正論理ボリュームに切り替え、

前記新たな正論理ボリュームの正側ページに割り当てられている前記プール側ページと前記新たな副論理ボリュームの副側ページに割り当てられている前記プール側ページが異なる階層に属するときには、前記新たな正論理ボリュームの正側ページに、前記新たな副論理ボリュームの副側ページに割り当てられている前記プール側ページと同じ階層に属するプール側ページを新たに割り当てる、ストレージシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ホスト計算機のアクセス対象となる論理ボリュームに対して、ページを動的に割り当てるための管理を行うボリューム管理装置及びストレージシステムに関する。

【背景技術】

【0002】

従来から、ホスト計算機に対して大規模なデータストレージサービスを提供する計算機システムが存在する。このシステムは、ホスト計算機と、ホスト計算機が接続するストレージ装置と、ストレージ装置の管理装置と、を備えたものとして知られている。

【0003】

ストレージ装置は、複数のハードディスクを RAID (Redundant Array of Independent/Inexpensive Disks) 方式で管理する。そして、多数のハードディスクが有する物理的な記憶領域を論理化し、これを論理ボリュームとしてホスト計算機に提供する。ホスト計算機

は、論理ボリュームにアクセスしてデータのリード・ライトを要求する。

【0004】

この種の論理化技術の一つにシン・プロビジョニングと称されるものがある。ストレージ装置は、物理的な記憶領域を有さない仮想的なボリュームをホスト計算機に対して設定する。ホスト計算機が、仮想的なボリュームにライトアクセスにすることに合わせて、ストレージ装置は、仮想ボリュームに対して記憶領域を順次割り当てていく。

【0005】

従って、この技術は、論理ボリュームに対して当初から大容量な記憶領域を割り当てる方式に比べて、ストレージ資源を有効に活用できる点で効果的である。この技術は、特開2003-15915号公報、及び、特開2006-338341号公報に記載されている。

10

【0006】

仮想ボリュームに記憶領域を提供するための手段は、それぞれ実の記憶領域が割り当てられた複数の論理ボリュームをプールと呼ばれる論理グループに纏めたものから構成される。複数の論理ボリュームはそれぞれプールボリュームと呼ばれる。仮想ボリュームに、ホスト計算機からライトアクセスがあると、仮想ボリュームのアクセス先に、プールボリュームの対応領域の記憶領域を割り当てる。

【0007】

ストレージ装置は、仮想ボリュームに対するホスト計算機からのアクセス領域を、プールボリュームに記憶容量が割り当てられた領域に対応させることにより、ホスト計算機からのライトデータを保存することができる。

20

【0008】

一方、ストレージ装置として、ストレージ装置への投資効率の向上のために、使用する記憶メディア（物理デバイス）を複数種類用意するとともに、データの価値を定義し、そのデータの価値と記憶メディアとを関連付けてデータを格納できるものが要求されている。そこで、米国特許2005/055603号公報では、シン・プロビジョニングにおいて、容量割当て単位であるページへの入出力アクセスに応じて、適当な記憶メディアをページに自動再配置することを開示している。

【先行技術文献】

【特許文献】

30

【0009】

【特許文献1】特開2003-15915号公報

【特許文献2】特開2006-338341号公報

【特許文献3】米国特許2005/055603号公報

【発明の開示】

【発明が解決しようとする課題】

【0010】

しかし、米国特許2005/055603号公報は、容量割当て単位であるページへの入出力アクセスに応じて、適当な記憶メディアをページに自動再配置することを開示しているものの、性能が異なる物理デバイスに対応して構成された論理ボリュームに、ページを割り当てるに際して、ペアを構成する論理ボリュームに関する状態遷移コマンドに応じて適切な物理デバイスを割り当てることについては開示していない。

40

【0011】

本発明は、前記従来技術の課題に鑑みて為されたものであり、その目的は、ペアを構成するボリュームに関する状態遷移コマンドに応じて各ボリュームのページに適切な物理デバイスを割り当てることができるボリューム管理装置およびストレージシステムを提供することにある。

【課題を解決するための手段】

【0012】

前記目的を達成するために、本発明は、性能が異なる複数の物理デバイスの論理的記憶

50

領域を形成する複数の論理デバイスを物理デバイスの性能に応じて階層化し、階層化された論理ボリュームを、各階層に対応したプールボリュームとして仮想化プールに格納し、正論理ボリュームに複数の正側ページを割り当てるとともに、副論理ボリュームに複数の副側ページを割り当て、仮想化プールに属する各階層のプールボリュームの記憶領域を複数のプール側ページに分割して管理し、正論理ボリュームと副論理ボリュームをペアとするペア構成に関するコマンドを受けたときに、正側ページにいずれかのプール側ページを割り当て、副側ページには正側ページに割り当てられたプール側ページと同じ階層に属するプール側ページを割り当て、コマンド発行元からペアに関する状態遷移コマンドを受けたときに、正側ページまたは副側ページのうち少なくとも一方に割り当てるプール側ページの階層を状態遷移コマンドに応じて選択することの特徴とする。

10

【発明の効果】

【0013】

本発明によれば、ペアを構成するボリュームに関する状態遷移コマンドに応じて各ボリュームのページに適切な物理デバイスを割り当てることができる。

【図面の簡単な説明】

【0014】

【図1】本発明が適用された計算機システムのハードウェアブロック図である。

【図2】図1に示した計算機システムにおいて、ストレージシステムが行う記憶領域の動的割当の動作を示す機能ブロック図である。

【図3】ストレージシステムのメモリブロック図である。

20

【図4】VDEV管理情報を説明するための構成図である。

【図5】LDEV管理情報を説明するためのブロック図である。

【図6】アドレス管理テーブルを説明するためのブロック図である。

【図7】POOL管理情報を説明するためのブロック図である。

【図8】ターゲットデバイス情報を説明するためのブロック図である。

【図9】仮想ボリュームとPSCBおよびプールボリュームの関係を説明するためのブロック図である。

【図10】仮想ボリュームとプールとの関係を説明するためのブロック図である。

【図11】仮想ボリュームとPSCBおよび階層化されたプールボリュームの関係を説明するためのブロック図である。

30

【図12】階層化されたプールボリュームに用いられるPSCBのブロック図である。

【図13A】プール作成処理を説明するためのフローチャートである。

【図13B】プール管理情報作成処理を説明するためのフローチャートである。

【図14A】仮想ボリューム作成処理を説明するためのフローチャートである。

【図14B】仮想ボリューム管理情報作成処理を説明するためのフローチャートである。

【図15】Tierの追加と削除を説明するための図である。

【図16】ペアクリエイト処理を説明するためのブロック図である。

【図17】ペアサスペンド処理を説明するためのブロック図である。

【図18】他のペアサスペンド処理を説明するためのブロック図である。

【図19】ページ移動を伴うペアサスペンド処理を説明するためのブロック図である。

40

【図20】スピンダウン指示を伴うペアサスペンド処理を説明するためのブロック図である。

【図21】ペアリシンク処理を説明するためのブロック図である。

【図22】リストア処理を説明するためのブロック図である。

【図23】クイックリシンク処理を説明するためのブロック図である。

【図24】ペア削除処理を説明するためのブロック図である。

【図25】スナップショット作成処理を説明するためのブロック図である。

【図26】スナップショット作成処理後のサスペンド処理を説明するためのブロック図である。

【図27A】ライトコマンド処理を説明するためのフローチャートである。

50

【図 27B】ライトアフタ処理を説明するためのフローチャートである。

【図 28】リードコマンド処理を説明するためのフローチャートである。

【図 29】イニシャルコピー処理を説明するためのフローチャートである。

【図 30】差分コピー処理を説明するためのフローチャートである。

【図 31】コピーアフタライト・コピーオンライト処理を説明するためのフローチャートである。

【図 32】リモートコピー処理を説明するためのブロック図である。

【図 33】リモートコピー後のペアサスペンド処理を説明するためのブロック図である。

【図 34】リモートコピー後のリバース処理を説明するためのブロック図である。

【図 35】リモートコピー後リバース処理を終了するときの処理を説明するためのブロック図である。

10

【発明を実施するための形態】

【0015】

本実施形態は、性能が異なる複数の物理デバイスの論理的記憶領域を形成する複数の論理デバイスを物理デバイスの性能に応じて階層化し、階層化された論理ボリュームを、各階層に対応したプールボリュームとして仮想化プールに格納し、正論理ボリュームに複数の正側ページを割り当てるとともに、副論理ボリュームに複数の副側ページを割り当て、仮想化プールに属する各階層のプールボリュームの記憶領域を複数のプール側ページに分割して管理し、ペア形成に関するコマンドを受けたときには、正側ページと副側ページに同じ階層に属するプール側ページを割り当て、その後、ペアに関する状態遷移コマンドを受けたときに、状態遷移コマンドでページ移動が指示されたときには、副側ページに割り当てられていたプール側ページを、正側ページに割り当てられていたプール側ページとは異なる階層のプール側ページに移動させるものである。

20

【0016】

以下、本発明の一実施形態を図面に基づいて説明する。図 1 は、本発明が適用された計算機システムのハードウェアブロック図である。計算機システムは、ホスト計算機 10 と、管理装置 20 と、これらが接続するストレージシステム 30 と、を備えている。ストレージシステム 30 は、ストレージ装置、あるいはストレージサブシステムとも呼ばれている。

【0017】

30

ホスト計算機 10 は、ストレージシステム 30 の論理的な記憶資源にアクセスする。管理装置 20 は、ストレージシステム 30 の記憶領域の構成を管理する。ストレージシステム 30 は、物理デバイス 34 に設定された記憶領域にデータを格納する。

【0018】

ホスト計算機 10 は、入力手段 110、出力手段 120、CPU (Central Processing Unit) 130、メモリ 140、ディスクアダプタ 150、ネットワークアダプタ 160 及びディスクドライブ 170 を備える。入力手段 110 は、ホスト計算機 10 を操作する管理者等から入力を受け付ける手段である。入力手段 110 は、例えば、キーボードで構成される。出力手段 120 は、ホスト計算機 10 の状態や設定項目を表示する手段である。出力手段 120 は、例えば、ディスプレイ装置で構成される。なお、ホスト計算機 10 は、1 または 2 台以上で構成される。

40

【0019】

CPU 130 は、ディスクドライブ 170 に格納されているプログラムをメモリ 140 に読み込んで、そのプログラムに規定された処理を実行する。メモリ 140 は、例えば、RAM 等で構成され、プログラムやデータ等を格納する。

【0020】

ディスクアダプタ 150 は、ストレージシステム 30 とストレージネットワーク 50 を介して接続し、ストレージシステム 30 にデータを送受信する。ストレージネットワーク 50 は、データ転送に適したプロトコル (例えば、Fibre Channel) で構成される。

【0021】

50

ネットワークアダプタ 160 は、ストレージシステム管理装置 20 又はストレージシステム 30 と管理ネットワーク 40 を介してデータを送受信する。管理ネットワーク 40 は、例えば Ethernet (登録商標) で構成される。

【0022】

ディスクドライブ 170 は、例えば、ハードディスク装置で構成され、データやプログラムを格納する。

【0023】

管理装置 20 は、入力手段 210、出力手段 220、CPU 230、メモリ 240、ネットワークアダプタ 250 及びディスクドライブ 260 を備える。

【0024】

入力手段 210 は、管理装置 20 を操作する管理者等の入力を受け付ける手段である。入力手段 210 は、例えば、キーボードで構成される。出力手段 220 は、管理装置 20 の状態や設定項目を表示する手段である。出力手段 220 は、例えば、ディスプレイ装置で構成される。

【0025】

CPU 230 は、ディスクドライブ 260 に格納されている管理プログラム (管理ソフト) をメモリ 240 に読み込んで、そのプログラムに基づくストレージシステム 30 に対する管理処理を実行する。メモリ 240 は、例えば、RAM 等で構成され、プログラムやデータ等を格納する。

【0026】

ネットワークアダプタ 250 は、ホスト計算機 10 又はストレージシステム 30 と管理ネットワーク 40 を介してデータを送受信する。

【0027】

ディスクドライブ 260 は、例えば、ハードディスク装置で構成され、データやプログラムを格納する。

【0028】

ストレージシステム 30 は、コントローラ 31、ストレージキャッシュメモリ 32、共有メモリ 33、物理デバイス (PDEV) 34、電源スイッチ 35 及び電源 36 を備える。

【0029】

コントローラ 31 は、物理デバイス (PDEV) 34 に構成された記憶領域へのデータの格納を制御する。

【0030】

ストレージキャッシュメモリ 32 は、物理デバイス (PDEV) 34 に読み書きされるデータを一時的に格納する。共有メモリ 33 は、コントローラ 31 や PDEV 34 の構成情報を格納する。物理デバイス (PDEV) 34 は、複数のディスク装置によって構成される。電源 36 は、ストレージシステム 30 の各部に電力を供給する。電源スイッチ 35 は、電源 36 からの電力の供給を ON/OFF するスイッチである。ディスク装置 (記憶デバイス) は、例えば、ハードディスクドライブで構成され、主としてユーザデータを格納する。記憶デバイスとしては、フラッシュメモリなどの半導体メモリからなるドライブでもよい。

【0031】

コントローラ 31 は、ホストアダプタ 310、ネットワークアダプタ 320、不揮発性メモリ 330、電源制御部 340、メモリ 350、プロセッサ 360、ストレージアダプタ 370 及び共有メモリアダプタ 380 によって構成される。

【0032】

ホストアダプタ 310 は、ストレージネットワーク 50 を介してホスト計算機 10 との間でデータを送受信する。ネットワークアダプタ 320 は、管理ネットワーク 40 を介してホスト計算機 10 又はストレージシステム管理装置 20 との間でデータを送受信する。

【0033】

不揮発性メモリ 330 は、ハードディスクやフラッシュメモリで構成され、コントローラ 31 で動作するプログラムや構成情報等を格納する。電源制御部 340 は、電源 36 か

10

20

30

40

50

ら供給される電力を制御する。

【0034】

メモリ350は、例えば、RAM(Random Access Memory)等で構成され、プログラムやデータ等を格納する。プロセッサ360は、不揮発性メモリ330に格納されているプログラムをメモリ350に読み込んで、そのプログラムに規定された処理を実行する。ストレージアダプタ370は、物理デバイス(PDEV)34及びストレージキャッシュメモリ32との間でデータを送受信する。共有メモリアダプタ380は、共有メモリ33との間でデータを送受信する。

【0035】

図2は、図1に示した計算機システムにおいて、ストレージシステム30が行う記憶領域の動的割当の動作を説明するための機能ブロック図である。

10

【0036】

図2において、複数の物理デバイス(PDEV)34は、例えば、パリティグループ342を構成し、パリティグループ342に属する物理デバイス(PDEV)34によって仮想デバイス(VDEV)400が構成される。仮想デバイス400は、第1種VDEVを構成し、パリティグループ342に属する物理デバイス34の論理的記憶領域を形成する複数の論理デバイス(LDEV)410によって分割される。

【0037】

各論理デバイス(LDEV)410は、第1種LDEVとして、システムプール420に格納され、ホスト計算機10のアクセス対象となる論理ボリュームまたは論理ユニット(LU)430と対応付けて管理される。すなわち、ホスト計算機10は、論理ユニット430をターゲットデバイスとしてアクセスすることができ、論理ユニット430は、論理デバイス410にホスト計算機10へのパスを定義して設定される。

20

【0038】

また、ストレージシステム30の外部に外部物理デバイス600が接続されている場合、外部物理デバイス600から仮想デバイス(VDEV)402が構成される。仮想デバイス(VDEV)402は、外部物理デバイス600の論理的記憶領域を形成する複数の論理デバイス(LDEV)412によって分割される。各論理デバイス(LDEV)412は、システムプール420に格納され、論理デバイス410と同様に、論理ボリュームまたは論理ユニット430に対応づけられている。

30

【0039】

また、各論理デバイス(LDEV)410、412は、階層管理型容量仮想化プール440に格納されたプールボリューム450に対応付けられている。仮想ボリューム460は、ホスト計算機10のアクセス対象となるターゲットデバイスを構成し、仮想デバイス(VDEV)で構成されている。この仮想デバイス(VDEV)は、第2種VDEVとして、複数の論理デバイス(LDEV)に分割され、各論理デバイス(LDEV)が、第2種LDEVとして、プールボリューム450に対応付けられている。

【0040】

この際、仮想ボリューム460の仮想記憶領域と、プールボリューム450の記憶領域および各論理デバイス410、412の記憶領域は、それぞれ複数の領域に分割され、各領域に、容量割当て単位であるページが割り当てられている。

40

【0041】

そして、複数の論理デバイス(LDEV)を有する仮想ボリューム460に対して、未割当領域への新規書込に関するアクセスが生じたときには(A1)、この新規書込アクセスに対して、仮想ボリューム460のページPが割り当てられ(A2)、仮想ボリューム460のページPに対して、プールボリューム450のページPが割り当てられ(A3)、プールボリューム450のページPに対して、論理ボリューム410のページPが割り当てられる(A4、A5)。論理デバイス410のページPには、パリティグループ342に属する物理デバイス(PDEV)34が割り当てられる(A6)。これにより、論理デバイス410のページPに対応する物理デバイス(PDEV)34にライトデータが書き込まれることになる

50

。

【0042】

ここで、物理デバイス（PDEV）34が、性能あるいはコストが異なる記憶デバイスであって、例えば、SSD（Solid State Drive）、SAS（Serial Attached SCSI）Disks（以下、SASと称することもある。）、SATA（Serial ATA）Disks（以下、SATAと称することもある。）、FC（Fibre Channel）Disks（以下、FCと称することもある。）などで構成され、これらの記憶デバイスが、その性能、例えば、リードアクセスやライトアクセスに対するレスポンスタイムに応じて複数のグループに分けられる場合、論理デバイス410、412は、各物理デバイス（PDEV）34の性能に対応して、階層化されて構築され、プールボリューム450も、各物理デバイス（PDEV）34の性能に対応して、階層化されて構築される。

10

【0043】

例えば、物理デバイス（PDEV）34が、性能が異なる3種類の物理デバイスとして、SSD、SAS、SATAで構成されていた場合、これら物理デバイス（PDEV）34は、SSDをTier0とし、SASをTier1とし、SATAをTier2として、3つのグループに分かれて階層化される。この場合、論理ボリューム410、412は、Tier0に属するSSDで構成される論理ボリュームと、Tier1に属するSASで構成される論理ボリュームと、Tier2に属するSATAで構成される論理ボリュームとして、3つの層に分かれて階層化されるとともに、プールボリューム450は、3つの層に分かれて階層化された論理ボリュームに対応して階層化されて、階層管理型容量仮想化プール440に格納されることになる。

20

【0044】

図3は、ストレージシステム30のメモリ350ブロック図である。メモリ350には、プロセッサ360によって読み込まれて実行される各種プログラムや、LDEVの設定に関する構成情報351及び容量仮想化プール440、システムプール420の設定に関するプール情報352が格納される。

【0045】

コマンド制御プログラム3501は、ホスト計算機10又はストレージシステム管理装置20からのコマンドを解釈し、そのコマンドに規定された処理を実行する。パス制御プログラム3502は、ホスト計算機10との間のパスを設定する。ホストアダプタ制御プログラム3504は、ホストアダプタ310の入出力データを制御する。構成制御プログラム3503は、ストレージシステム30の構成を制御する。ディスクI/Oプログラム3505は、物理デバイス（PDEV）34へのアクセスを制御する。

30

【0046】

ネットワーク制御プログラム3506は、ストレージネットワーク50又は管理ネットワーク40を介して送受信するデータを制御する。プール制御プログラム3507は、容量仮想化プール440やシステムプール420を設定する。電源制御プログラム3508は、電源36の供給する電力のON/OFFを制御する。キャッシュ制御プログラム3509は、ストレージキャッシュメモリ32の領域及びデータを制御する。ドライブ診断プログラム3510は、PDEV（物理デバイス34）の各ディスク装置の状態を診断する。

【0047】

構成情報351は、ストレージシステム30の仮想デバイス（VDEV）400及び論理デバイス（LDEV）410、412に関する設定を格納する。プール情報352は、容量仮想化プール440やシステムプール420に関する設定を格納する。構成情報351は、アドレス管理テーブル3511、LDEV管理情報3512、ターゲットデバイス情報3513、VDEV管理情報3514を含む。

40

【0048】

アドレス管理テーブル3511は、ターゲットデバイスとLDEVとVDEVおよび物理デバイスのアドレスのマッピング情報を格納する。アドレス管理テーブル3511は、ターゲットデバイス-LDEVマッピング情報35111、LDEV-VDEVマッピング情報35112及びVDEV-PDEVマッピング情報35113を含む。LDEV管理情報3512は、LDEV（論理デバ

50

イス410、412)に関する情報を格納する。ターゲットデバイス情報3513は、ターゲットデバイスに関する情報を格納する。VDEV管理情報3514は、VDEV(仮想デバイス400、402)に関する情報を格納する。

【0049】

階層管理型容量仮想化管理情報352は、POOL管理情報3521、POOL-VOL管理情報(容量仮想化プール用構成管理情報)3522、V-VOLDIR(容量仮想化領域管理情報)3523、PSCB(容量仮想化用プール領域管理情報)3524、SYS領域情報3525、階層管理情報3526を含む。

【0050】

POOL管理情報3521は、容量仮想化プール440の設定を格納する。POOL-VOL管理情報3522は、容量仮想化プール440のプールボリューム450の情報を格納する。V-VOLDIR 3523は、容量仮想化プール440のプールボリューム450のアドレスの割り当てを格納する。PSCB情報(容量仮想化用プール領域管理情報)3524は、容量仮想化プール440のプールボリューム450のアドレスの情報を格納する。SYS領域情報3525は、ストレージシステム30の構成情報を格納するLDEVの情報を格納する。階層管理情報3526は、容量仮想化プール440のプールボリューム450の階層に関する情報を格納する。

【0051】

図4は、VDEV管理情報の説明図である。VDEV管理情報3514は、VDEV固有情報35141から構成される。VDEV固有情報35141は、VDEV番号(VDEV#)35142、エミュレーションタイプ35143、総サイズ35144、残サイズ35145、デバイス属性35146、デバイス状態35147、設定LDEV数35148、LDEV番号35149、先頭VDEV-SLOT#35150及び終了VDEV-SLOT#35151で構成される。

【0052】

VDEV#35142は、VDEV(仮想デバイス400、402)の識別子である。エミュレーションタイプ35143は、VDEVのエミュレーションタイプの識別子である。総サイズ35144は、VDEVに設定されている総サイズである。残サイズ35145は、VDEVの未使用の領域のサイズである。

【0053】

デバイス属性35146は、VDEVに定義されている属性の識別子である。そのVDEVが第1種VDEV(論理ユニット430に割り当てられるVDEV)である場合には、第1種VDEVを示す識別子が格納され、そのVDEVが第2種VDEVであり、仮想ボリューム460に設定されている場合には、第2種VDEVを示す識別子が格納される。

【0054】

デバイス状態35147は、VDEVの状態を示す識別子である。VDEVの状態は、正常、閉塞、障害閉塞等がある。閉塞は、パンク閉塞など、障害の発生意外の要因によって閉塞されていることを示す。障害閉塞は、デバイスの何れかに障害が発生しておりそのために閉塞されていることを示す。

【0055】

設定LDEV数35148は、VDEVに設定されているLDEVの総数である。LDEV番号35149は、VDEVに設定されているLDEVの番号が格納される。先頭VDEV-SLOT#35150は、設定されているLDEVの物理的な先頭のスロット番号の識別子である。

【0056】

終了VDEV-SLOT#35151は、設定されているLDEVの物理的な最終のスロット番号である。このLDEV番号35149、先頭VDEV-SLOT#35150及び終了VDEV-SLOT#35151は、LDEV数と同じ数だけLDEV番号毎に設定される。

【0057】

図5は、LDEV管理情報の説明図である。LDEV管理情報3512は、VDEV固有情報35121から構成される。LDEV固有情報35121は、LDEV番号(LDEV#)35122、エミュレーションタイプ35123、サイズ35124、先頭スロット番号35125、終了

10

20

30

40

50

スロット番号 3 5 1 2 6、パス定義情報 3 5 1 2 7、デバイス属性 3 5 1 2 8、デバイス状態 3 5 1 2 9、プログラム使用状況 3 5 1 3 0 0 及び POOL-ID 3 5 1 3 0 1 で構成される。

【0058】

LDEV# 3 5 1 2 2 は、LDEV（論理デバイス 4 1 0、4 1 2）の識別子である。エミュレーションタイプ 3 5 1 2 3 は、LDEV のエミュレーションタイプの識別子である。サイズ 3 5 1 2 4 は、LDEV に設定されている総サイズである。

【0059】

先頭スロット番号 3 5 1 2 5 は、設定された LDEV の先頭のスロット番号の識別子である。終了スロット番号 3 5 1 2 6 は、設定された LDEV の最終のスロット番号である。パス定義情報 3 5 1 2 7 は、ホスト計算機 1 0 に定義されたパスの識別子である。

10

【0060】

デバイス属性 3 5 1 2 8 は、LDEV の属性の識別子である。LDEV が第 1 種 LDEV である場合には、第 1 種 LDEV を示す識別子が格納され、LDEV が第 2 種 LDEV である場合には、第 2 種 LDEV を示す識別子が格納される。また、LDEV が容量仮想化プール 4 4 0 に設定されている場合は、プール属性を示す識別子が格納される。

【0061】

デバイス状態 3 5 1 2 9 は、その LDEV が所属する VDEV の状態を示す識別子である。VDEV の状態は、正常、閉塞、障害閉塞等がある。閉塞は、パンク閉塞など、障害の発生意外の要因によって閉塞されていることを示す。障害閉塞は、デバイスの何れかに障害が発生しておりそのために閉塞されていることを示す。

20

【0062】

プログラム使用状況 3 5 1 3 0 0 は、LDEV が何れかのプログラムによって処理中である場合に、そのプログラムの識別子が格納される。POOL-ID 3 5 1 3 0 1 は、LDEV が容量仮想化プール 4 4 0 に設定されている場合に、その識別子が格納される。

【0063】

図 6 は、アドレス管理テーブルの説明図である。アドレス管理テーブル 3 5 1 1 は、ターゲットデバイスと LDEV と VDEV および物理デバイス 3 4 のアドレスのマッピング情報を格納する。アドレス管理テーブル 3 5 1 1 は、ターゲットデバイス－LDEV マッピング情報 3 5 1 1 1、LDEV－VDEV マッピング情報 3 5 1 1 2 及び VDEV－PDEV マッピング情報 3 5 1 1 3 を含む。

30

【0064】

ターゲットデバイス－LDEV マッピング情報 3 5 1 1 1 は、ターゲットデバイスのアドレスと LDEV のアドレスとの対応が格納される。LDEV－VDEV マッピング情報 3 5 1 1 2 は、LDEV（第 1 種 LDEV、第 2 種 LDEV）のアドレスと VDEV（第 1 種 VDEV、第 2 種 VDEV）のアドレスが格納される。

【0065】

VDEV－PDEV マッピング情報 3 5 1 1 3 は、VDEV のアドレスとその RAID グループ番号（又はパリティグループ）と PDEV（物理デバイス 3 4）のアドレスが格納される。

【0066】

40

ストレージシステム 3 0 は、このアドレス管理テーブル 3 5 1 1 を参照することによって、ターゲットデバイスのアドレスがどの LDEV のどのアドレスに対応するかを知ることができる。また、LDEV のアドレスがどの VDEV のどのアドレスに対応するかを知ることができる。また、VDEV のアドレスがどの RAID グループに属しており、どの PDEV のどのアドレスに対応するかを知ることができる。

【0067】

図 7 は、POOL 管理情報の説明図である。POOL 管理情報 3 5 2 1 は、POOL 固有情報 3 5 2 1 1 から構成される。POOL 固有情報 3 5 2 1 1 は、POOL-ID 3 5 2 1 2、属性／用途 3 5 2 1 3、エミュレーションタイプ 3 5 2 1 4、容量 3 5 2 1 5、空き容量 3 5 2 1 6、閾値 3 5 2 1 7、状態 3 5 2 1 8、POOL-VOL 数 3 5 2 1 9、POOL-VOL デバイス番号リスト 3

50

5 2 2 0、POOLを利用しているデバイス数 3 5 2 2 1 及びPOOLを利用しているデバイス番号 3 5 2 2 2 で構成される。

【0 0 6 8】

POOL-ID 3 5 2 1 2 は、POOL（容量仮想化プール 4 4 0）の識別子である。属性／用途 3 5 2 1 3 は、その容量仮想化プール 4 4 0 の属性及び用途を示す識別子である。用途は、例えば、スナップショットやSYS領域など、運用形態の用途である。

【0 0 6 9】

エミュレーションタイプ 3 5 2 1 4 は、容量仮想化プール 4 4 0 のエミュレーションタイプの識別子である。容量 3 5 2 1 5 は、容量仮想化プール 4 4 0 の総容量である、空き容量 3 5 2 1 6 は、容量仮想化プール 4 4 0 の未使用の領域のサイズである。

10

【0 0 7 0】

閾値 3 5 2 1 7 は、容量仮想化プール 4 4 0 が許容する最大のデータ格納容量である。状態 3 5 2 1 8 は、容量仮想化プール 4 4 0 の現在の状態である。例えば、定義中、拡張中、有効等である。POOL-VOL数 3 5 2 1 9 は、容量仮想化プール 4 4 0 として設定されているLDEVの総数である。

【0 0 7 1】

POOL-VOLデバイス番号リスト 3 5 2 2 0 は、容量仮想化プール 4 4 0 として設定されているLDEV番号の一覧である。POOLを利用しているデバイス数 3 5 2 2 1 は、容量仮想化プール 4 4 0 のLDEVが関連付けられている第 2 種LDEVの数である。POOLを利用しているデバイス番号 3 5 2 2 2 は、容量仮想化プール 4 4 0 のLDEVが関連付けられている第 2 種LDEV番号の一覧である。

20

【0 0 7 2】

図 8 は、ターゲットデバイス情報の説明図である。ターゲットデバイス情報 3 5 1 3 は、ターゲットデバイス固有情報 3 5 1 3 1 から構成される。ターゲットデバイス固有情報 3 5 1 3 1 は、ポート # 3 5 1 3 2、ターゲットデバイス番号 # 3 5 1 3 3、LDEV数 3 5 1 3 4、LDEV # 3 5 1 3 5、属性 3 5 1 3 6、状態 3 5 1 3 7、容量 3 5 1 3 8、許可ホスト情報 3 5 1 3 9 で構成される。

【0 0 7 3】

ポート # 3 5 1 3 2 は、ホストアダプタ 3 1 0 のポート番号である。ターゲットデバイス番号 # 3 5 1 3 3 は、ターゲットデバイスの番号、例えば、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 の番号を示す。LDEV数 3 5 1 3 4 は、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 を構成するLDEV（第 2 種LDEV）の数を示す。LDEV # 3 5 1 3 5 は、例えば、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 を構成するLDEV（第 2 種LDEV）の番号を示す。

30

【0 0 7 4】

属性 3 5 1 3 6 は、例えば、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 の属性を示す。状態 3 5 1 3 7 は、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 の状態、例えば、アクセス対象の状態にあることを示す。容量 3 5 1 3 8 は、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 を構成するLDEV（第 2 種LDEV）の容量を示す。許可ホスト情報 3 5 1 3 9 は、仮想ボリューム 4 6 0 をターゲットデバイスとしたときには、仮想ボリューム 4 6 0 に対するアクセスの許可の有無を示す情報である。

40

【0 0 7 5】

図 9 は、仮想ボリュームとPSCBおよびプールボリュームの関係を説明するためのブロック図である。VVOL-DIR 3 5 2 3 は、仮想ボリューム 4 6 0 の記憶領域である第 2 種LDEVの構成の情報である。PSCB（POOL-SLOT Control Block） 3 5 2 4 は、プールボリューム 4 5 0 に設定された第 1 種LDEVの構成の情報である。

【0 0 7 6】

ストレージシステム 3 0 は、物理デバイス（PDEV） 3 4 からRAID構成によって第 1 種VD

50

EV（仮想デバイス400）を構成する。この第1種VDEVを、記憶領域である第1種LDEV（論理デバイス410）に分割する。第1種LDEVがシステムプール420に設定される。このシステムプール420に設定された第1種LDEVによって構成されるボリュームをプールボリューム（POOL-VOL）450とする。

【0077】

また、ストレージシステム30は、仮想ボリューム（VVOL）460を設定し、第2種VDEV（仮想デバイス）を構成する。この第2種VDEVを記憶領域である第2種LDEV（論理デバイス）に分割する。

【0078】

ストレージシステム30は、仮想ボリューム460の第2種LDEVを、プールボリューム450の第1種LDEV（論理デバイス410）に割り当てる。これによって、ホスト計算機10がアクセスする仮想ボリューム460の記憶領域が、物理デバイス（PDEV）34から構成された第1種LDEV（論理デバイス410）に設定される。

【0079】

仮想ボリューム460の構成は、VVOL-DIR3523に格納される。VVOL-DIR3523は、LDEV番号（LDEV#）35231及びエントリ35232によって構成される。

【0080】

LDEV番号（LDEV#）35231は、第2種LDEV（論理デバイス）の識別子である。エントリ35232は、第2種LDEVの構成情報である。このエントリ35232は、第2種LDEVアドレス35233及びPSCBポインタ35234から構成される。

【0081】

第2種LDEVアドレス35233には、仮想ボリューム460の第2種LDEV（論理デバイス）のアドレスが格納される。

【0082】

PSCBポインタ35234には、第2種LDEV（論理デバイス）がプールボリューム450の第1種LDEV（論理デバイス410）に割り当てられている場合に、その第1種LDEV（論理デバイス410）の領域のポインタが格納される。なお、初期状態では、第2種LDEVは、第1種LDEVに割り当てられていないので、PSCBポインタ35234には、「NULL」が格納される。

【0083】

PSCB（Pool-Slot Control Block）3524は、プールボリューム450に設定されている第1種LDEV（論理デバイス410）の情報である。このPSCB3524は、プールボリューム450に設定されている第1種LDEV（論理デバイス410）のスロット毎に設定される。

【0084】

PSCB3524は、実ページ番号（実ページ#）35242、プールボリューム450のアドレス35243、ポインタA35244及びポインタB35245から構成される。

【0085】

実ページ番号（実ページ#）35242は、プールボリューム450における第1種LDEV（論理デバイス410）の実ページ番号である。プールボリューム450のアドレス35243は、プールボリューム450における第1種LDEV（論理デバイス410）のアドレスである。

【0086】

ポインタA35244及びポインタB35245は、プールボリューム450内の第1種LDEV（論理デバイス410）の前後のスロットの識別子である。

【0087】

また、プールボリューム450の領域のうち、未使用の領域は、その先頭がフリーPSCBキュー35240で示される。フリーPSCBキュー35240は、次のスロットを示すPSCB3524へのポインタを含む。

【0088】

10

20

30

40

50

ストレージシステム 30 は、フリーPSCBキュー 35240 に示されたポインタを参照して、次のPSCB 3524 を得る。さらに、次のPSCB 3524 のポインタB35245 を参照して、段階的にPSCB 3524 を辿る。そして、その未使用の領域の最終のスロットに対応するPSCB 3524 を得る。

【0089】

この最後のPSCB 3524 のポインタB35245 は、フリーPSCBキュー 35240 である。ストレージシステム 30 は、フリーPSCBキュー 35240 を辿り、PSCB 3524 のポインタによって連結された集合によって、容量仮想化プール 440 のプールボリューム 450 の未使用の領域を知ることができる。

【0090】

ストレージシステム 30 は、プールボリューム 450 に設定された第1種LDEV（論理デバイス 410）に対応するPSCB 3524 を設定する。具体的には、プールボリューム 450 に設定された第1種LDEV（論理デバイス 410）の各スロットに対応するPSCB 3524 を設定し、さらにフリーPSCBキュー 35240 を設定する。初期状態では、プールボリューム 450 は全て未使用であるため、フリーPSCBキュー 35240 によって連結される集合は、プールボリューム 450 に設定された第1種LDEV（論理デバイス 410）の全ての領域に対応する。

【0091】

そして、ストレージシステム 30 は、このプールボリューム 450 の領域を使用する場合に、必要なスロット分のPSCB 3524 を第2種LDEVであるVVOL-DIR 3523 に割り当てることで、当該領域が使用可能となる。複数のスロットの集合がページに相当する。ページは、複数のPSCBから特定される。ホスト計算機 10 から仮想ボリューム 460 へのアクセス、仮想ボリューム 460 のアクセス領域に対するプールボリューム 450 からの記憶領域の割り当ては、ページ単位で実行される。

【0092】

具体的には、ストレージシステム 30 は、フリーPSCBキュー 35240 を参照する。そして、第2種LDEVに割り当てる必要な領域分（ページ）のPSCB 3524 を取得する。この取得したPSCB 3524 を、それぞれVVOL-DIR 3523 のエントリに割り当てる。すなわち、VVOL-DIR 3523 の各エントリのPSCBポインタ 35234 に、対応するPSCB 3524 を示すポインタを格納する。なお、割り当て済みのPSCB 3524 は、フリーPSCBキュー 35240 の連結から外す。

【0093】

これによって、第2種LDEVの各ページ（スロット）が、VVOL-DIR 3523 の各エントリのPSCBポインタ 35234 で示されるPSCB 3424 に割り当てられる。PSCB 3524 は、第1種LDEV（論理デバイス 410）のスロットに対応しているので、結果として、第2種LDEVが第1種LDEV（論理デバイス 410）に割り当てられ、ホスト計算機 10 のアクセス対象である仮想ボリューム 460 が、ターゲットデバイスとして使用可能となる。

【0094】

図10は、仮想ボリュームとプールとの関係を説明するためのブロック図である。プール管理テーブル 35221 は、仮想ボリューム 460 の第2種LDEVとプールボリューム 450 の第1種LDEV（論理デバイス 410）との関係を管理するためのテーブルを構成する。

【0095】

プール管理テーブル 35221 には、仮想ボリューム 460 の第2種LDEVが、プールボリューム 450 に割り当てられていないときには、「FREE」の情報が、第2種LDEVに対応して格納され、仮想ボリューム 460 の第2種LDEVが、プールボリューム 450 に割り当てられたときには、「ALLOCATED」の情報が、第2種LDEVに対応して格納される。

【0096】

図11は、仮想ボリュームとPSCBおよび階層化されたプールボリュームの関係を説明するためのブロック図である。

10

20

30

40

50

【0097】

階層化されたプールボリューム450が容量仮想化プール440に格納された場合、階層毎にフリーPSCBをキュー管理し、1つのLDEV35231に対して、複数の階層にページ毎に領域を割り当てることとしている。この際、ページ単位の方法をPSCB単位の情報として管理する。PSCBポインタ35234は、ページデータがプールボリューム450内のどの領域に格納されているかを示す情報である。階層（Tier）番号35235は、プールボリューム450内のどの階層にページデータが格納されているかを示す階層の番号である。

【0098】

ここで、プールボリュームとして、例えば、Tier0に属するプールボリュームをプールボリューム452と、Tier1に属するプールボリュームをプールボリューム454とした場合、VVOL-DIR3523におけるLDEV1のあるページは、Tier0に属するプールボリューム452の領域が割り当てられ、PSCB3524のうちPSCB1にてアドレスが管理されているときに（S351）、このページをTier1に移行する場合、PSCB3524におけるPSCB3をフリーPSCBキューから取得し、PSCB1の内容をPSCB3へコピーする処理（S352）がコントローラ31によって実行される。

10

【0099】

この後、VVOL-DIR3523におけるLDEV1のあるページのマッピングをPSCB1からPSCB3へ変更し、PSCB1をフリーキューへ戻すために、PSCB1をフリーPSCBキュー35240へ接続する処理（S353）がコントローラ31によって実行される。

20

【0100】

なお、ページ単位の情報としては、他に、モニタから採取された情報がある。この場合、そのページへのアクセス頻度を定期的にモニタして管理することになる。また、プールボリューム450内に格納されたデータは、プール側のページ毎に情報を付け、どの仮想ボリュームのどのアドレスのデータが割り当てられているかを検索できる情報であってもよい。

【0101】

図12は、階層化されたプールボリュームに用いられるPSCBのブロック図である。

【0102】

容量仮想化プール440に階層化されたプールボリューム450を格納した場合、例えば、5階層に分けてプールボリューム450を容量仮想化プール440に格納したときには、Tier0のPSCB2524として、PSCB1401Fが用いられ、Tier1のPSCB2524として、PSCB1401Rが用いられ、Tier2のPSCB2524として、PSCB1401Aが用いられ、Tier3のPSCB2524として、PSCB1401Dが用いられ、Tier4のPSCB2524として、PSCB1401Uが用いられる。

30

【0103】

次に、プール作成処理を図13Aのフローチャートに従って説明する。ストレージシステム管理装置20のCPU230は、管理プログラム（管理ソフト）を起動し、プールID（Identifier）、閾値、用途、プールボリューム数、論理デバイス番号（LDEV#）、リザーブ（ターゲットデバイスとは異なる予備のデバイス）の割合を指定してプール作成を指示し（S11）、ストレージシステム30にプール作成指示を送信する（S12）。

40

【0104】

この後、ストレージシステム30のネットワークアダプタ320がプール作成指示を受信すると、コントローラ31は、プロセッサ360がコマンド制御プログラム3501を起動して、以下の処理を開始する。

まず、プロセッサ360は、プール作成指示を受領し（S13）、プール作成指示に付加されたコマンドの内容をチェックし（S14）、無効な内容があるか否かを判定し（S15）、コマンドに無効な内容がないときにはプール作成処理に移行し（S16）、プール作成処理に成功したか否かを判定し（S17）、プール作成処理に成功したときには、プール作成に成功した旨の応答をストレージシステム管理装置20に送信する（S18）。

50

【0105】

一方、プロセッサ360は、ステップS15において、コマンドの内容に無効な内容があると判定したときには、ステップS17において、プール作成処理に失敗したときとともに、プール作成に失敗した原因と、指示に対するエラー応答をストレージシステム管理装置20に送信し(S19)、このルーチンでの処理を終了する。

【0106】

ストレージシステム管理装置20のCPU230は、ストレージシステム30からプール作成処理の応答を受領すると(S20)、受領した応答の内容を基にプール作成処理が成功したか否かを判定し(S21)、プール作成処理に成功したときには、指示されたプールの作成が成功したことをログして、外部I/F(InterFace)やGUI(Graphical User Interface)に情報を提供し(S22)、このルーチンでの処理を終了する。 10

【0107】

ステップS21において、プール作成処理に失敗したと判定したときには、CPU230は、指示されたプールの作成が失敗したこと、及びその原因をログして、外部I/FやGUIに情報を提供し(S23)、このルーチンでの処理を終了する。

【0108】

次に、プール管理情報作成処理を図13Bフローチャートに従って説明する。

【0109】

この処理は、ステップS16におけるプール作成処理で実行するためのプール管理情報作成処理であって、プロセッサ360が、プール制御プログラム3507を起動することによって開始される。 20

【0110】

プロセッサ360は、プールIDのチェック、例えば、プールIDが有効で且つ未使用であるかをチェックし(S31)、プール(容量仮想化プール440)の状態を未定義から定義中に変更し(S32)、指示された論理デバイス(LDEV)の状態をチェックし(S33)、チェック対象の論理デバイスは、使用中、障害、閉塞、フォーマット中などのときは使用不可であるため、プール作成に失敗したとして処理を終了し、それ以外の論理デバイスについては以下の処理を実行する(S34)。

【0111】

次に、プロセッサ360は、LDEV情報にプールボリューム情報を設定し(S35)、プール管理情報の容量、空き容量、閾値、プールボリューム数を更新し(S36)、次に、プール空き管理キューを作成し(S37)、さらにリザーブキューを作成し(S38)、その後、プールの状態を有効に変更し(S39)、このルーチンでの処理を終了する。 30

【0112】

次に、仮想ボリューム(VOL)作成処理を図14Aフローチャートに従って説明する。ストレージシステム管理装置20のCPU230は、管理プログラム(管理ソフト)を起動し、処理を開始する。CPU230は、接続ホスト情報、ポート番号、LUN、論理デバイス番号(LDEV#)、エミュレーションタイプ、容量、仮想デバイス番号(VDEV#)、プールID、閾値、用途QoS(Quality of Services) LANEなどを指定して仮想ボリュームの作成を指示し(S41)、ストレージシステム30に仮想ボリュームの作成指示を送信する(S42)。 40

【0113】

一方、ストレージシステム30のネットワークアダプタ320が仮想ボリューム作成指示を受信すると、プロセッサ360によってコマンド制御プログラム3501が起動され、以下の処理が開始される。

【0114】

まず、プロセッサ360は、仮想ボリューム作成指示を受領し(S43)、仮想ボリューム作成指示に付加されたコマンドの内容をチェックし(S44)、コマンドの内容が無効な内容か否かを判定し(S45)、無効な内容がないときには仮想ボリューム作成処理を実行し(S46)、仮想ボリューム作成処理に成功したか否かを判定する(S47)。 50

【0115】

ステップS47で仮想ボリュームの作成に成功したと判定したときには、プロセッサ360は、パス定義処理、例えば、ポートに仮想ボリュームを論理ユニット（LU）として定義し、ホストグループや論理ユニット番号（LUN）セキュリティも設定し（S48）、仮想ボリュームの作成に成功した旨の応答をストレージシステム管理装置20に送信する（S49）。

【0116】

一方、プロセッサ360は、ステップS45において、コマンドの内容に無効な内容があると判定したときには、ステップS47において仮想ボリュームの作成に失敗したときとともに、仮想ボリュームの作成に失敗した原因と、指示に対してエラー応答をストレージ装置管理装置20に送信し（S50）、このルーチンでの処理を終了する。

10

【0117】

ストレージ装置管理装置20のCPU230は、ストレージシステム30からの仮想ボリューム作成応答を受領したときには（S51）、受領した応答の内容を基に仮想ボリュームの作成に成功したか否かを判定し（S52）、仮想ボリュームの作成処理に成功したと判定したときには、指示された仮想ボリュームの作成が成功したことをログして、外部I/FやGUIに情報を提供し（S53）、このルーチンでの処理を終了する。

【0118】

仮想ボリュームの作成処理に失敗したと判定したときには、CPU230は、指示された仮想ボリュームの作成が失敗したこと及びその原因をログして、外部I/FやGUIに情報を提供し（S54）、このルーチンでの処理を終了する。

20

【0119】

次に、仮想ボリューム管理情報作成処理を図14Bのフローチャートに従って説明する。この処理は、図14AのステップS46における仮想ボリューム作成処理で実行するための処理であって、コントローラ31のプロセッサ360が、プール制御プログラム3507を起動することによって開始される。

【0120】

プロセッサ360は、仮想ボリュームのチェックとして、例えば、仮想ボリュームが有効で且つ未使用であるかをチェックするとともに、仮想ボリュームの閾値などをチェックし（S61）、仮想ボリュームに属する仮想デバイス（VDEV）の番号が正しいか否かをチェックし（S62）、仮想デバイスの番号が正しいときには、仮想デバイス情報作成処理として、例えば、プール番号、容量、エミュレーションタイプ、RAIDレベル、状態（ノーマル）などを作成し（S63）、その後、論理デバイス（LDEV）番号が正しいか否かを判定する（S64）。

30

【0121】

プロセッサ360は、ステップS64で論理デバイス（LDEV）の番号が正しいと判定したときには、次に、論理デバイス情報作成処理として、例えば、プール番号、容量、属性（仮想ボリュームの属性）、状態（ノーマル）などの情報を作成し（S65）、その後、マップ情報作成処理として、特定のデータページを参照して、マップ情報を作成し（S66）、仮想ボリューム作成処理が正常に終了した旨の処理を実行し（S67）、このルーチンでの処理を終了する。

40

【0122】

一方、プロセッサ360は、ステップS62において、仮想デバイス（VDEV）の番号が正しくないと判定したとき、あるいはステップS64において、論理デバイス（LDEV）の番号が正しくないと判定したときには、仮想ボリュームの作成処理が異常で終了した旨の処理を実行し（S68）、このルーチンでの処理を終了する。

【0123】

次に、図15に、物理デバイス（PDEV）34を追加または削除するときの管理構成を示す。コントローラ31は、容量仮想化プール440に格納されたプールボリューム450を階層化して管理するに際して、Tier0に属する物理デバイス34として、SASを用い、T

50

ier 1 に属する物理デバイス 3 4 として SATA を用いているときには、(a) に示すように、これらの物理デバイス 3 4 よりも性能レベルの高い物理デバイス 3 4 として、SSD が追加された場合、Tier 0 に SSD を設定し、Tier 1 に SAS を設定し、Tier 2 に SATA を設定し、これら物理デバイス 3 4 の論理デバイス (LDEV) に対応したプールボリューム 4 5 0 を 3 つの階層に分けて階層化して管理することとしている。

【0124】

逆に、(b) に示すように、Tier 1 に属する SAS が削除された場合、コントローラ 3 1 は、Tier 2 に設定されていた SATA を、Tier 1 に繰り上げて管理することとしている。この場合、プールボリューム 4 5 0 は、2 つの階層に階層化されて管理されることになる。

【0125】

ここで、以下、ストレージシステム 3 0 において、当該ストレージシステム 3 0 内でデータをコピーすることをローカルコピーと称し、ストレージシステム 3 0 が他のストレージシステムに接続されている場合、ストレージシステム 3 0 内の論理ボリュームのデータの複製を他のストレージシステムにコピーすることをリモートコピーと称することとする。

【0126】

また、ローカルコピー及びリモートコピーにおいて、コピー元の論理ボリュームとコピー先の論理ボリュームの組をペアと称し、ペアを構成する論理ボリュームのうち、例えば、ホスト計算機 1 0 のアクセス対象となる論理ボリュームを正論理ボリューム (P-VOL) と称し、正論理ボリューム (P-VOL) のペアの対象となる論理ボリュームを副ボリューム (S-VOL) と称することとする。

【0127】

新たなペアを形成するときには、形成コピー処理が実行される。この場合、新たに形成されるペアのコピー元 (正論理ボリューム) とコピー先 (副ボリューム) が指定される。この時点では、副ボリュームには、正論理ボリュームと同じデータは格納されていないが、その後、コピー元 (正論理ボリューム) のデータを順次コピー先 (副ボリューム) にコピーする、形成コピーを実行することで、副ボリューム (S-VOL) には、正論理ボリューム (P-VOL) と同一のデータが格納される。形成コピーを実行している状態は、「形成コピー中」状態である。

【0128】

データがコピーされ、正論理ボリューム (P-VOL) と副ボリューム (S-VOL) に同一のデータが格納された状態、すなわち、データが二重化された状態を「同期」状態という。「同期」状態において、正論理ボリューム (P-VOL) のデータが更新されると、更新されたデータは、副ボリューム (S-VOL) にコピーされる。このコピーを更新コピーという。

【0129】

データのコピーが停止した状態は、「サスペンド (Suspend)」状態である。「サスペンド (Suspend)」状態のペアにおいては、副ボリューム (S-VOL) には、正論理ボリューム (P-VOL) とは異なるデータが格納されている場合もある。「サスペンド (Suspend)」状態及び「形成コピー中」状態においては、差分ビットマップを用いた差分管理が実行される。

【0130】

例えば、正論理ボリューム (P-VOL) が属するペアが「同期」状態であるとき、差分ビットマップの全ビットは「OFF (無効) または「0」である。正論理ボリューム (P-VOL) が属するペアが「サスペンド」状態であるときに、ホスト計算機 1 0 からのアクセスにより、正論理ボリューム (P-VOL) のデータが更新されたときには、そのデータが書き込まれた記憶領域に対応するビットは、「ON (有効) または「1」に更新される。

【0131】

一方、ペアに関連する状態遷移コマンドとして、「スプリット (SPRIT)」のコマンドが発行されたときには、「同期」状態のペアを「サスペンド」状態に変更するための処理が実行される。スプリットの処理が実行されると、正論理ボリューム (P-VOL) から副ボ

10

20

30

40

50

リューム (S-VOL) へのデータのコピーは実行されない。

【0132】

また、状態遷移コマンドとして、「リシンク (RESYNC)」のコマンドが発行されたときには、「サスペンド」状態のペアを「同期」状態に変更するための処理が実行される。リシンクの処理が実行されると、正論理ボリューム (P-VOL) に格納されたデータのうち、少なくとも副ボリューム (S-VOL) に格納されたデータとは異なるデータが副ボリューム (S-VOL) にコピーされる。この結果、副ボリューム (S-VOL) には、正論理ボリューム (P-VOL) と同一のデータが格納される。

【0133】

状態遷移コマンドとして、「リストア (RESTORE)」または「リバースリシンク (REVERSE-RESYNC)」のコマンドが発行されたときには、「サスペンド」状態のペアを「同期」状態に変更するための処理が実行される。リストアまたはリバースリシンクの処理が実行されると、副ボリューム (S-VOL) に格納されたデータのうち、少なくとも正論理ボリューム (P-VOL) に格納されたデータとは異なるデータが正論理ボリューム (P-VOL) にコピーされる。この結果、正論理ボリューム (P-VOL) には、副ボリューム (S-VOL) と同一のデータが格納される。

【0134】

次に、ホスト計算機 10 からストレージシステム 30 に対して、ペアに関するコマンドが発行されたときの処理を図 16 乃至図 26 に従って説明する。まず、コントローラ 31 は、図 16 に示すように、仮想ボリューム 460 を構成するボリュームであって、ホスト計算機 10 のアクセス対象となる運用ボリュームを正論理ボリューム (PVOL) 462 とし、正論理ボリューム 462 のペアの対象となる仮想ボリュームを副論理ボリューム (SVOL) 464 として管理し、正論理ボリューム 462 の仮想記憶領域に複数の正側ページ P1、P2、・・・を割り当て、副論理ボリューム 464 の仮想記憶領域に複数の副側ページ S1、S2、・・・を割り当てて管理することとしている。

【0135】

また、コントローラ 31 は、容量仮想化プール 440 に格納されるプールボリュームを 3 つの階層に階層化して管理することとしている。例えば、Tier 0 に属する SSD で構成される論理デバイス (LEDV) 410 に対応したプールボリュームをプールボリューム 452 とし、Tier 1 に属する SAS で構成される論理デバイス (LDEV) 410 に対応したプールボリュームをプールボリューム 454 とし、Tier 2 に属する SATA で構成される論理デバイス (LDEV) 410 に対応したプールボリュームをプールボリューム 456 として、階層化して管理し、プールボリューム 452 に複数のプール側ページ P01、P02、・・・を割り当て、プールボリューム 454 に複数のプール側ページ P11、P12、・・・を割り当て、プールボリューム 456 に複数のプール側ページ (P21、P22、・・・) を割り当てることとしている。

【0136】

ここで、ホスト計算機 10 からペアに関するコマンドとして、正論理ボリューム 462 と副論理ボリューム 464 とをペアとするペア形成 (ペアクリエイト) に関するコマンドをコントローラ 31 が受けたときには、コントローラ 31 は、コマンドで指定されたページ、例えば、正側ページ P1、P2 にプール側ページ P01、P11 を割り当てるとともに、副論理ボリューム 464 の副側ページ S1、S2 にプール側ボリューム P02、P12 を割り当てる。

【0137】

すなわち、正側ページ P1 と副側ページ S1 はペアとなるページであり、正側ページ P2 と副側ページ S2 はペアとなるページであるため、正側ページ P1、P2 とそれぞれペアとなる副側ページ S1、S2 には、正側ページ P1、P2 に割り当てられているプール側ページと同じ階層に属するプール側ページを割り当てる。

【0138】

この後、正論理ボリューム 462 と副論理ボリューム 464 に対するリードアクセスま

10

20

30

40

50

たはライトアクセスが禁止された状態となり、正側ページP1、P2のデータが、更新前のデータとして、それぞれ副側ページS1、S2にコピーされる。これにより、正側ページP1、P2と副側ページS1、S2のデータの内容が同一となる。

【0139】

次に、ペアクリエイトに関するコマンドを受けた後、ペアに関する状態遷移コマンドとして、ペアサスペンドのコマンドを受けたときには、図17に示すように、コントローラ31は、正論理ボリューム462と副論理ボリューム464との間の更新コピーを停止する。このとき、ホスト計算機10は、正論理ボリューム462または副論理ボリューム462に対して、リードアクセスまたはライトアクセスが可能となる。

【0140】

ペアサスペンドのコマンドを受けたときには、副側ページS1、S2のデータを、Tier0のSSDやTier1のSASに格納しておく必要はないので、SSDやSASよりも低コストのSATAへの移動が、ペアサスペンドのコマンドに指示されているときには、コントローラ31は、プールボリューム452のプール側ページP02をプールボリューム456に移動して、プール側ページP21とし、プールボリューム454のプール側ページP12をプールボリューム456に移動して、プール側ページP22とし、副側ページS1とプール側ページP21とを対応付け、副側ページS2とプール側ページP22とを対応付ける。

【0141】

これにより、状態遷移コマンドに応じてTier0に属するプール側ページP02とTier1に属するプール側ページP12をそれぞれTier2に属するプール側ページP21、P22に移動させることができる。この場合、ホスト計算機10が、副論理ボリューム464の副側ページS1、S2をライトアクセスしたときには、ライトデータがSATAに格納されることになる。

【0142】

また、ホスト計算機10は、ストレージシステム30に対して、ペアクリエイトのコマンドを発行した後、ペアサスペンドのコマンドを発行したときには、ストレージシステム30の正論理ボリューム462と副論理ボリューム464に対して、それぞれリードアクセスまたはライトアクセスが可能となる。

【0143】

この際、図18に示すように、ペアサスペンドのコマンドに、外部装置、例えば、テープ装置70に対するバックアップが指示されているときには、コントローラ31は、副側ページS1、S2に格納されているデータをテープ装置70にバックアップするための処理を実行する。

【0144】

次に、状態遷移コマンドとして、ペアサスペンドのコマンドを受けたときに、ペアサスペンドのコマンドにページ移動指示が含まれているときには、コントローラ31は、図19に示すように、ページ移動指示に従ってページ移動の処理を実行する。

【0145】

例えば、コントローラ31は、プールボリューム452のプール側ページP02をプールボリューム456に移動して、プール側ページP21とし、プールボリューム454のプール側ページP12をプールボリューム456に移動して、プール側ページP22とし、副側ページS1とプール側ページP21とを対応付け、副側ページS2とプール側ページP22とを対応付ける。

【0146】

これにより、状態遷移コマンドに応じてTier0に属するプール側ページP02とTier1に属するプール側ページP12をそれぞれTier2に属するプール側ページP21、P22に移動させることができる。

【0147】

次に、状態遷移コマンドとして、ペアサスペンドのコマンドを受けた後、スピンドウン指示を含むコマンドとして、外部物理デバイス600に属する物理デバイスをスピンドウ

10

20

30

40

50

ンさせるコマンドを受けたときには、コントローラ 31 は、図 20 に示すように、コマンドの指示に従って処理を実行する。

【0148】

この際、コントローラ 31 は、スピンドウン指示として、N 世代、例えば、5 世代より前の古いデータを外部物理デバイス 600 に格納するとともに、その外部物理デバイスをスピンドウンさせる指示を受けたときには、容量仮想化プール 440 に、プールボリュームとして、外部物理デバイス 600 に対応したプールボリューム 458 を格納し、このプールボリューム 458 にプール側ページ P31、P32 を割り当て、このプール側ページ P31、P32 を、外部物理デバイス 600 の論理デバイスに設定されたページ P601、P602 に対応付ける。

10

【0149】

この後、コントローラ 31 は、副側ページ S1、S2 が Tier 2 のプール側ページ P21、P22 と対応付けられているときには、副側ページ S1、S2 をそれぞれプール側ページ P31、P32 に対応付け、副側ページ S1、S2 に格納された 5 世代前のデータを、プール側ページ P31、P32 を介して外部物理デバイス 600 のページ P601、P602 に移動させ、その後、5 世代前のデータが格納された外部物理デバイス 600 をスピンドウンさせ、消費電力を低減させるための処理を実行する。

【0150】

次に、ペアサスペンドのコマンドを受けたあと、状態遷移コマンドとしてペアリシンク（再同期）のコマンドを受けたときには、コントローラ 31 は、図 21 に示すように、ペアリシンクのコマンドに従って処理を実行する。

20

【0151】

この場合、例えば、ペアサスペンドのコマンドを受けた後、正論理ボリューム 462 のデータが更新され、差分データを格納するページとして正側ページ P3 が追加され、正側ページ P3 がプール側ページ P12 に対応付けられているときには、コントローラ 31 は、正側ページ P3 とペアとなる副側ページとして、副論理ボリューム 464 に副側ページ S3 を設定し、副側ボリューム S3 に対して、プール側ページ P13 と同一階層に属するプール側ページ P13 を割り当て、正側ページ P3 のデータを副側ページ S3 にコピーする。なお、副側ページ S1、S2 のデータはプール側ページ P21、P22 に格納されているものとする。

30

【0152】

次に、コントローラ 31 が、ペアリシンク リバース（リストア）のコマンドを受けたときの処理を図 22 に従って説明する。ペアサスペンドのコマンドを受けた後、状態遷移コマンドとして、ペアリシンクリバース（リストア）のコマンドを受けたときには、コントローラ 31 は、副論理ボリューム 464 にバックアップされていたデータを正論理ボリューム 462 に戻すためのリストア処理を行う。

【0153】

このとき、例えば、副論理ボリューム 464 に、副側ページとして S1、S2、S3 が存在し、副側ページ S1 がプール側ページ P21 に、副側ページ S2 がプール側ページ P22 に、副側ページ S3 がプール側ページ P13 にそれぞれ対応付けられているものとする。この状態で副側ページ S1、S2、S3 のデータを正側ボリューム 462 にコピーすると、副側ページ P1 のデータが正側ページ P1 に、副側ページ S2 のデータが正側ページ P2 に、副側ページ S3 のデータが正側ページ P3 にそれぞれコピーされる。

40

【0154】

この際、正側ページ P1 は、ペアクリエイトのコマンドをコントローラ 31 が受けたときには、Tier 0 の SSD に対応したプールボリューム 452 に属していた場合、正側ページ P1 をプール側ページ P01 に対応付ける処理がコントローラ 31 によって実行される。また、正側ページ P2 は、ペアクリエイトのコマンドをコントローラ 31 が受けたときには、Tier 1 の SAS に対応したプールボリューム 454 に属していた場合、副側ページ P2 をプール側ページ P11 に対応付ける処理がコントローラ 31 によって実行される。

50

【0155】

また、正側ページP4として、差分データが格納されていたページがリストア後に、正論理ボリューム462に存在していても、正側ページP4に格納されていた差分データは、リストア処理によって時間的に前の状態に戻るため、不要となる。このため、コントローラ31は、正側ページP4に割り当てられていたプール側ページP02をフリー化して、割り当てを解除する。

【0156】

次に、状態遷移コマンドとして、クイックリストアのコマンドをコントローラが実行するときの処理を図23に従って説明する。

【0157】

コントローラ31は、ホスト計算機10からクイックリストアのコマンドを受けたときには、正論理ボリューム462を副論理ボリュームとし、副論理ボリューム464を正論理ボリュームとする設定変更を実行し、その後、ホスト計算機10からのコマンドに対して、設定変更されたボリュームをアクセス先として処理を実行する。

【0158】

この際、ホスト計算機10は、クイックリストアのコマンドを発行した後は、ストレージシステム30の正論理ボリューム464または副論理ボリューム464にアクセスすることで、コントローラ31がペアクリエイイトに伴うペアコピーを実行することなく、コントローラ31がリストア処理を実行したときと同じデータを得ることができ、ストレージシステム30に対するアクセスを高速化することができる。

【0159】

この場合、新たな正側ページP1、P2、P3には、それぞれプール側ページP02、P12、P14が割り当てられ、新たな副側ページS1、S2、S3、S4には、プール側ページP01、P11、P13、P03が割り当てられる。また、プール側ページP22は、プールボリューム454に移動して、プール側ページ12として管理され、プール側ページP21は、プールボリューム452に移動して、プール側ページP02として管理される。

【0160】

次に、状態遷移コマンドとして、ペア削除のコマンドをコントローラが受けたときの処理を、図24に従って説明する。

【0161】

状態遷移コマンドとして、ペア削除のコマンドを受けたときには、コントローラ31は、正論理ボリューム462と副論理ボリューム464をペアとする関係を削除するとともに、正側ページP1、P2、P3と副側ページS1、S2、S3をそれぞれペアとする関係を削除する。この際、コントローラ31は、正側ページP1、P2、P3、P4とプール側ページP01、P11、P13、P02との対応関係を保持し、副側ページS1、S2、S3とプール側ページP21、P22、P13との対応関係を保持した状態で、その後のコマンドに従って処理を実行することになる。

【0162】

次に、状態遷移コマンドとしてスナップショット作成コマンドを受けた時の処理を図25に従って説明する。スナップショット作成コマンドを受けたときには、コントローラ31は、副論理ボリューム464の他に副論理ボリューム466を作成し、副論理ボリューム466に副側ボリュームS11、S12を割り当て、正側ページP1と副側ページS1にそれぞれプール側ページP01が割り当てられているときには、副側ページS11にプール側ページP01を割り当て、正側ページP2と副側ページS2にプール側ページP11が割り当てられているときには、副側ページS12に、同じ階層のプール側ページP12を割り当てる。

【0163】

次に、コントローラ31は、正側ページP1、P2にそれぞれスナップショットで得られたデータを副側ページS1、S2、S11にコピーする。

【0164】

この後、ペアサスペンドのコマンドを受けて、正論理ボリューム462と副論理ボリューム464がペアサスペンド状態となったときには、ホスト計算機10は、正論理ボリューム462または副論理ボリューム464、466に対してそれぞれライトアクセスが可能となる。

【0165】

具体的には、図26に示すように、スナップショット作成処理の後、ペアサスペンドのコマンドが入力されたときには、コントローラ31は、正論理ボリューム462または副論理ボリューム464、466に対するライト要求を受け付ける。この際、例えば、正側ページP1にライトアクセスがあった場合、ライトアクセス前のデータが副側ページS1に格納されていたときには、副側ページS1のデータを、正側ページP1に割り当てられてい

10

【0166】

なお、このとき、ペアサスペンドのコマンドに、ライト前のデータをTier1またはTier2へ移動させる指示があったときには、ライト前のデータを副側ページS1からTier1に属するプールボリューム454のプール側ページまたはTier2に属するプールボリューム456のプール側ページにコピーすることもできる。

【0167】

次に、正論理ボリューム・副論理ボリュームライト処理を図27のフローチャートに従って説明する。

20

【0168】

この処理は、コントローラ31のプロセッサ360がコマンド制御プログラム3501を起動することによって開始される。まず、プロセッサ360は、ホスト計算機10からのライトコマンドを受領し(S81)、受領したライトコマンドのパラメータをチェックし(S82)、チェック結果を基に転送可否に関する応答をホストコマンド応答としてホスト計算機10に送信する(S83)。

【0169】

次に、プロセッサ360は、ホスト計算機10から、ライトデータを受領したときには(S84)、正論理ボリューム462と副論理ボリューム464のペア状態を確認し(S85)、書込先アドレスをロックし(S86)、ライトデータがストレージキャッシュメモリ32に存在するか否かの判定処理、すなわち、キャッシュヒット(H)またはキャッシュミス(M)の判定処理を行い(S87)、キャッシュヒットしたか否かを判定する(S88)。

30

【0170】

次に、プロセッサ360は、キャッシュヒットしないときには、ストレージキャッシュメモリ32に書込領域を確保し(S89)、その後、ステップS90の処理に移行し、ステップS88でキャッシュヒットしたと判定したときには、ストレージキャッシュメモリ32のスロット状態を確定し(S90)、ライトデータをストレージキャッシュメモリ32に転送し(S91)、差分ビットマップを管理するための差分ビットマップ処理を行い(S92)、書込先アドレスをアンロックとし(S93)、ホスト計算機10に対してライト

40

【0171】

次に、正論理ボリューム・副論理ボリュームのライト処理として、ライトアフタ処理を図27Bのフローチャートに従って説明する。

【0172】

この処理は、コントローラ31のプロセッサ360がディスクI/Oプログラム3505を起動することによって開始される。まず、プロセッサ360は、ホストダーティキューサーチを行い(S101)、仮想ボリューム、例えば、正論理ボリューム462のキャッシュか否かを判定し(S102)、正論理ボリューム462のキャッシュであるときには、デフォルトページか否かを判定し(S103)、デフォルトページであるときには、空

50

きページがあるか否かを判定する（S104）。

【0173】

プロセッサ360は、空きページがあるときには、新規割当処理として、割当ページをフリーキューから選択し（S105）、空き容量カウンタを減算し（S106）、減算結果が閾値か否かを判定し（S107）、減算結果が閾値であるときには、空き容量が少ない旨の警告を行い（S108）、その後、ステップS109の処理に移り、減算値が閾値でないと判定したときとともに、未割当領域とプールフリーの領域とをリンクし（S109）、ステップS114の処理に移行する。

【0174】

一方、ステップS103において、デフォルトページでないと判定したときには、プロセッサ360は、割当済みページを算出し（S110）、その後、ステップS114の処理に移行する。

また、ステップS104において、空きページがないと判定したときには、プロセッサ360は、空き容量枯渇の警告を行い（S111）、その後、プール書込禁止の警告を行い（S112）、次に、リザーブを使用可能に設定できるか否かを判定し（S113）、設定できないときにはステップS114の処理に移り、設定できるときには、このルーチンでの処理を終了する。

【0175】

プロセッサ360は、ステップS114においては、RAIDレベルに応じてパリティデータなどを作成し、作成したパリティデータなどをディスク書き込み用のキューに繋ぎ換える処理を行い（S115）、ライトアフタ処理を終了する。

次に、正論理ボリューム・副論理ボリュームに対するリードコマンド処理を図28のフローチャートに従って説明する。

【0176】

この処理は、コントローラ31のプロセッサ360が、コマンド制御プログラム3501とディスクI/Oプログラム3505を起動することによって開始される。

まず、プロセッサ360は、リードコマンドを受領し（S121）、受領したリードコマンドのパラメータをチェックし（S122）、次に、コマンドパラメータのチェック結果を基にペア状態、例えば、正論理ボリューム462と副論理ボリューム464のペア状態を確認し（S123）、その後、読み込み先をロックし（S124）、ストレージキャッシュメモリ32にリードデータが存在するかをチェックするためのキャッシュヒットまたはキャッシュミス処理を実行し（S125）、キャッシュヒットとしたか否かを判定する（S126）。

【0177】

キャッシュヒットしたときには、プロセッサ360は、ストレージキャッシュメモリ32に領域割当があるかを確認し（S127）、ストレージキャッシュメモリ32に存在する領域が割当済みか否かを判定し（S128）、割当済みのときには、仮想プールアドレス変換を行う（S129）。

【0178】

ストレージキャッシュメモリ32に領域が割当済みでないときには、プロセッサ360は、ページ割当の要否を判定し（S130）、ページ割当が必要であると判定したときには、ストレージキャッシュメモリ32にページを動的に割り当て、ステップS132の処理に移行する。

【0179】

プロセッサ360は、ステップS126において、キャッシュヒットでないと判定したときには、ストレージキャッシュメモリ32にスロット状態を確定し（S132）、ステップS130において、ページ割当が必要でないと判定したときには、デフォルトページのデータ領域をリードし（S133）、その後、リードデータをストレージキャッシュメモリ32に転送する（S134）。

【0180】

10

20

30

40

50

この後、プロセッサ360は、ストレージキャッシュメモリ32のリードデータをホスト計算機10に転送し(S135)、読み込み先をアンロックとし(S136)、ホスト計算機10に対して、リードコマンド処理が終了した旨の完了報告を行い(S137)、リードコマンド処理を終了する。

【0181】

次に、イニシャルコピー処理を図29のフローチャートに従って説明する。この処理は、ホスト計算機10からの状態遷移コマンドとして、例えば、ペアクリエイトのコマンドをコントローラ31が受信したときに開始される。

【0182】

まず、ペアクリエイトのコマンドをコントローラ31が受信すると、プロセッサ360は、コピー対象範囲全体にコピーするためのビットマップをオンとし(S141)、コピー元処理開始アドレスとして、例えば、正論理ボリューム462の処理開始アドレスを設定し(S142)、コピービットマップをチェックし(S143)、コピーの可否を判定するための処理を行い(S144)、コピーが必要か否かを判定し(S145)、コピーが必要でないときにはステップS146の処理に移行し、コピー処理判定対象アドレスを1つ先に進め、ステップS143の処理に戻り、ステップS143～S145の処理を繰り返す。

10

【0183】

一方、ステップS145でコピーが必要と判定したときには、プロセッサ360は、コピー元アドレスとコピー先アドレス、例えば、正論理ボリューム462をコピー元とし、副論理ボリューム464をコピー先としたときには、正論理ボリューム462と副論理ボリューム464のアドレスを確認し(S147)、コピー元アドレスをロックし(S148)、コピー元のデータがストレージキャッシュメモリ32に存在するかを判定するためのコピー元ヒット(H)またはヒットミス(M)の処理を行い(S149)、キャッシュヒットしたか否かを判定する(S150)。

20

【0184】

キャッシュヒットしたときには、プロセッサ360は、コピー元キャッシュ領域を確定し(S151)、コピー先アドレスをロックし(S152)、コピー先に関するヒットまたはミスをチェックするための処理を行い(S153)、コピー先のストレージキャッシュメモリ32に対してキャッシュヒットしたか否かを判定する(S154)。

30

【0185】

キャッシュヒットしないときには、プロセッサ360は、コピー先にキャッシュ領域を確保し(S155)、その後、コピー先にキャッシュ領域を確定し(S156)、コピー元のストレージキャッシュメモリ32に存在するデータをコピー先のストレージキャッシュメモリ32にコピーし(S157)、ステップS165の処理に移行する。

【0186】

一方、ステップS150において、キャッシュヒットしないと判定したときには、プロセッサ360は、コピー元のディスク領域を確保し(S158)、コピー先アドレスをロックし(S159)、コピー先に対するヒットまたはヒットミスをチェックするための処理を行い(S160)、次に、キャッシュヒットしたか否かを判定する(S161)。キャッシュヒットしないときには、プロセッサ360は、コピー先のストレージキャッシュメモリ32に領域を確保し(S162)、キャッシュヒットしたときには、コピー先ストレージキャッシュメモリ32の領域にキャッシュ領域を確定し(S163)、コピー元ディスクのデータをコピー先のストレージキャッシュメモリ32にコピーし(S164)、ステップS165の処理に移行する。

40

【0187】

プロセッサ360は、ステップS165において、コピービットマップをオフとし、その後、コピー先アドレスをアンロックとし(S166)、次に、コピー元アドレスをアンロックとし(S167)、コピー処理対象アドレスが終端であるか否かを判定し(S168)、終端でないときにはステップS146の処理に戻り、コピー処理判定対象アドレスを

50

1つ先に進めたあと、ステップS143の処理に戻り、ステップS143～S168の処理を繰り返し、コピー処理対象アドレスが終端であると判定したときには、イニシャルコピーを終了する。

【0188】

次に、差分コピー処理を図30のフローチャートに従って説明する。この処理は、ホスト計算機10からの状態遷移コマンドとして、例えば、ペアリスINKのコマンドをコントローラ31が受信したときに開始される。

【0189】

まず、コントローラ31のプロセッサ360は、ペアリスINKのコマンドを受信したときには、コピー元処理開始アドレスとして、例えば、正論理ボリューム462の処理開始アドレスを設定し（S181）、次に、差分ビットマップ（図27AのステップS92において処理した差分ビットマップ）をチェックし（S182）、その後、コピーの要否を判定するための処理を行い（S183）、コピーが必要か否かを判定する（S184）。プロセッサ360は、ステップS184でコピーが必要でないと判定したときには、ステップS185の処理に移行し、コピー処理判定対象アドレスを1つ先に進めて、ステップS182の処理に戻り、ステップS182～S184の処理を繰り返す。

【0190】

一方、ステップS184でコピーが必要と判定したときには、プロセッサ360は、コピー元アドレスとコピー先アドレス、例えば、正論理ボリューム462をコピー元とし、副論理ボリューム464をコピー先としたときには、正論理ボリューム462と副論理ボリューム464のアドレスを確認し（S186）、コピー元アドレスをロックし（S187）、コピー元のデータがストレージキャッシュメモリ32に存在するかを判定するためのコピー元ヒット（H）またはヒットミス（M）の処理を行い（S188）、キャッシュヒットしたか否かを判定する（S189）。

【0191】

キャッシュヒットしたときには、プロセッサ360は、コピー元キャッシュ領域を確定し（S190）、コピー先アドレスをロックし（S191）、コピー先に関するヒットまたはミスをチェックするための処理を行い（S192）、コピー先のストレージキャッシュメモリ32に対してキャッシュヒットしたか否かを判定する（S193）。

【0192】

キャッシュヒットしないときには、プロセッサ360は、コピー先にキャッシュ領域を確保し（S194）、その後、コピー先にキャッシュ領域を確定し（S195）、コピー元のストレージキャッシュメモリ32に存在するデータをコピー先のストレージキャッシュメモリ32にコピーし（S196）、ステップS204の処理に移行する。

【0193】

一方、ステップS189において、キャッシュヒットしないと判定したときには、プロセッサ360は、コピー元のディスク領域を確保し（S197）、コピー先アドレスをロックし（S198）、コピー先に対するヒットまたはヒットミスをチェックするための処理を行い（S199）、次に、キャッシュヒットしたか否かを判定する（S200）。

キャッシュヒットしないときには、プロセッサ360は、コピー先のストレージキャッシュメモリ32に領域を確保し（S201）、キャッシュヒットしたときには、コピー先ストレージキャッシュメモリ32の領域にキャッシュ領域を確定し（S202）、コピー元ディスクのデータをコピー先のストレージキャッシュメモリ32にコピーし（S203）、ステップS204の処理に移行する。

【0194】

プロセッサ360は、ステップS204において、コピービットマップをオフとし、その後、コピー先アドレスをアンロックとし（S205）、次に、コピー元アドレスをアンロックとし（S206）、コピー処理対象アドレスが終端であるか否かを判定し（S207）、終端でないときにはステップS185の処理に戻り、コピー処理判定対象アドレスを1つ先に進めたあと、ステップS182の処理に戻り、ステップS182～S207の処理

を繰り返し、コピー処理対象アドレスが終端であると判定したときには、差分コピーを終了する。

【0195】

次に、コピーアフタライト (CAW) ・コピーオンライト (COW) の処理を図31のフローチャートに従って説明する。

【0196】

この処理は、ホスト計算機10から、正論理ボリューム462に対してライトアクセスが実行されたとき、あるいは、副論理ボリューム464に対してリードアクセスまたはライトアクセスが実行されたときに、コントローラ31によって開始される。

【0197】

まず、プロセッサ360は、コピー元として、例えば、正論理ボリューム462を確認するための処理と、コピー先として、例えば、副論理ボリューム464を確認するための処理を行い (S211)、コピー元アドレスをロックし (S212)、コピー元に関してキャッシュヒットまたはキャッシュミスをチェックするための処理を開始し (S213)、コピー元のストレージキャッシュメモリ32にコピーすべきデータが存在するか否かの判定を行う (S214)。

【0198】

キャッシュヒットしたときには、プロセッサ360は、コピー元キャッシュ領域を確定し (S215)、コピー先アドレスをロックし (S216)、コピー先に関するヒットまたはミスをチェックするための処理を行い (S217)、コピー先のストレージキャッシュメモリ32に対してキャッシュヒットしたか否かを判定する (S218)。

【0199】

キャッシュヒットしないときには、プロセッサ360は、コピー先にキャッシュ領域を確保し (S219)、その後、コピー先にキャッシュ領域を確定し (S220)、コピー元のストレージキャッシュメモリ32に存在するデータをコピー先のストレージキャッシュメモリ32にコピーし (S221)、ステップS229の処理に移行する。

【0200】

一方、ステップS214において、キャッシュヒットしないと判定したときには、プロセッサ360は、コピー元のディスク領域を確保し (S222)、コピー先アドレスをロックし (S223)、コピー先に対するヒットまたはヒットミスをチェックするための処理を行い (S224)、次に、キャッシュヒットしたか否かを判定する (S225)。キャッシュヒットしないときには、プロセッサ360は、コピー先のストレージキャッシュメモリ32に領域を確保し (S226)、キャッシュヒットしたときには、コピー先ストレージキャッシュメモリ32の領域にキャッシュ領域を確定し (S227)、コピー元ディスクのデータをコピー先のストレージキャッシュメモリ32にコピーし (S228)、ステップS229の処理に移行する。

【0201】

プロセッサ360は、ステップS229において、コピー先アドレスをアンロックとし、次に、コピー元アドレスをアンロックとし (S230)、コピーアフタライト (CAW) ・コピーオンライト (COW) の処理を終了する。

【0202】

次に、計算機システムにおけるディザスタリカバリ (DR) の処理を図32に従って説明する。図32は、ストレージシステム30と同一構成のストレージシステムをローカル側ストレージシステム30Aとし、ストレージシステム30と同一構成の他のストレージシステムをリモート側ストレージシステム30Bとし、ローカル側ストレージシステム30Aとリモート側ストレージシステム30Bとをストレージネットワーク51で接続したときの計算機システムの構成を示す。

【0203】

ここで、ディザスタリカバリ (DR) の処理を開始するに際して、ローカル側ストレージシステム30Aからローカル側ストレージシステム30Bに対して、ストレージネットワー

10

20

30

40

50

ク51を介してリモートコピーのコマンドが送信されたときには、正論理ボリューム462Aのデータが、リモート側ストレージシステム30Bの正論理ボリューム462Bにコピーされ、その後、正論理ボリューム462Bのデータが、リモート側ストレージシステム30Bの副論理ボリューム464Bにコピーされる。

【0204】

この際、ローカル側ストレージシステム30Aにおいて、正側ページP1、P2に対して、プール側ページP01、P11が対応付けられ、副側ページS1、S2に対して、プール側ページP02、P12が割り当てられているときには、リモート側ストレージシステム30Bにおいても、正側ページP1、P2には、プール側ページP01、P11が割り当てられ、副側ページS1、S2には、プール側ページP02、P12がそれぞれ割り当てられる。

10

【0205】

すなわち、ローカルストレージシステム30Aからリモート側ストレージシステム30Bにデータをバックアップするときには、バックアップ元のページとバックアップ先のページは同じ階層のものが対応づけられる。

【0206】

次に、リモート側ストレージシステム30Bにローカル側ストレージシステム30Aのデータがバックアップされた後、リモート側ストレージシステム30Bに、ペアサスペンドのコマンドが入力されたときの処理を図33に従って説明する。

【0207】

リモート側ストレージシステム30Bに、ローカル側ストレージシステム30Aのデータがバックアップされた後、ローカル側ストレージシステム30Aからリモート側ストレージシステム30Bに対して、ペアサスペンドのコマンドが転送されたときには、ローカル側ストレージシステム30Aとリモート側ストレージシステム30Bのコントローラ31は、リモートコピーの停止に伴って、それぞれホスト計算機10からのIO処理を個別に実行することになる。

20

【0208】

例えば、リモート側ストレージシステム30Bにおいては、ペアクリエイトの状態で、ホスト計算機10からのIO処理を受け付けたときには、ホスト計算機10からのIO処理を実行することになる。

【0209】

次に、ローカル側ストレージシステム30Aが被災から復旧したときのリカバリ処理を図34に従って説明する。

30

【0210】

ローカル側ストレージシステム30Aがリモート側ストレージシステム30Bにデータをバックアップした後、被災から復旧し、リモート側ストレージシステムBのデータをローカル側ストレージシステム30Aにコピーするに際しては、リモート側ストレージシステム30Bからローカル側ストレージシステム30Aに対して、リバース処理としてのコマンドが転送され、リモート側ストレージシステム30Bにおける正論理ボリューム462Bのデータが、ローカル側ストレージシステム30Aにおける正論理ボリューム462Aにコピーされる。この後、ローカル側ストレージシステム30Aにおいては、正論理ボリューム462Aのデータが副論理ボリューム464Aにコピーされる。

40

【0211】

リモートコピーが完了したときには、図35に示すように、ローカル側ストレージシステム30Aからストレージネットワーク51を介して、リモート側ストレージシステム30Bに、リカバリ処理に伴うリモートコピーが完了した旨の情報が転送され、リカバリ処理が完了することになる。

【0212】

また、ローカル側ストレージシステム30Aとリモート側ストレージシステム30Bにおいては、それぞれホスト計算機10からのコマンドに従って各種の処理が実行される。例えば、ホスト計算機10から状態遷移コマンドが発行されたときには、状態遷移コマンド

50

に従った処理として、図 16～図 26 に示す処理が、ローカル側ストレージシステム 30A のコントローラ 31 とリモート側ストレージシステム 30B のコントローラ 31 によってそれぞれ実行される。

【0213】

本実施例によれば、ホスト計算機 10 から、ペア形成に関するコマンドを受けたときには、正側ページと副側ページに同じ階層に属するプール側ページを割り当て、その後、ペアに関する状態遷移コマンドを受けたときに、状態遷移コマンドでページ移動が指示されたときには、副側ページに割り当てられていたプール側ページを、正側ページに割り当てられていたプール側ページとは異なる階層のプール側ページに移動させ、副側ページには、移動したプール側ページを割り当てるようにしたため、状態遷移コマンドに応じて副論理ボリューム 464 の副側ページに適切な物理デバイス 34 を割り当てることができる。

10

【0214】

また、本実施例によれば、状態遷移コマンドに応じて副論理ボリューム 464 の副側ページに適切な物理デバイス 34 を割り当てることができるので、論理ボリュームの用途に応じて、性能、信頼性を高めることができる。

【0215】

また、本実施例においては、正論理ボリューム 462 の正側ページに割り当てられたプール側ページを、状態遷移コマンドに応じて、階層の異なるプール側ページに移動させることもできる。

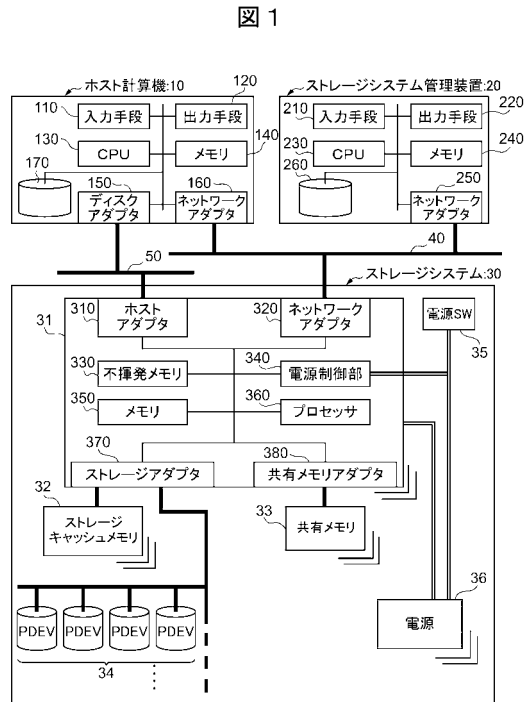
【符号の説明】

20

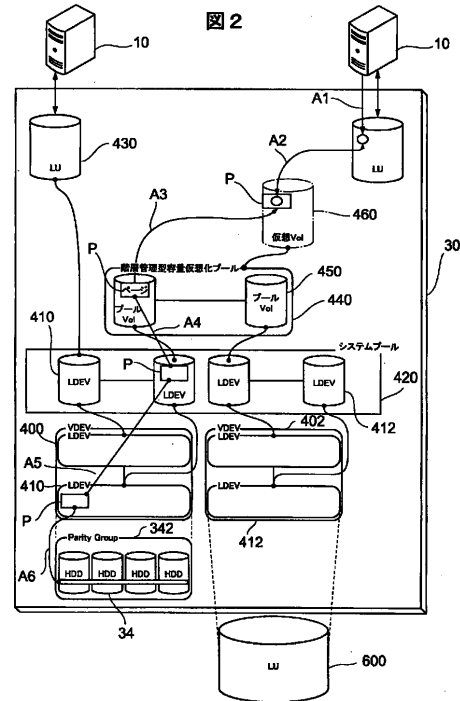
【0216】

10 ホスト計算機、20 ストレージシステム管理装置、30 ストレージシステム、30A ローカル側ストレージシステム、30B リモート側ストレージシステム、31 コントローラ、32 ストレージキャッシュメモリ、33 共有メモリ、34 物理デバイス、40 管理ネットワーク、50 ストレージネットワーク、360 プロセッサ、400 仮想デバイス、410 論理デバイス、420 システムプール、440 階層管理型容量仮想化プール、460 仮想ボリューム、462 正論理ボリューム、464 副論理ボリューム

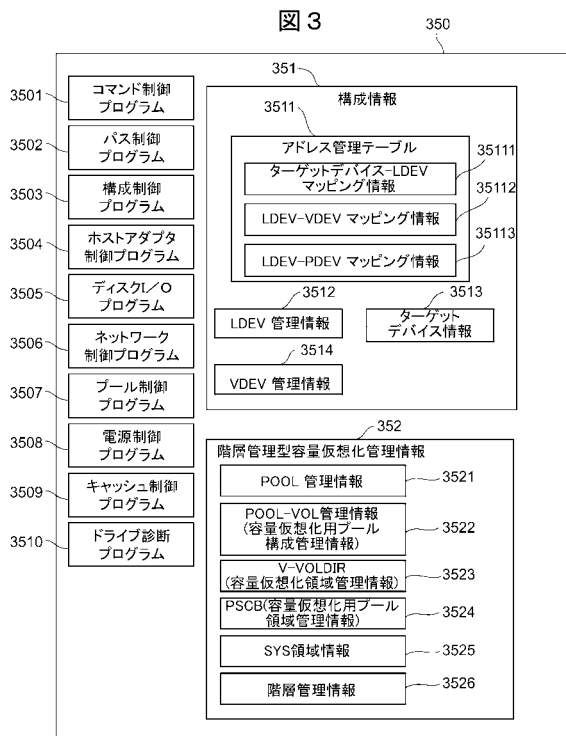
【図 1】



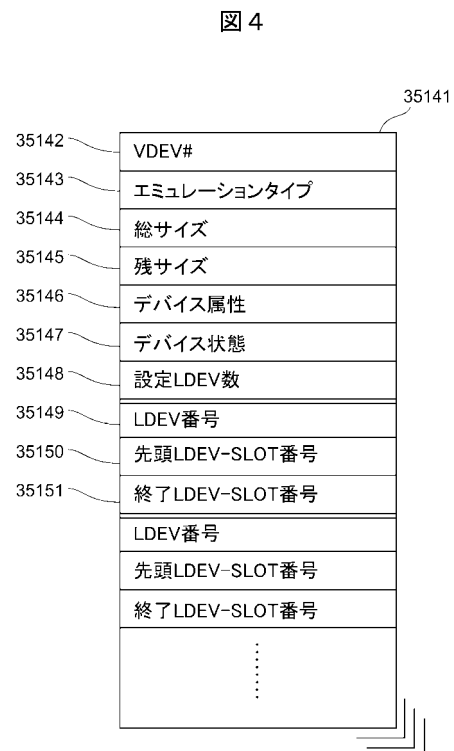
【図 2】



【図 3】



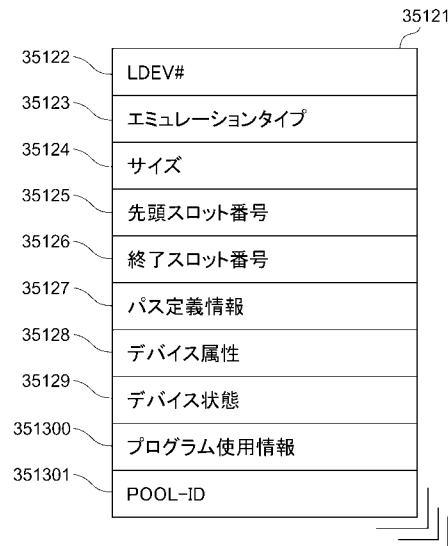
【図 4】



VDEV管理情報3514

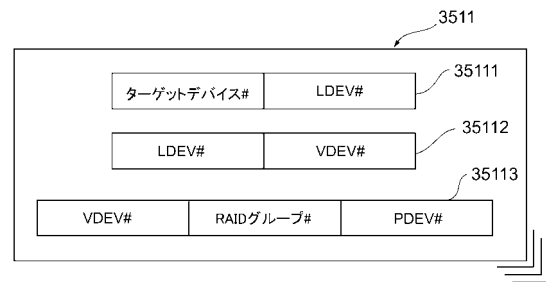
【図 5】

図 5

LDEV管理情報3512

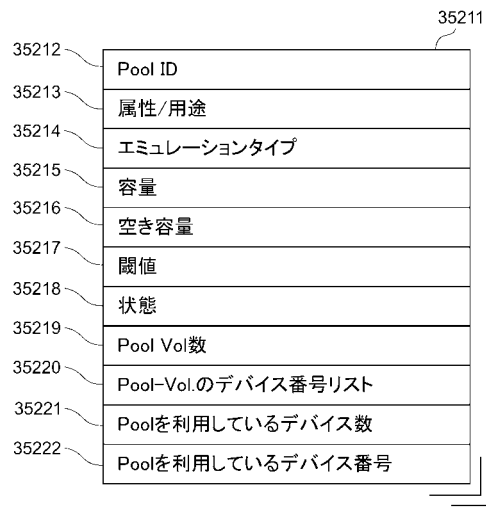
【図 6】

図 6



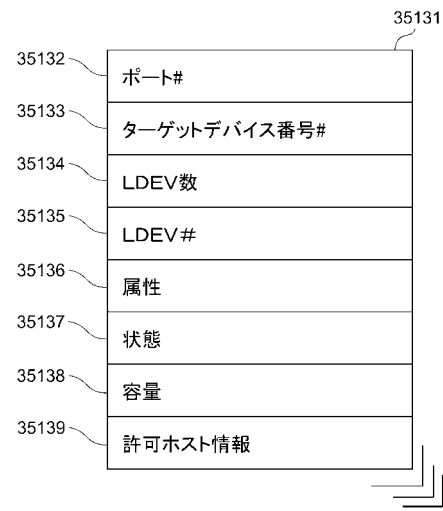
【図 7】

図 7

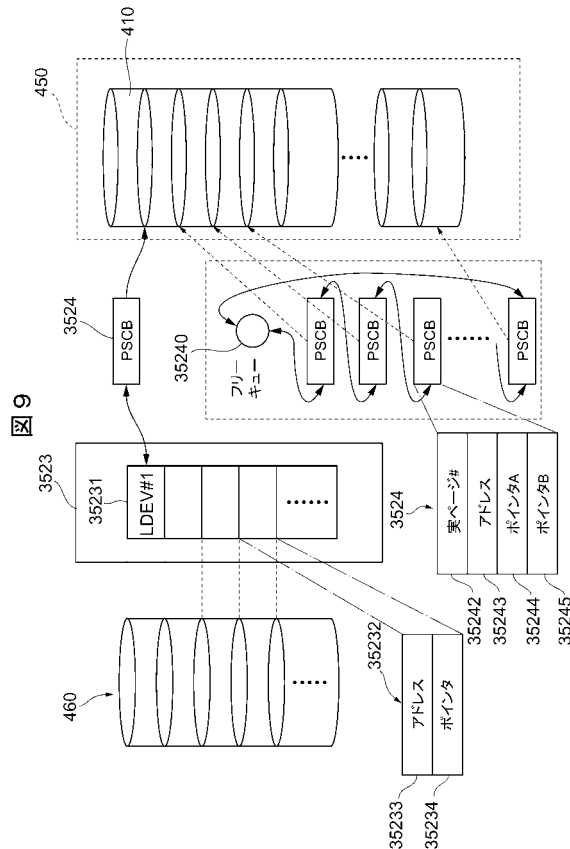
POOL管理情報3521

【図 8】

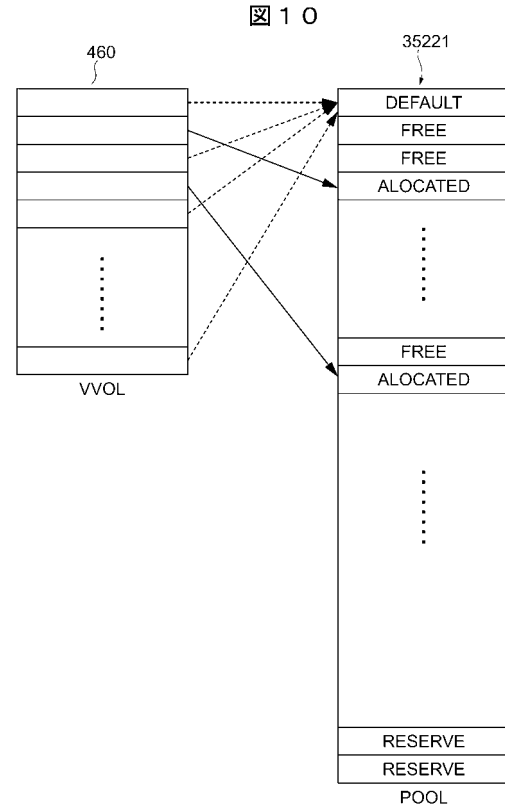
図 8

ターゲットデバイス情報3513

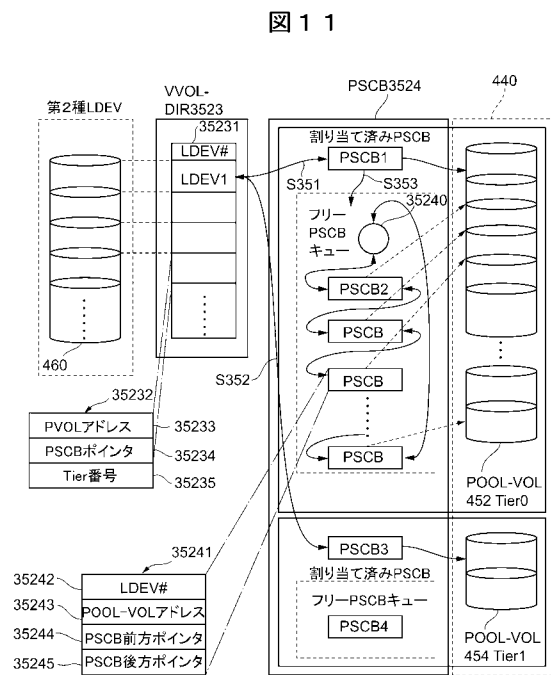
【図 9】



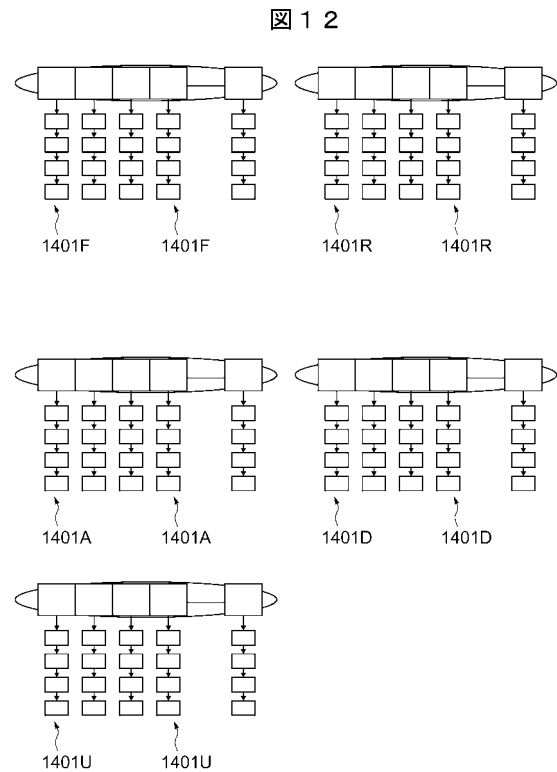
【図 10】



【図 11】

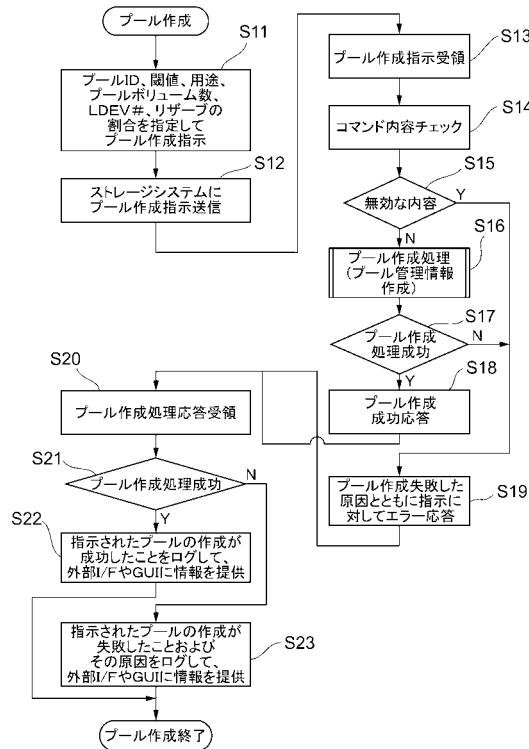


【図 12】



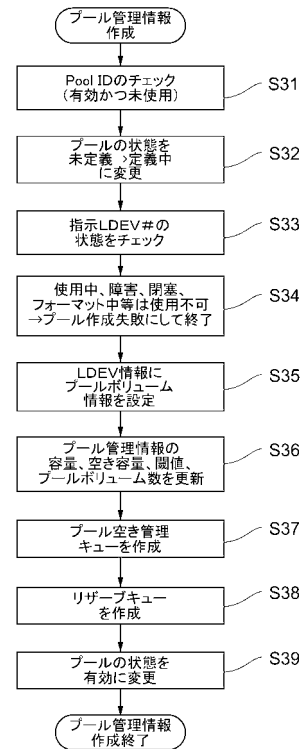
【図 1 3 A】

図 1 3 A



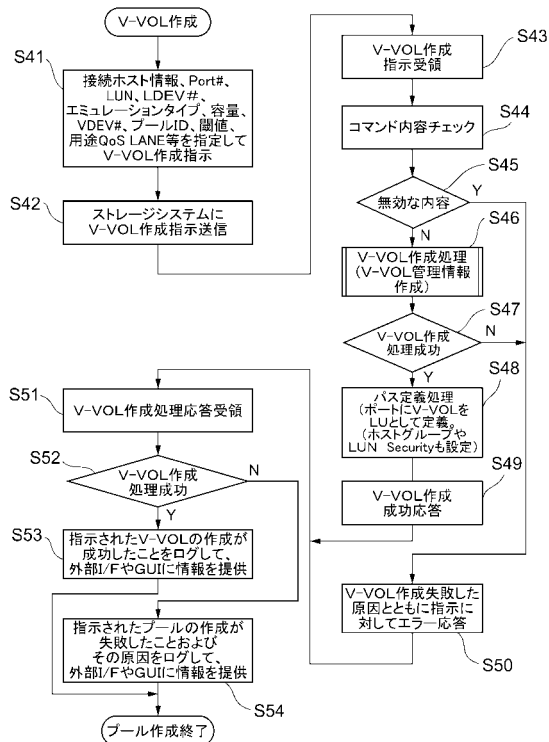
【図 1 3 B】

図 1 3 B



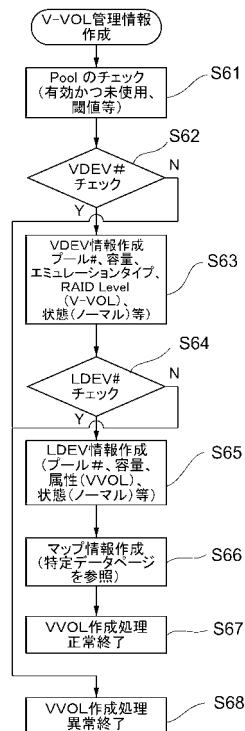
【図 1 4 A】

図 1 4 A

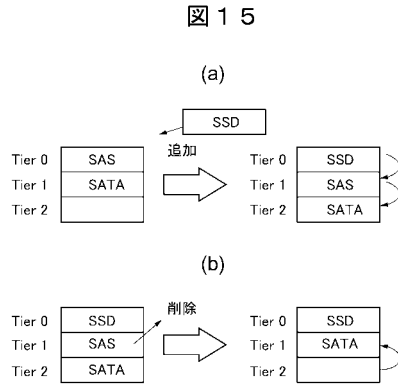


【図 1 4 B】

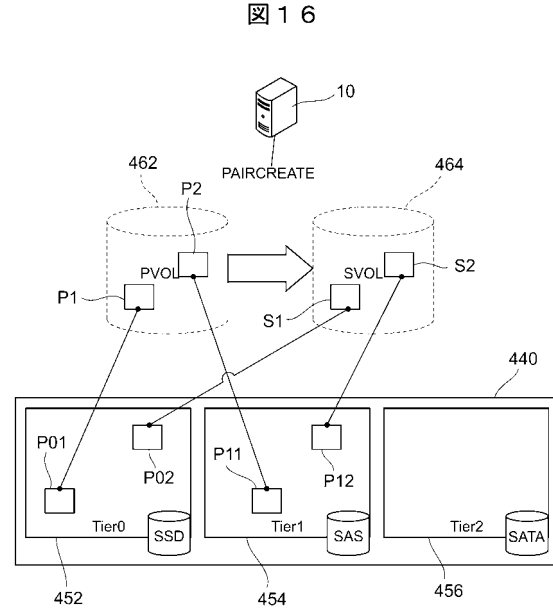
図 1 4 B



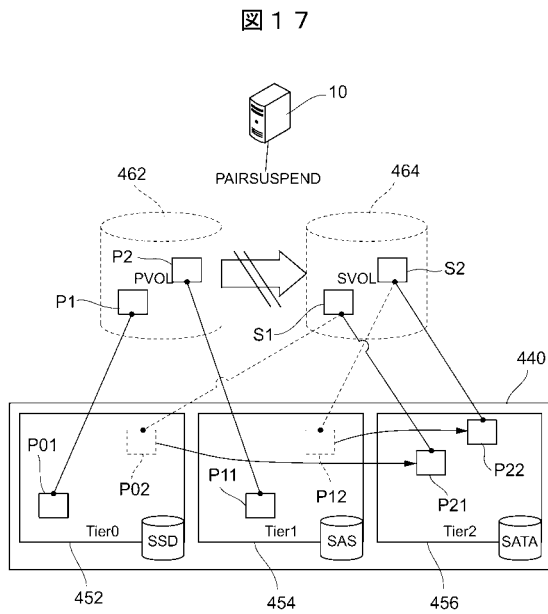
【図 15】



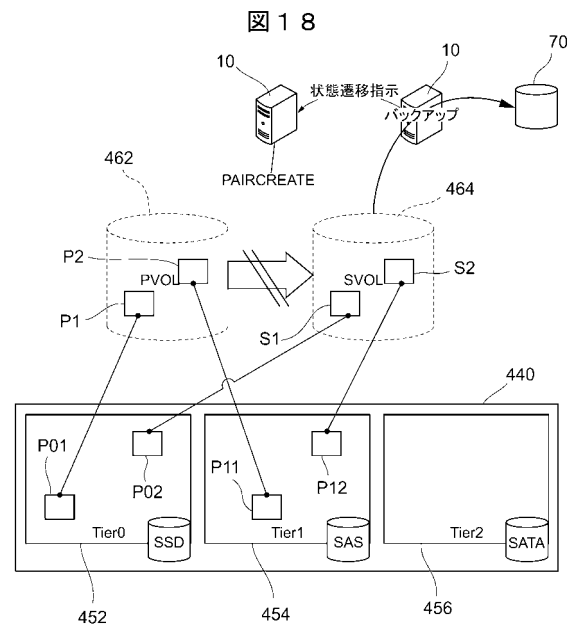
【図 16】



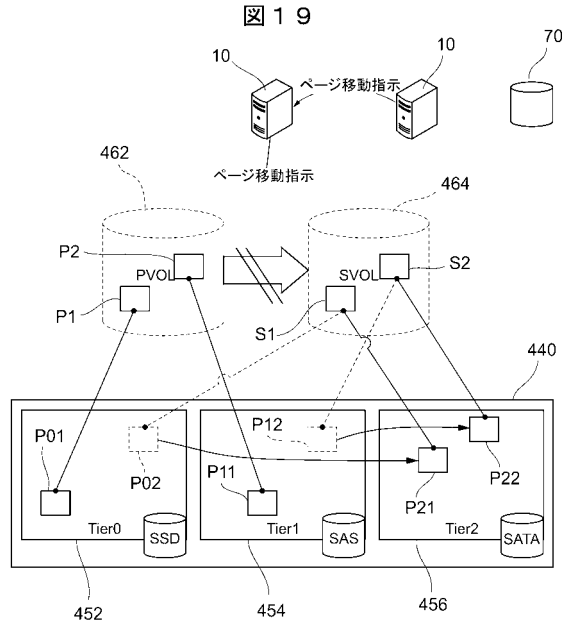
【図 17】



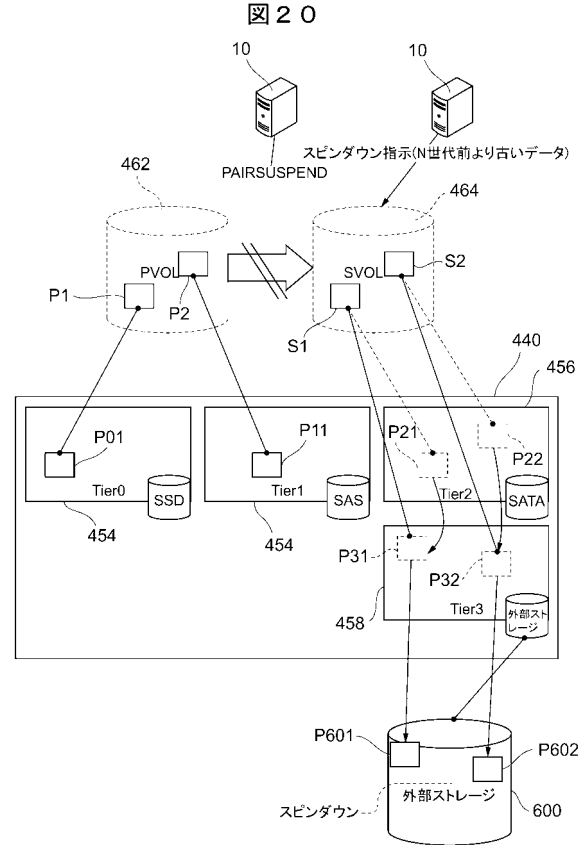
【図 18】



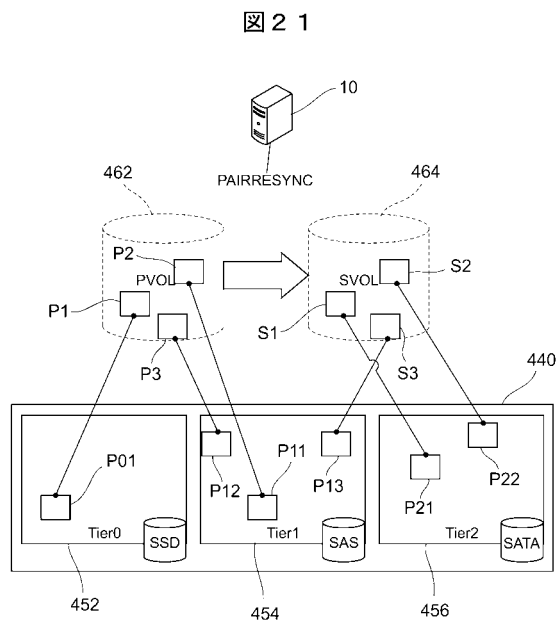
【図 19】



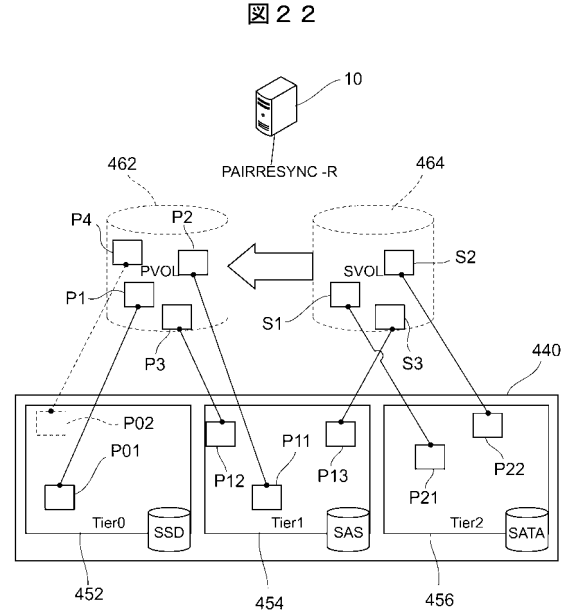
【図 20】



【図 21】

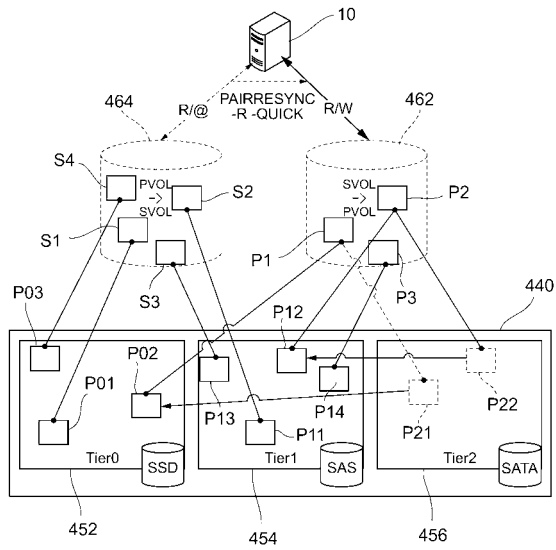


【図 22】



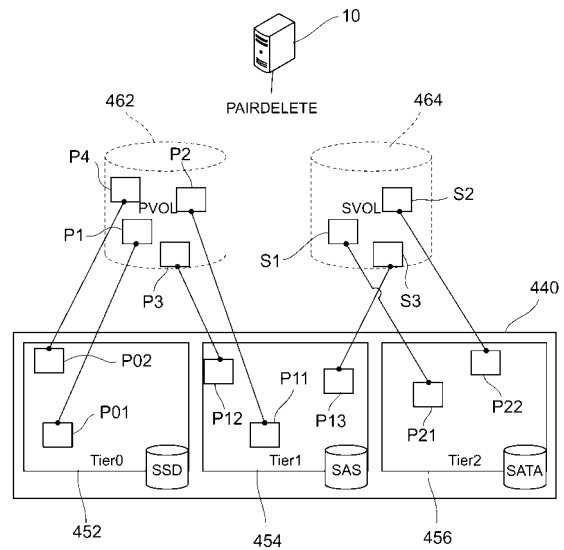
【図 2 3】

図 2 3

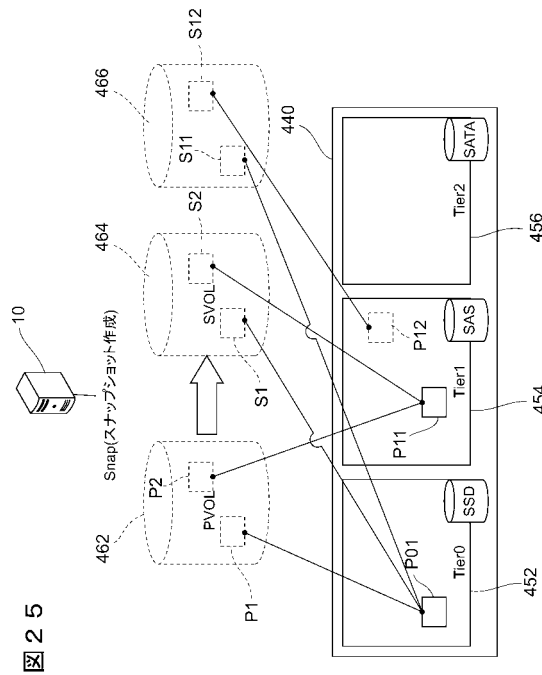


【図 2 4】

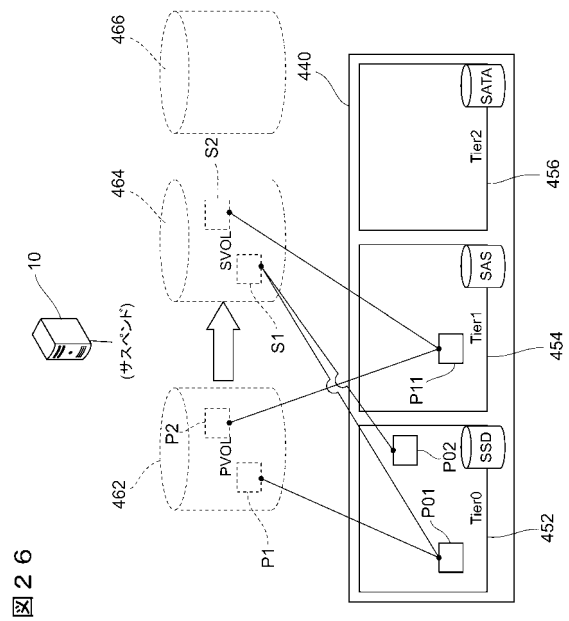
図 2 4



【図 2 5】

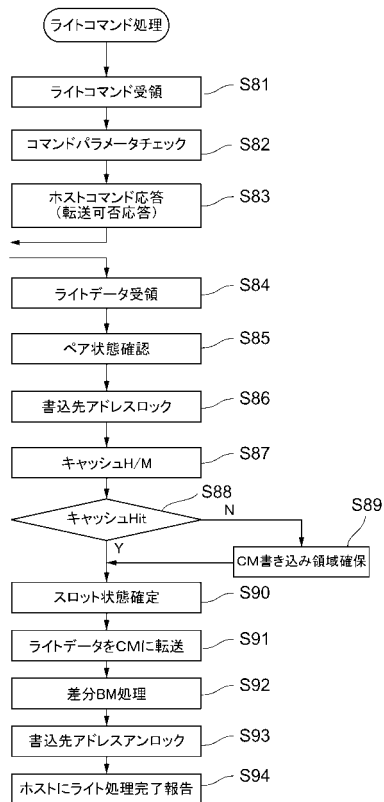


【図 2 6】



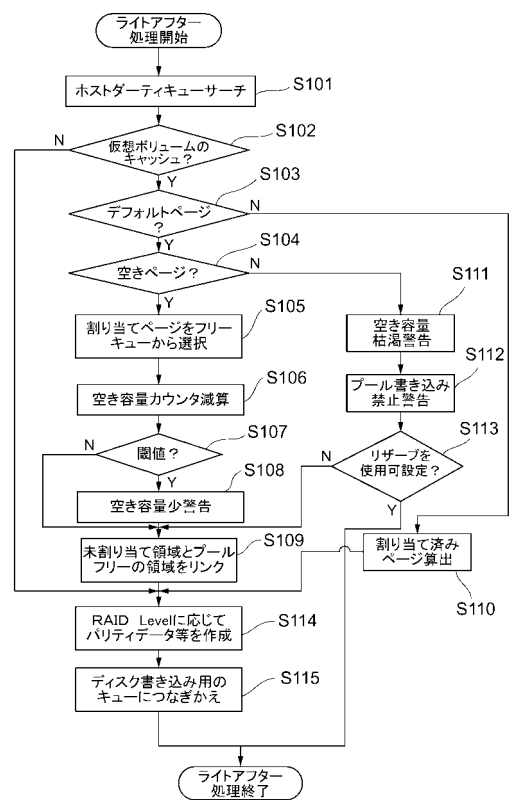
【図 27 A】

図 27 A



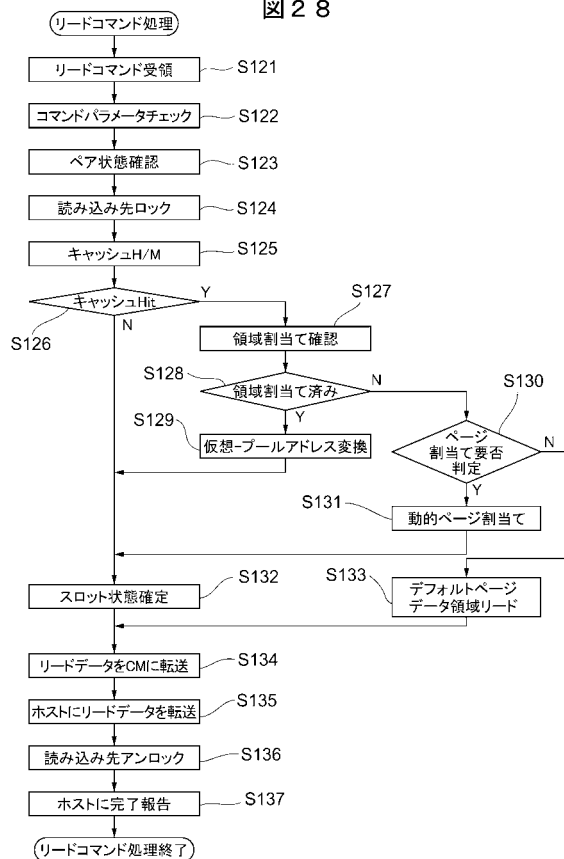
【図 27 B】

図 27 B



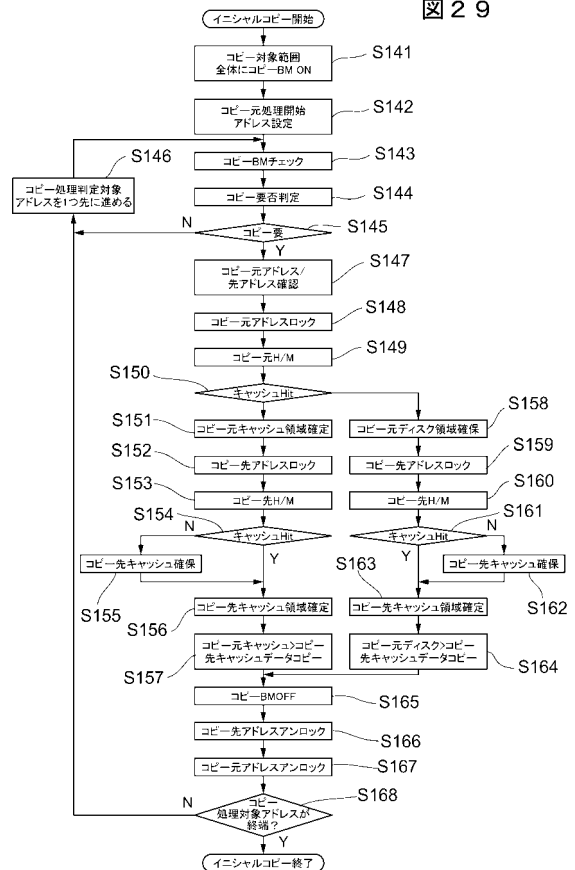
【図 28】

図 28

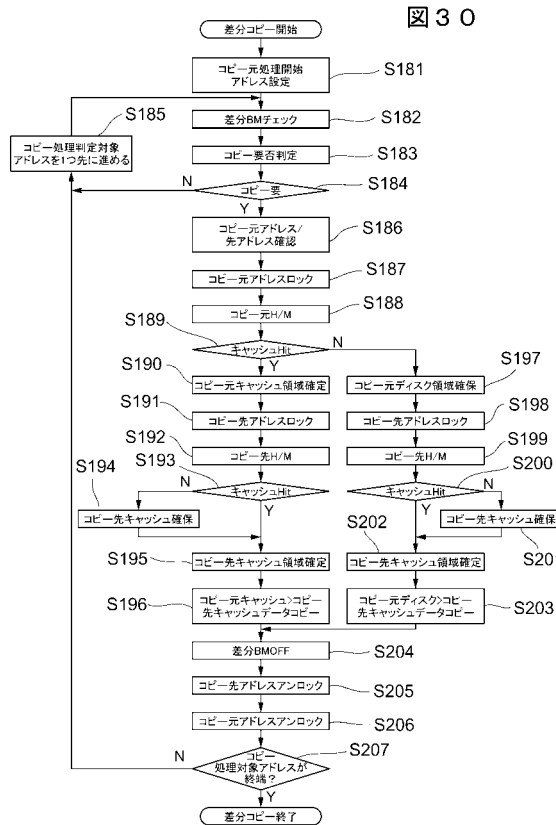


【図 29】

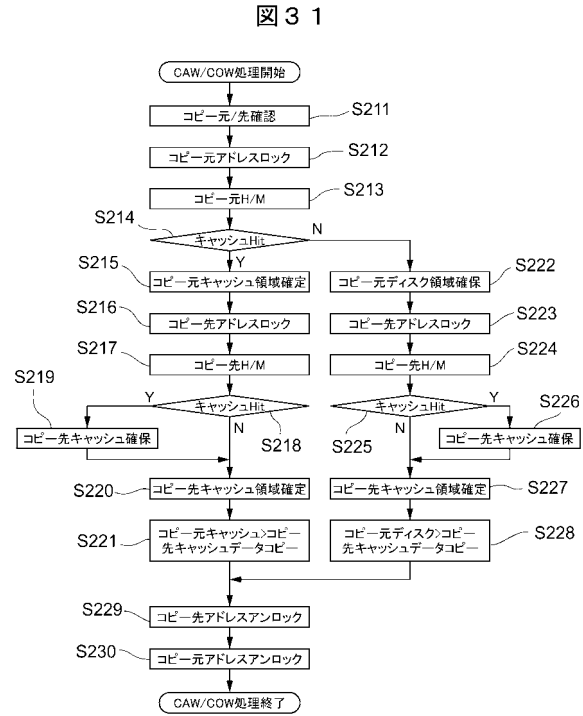
図 29



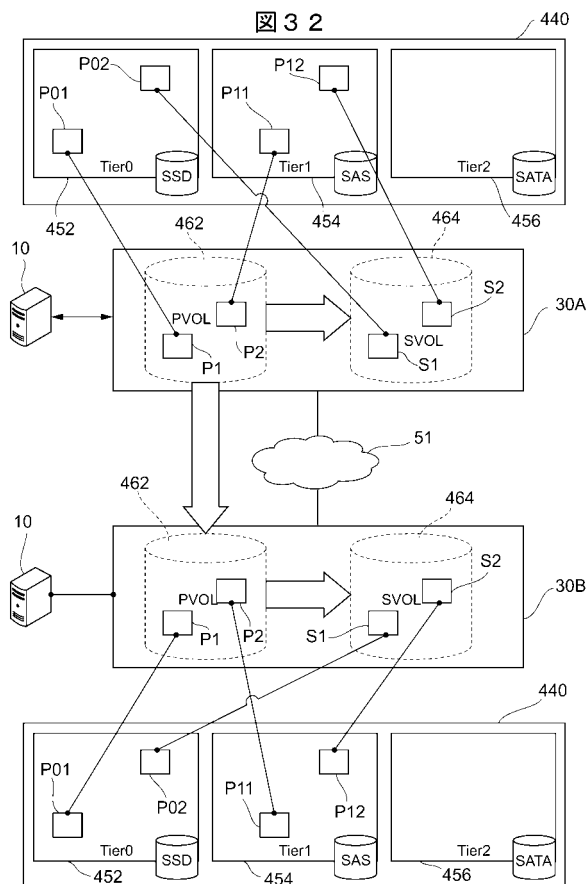
【図 30】



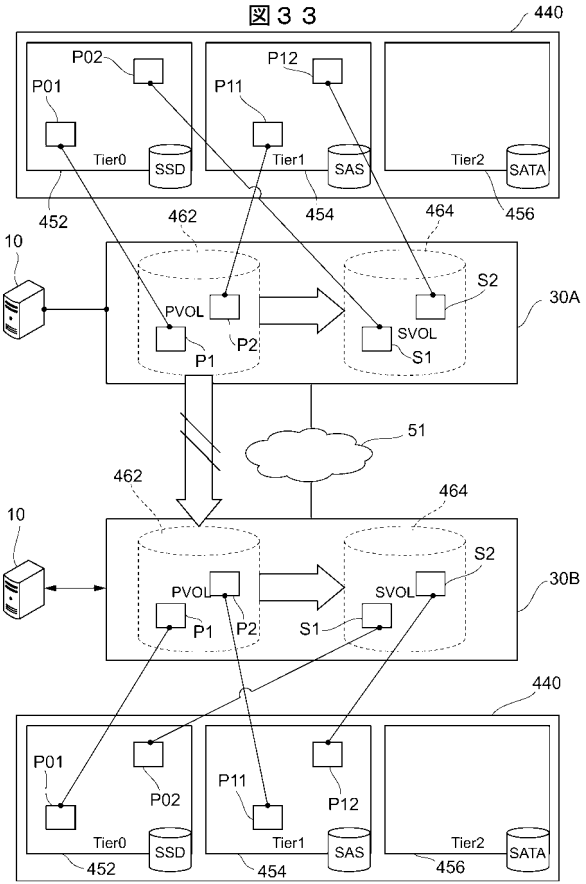
【図 3 1】



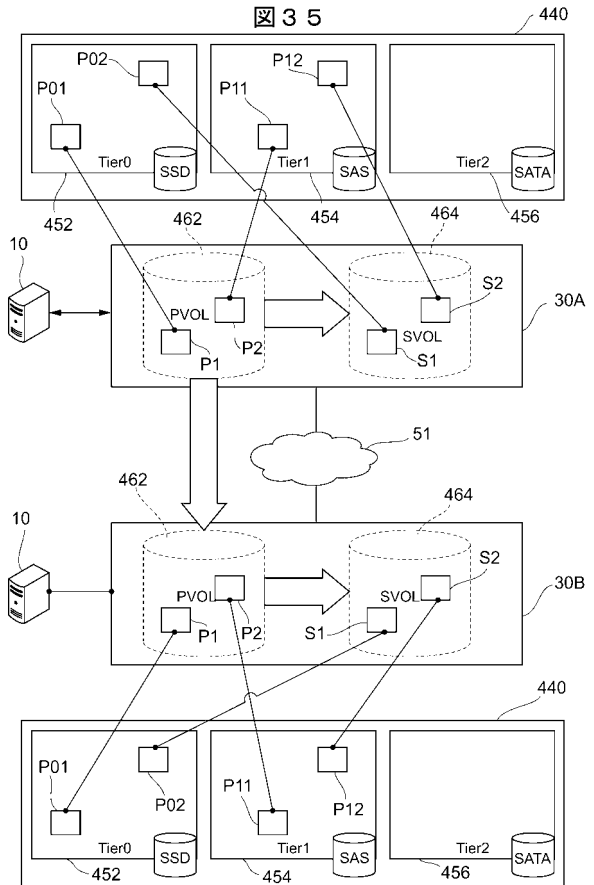
【図 3 2】



【図 3 3】



【図 3 5】



フロントページの続き

(51)Int.Cl. F I
G 0 6 F 13/10 3 4 0 A

(56)参考文献 特開 2 0 0 8 - 1 3 4 9 8 7 (J P, A)
特開 2 0 0 9 - 2 5 1 9 7 0 (J P, A)

(58)調査した分野(Int.Cl., D B名)
G 0 6 F 1 2 / 0 0
G 0 6 F 3 / 0 6
G 0 6 F 1 3 / 1 0