

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5830599号
(P5830599)

(45) 発行日 平成27年12月9日(2015.12.9)

(24) 登録日 平成27年10月30日(2015.10.30)

(51) Int. Cl.		F I	
G06F 13/14	(2006.01)	G06F 13/14	310H
G06F 13/10	(2006.01)	G06F 13/10	340A
G06F 3/06	(2006.01)	G06F 3/06	302Z
		G06F 13/14	330B

請求項の数 10 (全 42 頁)

(21) 出願番号	特願2014-507271 (P2014-507271)	(73) 特許権者	000005108
(86) (22) 出願日	平成23年10月28日(2011.10.28)		株式会社日立製作所
(65) 公表番号	特表2014-529774 (P2014-529774A)		東京都千代田区丸の内一丁目6番6号
(43) 公表日	平成26年11月13日(2014.11.13)	(74) 代理人	110001678
(86) 国際出願番号	PCT/JP2011/006043		特許業務法人藤央特許事務所
(87) 国際公開番号	W02013/061380	(72) 発明者	永見 明久
(87) 国際公開日	平成25年5月2日(2013.5.2)		東京都千代田区丸の内一丁目6番6号 株
審査請求日	平成26年2月13日(2014.2.13)		式会社日立製作所内
前置審査		(72) 発明者	兼田 泰典
			東京都千代田区丸の内一丁目6番6号 株
			式会社日立製作所内
		審査官	小林 義晴

最終頁に続く

(54) 【発明の名称】 計算機システム及びその管理システム

(57) 【特許請求の範囲】

【請求項1】

複数のホストプログラムが実行される複数のホスト計算機と、
前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステムと、
管理システムと、を含み、
前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、
前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、
前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリュームを含み、
前記管理システムは、
前記複数のストレージシステムにおいて、前記第1ボリュームの新たな配置先候補ポートを決定し、
前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、
前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、
前記管理システムは、さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における接続装置数についての情報及び前記第1ボリュームのアクセス状況

についての情報を取得し、

前記管理システムは、前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況についての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記帯域についての情報は、前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれにおける最小帯域の情報を含み、

前記管理システムは、前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれにおける最小帯域に基づいて、前記新たな配置先候補ポートそれぞれに対する前記優先度を決定する、計算機システム。

【請求項2】

前記管理システムは、さらに、前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれにおける最小帯域経路の利用状況の情報を取得し、

前記管理システムは、前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれの前記利用状況に基づいて、前記新たな配置先候補ポートそれぞれに対する前記優先度を決定する、請求項1に記載の計算機システム。

【請求項3】

前記管理システムは、さらに、前記第1ボリュームの前記第1ポートにおけるI/O量に関する情報を取得し、

前記管理システムは、前記新たな配置先候補ポートそれぞれに対する前記優先度を前記I/O量に関する情報に基づいて決定する、請求項1に記載の計算機システム。

【請求項4】

複数のホストプログラムが実行される複数のホスト計算機と、

前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステムと、

管理システムと、を含み、

前記複数のストレージシステムから複数のシングルストレージプールが構成され、

前記複数のシングルストレージプールのそれぞれは、前記複数のストレージシステムのうちの一部のストレージシステムであって、ネットワークを介してボリューム移動が可能である2以上のストレージシステムで構成され、

前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、

前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、

前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリュームを含み、

前記管理システムは、

前記複数のシングルストレージプールそれぞれを構成するストレージシステム及び前記複数のストレージシステムそれぞれのポートを管理する情報と、前記複数のストレージシステムの容量の利用状況を管理する情報とを参照して、前記第1ボリュームの新たな配置先候補ポートを、前記第1ボリュームのポートが属するシングルストレージプール内のポートであって、前記第1ボリュームを格納できる残り容量を有し、前記第1ストレージシステムと異なる他のストレージシステムのポートから決定し、

前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、

前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記管理システムは、

さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における接続装置数についての情報及び前記第1ボリュームのアクセス状況についての情報を取得し、

前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況につい

10

20

30

40

50

ての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記複数のボリュームは、ホスト計算機の第2ポートに配置された第1ホストプログラムに提供される第2ボリュームを含み、

前記帯域についての情報は、前記第2ポートと前記第1ホストプログラムの新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記管理システムは、前記第2ポートと前記第1ホストプログラムの新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を参照して、前記第1ホストプログラムの新たな配置先候補ポートそれぞれに対する前記優先度を決定する、計算機システム。

【請求項5】

複数のホストプログラムが実行される複数のホスト計算機と、

前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステムと、

管理システムと、を含み、

前記複数のストレージシステムから複数のシングルストレージプールが構成され、

前記複数のシングルストレージプールのそれぞれは、前記複数のストレージシステムのうちの一部のストレージシステムであって、ネットワークを介してボリューム移動が可能である2以上のストレージシステムで構成され、

前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、

前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、

前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリュームを含み、

前記管理システムは、

前記複数のシングルストレージプールそれぞれを構成するストレージシステム及び前記複数のストレージシステムそれぞれのポートを管理する情報と、前記複数のストレージシステムの容量の利用状況を管理する情報とを参照して、前記第1ボリュームの新たな配置先候補ポートを、前記第1ボリュームのポートが属するシングルストレージプール内のポートであって、前記第1ボリュームを格納できる残り容量を有し、前記第1ストレージシステムと異なる他のストレージシステムのポートから決定し、

前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、

前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記管理システムは、

さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における接続装置数についての情報及び前記第1ボリュームのアクセス状況についての情報を取得し、

前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況についての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記複数のボリュームは、第1ホストプログラムに提供される第2ボリュームを含み、

前記第1ホストプログラムは第2ポートに配置されており、

前記帯域についての情報は、前記第2ポートと前記第2ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記管理システムは、前記第2ポートと前記第2ボリュームの新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を参照して、前記第2ボリュームの新たな配置先候補ポートそれぞれに対する優先度を決定する、計算機システム。

【請求項6】

複数のホストプログラムが実行される複数のホスト計算機と、前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステムと、を含む計算機システムを管理する、管理システムであって、

10

20

30

40

50

前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、
前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、
前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリ

ュームを含み、
前記管理システムは、

前記複数のストレージシステムにおいて、前記第1ボリュームの新たな配置先候補ポ
ートを決定し、

前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネ
ットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、

前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける
前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域について
の情報を含み、

さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における
接続装置数についての情報及び前記第1ボリュームのアクセス状況についての情報を取得
し、

前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況につ
いての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記帯域についての情報は、前記第1ポートと前記新たな配置先候補ポートとの間の経
路それぞれにおける最小帯域の情報を含み、

前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれにおける最小帯域
に基づいて、前記新たな配置先候補ポートそれぞれに対する前記優先度を決定する、管理
システム。

【請求項7】

さらに、前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれにおける
最小帯域経路の利用状況の情報を取得し、

前記第1ポートと前記新たな配置先候補ポートとの間の経路それぞれの前記利用状況に
基づいて、前記新たな配置先候補ポートそれぞれに対する前記優先度を決定する、請求項
6に記載の管理システム。

【請求項8】

さらに、前記第1ボリュームの前記第1ポートにおけるI/O量に関する情報を取得し
、

前記新たな配置先候補ポートそれぞれに対する前記優先度を前記I/O量に関する情報
に基づいて決定する、請求項6に記載の管理システム。

【請求項9】

複数のホストプログラムが実行される複数のホスト計算機と、前記複数のホストプログ
ラムに提供される複数のボリュームを含む複数のストレージシステムと、を含む計算機シ
ステムを管理する、管理システムであって、

前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、

前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、

前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリ
ュームを含み、

前記複数のストレージシステムから複数のシングルストレージプールが構成され、

前記複数のシングルストレージプールのそれぞれは、前記複数のストレージシステムの
うちの一部のストレージシステムであって、ネットワークを介してボリューム移動が可能
である2以上のストレージシステムで構成され、

前記管理システムは、

前記複数のシングルストレージプールそれぞれを構成するストレージシステム及び前記
複数のストレージシステムそれぞれのポートを管理する情報と、前記複数のストレージシ
ステムの容量の利用状況を管理する第2情報と、を参照して、前記第1ボリュームの新た
な配置先候補ポートを、前記第1ボリュームのポートが属するシングルストレージプール

10

20

30

40

50

内のポートであって、前記第1ボリュームを格納できる残り容量を有し、前記第1ストレージシステムと異なる他のストレージシステムのポートから決定し、

前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、

前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における接続装置数についての情報及び前記第1ボリュームのアクセス状況についての情報を取得し、

前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況についての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記複数のボリュームは、ホスト計算機の第2ポートに配置された第1ホストプログラムに提供される第2ボリュームを含み、

前記帯域についての情報は、前記第2ポートと前記第1ホストプログラムの新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記第2ポートと前記第1ホストプログラムの新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を参照して、前記第1ホストプログラムの新たな配置先候補ポートそれぞれに対する前記優先度を決定する、管理システム。

【請求項10】

複数のホストプログラムが実行される複数のホスト計算機と、前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステムと、を含む計算機システムを管理する、管理システムであって、

前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置され、

前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置され、

前記複数のボリュームは、第1ストレージシステムの第1ポートに配置された第1ボリュームを含み、

前記複数のストレージシステムから複数のシングルストレージプールが構成され、

前記複数のシングルストレージプールのそれぞれは、前記複数のストレージシステムのうちの一部のストレージシステムであって、ネットワークを介してボリューム移動が可能である2以上のストレージシステムで構成され、

前記管理システムは、

前記複数のシングルストレージプールそれぞれを構成するストレージシステム及び前記複数のストレージシステムそれぞれのポートを管理する情報と、前記複数のストレージシステムの容量の利用状況を管理する第2情報と、を参照して、前記第1ボリュームの新たな配置先候補ポートを、前記第1ボリュームのポートが属するシングルストレージプール内のポートであって、前記第1ボリュームを格納できる残り容量を有し、前記第1ストレージシステムと異なる他のストレージシステムのポートから決定し、

前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得し、

前記帯域についての情報は、前記第1ポートと前記複数のストレージシステムにおける前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

さらに、前記第1ポートと前記新たな配置先候補ポートそれぞれとの間の経路における接続装置数についての情報及び前記第1ボリュームのアクセス状況についての情報を取得し、

前記帯域についての情報、前記接続装置数についての情報及び前記アクセス状況についての情報を参照して、前記新たな配置先候補ポートそれぞれに対する優先度を決定し、

前記複数のボリュームは、第1ホストプログラムに提供される第2ボリュームを含み、

前記第1ホストプログラムは第2ポートに配置されており、

10

20

30

40

50

前記帯域についての情報は、前記第2ポートと前記第1ボリュームの前記新たな配置先候補ポートそれぞれとの間の経路の帯域についての情報を含み、

前記第2ポートと前記第2ボリュームの新たな配置先候補ポートそれぞれとの間の経路の帯域について前記情報を参照して、前記第2ボリュームの新たな配置先候補ポートそれぞれに対する前記優先度を決定する、管理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システム及びその管理システムに関し、特に、計算機システムにおけるボリューム又はホストプログラムの配置に関する。

10

【背景技術】

【0002】

近年、計算機で取り扱われる情報量は飛躍的に増大し、これに応じて情報処理を行うホストやストレージをホスティングするデータセンタの大規模化が進んでいる。データセンタに設置されるサーバ、ストレージシステム、ネットワークは、それぞれ個々の装置毎に管理コストを削減する技術進化がなされてきた。

【0003】

サーバは単体性能の高性能化と併せて、高集積なブレード型の装置に仮想化機構を搭載し、複数サーバ間で負荷分散可能なアーキテクチャが採用され始めている。ストレージシステムは、多数の小型ストレージに大容量データを格納するのではなく、高性能、高信頼、高可用性な大型ストレージシステムに大容量データを集約して格納する設計が行われている。ネットワーク機器についても、装置あたりの接続数を増大し、処理性能を向上することで、増大するシステム間のトラフィックに対応している。

20

【0004】

このような状況の中、データセンタでは、個々の装置を調達し、大規模システムを構築・運用するための管理コストが新たな課題として発生してきた。このコストを削減するため、サーバ、ストレージ、ネットワークのシステムを構築済みの統合型システムが提案され始めている。

【0005】

この統合型システムにおいては、ある決まったスペックのサーバやストレージシステム、ネットワークのコンポーネントが構築済みであるため、その性能の範囲内で活用することを想定した場合には構築・運用コストを削減することが可能である。

30

【0006】

システム性能の向上の課題の一つとして、ストレージシステムの管理が挙げられる。性能や容量の設計の難しさやストレージエリアネットワーク(SAN)の構築の難易度が高いことから、大規模なデータセンタではストレージシステム専門の管理者が配置されており、単純化が特に求められており、技術開発がなされている。

【0007】

大規模ストレージシステムを実現する技術の一つに、ストレージエリアネットワークがある。1台以上のストレージシステムと複数のホスト計算機をストレージエリアネットワークにより接続し、複数のホスト計算機が、1台以上のストレージシステムからなる大容量記憶システムを共有する計算機システムが広く普及している。

40

【0008】

このような計算機システムでは、記憶資源や計算機資源を後から追加・削除・交換することが容易であり、拡張性に優れるという利点がある。SANに接続するストレージシステムには、RAID(Redundant Array of Independent Disks)構成された記憶装置を有するストレージシステムが一般的によく利用される。

【0009】

また、SANに含まれるファイバチャネルスイッチは、ファイバチャネルネットワーク

50

の論理的な分割単位であるゾーニング機能を提供する。このファイバチャネルゾーニングのひとつの形態は、ひとつのゾーンをネットワークインタフェースの集合で構成する。

【0010】

ゾーンに登録されたネットワークインタフェースは、ゾーンに登録されていないネットワークインタフェースとのデータ転送を受け付けない。すなわち、データ転送は、ゾーンに登録されたネットワークインタフェースの間でのみ可能となる。また、ファイバチャネルスイッチは、SANファブリックに接続されている機器の情報を管理し、新規機器のログインやログアウトに応じて状態を更新する、又は、クエリに応じて接続状態を応答することが可能である。

【0011】

特許文献1は、複数の仮想ボリュームと複数のプールを有するストレージシステムを開示する。各プールは、性能が異なる複数の階層に階層化されている。ストレージシステムは、各仮想ボリュームと実領域との関係及び実領域のアクセス頻度を表す情報を管理する。管理システムは、各仮想ボリュームの要求性能と性能を表す管理情報を有する。管理システムは、管理情報を基に、状況が不適切である仮想ボリュームを特定し、その仮想ボリュームの状況を適切状況に変えるために、その仮想ボリュームのプールに関わるマイグレーションを行う。

【先行技術文献】

【特許文献】

【0012】

【特許文献1】国際公開第2011/092738号パンフレット

【発明の概要】

【発明が解決しようとする課題】

【0013】

ストレージシステムに対する必要なシステム性能が拡大し、単一のストレージシステムの性能を超える場合には、複数のストレージシステムを一つのリソースとして利用できることが、システムの性能の拡張及び管理コストの削減に重要である。

【0014】

しかし、複数のストレージシステムが互いにデータ移動用のSANで接続されており、そのSANを介してボリュームの移動が可能である構成(シングルストレージプール構成)において、ホスト計算機と複数ストレージシステムとの間のSANが、均一でないことがある。

【0015】

このようなシステムにおいて、一つのストレージシステムから、管理システムにより提示されたボリューム移動可能範囲における他のストレージシステムに、ボリュームを移動すると、ホスト計算機に対するボリュームのI/O(Input/Output)性能が大きく変化してしまうことがある。特に、移動先のストレージシステムとホスト計算機との間パスのI/O帯域が狭い場合、ホスト計算機に対するボリュームのI/O性能が低下する。

【0016】

若しくは、複数のストレージシステムによるシングルストレージプールが構成されており、それらの内の一つのストレージシステムに仮想マシンをデプロイする場合に、ホスト計算機と複数ストレージシステムとの間のSANが均一でないため、仮想マシンをデプロイするストレージシステムによって、I/O性能が異なる可能性がある。特に、ストレージシステムとホスト計算機との間パスのI/O帯域が狭い場合、デプロイされる仮想マシンに対するボリュームのI/O性能が低いものになってしまう。

【課題を解決するための手段】

【0017】

本発明の一態様は、複数のホストプログラムが実行される複数のホスト計算機と、前記複数のホストプログラムに提供される複数のボリュームを含む複数のストレージシステム

10

20

30

40

50

と、管理システムと、を含む計算機システムである。前記複数のホストプログラムのそれぞれは、ホスト計算機のポートに配置される。前記複数のボリュームのそれぞれは、ストレージシステムのポートに配置される。前記管理システムは、前記複数のストレージシステム及び前記複数のホスト計算機の間におけるデータ通信ネットワークに含まれる複数の経路のそれぞれの帯域についての情報を取得する。前記管理システムは、前記情報を参照して、第1ホストプログラム及び第1ボリュームの少なくとも一方について、新たな配置先候補ポートそれぞれに対する優先度を決定する。

【発明の効果】

【0018】

本発明の一態様によれば、ボリュームとホストプログラムの位置関係を適切に選択することができる。他の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

10

【図面の簡単な説明】

【0019】

【図1】第1実施形態の計算機システム全体の構成図である。

【図2】第1実施形態の管理計算機のハードウェア構成の一例を示す図である。

【図3】第1実施形態のストレージシステムのハードウェア構成の一例を示す図である。

【図4】第1実施形態のホスト計算機のハードウェア構成の一例を示す図である。

【図5】第1実施形態の接続装置のハードウェア構成の一例を示す図である。

【図6】第1実施形態のストレージシステムの不揮発記憶装置のボリューム構成及び格納情報の一例を示す図である。

20

【図7】第1実施形態の管理計算機の不揮発記憶装置が格納する情報の一例を示す図である。

【図8】第1実施形態の管理計算機におけるストレージシステム管理テーブルの構成例を示す図である。

【図9】第1実施形態の管理計算機における接続装置管理テーブルの構成例を示す図である。

【図10】第1実施形態の管理計算機における移動範囲管理テーブルの構成例を示す図である。

【図11】第1実施形態の管理計算機におけるストレージトポロジ管理の閾値のテーブルの構成例を示す図である。

30

【図12】第1実施形態の管理計算機におけるボリューム管理テーブルの構成例を示す図である。

【図13】第1実施形態の管理計算機におけるストレージシステム使用状況管理テーブルの構成を示す図である。

【図14】第1実施形態の管理計算機におけるシンプロビジョニングプール使用状況管理テーブルの構成を示す図である。

【図15】第1実施形態の管理計算機における移動先推奨度テーブルの構成例を示す図である。

【図16】第1実施形態のストレージシステムの管理情報格納部の格納情報の一例を示す図である。

40

【図17】第1実施形態のストレージシステムにおけるストレージトポロジ管理テーブルの構成例を示す図である。

【図18】第1実施形態のホスト計算機の不揮発性記憶装置の格納情報の一例を示す図である。

【図19】第1実施形態のホスト計算機における接続先ストレージシステムボリューム管路テーブルの構成例を示す図である。

【図20】第1実施形態の接続装置の不揮発性記憶装置が格納する情報の一例を示す図である。

【図21】第1実施形態の接続装置における接続装置管理テーブルの構成例を示す図であ

50

る。

【図 2 2】第 1 実施形態のボリュームマイグレーションの例を説明するシーケンス図である。

【図 2 3】第 1 実施形態の移動先候補ポートの推奨度の算出方法の例を示すフローチャートである。

【図 2 4】第 1 実施形態のストレージポロジ管理テーブルの更新の例を説明するシーケンス図である。

【図 2 5】第 2 実施形態のストレージシステムの管理情報格納部の格納情報の一例を示す図である。

【図 2 6】第 2 実施形態のストレージシステムにおけるボリュームアクセス状況管理テーブルの構成例を示す図である。

10

【図 2 7 A】第 2 実施形態の移動先候補ポートの推奨度の算出方法の例を示すフローチャートの一部を示す図である。

【図 2 7 B】第 2 実施形態の移動先候補ポートの推奨度の算出方法の例を示すフローチャートの一部を示す図である。

【図 2 8】第 3 実施形態の管理計算機の不揮発記憶装置が格納する情報の一例を示す図である。

【図 2 9】第 3 実施形態の管理計算機におけるホスト管理テーブルの構成例を示す図である。

【図 3 0】第 3 実施形態の管理計算機におけるホストポロジ管理の閾値のテーブルの構成例を示す図である。

20

【図 3 1】第 3 実施形態のストレージシステムの管理情報格納部の格納情報の一例を示す図である。

【図 3 2】第 3 実施形態のストレージシステムにおけるホストポロジ管理テーブルの構成例を示す図である。

【図 3 3】第 3 実施形態において、仮想マシンの配置先候補ポートの推奨度の算出方法の例を示すフローチャートである。

【図 3 4】第 3 実施形態において、ボリュームの接続ポートの推奨度の算出方法の例を示すフローチャートである。

【発明を実施するための形態】

30

【0020】

以下、図面を参照しつつ、本発明を実施するための複数の実施形態を説明する。説明の明確化のため、以下の記載及び図面は、適宜、省略及び簡略化がなされている。各実施形態の構成の一部を他の実施形態の構成に置き換えることが可能であり、各実施形態の構成に他の実施形態の構成を加えることも可能である。実施形態は本発明の態様を示すが、本発明が実施形態のみに限定されるものではない。

【0021】

第 1 実施形態

第 1 実施形態の情報記憶システムは、複数のストレージシステムを含み、その複数のストレージシステム内において、ボリューム（データ）を移動することができる。好ましい構成において、情報記憶システムは、ホストからの I/O を停止することなく、ボリュームをストレージシステム間で移動することができる（ボリュームのライブマイグレーション）。このように、ストレージシステム間でボリュームを移動できる情報記憶システムをシングルストレージプールとも呼ぶ。

40

【0022】

ホスト計算機は、ネットワークを介して、情報記憶システムにおける複数のストレージシステムのボリュームにアクセスすることができる。ホスト計算機上において、ホストプログラムであるオペレーティングシステム（OS）又は仮想マシン（ゲスト OS を含む）が動作している。

【0023】

50

第1実施形態は、ホスト計算機とストレージシステムと間のデータネットワークにおける、ストレージシステム間の経路の構成及び帯域特性の情報を参照して、ボリュームの移動先候補ポートの優先度を決定する。移動先候補ポートは、現在のストレージシステムと同一又は異なるストレージシステムのポートである。これにより、ホスト計算機からボリュームへのアクセス性能のボリューム移動前後の変化を小さくすることができる。

【0024】

本実施形態は、移動先候補ポートの優先度の算出において、移動元のストレージシステムポートと移動先のストレージシステムポートとの間の経路の構成及び帯域特性の情報を使用する。これにより、ホスト計算機とストレージシステムとの間の経路についての情報を有していなくとも、ボリュームの適切な移動先を決定することができる。

10

【0025】

第1実施形態は、経路の構成情報の例として経路の経由接続装置（スイッチ）の数を参照し、経路の帯域特性の例として経路における最小帯域を参照する。第1実施形態は、一例において、さらに、経路の利用状況を参照して優先度を決定する。これにより、適切に移動先候補のストレージシステムの優先度を決定する。以下において、第1実施形態のより具体的な構成を、図を参照して説明する。

【0026】

図1は、第1実施形態における、計算機システム全体の構成例を説明するための図である。本図に示すように、計算機システムは管理計算機1と、複数の（図1の例において3台の）ストレージシステム2と、1台以上（図1の例において2台）のホスト計算機3と、複数の接続装置40からなる接続装置群4、そして接続装置5、6を備える。本計算機システムの接続装置は、入力ポートで受信したデータを解析してそのデータを送信する出力ポートを選択し、その出力ポートから受信データを送信するスイッチである。接続装置は、例えば、ファイバチャネルスイッチである。

20

【0027】

管理計算機1、ストレージシステム2、ホスト計算機3、接続装置群4の接続装置40、接続装置5、6は、それぞれ、1以上のポートを有しており、各装置（管理計算機1、ストレージシステム2、ホスト計算機3を含む）のポートが、他の装置のポート（対向ポート）に接続される。対向ポートは、一般に、ケーブルで接続される。

【0028】

接続装置群4は、ホスト計算機3とストレージシステム2とを通信可能に接続し、それらの間のデータ（ホストデータ）を転送する。接続装置群4における接続装置40は複数のポートを有し、接続装置40のポートは、ホスト計算機3、ストレージシステム2又は接続装置群4における他の接続装置40のポートと接続する。

30

【0029】

接続装置群4は、ホスト計算機3からのデータI/O要求（データの入力又は出力の要求）及びライトデータを、ストレージシステム2に転送する。さらに、ストレージシステム2からのリードデータをホスト計算機3に転送する。接続装置群4は、さらに、ストレージシステム2間を接続する経路を含む。ただし、本構成において、ストレージシステム2間のボリューム（データ）の移動は、接続装置5を介して行われる。

40

【0030】

接続装置5は、ストレージシステム2間のデータを転送する。接続装置5は複数のポートを有し、そのポートはストレージシステム2のポートと接続する。接続装置5は、ストレージシステム2間のデータI/O要求及び要求されているデータを転送する。本実施形態のストレージシステム2は、ストレージシステム2間においてボリューム（データ）を移動する。ボリュームマイグレーションは、接続装置5を介して行われる。

【0031】

接続装置6は、管理計算機1と他の装置、つまり、ストレージシステム2、ホスト計算機3、接続装置群4の各接続装置40及び接続装置5とを通信可能に接続し、管理計算機1とそれらとの間のデータを転送する。接続装置6は複数のポートを有し、そのポートは

50

、管理計算機 1 のポート又は他の装置のポートと接続する。接続装置 6 は、管理計算機 1 からの管理用通信要求を、ストレージシステム 2、ホスト計算機 3、接続装置群 4 の接続装置 40、接続装置 5 へ転送し、それらからの応答を管理計算機 1 に転送する。

【0032】

本計算機システムにおいて、接続装置群 4 を使用してホスト計算機 3 とストレージシステム 2 との間に SAN (Storage Area Network) が構成されている。接続装置 5 を使用して、ストレージシステム 2 の間に SAN が構成されている。ストレージシステム 2 は、接続装置 5 を介することなく、直接接続されてもよい。本例において、接続装置 6 により、計算機システム全体が管理計算機 1 から IP ネットワークにて接続可能である。

10

【0033】

接続装置群 4 によるネットワーク及び接続装置 5 によるネットワークは、物理的に同一の接続装置で提供され、論理的に独立なネットワークでもよい。また、接続装置群 4 によるネットワーク及び接続装置 5 によるネットワークは、iSCSI (Internet Small Computer System Interface) で構成される IP ネットワークでもよく、接続装置 6 で提供される IP ネットワークと同一の装置で提供され、論理的に独立なネットワークでもよい。

【0034】

さらに、接続装置群 4、接続装置 5、接続装置 6 によるネットワークは、FCoE (Fibre Channel over Ethernet) で構成されていてもよく、これらは、同一の装置で提供され、論理的に独立なネットワークでもよい。

20

【0035】

図 2 は、管理計算機 1 のハードウェア構成例を示す。管理計算機 1 はキャッシュメモリ 11、不揮発記憶装置 12、プロセッサ 13、プログラムメモリ 14、管理用通信 I/F 16、内部バス 17、画像出力装置 18、入力装置 19 などを有する計算機である。不揮発記憶装置 12 は、ネットワークを介して接続された外部記憶装置でもよい。

【0036】

管理計算機 1 は、不揮発記憶装置 12 に格納された OS やアプリケーションプログラムなどのソフトウェアをプログラムメモリ 14 にロードし、プロセッサ 13 がプログラムメモリ 14 からそれを読みだして実行することで、所定の機能を達成する。プロセッサ 13 は、複数のチップ及び複数のパッケージを含むことができる。この点は、他の装置 (ストレージシステム 2 を含む) において同様である。

30

【0037】

管理計算機 1 は、頻繁に用いるデータをキャッシュメモリ 11 に一時的に格納し、処理を高速化する。管理計算機 1 は、キーボードやマウスなどの入力装置 19 やディスプレイなどの画像出力装置 18 を有し、入力装置 19 が管理者 (ユーザ) からの入力を受け付け、画像出力装置 18 がプロセッサ 13 から指示された情報を出力する。管理計算機 1 は、さらに、接続装置 6 と接続する管理用通信 I/F 16 を有する。

【0038】

図 1 の構成例において、管理システムは管理計算機 1 で構成されているが、管理システムは複数の計算機で構成してもよい。複数の計算機の一つは表示用計算機であってもよく、管理処理の高速化や高信頼化のために、複数の計算機が管理計算機と同等の処理を実行してもよい、管理システムは、ストレージシステム又は接続装置の一部の機能を含むことができる。

40

【0039】

図 3 は、ストレージシステム 2 のハードウェア構成例を示す。ストレージシステム 2 はキャッシュメモリ 21、不揮発記憶装置 22、プロセッサ 23、プログラムメモリ 24、データ I/O 用通信 I/F 25、管理用通信 I/F 26、内部バス 27、画像出力装置 28、入力装置 29 を有する。

【0040】

50

ストレージシステム 2 は、不揮発記憶装置 2 2 に格納された OS や制御プログラムなどのソフトウェアをプログラムメモリ 2 4 にロードし、プロセッサ 2 3 がプログラムメモリ 2 4 から読みだしてそれを実行することで、所定の機能を達成する。プロセッサ 2 3 は、ストレージシステム 2 のコントローラとして機能する。ストレージシステム 2 は、頻繁に用いるデータをキャッシュメモリ 2 1 に格納する。

【 0 0 4 1 】

不揮発性記憶装置 2 2 は、複数の記憶ドライブを含む。記憶ドライブは、不揮発性の磁気ディスクを有するハードディスクドライブ (HDD)、不揮発半導体メモリ (例えばフラッシュメモリ) を搭載した Solid State Drive (SSD) を含む。記憶ドライブは、ホスト計算機 3 から送られたデータを格納する。複数の記憶ドライブが RAID 演算によるデータの冗長化を行うことで、1つの記憶ドライブに障害が発生した場合のデータ消失を防ぐことができる。不揮発性記憶装置 2 2 におけるボリューム構成については後述する。

10

【 0 0 4 2 】

ストレージシステム 2 は、キーボードやマウスなどの入力装置 2 9 及びディスプレイなどの画像出力システム 2 8 を有し、入力装置 2 9 が管理者 (ユーザ) からの入力を受け付け、画像出力システム 2 8 がプロセッサ 2 3 から指示された情報を出力する。ストレージシステム 2 は、それぞれ接続装置群 4 又は接続装置 5 と接続するデータ I/O 用通信 I/F 2 5 を有し、さらに、接続装置 6 と接続する管理用通信 I/F 2 6 を有する。管理者は、管理計算機 1 から接続装置 6 を介してストレージシステム 2 にアクセスし、ストレージシステム 2 を管理、制御することができる。

20

【 0 0 4 3 】

図 1 は、簡単のため 3 つのストレージシステム 2 を例示しているが、その数はシステム設計に依存する。データ I/O 用通信 I/F 2 5 には計算機システムで一意的な物理ポート ID が割り当てられており、NP IV (N__Port ID Virtualization) 又はその他の仮想化技術により、複数の仮想ポート ID を割り当てることが可能である。

【 0 0 4 4 】

例えば、NP IV は、N__Port の ID を仮想化し、一つの物理ポート (N__Port) 上に複数の仮想ポートを割り当てる。NP IV は、一つの物理ポートに対して、複数のホスト計算機 3 又は複数の仮想マシンのそれぞれの仮想ポートを割り当てることができる。仮想ポートが使用される場合、仮想ポート ID と物理ポート ID との関連付けが計算機システム内においてなされる。以下の説明において、計算機、ストレージシステム又は接続装置のポート ID という語は、特に指摘のない場合、物理ポートの ID を意味する。

30

【 0 0 4 5 】

図 4 は、ホスト計算機 3 のハードウェア構成例を示す。ホスト計算機 3 は、キャッシュメモリ 3 1、不揮発記憶装置 3 2、プロセッサ 3 3、プログラムメモリ 3 4、データ I/O 用通信 I/F 3 5、管理用通信 I/F 3 6、内部バス 3 7、画像出力装置 3 8、入力装置 3 9 などを有する計算機である。

【 0 0 4 6 】

ホスト計算機 3 は、不揮発記憶装置 3 2 に格納された OS やアプリケーションプログラムなどのソフトウェアをプログラムメモリ 3 4 にロードし、プロセッサ 3 3 がプログラムメモリ 3 4 から読みだして実行することで、所定の機能を達成する。ホスト計算機 3 は、頻繁に用いるデータをキャッシュメモリ 3 1 に格納する。

40

【 0 0 4 7 】

ホスト計算機 3 は、キーボードやマウスなどの入力装置 3 9 やディスプレイなどの画像出力装置 3 8 を有し、入力装置 3 9 が管理者 (ユーザ) からの入力を受け付け画像出力装置 3 8 がプロセッサ 3 3 から指示された情報を出力する。ホスト計算機 3 は、接続装置群 4 と接続する 1 つ以上のデータ I/O 用通信 I/F 3 5 を有し、接続装置 6 と接続する管理用通信 I/F 3 6 を有する。

50

【0048】

図1は、簡単のため2つのホスト計算機3を例示しているが、システムはいくつのホスト計算機3を有していてもよい。データI/O用通信I/F35には計算機システムで一意のポートID割り当てられており、NPIV又はその他の仮想化技術により、複数の仮想ポートIDを割り当てることが可能である。例えば、NPIVは、ホスト計算機3の物理ポートに、複数の仮想マシンの仮想ポートを割り当てることができる。

【0049】

図5は、接続装置群4における接続装置40のハードウェア構成例を示す。接続装置は、キャッシュメモリ41、不揮発記憶装置42、プロセッサ43、プログラムメモリ44、データI/O用通信I/F45、管理用通信I/F46、内部バス47を有する。

10

【0050】

接続装置40は、不揮発記憶装置42に格納されたOSや制御プログラムなどのソフトウェアをプログラムメモリ44にロードし、プロセッサ43がプログラムメモリ44から読みだして実行することで、所定の機能を達成する。接続装置40は、頻繁に用いるデータをキャッシュメモリ41に格納する。接続装置5及び接続装置6は、これと同様のハードウェア構成を有することができる。

【0051】

図6は、ストレージシステム2の不揮発記憶装置22のボリューム構成を示す。不揮発記憶装置22は、ストレージシステム管理情報格納部221と不揮発記憶媒体を備える異なる種類のメディア222A~222Cを含む。本実施形態において、1又は複数の記憶ドライブが提供する記憶領域及びそのインターフェースを含む構成をメディアと呼ぶ。典型的には、メディアは、RAIDグループが提供する記憶領域及びそのインターフェースを含む構成である。

20

【0052】

図6の例において、半導体メモリであるフラッシュメモリを利用したSSDメディア222A、Serial Attached SCSIのインターフェースを備えたSAS (Serial Attached SCSI)メディア222B、Serial Advanced Technology Attachmentのインターフェースを備えたSATAメディア222Cが、例示されている。不揮発記憶装置22は、典型的には、複数のSSDメディア222A、複数のSASメディア222B、複数のSATAメディア222Cを含む。

30

【0053】

ストレージシステム2は、メディアが提供する実記憶領域により構成され、ホスト計算機3やシンプロビジョニングプール223から読み書き可能な、1つ以上のボリューム(実ボリューム)2221を提供する。単一種類の記憶ドライブの記憶領域によってボリューム2221を構成してもよいし、複数種類の記憶ドライブによりRAID (Redundant Arrays of Independent Disks)を構成し、そのRAIDによりボリューム2221を提供してもよい。以下においては、単一種類の記憶ドライブの記憶領域によってボリューム2221を構成する例を説明する。

【0054】

シンプロビジョニングプール223において、メディア222A~222Cのボリューム2221群の中から管理者(ユーザ)により選択されたボリューム2221が、複数ページ2231に分割される。シンプロビジョニングプール223は、ページ単位で管理される。シンプロビジョニングボリューム224が格納すべきデータ量に応じて、ページがシンプロビジョニングボリューム224に割り当てられる。

40

【0055】

シンプロビジョニングボリューム224は、ホスト計算機3に提供される仮想ボリュームであって、その容量は仮想化されている。シンプロビジョニングボリューム224にホスト計算機3から書き込みがありデータ格納領域が必要になると、ストレージシステム2は、ページ2231をシンプロビジョニングボリューム224に割り当てる。

50

【 0 0 5 6 】

シンプロビジョニングは、ホスト計算機 3 により認識されるシンプロビジョニングボリューム 2 2 4 の容量を、シンプロビジョニングボリューム 2 2 4 に割り当てられている実容量（全ページの容量）よりも大きくすることができ、ホスト計算機 3 に割り当てられるボリューム容量を実現するために必要な実用量を、それよりも小さくすることができる。

【 0 0 5 7 】

図 6 において、シンプロビジョニングプール 2 2 3 は、複数の実ボリューム（プールボリューム） 2 2 2 1 を含む。シンプロビジョニングプール 2 2 3 は、一般に、ストレージシステム 2 内の性能の異なる複数のメディアを含み、アクセス性能によって複数の階層に階層化される。各階層は、1 又は複数の実ボリューム 2 2 2 1 で構成されている。

10

【 0 0 5 8 】

ストレージシステム 2 は、シンプロビジョニングボリューム 2 2 4 に対するホスト計算機 3 からの I / O を分析し、I / O 負荷の高いページを、性能の高い高価なメディアで構成されたリソースから成る上位階層に、そうでないページを性能の低い安価なメディアで構成されたリソースから成る下位階層に自動配置する。これにより、仮想ボリュームへのアクセス性能を維持しつつ、システムコストを低減することができる。

【 0 0 5 9 】

図 6 は示していないが、ストレージシステム 2 は、容量が仮想化されているシンプロビジョニングボリュームの他、容量が仮想化されずホストが認識する容量と実容量が一致する通常属性のボリュームも提供することができる。

20

【 0 0 6 0 】

図 7 は、管理計算機 1 の不揮発記憶装置 1 2 が格納している情報を示している。ストレージシステム管理プログラム 1 2 1 は、計算機システム内のストレージシステム 2 と、管理用通信 I / F 1 6 に接続された接続装置 6 を経由して通信し、その状態の参照、構成情報の設定を行う。接続装置管理プログラム 1 2 2 は、接続装置群 4 の各接続装置 4 0 及び接続装置 5 と、管理用通信 I / F 1 6 に接続された接続装置 6 を経由して通信し、その状態の参照、構成情報の設定を行う。

【 0 0 6 1 】

ホスト計算機管理プログラム 1 2 3 は、計算機システム内のホスト計算機 3 のエージェント 3 2 2 と、管理用通信 I / F 1 6 に接続された接続装置 6 を経由して通信し、その状態の参照、構成情報の設定を行う。運用管理プログラム 1 2 4 は、管理計算機 1 の不揮発記憶装置 1 2 内にある各種テーブルを参照し、ボリューム配置の計画及びその他の処理を行う。運用管理プログラム 1 2 4 は、各種プログラムと連携して計算機システム内の運用管理を行い、管理者に対してユーザインタフェースを提供する。オペレーティングシステム 1 2 5 は、各種プログラムの実行基盤であり、各種テーブルの管理基盤である。

30

【 0 0 6 2 】

各プログラムは、プロセッサ 1 3 によって実行されることで、定められた処理を記憶装置及び通信インタフェースを用いながら行う。従って、本実施形態においてプログラムを主語とする説明は、プロセッサを主語とした説明でもよい。若しくは、プログラムが実行する処理は、そのプログラムが動作する計算機及び計算機システムが行う処理である。

40

【 0 0 6 3 】

プロセッサ 1 3 は、プログラムに従って動作することによって、所定の機能を実現する機能部として動作する。例えば、プロセッサ 1 3 は、ストレージシステム管理プログラム 1 2 1 に従って動作することでストレージシステム管理部として機能し、接続装置管理プログラム 1 2 2 に従って動作することで接続装置管理部として機能する。他のプログラムについても同様である。

【 0 0 6 4 】

さらに、プロセッサ 1 3 は、各プログラムが実行する複数の処理のそれぞれを実現する機能部としても動作する。プロセッサ 1 3 及びプログラムを含む管理計算機 1 は、これらの機能部を含む装置である。上記プログラムとプロセッサの関係、さらに、これらとこれ

50

らを含む装置との関係の説明は、他の装置及びシステム、つまり、ホスト計算機 3、接続装置群 4 の接続装置、接続装置 5、6、ストレージシステム 2 において同様である。

【0065】

不揮発記憶装置 12 は、この他、ストレージシステム管理テーブル 100、接続装置管理テーブル 110、移動範囲管理テーブル 130、ストレージトポロジ管理の閾値のテーブル 140、ボリューム管理テーブル 150、ストレージシステム使用状況管理テーブル 160、シンプロビジョニングプール使用状況管理テーブル 170、移動先推奨度テーブル 180 を格納する。

【0066】

本実施形態において、管理計算機 1 が使用する情報は、データ構造に依存せずどのようなデータ構造で表現されていてもよい。例えば、テーブル、リスト、データベース又はキューから適切に選択したデータ構造体が、情報を格納することができる。この点は、他の装置（ストレージシステムを含む）において同様である。

10

【0067】

図 8 は、管理計算機 1 が有するストレージシステム管理テーブル 100 の構成例を示す。ストレージシステム管理テーブル 100 は、計算機システム内のストレージシステム 2 を管理するための一意な識別子を格納するストレージシステム ID のコラム 1001 及び管理ポート IP アドレスのコラム 1002 を有する。管理ポート IP アドレスは、管理計算機 1 がストレージシステム 2 と通信を行うための IP アドレスである。

【0068】

20

これらの値は、例えば、管理者によって入力される。本実施形態において、対象を識別するための情報として、ID、名称及び番号などが使用されるが、これらの呼び名は互いに置換可能である。ストレージシステム管理テーブル 100 のエントリは、計算機システムへのストレージシステム 2 の追加又は削除に応じて、追加又は削除される。

【0069】

図 9 は、管理計算機 1 が有する接続装置管理テーブル 110 の構成例を示す図である。接続装置管理テーブル 110 は、接続装置 ID のコラム 1101、接続装置の管理ポート IP アドレスのコラム 1102 を有する。各コラムの値は、管理者により運用管理プログラム 124 から入力される。接続装置 ID は、計算機システム内で一意の識別子である。管理ポート IP アドレスは、管理計算機 1 が接続装置と通信を行うための IP アドレスである。管理計算機 1 は、接続装置群 4 及び接続装置 5、6 との管理通信のために、これを参照する。

30

【0070】

図 10 は、管理計算機 1 が有する移動範囲管理テーブル 130 の構成例を示している。移動範囲管理テーブル 130 は、ボリュームを移動することができる範囲、具体的には、ボリュームの移動先となることができる物理ポートを定義する。物理ポートは、計算機システム内において、ストレージシステム ID とポート ID とで一意に識別される。移動範囲管理テーブル 130 は、シングルストレージプール並びにそれに含まれるストレージシステム 2 及びそのポートの情報を格納するテーブルである。

【0071】

40

本例において、ボリュームは、そのボリュームが属するシングルストレージプール内で移動することができる。ホスト計算機 3 は、シングルストレージプール内の任意のポートにアクセスすることができ、シングルストレージプール内でのボリュームのライブマイグレーション（システムの無停止でのボリューム移動）を行うことができる。設計により、これとは異なるボリュームの移動範囲を定義してもよい。

【0072】

図 10 に示すように、移動範囲管理テーブル 130 は、シングルストレージプール ID のコラム 1301、ストレージシステム ID のコラム 1302、ポート ID のコラム 1303 を有する。図 10 の移動範囲管理テーブル 130 において、例えば、シングルストレージプール S P o o l 1 は、ストレージシステム S T 1、S T 2 及び S T 3 で構成されて

50

いる。シングルストレージプール S P o o l 1 のいずれかのポートが割り当てられているボリュームは、シングルストレージプール S P o o l 1 内の他のポートに移動することができる。

【 0 0 7 3 】

ボリュームは、シングルストレージプール内において、そのボリュームが格納されているストレージシステム 2 から他のストレージシステム 2 に移動することができ、また、そのボリュームが格納されているストレージシステム 2 内で移動することができる。同一ストレージシステム 2 内での移動は、割り当てるポートの変更である。

【 0 0 7 4 】

図 1 1 は、管理計算機 1 が使用するストレージトポロジ管理の閾値テーブル 1 4 0 の一例を示している。ストレージトポロジは、ストレージシステム 2 とホスト計算機 3 とをつなぐネットワーク（接続装置群 4 によるネットワーク）において、ストレージシステム 2 をつなぐネットワークのトポロジである。具体的には、ストレージシステム 2 間の経路の構成及び帯域の情報を含む。本例において、経路の構成情報として、経路における経由接続装置の数が使用され、経路の帯域情報として、経路における最小帯域が使用される。

【 0 0 7 5 】

ストレージトポロジ管理の閾値のテーブル 1 4 0 は、最小帯域閾値のカラム 1 4 0 1、経由接続装置数閾値のカラム 1 4 0 2、最小帯域経路の平均帯域利用率閾値のカラム 1 4 0 3、最小帯域経路の最大帯域利用率閾値のカラム 1 4 0 4、遅延時間閾値のカラム 1 4 0 5 を有する。運用管理プログラム 1 2 4 は、ボリュームの移動先ポートの推奨度の算出において、ストレージトポロジ管理の閾値テーブル 1 4 0 を参照する。推奨度の算出方法の詳細は後述する。ストレージトポロジ管理の閾値テーブル 1 4 0 の値は、管理者により予め設定される。

【 0 0 7 6 】

最小帯域閾値は、ストレージシステム 2 のポート間の経路（1 又は複数の接続装置 4 0 を含む）における最小帯域の閾値である。経由接続装置数閾値は、ストレージシステム 2 のポート間の経路における経由接続装置数の閾値である。最小帯域経路は、ストレージシステム 2 のポート間の経路に含まれる経路において最も帯域が小さい経路である。

【 0 0 7 7 】

ストレージシステムポートの間の経路は、接続装置のポートを含み、ストレージシステムポートと接続装置ポートとの間の経路、接続装置ポート間の経路を含む。例えば、ストレージシステムポート間の経路が、複数の接続装置 4 0 を通り、ある接続装置ペアのポート間の経路の帯域が最も小さい場合、その接続装置間の経路が最小帯域経路である。経路の最小帯域は、それを挟むポートペアのより小さい帯域の値により決まる。

【 0 0 7 8 】

平均帯域利用率は、所定時間、例えば 1 時間における経路の帯域利用率の平均値である。具体的には、所定時間内の利用帯域（例えば b p s で表される）の平均値を、経路の規定帯域値で割った値である。最大帯域利用率は、所定時間内において最も高い帯域利用率を示す。例えば、経路の規定帯域値が b p s で表される場合に、所定時間内の最も高い b p s の値を経路の規定帯域値で割った値である。

【 0 0 7 9 】

図 1 2 は、管理計算機 1 が有するボリューム管理テーブル 1 5 0 の構成例を示している。ボリューム管理テーブル 1 5 0 は、計算機システム内に存在するボリュームを、それぞれに一意の番号を割り当てて管理する。ボリューム管理テーブル 1 5 0 は、ボリューム番号のカラム 1 5 0 1 を有し、各ボリュームは、ボリューム番号が割り当てられ、それに対応して他の項目情報が管理される。

【 0 0 8 0 】

容量のカラム 1 5 0 2 は各ボリュームのボリューム容量を、使用量のカラム 1 5 0 3 は各ボリュームの使用量を格納する。ボリュームタイプのカラム 1 5 0 4 は、各ボリュームが通常ボリューム、シンプロビジョニングボリューム、プールボリューム（シンプロビジ

10

20

30

40

50

ョングプールに実容量を提供するボリューム)のいずれなのかを示す。

【0081】

ボリューム管理テーブル150は、さらに、ボリュームが存在するストレージシステムIDのカラム1505、そのストレージシステム内(のみ)でボリュームを一意に特定する識別情報であるボリュームIDのカラム1506、そして、そのボリュームに割り当てられているポートのポートIDのカラム1507を有する。ポートIDは、各ストレージシステム2内で一意の値である。

【0082】

ボリューム管理テーブル150のエントリは、例えば、ボリューム作成時に、管理者が運用管理プログラム124を使用して作成することができる。ストレージシステム管理プログラム121は、ストレージシステム2から、ボリューム管理テーブル150に格納する情報を得て、それを格納する。

【0083】

図13は、管理計算機1が有する、ストレージシステム使用状況管理テーブル160の構成例を示す図であり、各ストレージシステム2の容量の利用状況を管理する。ストレージシステム使用状況管理テーブル160は、ストレージシステムIDのカラム1601を有する。

【0084】

ストレージシステム使用状況管理テーブル160は、さらに、ストレージシステムIDにより特定されるストレージシステム2の、総容量のカラム1602、容量閾値のカラム1603、使用量のカラム1604の有する。総容量の値は、管理者又はストレージシステム2から得られ、容量閾値は管理者により設定される。使用量の値を取得するため、定期的に管理計算機1のストレージシステム管理プログラム121が、ストレージシステム2に接続装置6を経由して問い合わせ、取得した値をテーブル160に格納する。

【0085】

図14は、管理計算機1が有する、シンプロビジョニングプール使用状況管理テーブル170の例を示す図である。シンプロビジョニングプール使用状況管理テーブル170は、計算機システム内のシンプロビジョニングプール223の使用状況を管理する。図14は、一つのシンプロビジョニングプールT P o o l 1の情報のみを例示している。シンプロビジョニングプール使用状況管理テーブル170は、計算機システムにおいて共通して設定されるプールIDのカラム1701、1つ以上のストレージシステムにより提供される全体容量のカラム1702、その使用量のカラム1703を有する。

【0086】

さらに、シンプロビジョニングプール使用状況管理テーブル170は、プールを構成する個々のストレージシステム2のIDをカラム1704に格納し、さらに、各ストレージシステムの容量のカラム1705、使用量のカラム1706、I O P S (I n p u t O u t p u t P e r S e c o n d) のカラム1708を有する。

【0087】

さらに、シンプロビジョニングプール使用状況管理テーブル170は、ストレージシステム2において使用量、I O P S が超えた場合に調整を行う値として、容量閾値のカラム1707、性能閾値のカラム1709を有する。シンプロビジョニングプール使用状況管理テーブル170のエントリは、例えば、管理者によりシンプロビジョニングプール223の作成指示が行われたときに追加され、削除指示が行われたときに削除される。

【0088】

容量閾値、性能閾値の値は、管理者が運用管理プログラム124を使用して設定することができる。使用量のカラム1706及びI O P S のカラム1708の値を取得するため、管理計算機1のストレージシステム管理プログラム121は、定期的に、ストレージシステム2に、接続装置6を経由して問い合わせ、ストレージシステム2から取得した値を、それらのフィールドに格納する。

【0089】

10

20

30

40

50

図15は、管理計算機1が有する移動先推奨度テーブル180の例を示している。移動先推奨度テーブル180は、各ポートに接続されているボリュームの移動先ポートとしての他のポートの推奨度を示すテーブルである。図15の移動先推奨度テーブル180は、ストレージシステムST1におけるポートPT1に対する移動先ポートの推奨度を示すテーブルの例である。推奨度は、移動先ポートの優先度を表す値である。

【0090】

管理計算機1の運用管理プログラム124は、ストレージシステム2から取得する経路情報から、各ポートの移動先推奨度を算出することができる。典型的に、運用管理プログラム124は、各ストレージシステム2の各ポートの移動先推奨度テーブル180を、ボリューム移動先決定の際に作成する。推奨度の算出については後述する。

10

【0091】

ボリューム移動において、移動先推奨度テーブル180における推奨度が、管理者に提示される。管理者は、提示された推奨度を参照して、ボリュームの移動先を決定する。又は、管理計算機1は、算出した推奨度(優先度)を参照して、自動でボリュームの移動先を決定してもよい。

【0092】

図15の移動先推奨度テーブル180は、シングルストレージプールIDのカラム1801、ストレージシステムIDのカラム1802、移動先ポートIDのカラム1803、移動先推奨度のカラム1804を有する。本例において、移動先推奨度のカラム1804における星の数が多い程、推奨度が高いことを示している。

20

【0093】

図15の例において、ストレージシステムST1におけるポートPT1に接続するボリュームを移動することができるポートは、シングルストレージプールSPool1内における、当該ポートと異なる他のポートである。図15が示すように、移動先ポートは、同一ストレージシステムST1内の他のポートを含む。一つのストレージシステム2の異なるポートは、ホスト計算機3との間において、同一の接続装置を経由する又は異なる接続装置を経由する。二つの経路における接続装置40のポートが異なれば、それらの帯域特性は異なりえる。

【0094】

図16は、各ストレージシステム2のストレージシステム管理情報格納部221が格納する情報を示している。データI/Oプログラム2211は、計算機システム内のホスト計算機3とデータI/O用通信I/F25の接続された接続装置4群を経由して通信を行い、ボリューム2221、シンプロビジョニングボリューム224のデータを送受信し、書き込みを行う。

30

【0095】

プール管理プログラム2212は、シンプロビジョニングプール223のページ2231の管理、プールボリュームとしてのボリューム2221の管理、そしてシンプロビジョニングボリューム224へのページ2231割り当て管理を行う。ボリューム管理プログラム2213は、ボリューム2221及びシンプロビジョニングボリューム224を管理する。

40

【0096】

運用管理プログラム2214は、各種プログラムと連携してストレージシステム2の運用管理を行う。オペレーティングシステム2215は、各種プログラムの実行基盤であり、各種テーブルの管理基盤である。ストレージトポロジ管理プログラム2216は、それが動作するストレージシステム2と他のストレージシステム2と間を接続するネットワーク(ストレージシステムポート間の経路)のトポロジ情報を管理する。

【0097】

このネットワークは、ホスト計算機3とストレージシステム2とをつなぐネットワークであり、接続装置群4により形成されるデータネットワークである。ストレージシステムトポロジ管理情報は、ストレージシステム2のポート間の経路の構成情報を含み、具体的

50

には、経路における接続装置構成及び帯域の情報を含む。

【 0 0 9 8 】

ストレージシステム管理情報格納部 2 2 1 は、この他、ストレージトポロジ管理テーブル 2 0 0 及び管理用通信 I / F アドレス管理テーブル 2 1 0 を格納している。管理用通信 I / F アドレス管理テーブル 2 1 0 は、管理用通信 I / F 2 6 の IP アドレスを格納している。ストレージシステム 2 を計算機システムに導入するときに、管理者がこの値を設定する。

【 0 0 9 9 】

ストレージトポロジ管理テーブル 2 0 0 は、当該ストレージシステム 2 と他のストレージシステム 2 との間のネットワークのトポロジ情報を含む。このネットワークは、同一ストレージシステム 2 の異なるポートを接続する経路を含む。上述のように、ストレージトポロジ管理テーブル 2 0 0 は、当該ストレージシステム 2 の各ポートと他のポート（同一ストレージシステム 2 の他のポートを含む）との間の経路の構成情報を格納する。

10

【 0 1 0 0 】

図 1 7 は、ストレージシステム 2 が有するストレージトポロジ管理テーブル 2 0 0 の一例を示す。各ストレージシステム 2 が、ストレージトポロジ管理テーブル 2 0 0 を有する。図 1 7 は、ストレージ ID が S T 1 であるストレージシステム 2 におけるストレージトポロジ管理テーブル 2 0 0 を例示している。

【 0 1 0 1 】

ストレージトポロジ管理テーブル 2 0 0 は、ポート ID のカラム 2 0 0 1、対向ストレージシステム ID のカラム 2 0 0 2、対向ポート ID のカラム 2 0 0 3、最小帯域のカラム 2 0 0 4、経路接続装置数のカラム 2 0 0 5、最小帯域経路の平均帯域利用率のカラム 2 0 0 6、最小帯域経路の最大帯域利用率のカラム 2 0 0 7、遅延時間のカラム 2 0 0 8 を有する。

20

【 0 1 0 2 】

ポート ID のカラム 2 0 0 1 は、テーブル 2 0 0 が格納されているストレージシステム 2 内の各ポートの ID を格納する。対向ストレージシステム ID のカラムは、ポート ID のカラム 2 0 0 1 におけるポート ID で特定されるポートが接続する対向のストレージシステム 2 の ID を格納する。上述のように、本例においてはボリュームはシングルストレージプール内で移動可能であるので、ストレージシステム S T 1 と共にシングルストレージプールを構成するストレージシステム S T 2、S T 3 がテーブル 2 0 0 内に登録されている。

30

【 0 1 0 3 】

対向ポート ID のカラム 2 0 0 3 は、ポート ID で特定されるポートが接続する、対向ストレージシステム 2 のポートの ID を格納する。最小帯域のカラム 2 0 0 4 は、各エントリのストレージシステムポート間の経路にふくまれる経路（それぞれがストレージシステムの間経路の一部）の最小帯域の値を格納する。経路接続装置数のカラム 2 0 0 5 は、各エントリのストレージシステムポート間の経路における接続装置数を格納する。

【 0 1 0 4 】

最小帯域経路の平均帯域利用率及び最小帯域経路の最大帯域利用率は、図 1 1 のストレージトポロジ管理の閾値のテーブル 1 4 0 の説明において述べた通りである。遅延時間のカラム 2 0 0 8 は、各エントリのストレージシステムポート間の遅延時間の値を格納する。遅延時間のカラム 2 0 0 8 において、「 > 1 m s 」は、1 m s 未満の値を意味する。

40

【 0 1 0 5 】

図 1 7 に示すストレージトポロジ管理テーブル 2 0 0 の例において、各経路（エントリ）の構成情報は経路上の接続装置の数を示し、各経路の帯域の情報は経路の最小帯域の値を示す。また、使用状況における帯域についての情報として、最小帯域経路の平均帯域利用率及び最大帯域利用率が示されており、遅延時間が、使用状況を示す情報として含まれている。

【 0 1 0 6 】

50

例えば、ストレージシステム 2 が計算機システムに追加されると、そのストレージシステム 2 におけるストレージトポロジ管理テーブル 2 0 0 が、ストレージトポロジ管理プログラム 2 2 1 6 により作成される。例えば、管理者は、ポート ID、シングルストレージプールに含まれる対向ストレージシステム及びそれらのポートの ID を、ポート ID のコラム 2 0 0 1、対向ストレージシステム ID のコラム 2 0 0 2、対向ポート ID のコラム 2 0 0 3 に設定する。

【 0 1 0 7 】

ストレージトポロジ管理プログラム 2 2 1 6 は、接続装置群 4 の各接続装置 4 0 から、最小帯域のコラム 2 0 0 4、経路接続装置数のコラム 2 0 0 5、最小帯域経路の平均帯域利用率のコラム 2 0 0 6、最小帯域経路の最大帯域利用率のコラム 2 0 0 7 に格納する情報を取得する。接続装置群 4 の各接続装置 4 0 が有するネットワークの構成情報の詳細は後述する。

10

【 0 1 0 8 】

ストレージトポロジ管理プログラム 2 2 1 6 は、利用率の値を所定のタイミングで更新する。また、新たなストレージシステム 2 が追加される又はストレージシステムのポートがネットワーク装置と接続 / 切断されると、他のストレージシステム 2 のストレージトポロジ管理プログラム 2 2 1 6 は、追加、削除されたストレージシステム 2 2 についての情報をストレージトポロジ管理テーブル 2 0 0 に追加、削除する。

【 0 1 0 9 】

ストレージトポロジ管理プログラム 2 2 1 6 は、遅延時間のコラム 2 0 0 8 に格納する値を、例えば、ネットワークにおける ping コマンド (例えば f c p i n g コマンド) により取得することができる。ストレージトポロジ管理プログラム 2 2 1 6 は、遅延時間のコラム 2 0 0 8 の値を、所定のタイミングで更新する。

20

【 0 1 1 0 】

図 1 8 は、各ホスト計算機 3 の不揮発記憶装置 3 2 が格納している情報を示している。アプリケーション 3 2 1 は、ホスト計算機 3 が提供する機能やサービスを実行するために必要なプログラムである。エージェント 3 2 2 は、管理計算機 1 のホスト計算機管理プログラム 1 2 3 と通信し、ホスト計算機 3 の状態及び構成情報の設定を行う。オペレーティングシステム 3 2 3 は、アプリケーション 3 2 1 の各種プログラムの実行基盤であり、各種テーブルの管理基盤である。

30

【 0 1 1 1 】

不揮発記憶装置 3 2 は、これらの他、接続先ストレージシステムボリューム管理テーブル 3 0 0、管理用通信 I / F アドレス管理テーブル 3 1 0 を格納している。管理用通信 I / F アドレス管理テーブル 3 1 0 は、管理用通信 I / F 3 6 の IP アドレスを格納している。ホスト計算機 3 を計算機システムに導入するときに、管理者がこの値を設定する。

【 0 1 1 2 】

図 1 9 は、ホスト計算機 3 が有する、接続先ストレージシステムボリューム管理テーブル 3 0 0 の一例を示している。ホスト計算機 3 がこのテーブル 3 0 0 を有し、このテーブル 3 0 0 は、それが格納されているホスト計算機 3 に割り当てられ、読み書き可能なボリュームの情報を格納している。このテーブル 3 0 0 は、ストレージシステム ID のコラム 3 0 0 1、ボリューム ID のコラム 3 0 0 2、ポート ID のコラム 3 0 0 3 を有する。

40

【 0 1 1 3 】

ストレージシステム ID は、ボリュームが格納されているストレージシステムの ID である。ボリューム ID は、該当ボリュームのストレージシステム 2 における ID である。ポート ID はアクセス先のストレージシステム 2 のポート ID であり、その値は管理者が設定してもよく、管理計算機 1 がエージェント 3 2 2 に通知した値を格納してもよい。ホスト計算機 3 のオペレーティングシステム 3 2 3 は、ポート ID を参照しストレージシステム 2 にアクセスする。

【 0 1 1 4 】

図 2 0 は、接続装置 4 群における各接続装置 4 0 の不揮発記憶装置 4 2 が格納している

50

情報を示している。データI/Oプログラム421は、データI/O用通信I/F45に接続された計算機システム内の計算機及び装置からのI/O要求を受け付け、不揮発記憶装置42内のテーブルを参照し、それらと通信を行う。管理プログラム422は、不揮発記憶装置内42内のテーブルの値を設定するインタフェースを提供する。オペレーティングシステム423は、各種プログラムの実行基盤であり、各種テーブルの管理基盤である。

【0115】

不揮発記憶装置42は、この他、接続装置管理テーブル400、管理用通信I/Fアドレス管理テーブル410を格納している。管理用通信I/Fアドレス管理テーブル410は、管理用通信I/F46のアドレスを格納している。計算機システムに接続装置を導入するとき管理者が、このテーブル420に値を設定する。接続装置5、6も同様のプログラム及びテーブルを有する。

10

【0116】

図21は、接続装置管理テーブル400の構成例を示している。接続装置管理テーブル400は、接続装置の各ポートについての情報を管理する。具体的には、各ポートに接続されている対向装置及びその対向ポート並びに各ポートの帯域及び使用状況の情報を格納する。本例において、接続装置管理テーブル400は、ポートIDのカラム4001、対向装置IDのカラム4002、接続ポートIDのカラム4003、物理帯域のカラム4004、平均トラフィック量のカラム4005、最大トラフィック量のカラム4006を有する。

20

【0117】

ポートIDのカラム4001は、接続装置の各ポートのIDを格納する。対向装置IDのカラム4002は、各ポートが接続する対向の装置のIDを格納する。ポートが接続する対向の装置は、ストレージシステム2、ホスト計算機3又は他の接続装置である。図21の例においては、ストレージシステム2又は接続装置のIDが各フィールドに格納されている。

【0118】

接続ポートIDのカラム4003は、対向装置において接続されるポートのIDを格納する。物理帯域のカラム4004は、各ポートの予め規定されている帯域の値を格納する。平均トラフィック量は、所定時間、例えば1時間におけるトラフィック量の平均値であり、本例において、その値はbpsで表されている。最大トラフィック量は、上記所定時間におけるトラフィック量の最大値であり、本例において、その値はbpsで表されている。

30

【0119】

接続装置の管理プログラム422は、新たな装置がそのポートに接続されると、接続された装置からそのID及び接続ポートIDの値を取得し、接続装置管理テーブル400における接続されたポートのエントリに、取得したIDを格納する。物理帯域の値は、各ポートに対して予め設定されている。データI/Oプログラム421は、各ポートのトラフィック量を測定し、所定時間毎に、接続装置管理テーブル400において、平均トラフィック量のカラム4005の値と最大トラフィック量のカラム4006の値を更新する。

40

【0120】

上述のように、ストレージシステム2のストレージトポロジ管理プログラム2216は、ストレージトポロジ管理テーブル200の情報を接続装置群4の各接続装置に問い合わせる。ストレージシステム2から問い合わせを受けた管理プログラム422は、接続装置管理テーブル400から必要な情報を得て、ストレージシステム2に送信する。

【0121】

例えば、ストレージシステム2が新たに接続装置に接続されると、ストレージトポロジ管理プログラム2216は、接続装置群4の各接続装置40に対して、ネットワークの接続状態を示す構成情報を要求する。この要求に回答して、各接続装置40は、接続装置管理テーブル400に格納されている全ての情報、つまり、カラム4001からカラム40

50

06に格納されている値を、問い合わせ元のストレージシステム2に送信する。

【0122】

ストレージトポロジ管理プログラム2216は、接続装置群4の各接続装置40から取得した情報から、ストレージトポロジ管理テーブル200のカラムに格納する値を得ることができる。例えば、ストレージトポロジ管理プログラム2216は、各接続装置40のポートとその接続先ポートとの関係から、ストレージシステム2のポートと対向ポートとの間の経路接続装置及びそれらの経路ポートを特定することができる。

【0123】

ストレージシステムポート間の経路の接続装置構成から、ストレージトポロジ管理プログラム2216は、その経路における経路接続装置数の値を得ることができる。さらに、経路上の接続装置40の各ポートが特定されているため、ストレージトポロジ管理プログラム2216は、経路上のポートの物理帯域の値から、経路の最小帯域の値を得ることができる。

10

【0124】

ストレージトポロジ管理プログラム2216は、経路上のポートの物理帯域の値における最小値を、その経路の最小帯域の値と決定することができる。複数ポートが纏められて一つの物理ポートを構成している場合、その複数のポートからなる統合ポートの物理帯域が、経路の最小帯域を決めるときに参照される。

【0125】

ストレージトポロジ管理プログラム2216は、接続装置管理テーブル400における平均トラフィック量及び最大トラフィック量から、ストレージポート間における最小帯域経路の平均帯域利用率及び最大帯域利用率を算出することができる。ストレージトポロジ管理プログラム2216は、これらの値の更新のみを行う場合、各ストレージポート間経路における最小帯域経路の接続装置に平均トラフィック量及び最大トラフィック量の値を要求し、取得した値により平均帯域利用率及び最大帯域利用率を更新することができる。

20

【0126】

図22のフローチャートを参照して、計算機システムにおけるリソース平準化のために移動するボリューム及びボリュームの移動先を選択する処理の例を説明する。本例において、管理計算機1は、ストレージシステム2の使用状況に基づいて、移動するボリューム候補を選択し、さらに、ボリュームの移動先ポートの推奨度を、ストレージシステムポート間の経路に基づいて算出する。管理者は、この推奨度を参照して、ボリュームの移動先を決定することができる。本例では、推奨度を算出するボリュームの移動先候補のポートは、現在のストレージシステム2とは異なるストレージシステム2のポートである。

30

【0127】

管理計算機1の運用管理プログラム124は、マイグレーション対象のボリュームを検索する。具体的には、ストレージシステム使用状況管理テーブル160を参照し、ストレージシステムIDカラム1601のフィールド毎に、総容量カラム1602のフィールドの値に対する使用量カラム1604のフィールドの値の割合が、容量閾値カラム1603のフィールドの値を超えているエントリを検索する(S101)。

【0128】

使用量の割合が容量閾値を超えているストレージシステム2が存在する場合、運用管理プログラム124は、ボリューム管理テーブル150のストレージシステムIDのカラム1505において、そのストレージシステムIDのエントリを検索する。運用管理プログラム124は、該当するエントリのボリュームの中で、ボリュームタイプ1504の値が通常であるエントリから、一つのボリュームをマイグレーションボリュームとして選択する(S102)。ストレージシステム管理プログラム121は、例えば、使用量が最も多いボリュームを選択する。

40

【0129】

次に、運用管理プログラム124は、シンプロビジョニングの対象となっているボリュームについての選択を行う。シンプロビジョニングプール使用状況管理テーブル170を

50

参照し、容量カラム 1705 のフィールドの値に対する使用量カラム 1706 のフィールドの値の割合が容量閾値カラム 1707 のフィールドの値を超えているエントリのストレージシステム ID を検索する。加えて、運用管理プログラム 124 は、IOPS カラム 1708 のフィールドの値が性能閾値カラム 1709 のフィールド値を超えているエントリのストレージシステム ID を検索する (S103)。

【0130】

次に、運用管理プログラム 124 は、ボリューム管理テーブル 150 を参照し、ストレージシステム ID のカラム 1505 から、ステップ S103 で特定された使用量又は IOPS が閾値を超えているストレージシステムのエントリを検索する。そのストレージシステムのボリューム (エントリ) の中で、ボリュームタイプ 1504 がシンプロビジョニング 10
となっているエントリから、一つのボリュームをマイグレーションボリュームとして選択する (S104)。運用管理プログラム 124 は、例えば、使用量が最も多いボリュームを選択する。尚、本フローでは一連の処理の中でシンプロビジョニングの適応されているボリュームとそれ以外の普通のボリュームからの選択を行っているが、S101 と S102 又は、S103 と S104 の何れかのみを行うようにしても良い。

【0131】

次に、運用管理プログラム 124 は、ボリューム管理テーブル 150 において、選択したボリュームのエントリのポート ID のカラム 1507 を参照して、当該ボリュームのポート ID を特定する (S105)。

【0132】

さらに、運用管理プログラム 124 は、選択したボリュームを移動することができるポートを、当該ボリュームのポート ID から特定する (S106)。本例において、運用管理プログラム 124 は、選択したボリュームのポートが属するシングルストレージプール内のポートであって、選択したボリュームを格納できる残り容量を有する他のストレージシステム 2 のポートを、移動可能なポートとして選択する。 20

【0133】

運用管理プログラム 124 は、移動範囲管理テーブル 130 を参照して、移動するボリュームのポートが属するシングルストレージプール並びにそれに含まれる他のストレージシステム及びポートを特定することができる。

【0134】

例えば、運用管理プログラム 124 は、ストレージシステム使用状況管理テーブル 160 を参照して、通常タイプのボリュームの移動先ストレージシステム 2 (ポート) 候補を決定し、シンプロビジョニングプール使用状況管理テーブル 170 を参照して、シンプロビジョニングボリュームの移動先ストレージシステム 2 (ポート) 候補を決定する。 30

【0135】

運用管理プログラム 124 は、ストレージシステム使用状況管理テーブル 160 を参照することでストレージシステムの容量及び使用量を知ることができ、また、シンプロビジョニングプール使用状況管理テーブル 170 を参照することで、シンプロビジョニングプールにおける各ストレージシステムの容量及び使用量を知ることができる。シンプロビジョニングボリュームの移動先は、同一のシンプロビジョニングプールに記憶領域を提供するストレージシステムに限定してもよいし、他のシンプロビジョニングプールに移動可能であってもよい。 40

【0136】

次に、運用管理プログラム 124 は、特定した移動先候補ポートの推奨度を算出する (S107)。推奨度の算出方法の詳細は、図 23 を参照して後述する。複数のボリュームが選択されている場合にはそれぞれのボリュームについて推奨度を算定する。ストレージシステム管理プログラム 121 は、算出した推奨度を、移動先推奨度テーブル 180 に格納する。さらに、運用管理プログラム 124 は、移行推奨度テーブル 180 から取得した推奨度を含む情報を、画像出力装置 18 に出力する (S108)。

【0137】

10

20

30

40

50

例えば、移動先推奨度テーブル180と同様の内容のテーブルが表示される。運用管理プログラム124は、この他、ボリュームがシンプロビジョニングボリュームの場合に、現在のシンプロビジョニングプール及び移動先候補のシンプロビジョニングプールの情報を表示してもよい。

【0138】

管理者は、表示された移動先ポートの候補及びそれらの推奨度を参照して、入力装置19を使用してボリュームを移動するポートを選択する(S109)。運用管理プログラム124は、管理者に選択されたポートをボリュームの移動先ポートと決定し、ボリュームの移動元ストレージシステム及び/又は移動先ストレージシステムに、ボリュームマイグレーションを指示する。図22の例において、例えば、ストレージシステム2Aがボリュームの移動元であり、ストレージシステム2Bが移動先である。

10

【0139】

ストレージシステム2A、2Bは、選択されたボリュームが存在するストレージシステム2Aから、選択されたマイグレーション先のストレージシステム2Bへ、当該ボリュームのマイグレーションを実施する(S110)。ストレージシステム管理プログラム121は、ストレージシステム2Bに、ボリュームの移動を指示する。ストレージシステム2Bのボリューム管理プログラム2213は、指示されたボリューム(のデータ)の転送を、ストレージシステム2Aに要求する。

【0140】

ストレージシステム2A、2Bのボリューム管理プログラム2213は、ストレージシステム2Aの対象のボリュームのデータを、ストレージシステム2Bのボリュームに移動する。さらに、ストレージシステム2A、2Bのボリューム管理プログラム2213は、それぞれ、ボリュームの管理情報(不図示)を更新する。管理計算機1の運用管理プログラム124は、ストレージシステム2Bから、指示したボリュームのマイグレーションの終了通知を受けると、ボリューム管理テーブル150の該当エントリの値を更新する。

20

【0141】

上記フローにより、ボリューム移動において、ネットワークを介したボリュームのアクセス性能を考慮して、より適切な移動先ポートを選択することができる。移動先候補の推奨度の算出は、ストレージシステムポート間の経路に基づくため、ホスト計算機3の接続情報を管理していないシステムであっても、推奨度を算出することができる。

30

【0142】

上記例は、移動するボリュームを自動的に管理システムが選択するが、管理者が移動するボリュームを指定してもよい。推奨度を算出するボリューム移動先候補ポートは、ボリュームが存在する現在のストレージシステム2のポートを含むことができる。管理計算機1は、指定されたボリュームについて、上記フローを実行する。管理システムは、管理者に移動先候補ポートの推奨度を提示することなく、推奨度に応じて移動先ポートを決定してもよい。

【0143】

次に、図23のフローチャートを参照して、移動先候補ポートの推奨度の算出方法を説明する。図22を参照して説明したように、運用管理プログラム124は、移動先候補として選択したポートの推奨度を算出する(S107)。ストレージシステム管理プログラム121は、移動元ポートと移動先ポートとの間の経路の構成(接続装置構成及び帯域)及び使用状況に基づいて、推奨度を算出する。運用管理プログラム124は、移動先ポート候補のそれぞれについて、図23のフローを実行する。

40

【0144】

図23のフローチャートにおいて、運用管理プログラム124は、推奨度項目数カウントを、項目数に合わせる(S201)。本例において、項目数は5である。次に、運用管理プログラム124は、推奨度を算出する移動先候補ポートを選択し、ボリューム移動元のストレージシステム2から、ストレージトポロジ管理テーブル200における、ボリュームの移動元ポートと選択した移動先候補ポートのエントリの情報を取得して、キャッシ

50

メモリ 11 又は不揮発性記憶装置に格納する (S 202)。

【0145】

運用管理プログラム 124 は、ストレージポロジ管理テーブル 200 から取得した各値と、ストレージポロジ管理の閾値のテーブル 140 に格納されている各閾値とを比較して、推奨度を算出する。具体的には、運用管理プログラム 124 は、当該経路の最小帯域の値と最小帯域閾値とを比較する (S 203)。最小帯域の値が閾値以下である場合 (S 203: NO)、運用管理プログラム 124 は、カウントを 1 だけ減らす (S 204)。最小帯域の値が閾値より大きい場合 (S 203: YES)、カウントは維持される。

【0146】

次に、ストレージシステム管理プログラム 121 は、当該経路の経由接続装置数の値と経由接続装置数閾値とを比較する (S 205)。経由接続装置数の値が閾値以上である場合 (S 205: NO)、運用管理プログラム 124 は、カウントを 1 だけ減らす (S 206)。経由接続装置数の値が閾値より小さい場合 (S 205: YES)、カウントは維持される。

10

【0147】

次に、運用管理プログラム 124 は、当該経路における最小帯域経路の平均利用率の値と平均帯域利用率閾値とを比較する (S 207)。平均帯域利用率の値が閾値以上である場合 (S 207: NO)、ストレージシステム管理プログラム 121 は、カウントを 1 だけ減らす (S 208)。平均帯域利用率の値が閾値より小さい場合 (S 207: YES)、カウントは維持される。

20

【0148】

次に、運用管理プログラム 124 は、当該経路における最小帯域経路の最大帯域利用率の値と最大帯域利用率閾値とを比較する (S 209)。最大帯域利用率の値が閾値以上である場合 (S 209: NO)、運用管理プログラム 124 は、カウントを 1 だけ減らす (S 210)。最大帯域利用率の値が閾値より小さい場合 (S 209: YES)、カウントは維持される。

【0149】

次に、運用管理プログラム 124 は、当該経路における遅延時間の値と遅延時間閾値とを比較する (S 211)。遅延時間の値が閾値以上である場合 (S 211: NO)、運用管理プログラム 124 は、カウントを 1 だけ減らす (S 212)。遅延時間の値が閾値より小さい場合 (S 211: YES)、カウントは維持される。

30

【0150】

最後に、運用管理プログラム 124 は、カウントの値を所定の計算式に代入して推奨度を算出する (S 213)。一例において、推奨度の計算は、推奨度の最大値 (本例において 5) とカウント数の積を項目数で割った値を算出し、それを四捨五入により整数化する。運用管理プログラム 124 は、この値を、当該経路の推奨度の値として移行先推奨度テーブル 180 に格納する。

【0151】

本例は、ストレージシステムポート間の経路構成及び使用状況に基づいて、移行先ポートの推奨度を算出する。ホストストレージ間のネットワークについて、現在のボリューム位置 (ポート) からネットワーク的に近い新たな位置 (ポート) を選択することで、ホスト計算機 3 からのアクセス性能の変化が小さいボリューム配置先を選択することができる。

40

【0152】

上記推奨度の算出方法は一例であって、運用管理プログラム 124 は、これと異なる方法により上記項目を考慮した推奨度を算出することができる。運用管理プログラム 124 は、上記 5 項目の一部のみ参照して推奨度を算出してもよい。例えば、運用管理プログラム 124 は、最小帯域の項目のみで推奨度を算出する、若しくは、最小帯域と経由接続装置数により推奨度を算出する。

【0153】

50

各項目に重みが加えられており、変化されるカウント数が項目によって異なってもよい。例えば、最小帯域と経路接続装置数に最も大きな重み係数を与え、平均帯域利用率と最大帯域利用率に次の大きな重み係数を与え、遅延時間に最も小さい重み係数を与える。運用管理プログラム 124 は、上記 5 項目と異なる他の項目を参照してもよい。上記例において、各項目の閾値は一つのみであるが、運用管理プログラム 124 は、各項目において複数の閾値を使用してもよい。例えば、運用管理プログラム 124 は、ボリューム、経路又はストレージシステムによって閾値を使い分けてもよい。

【 0 1 5 4 】

上記構成例において、管理計算機 1 がストレージトポロジの閾値のテーブル 210 を有する。各ストレージシステム 2 は、そのストレージシステムのストレージトポロジの閾値のテーブル 210 を有し、管理計算機 1 からの要求に应答して、それら閾値を管理計算機 1 に送信してもよい。

10

【 0 1 5 5 】

上記例における推奨度は 5 レベルを有するが、そのレベル数は 2 以上の任意の値でよい。例えば、推奨度が 2 レベルで規定されており、上記 5 項目におけるいずれか一つの項目がネガティブである場合（カウントを減ずる状態である場合）、その推奨度は低であり、全ての項目がポジティブである場合に推奨度が高であってもよい。

【 0 1 5 6 】

管理計算機 1 は、各ストレージシステム 2 から、ストレージトポロジ管理テーブル 200 に格納されている情報を取得し、上記処理及びその他の処理のためにそれ自体が参照するためのストレージトポロジ管理テーブルを作成、保持してもよい。そのストレージトポロジ管理テーブルは、例えば、全てのストレージシステム 2 のストレージトポロジ管理テーブル 200 を統合したものである。管理計算機 1 は、所定のイベントに应答して更新情報を各ストレージシステム 2 から取得して、自身のストレージトポロジ管理テーブルを更新する。これは、後述するホストトポロジ管理テーブルについても同様である。

20

【 0 1 5 7 】

次に、図 24 を参照して、ストレージトポロジ管理テーブル 200 の更新（作成）の例を説明する。図 24 のシーケンス図は、新しいストレージシステムの計算機システムへの組み込みに対する、ストレージトポロジ管理テーブル 200 の更新（作成）の例を示している。

30

【 0 1 5 8 】

図 24 において、まず、新しいストレージシステム 2C のポートが接続装置 40 のポートに接続される。ストレージシステム 2C が新たに接続された接続装置 40 の管理プログラム 422 は、ストレージシステム 2C の運用管理プログラム 2214 とのネゴシエーションにより、ストレージシステム 2C 及び接続ポートの ID を取得し、接続装置管理テーブル 400 に追加する（S401）。ストレージシステム 2C の複数のポートが異なる接続装置 40 に接続される場合、それぞれの接続装置 40 が接続装置管理テーブル 400 に情報を登録する。

【 0 1 5 9 】

接続装置管理テーブル 400 への登録が完了すると、接続装置 40 の管理プログラム 422 は、その完了をストレージシステム 2C に通知する。ストレージシステム 2C の運用管理プログラム 2214 は、接続装置群 4 の各接続装置 40 に、トポロジ情報を要求する。トポロジ情報は、接続装置管理テーブル 400 に含まれている情報である。運用管理プログラム 2214 は、接続装置群 4 の各接続装置 40 から、接続装置管理テーブル 400 に含まれている全ての情報を取得する。

40

【 0 1 6 0 】

接続されたストレージシステム 2C の運用管理プログラム 2214 は、取得した情報により、ストレージトポロジ管理テーブル 200 を更新する（S402）。ここでは、更新は、新たなデータの格納である。接続装置群 4 の各接続装置 40 から取得した情報から、ストレージトポロジ管理テーブル 200 のカラム 2004 からカラム 2007 の値を取得

50

して格納する。運用管理プログラム 2214 は、さらに、各経路の遅延時間を測定して遅延時間のカラム 2008 に格納する。

【0161】

図 24 は示していないが、他のストレージシステム 2 は、例えば、管理計算機 1 からの指示若しくは物理配線の構成変更に応答して又は定期的に、ストレージシステム 2C についての情報、つまり、ストレージシステム 2C の接続ポートの情報を、ストレージシステム 2C のポートと接続された接続装置 40 から取得し、ストレージトポロジ管理テーブル 200 を更新する。

【0162】

管理計算機 1 の運用管理プログラム 124 は、例えばボリュームマイグレーションの要求を受けると、当該ボリュームのストレージシステム 2、本例において、ストレージシステム 2C に、推奨度の算出のため、当該ボリュームのストレージシステムポートと他のストレージシステムポートと間のトポロジ情報を要求する。トポロジ情報は、ストレージトポロジ管理テーブル 200 に格納されている経路の情報である。

【0163】

ストレージシステム 2C の運用管理プログラム 2214 は、管理計算機 1 からの要求に応答して、各接続装置 40 にトポロジ情報（接続装置管理テーブル 400 における情報）における更新情報を要求する。更新情報は、例えば、ポートの対向装置又は対向ポートにおける更新又はトラフィック量の更新測定値についての情報である。

【0164】

運用管理プログラム 2214 は、各接続装置 40 から取得した更新情報により、自身のストレージトポロジ管理テーブル 200 を更新し（S403）、ストレージトポロジ管理テーブル 200 から取得したトポロジ情報を、管理計算機 1 に送信する。管理計算機 1 の運用管理プログラム 124 は、取得した経路情報により、経路（移動先ポート）の推奨度を算出する。

【0165】

上記例は推奨度算出において、最小帯域を参照する。本実施形態の推奨度算出方法は、これと異なる経路の帯域についての情報を参照してもよい。例えば、管理計算機 1 は、経路における帯域の平均値を参照して推奨度を決定してもよい。

【0166】

上述のように、管理計算機 1 は、ストレージトポロジを管理するテーブルを保持、管理してもよい。ストレージシステム管理プログラム 121 は、所定のイベントに応答して、各ストレージシステム 2 から、ストレージトポロジ管理テーブル 200 の全ての情報又は更新された情報を受信する。ストレージシステム管理プログラム 121 は、受信した情報により、ストレージトポロジ管理テーブルを作成、更新する。

【0167】

上記例において、管理計算機 1 は、ストレージシステム 2 を介して、ストレージトポロジの情報を収集する。管理計算機 1 は、接続装置 40 から直接にストレージトポロジの情報を収集してもよい。例えば、管理計算機 1 は、ストレージトポロジ管理テーブル 200 における遅延時間の情報をストレージシステム 2 から取得し、その他の情報を接続装置 40 から取得する。推奨度算出の他の方法についての説明は、以下に説明する他の実施形態において同様である。

【0168】

上記例は、ボリュームのライブマイグレーションにおいて移動先候補ポートの推奨度を算出する。推奨度を算出するボリューム配置先ポート候補は、既存ボリュームの移動先候補ポートに限らず、新たに形成するボリュームの配置先候補ポートでもよい。

【0169】

第 2 実施形態

以下において、本発明の第 2 実施形態を説明する。第 2 実施形態は、移動するボリュームの I/O 量及び経路の使用状況に基づき、移動先ポートの推奨度を算出する。これによ

10

20

30

40

50

り、移動するボリュームのアクセス状況に応じて適切な移動先ポートを選択することができる。以下では、第2実施形態において第1実施形態と異なる点について主に説明する。

【0170】

図25は、第2実施形態のストレージシステム管理情報格納部221が含む情報を示す。第1実施形態と異なる点として、ボリュームアクセス状況管理テーブル220が追加されている。ボリュームアクセス状況管理テーブル220は、各ボリュームへのアクセス状況を管理するためのテーブルである。

【0171】

図26は、ボリュームアクセス状況管理テーブル220の構成例を示す。ボリュームアクセス状況管理テーブル220は、ポートIDのカラム2201、ボリュームIDのカラム2202、ボリュームアクセスの平均利用帯域のカラム2203、ボリュームアクセスの最大利用帯域のカラム2205を有している。ポートID及びボリュームIDは第1実施形態で説明した通りであり、ボリュームIDのカラム2202は、ポートIDにより識別されるポートに接続するボリュームを示している。

【0172】

ボリュームアクセスの平均利用帯域及び最大利用帯域は、ボリュームのI/O量を示す値である。ボリュームアクセスの平均利用帯域は、所定時間（例えば1時間）における、ボリュームへのアクセスにおける利用帯域の平均値である。ボリュームアクセスの最大利用帯域は、上記所定時間における、ボリュームへのアクセスの帯域の最大値である。本例において、これらはbpsで表されている。ストレージシステム2のデータI/Oプログラム2211は、各ボリュームのこれらの値を測定し、所定時間経過毎に、ボリュームアクセス状況管理テーブル220の値を更新する。

【0173】

図27A、27Bのフローチャートを参照して、第2実施形態における推奨度の決定方法を説明する。図27A、27Bはそれぞれフローチャートの一部を示し、結合子Aにおいて結合する。以下においては、第1実施形態の図23のフローチャートとの相違点を主に説明する。まず、管理計算機1の運用管理プログラム124は、推奨度項目数カウントを、項目数に合わせる（S501）。本例において、項目数は7である。

【0174】

次に、運用管理プログラム124は、推奨度を算出する移動先候補ポートを選択し、ボリューム移動元のストレージシステム2から、ストレージトポロジ管理テーブル200における、ボリュームの移動元ポートと選択した移動先候補ポートのエントリの情報を取得する。さらに、運用管理プログラム124は、移動するボリュームのアクセス状況情報を、当該ボリュームが属するストレージシステム2から取得する。具体的には、運用管理プログラム124は、ボリュームアクセス状況管理テーブル220に格納されている、ボリュームの平均利用帯域の値と最大利用帯域の値を取得する（S502）。

【0175】

運用管理プログラム124は、経路における平均帯域利用率の値と最小帯域の値との積を算出し、それとボリュームのアクセス平均利用帯域の値との和を計算する。運用管理プログラム124は、算出した和と経路の最小帯域閾値とを比較する（S503）。最小帯域閾値は、最小帯域の値と平均帯域利用率閾値との積である。閾値テーブル140は、これとは異なる最小帯域閾値を格納していてもよい。

【0176】

算出した和が、最小帯域閾値より大きい場合（S503：YES）、運用管理プログラム124は、カウントを1だけ減らす（S504）。算出した和が、最小帯域閾値以下である場合（S503：NO）、カウントは維持される。

【0177】

運用管理プログラム124は、経路における最大帯域利用率の値と最小帯域の値との積を算出し、それとボリュームのアクセス最大利用帯域の値との和を計算する。運用管理プログラム124は、算出した和と経路の最大帯域閾値とを比較する（S505）。最大帯

10

20

30

40

50

域閾値は、最小帯域の値と最大帯域利用率閾値との積である。閾値テーブル140は、これとは異なる最大帯域閾値をの情報を格納していてもよい。

【0178】

算出した和が、最大帯域閾値より大きい場合(S505: YES)、運用管理プログラム124は、カウントを1だけ減らす(S506)。算出した和が、最大帯域閾値以下である場合(S505: NO)、カウントは維持される。

【0179】

図27A、27BのフローチャートにおけるステップS507からステップS516は、図23のフローチャートにおけるステップS203からステップS212までと同様であり、説明を省略する。図27BのステップS517において、運用管理プログラム124は、推奨度を所定の数式により算出する。算出式は、第1実施形態における算出式(図23におけるステップS213)と同様である。ただし、第2実施形態においては、項目数が7であり、推奨度の最大値は7である。

10

【0180】

本実施形態は、経路の現在の使用状況(上記例で平均利用帯域及び最大利用帯域で規定)に、移動するボリュームのI/O量が追加されることによる影響を考慮して推奨度を算出する。これにより、ボリュームのI/O特性に応じてより適切な移動先ポートを選択することができる。

【0181】

本例は、推奨度の算出において、第1実施形態の項目数にボリュームのアクセス状況を示す二つの項目を追加している。第1実施形態と同様に、運用管理プログラム124は、これとは異なる方法により推奨度を算出することができる。例えば、運用管理プログラム124は、ボリュームのアクセス状況を示す2項目の一方のみを使用してもよく、他の5項目の一部のみを参照する又はそれらを参照しなくともよい。

20

【0182】

第3実施形態

以下において、本発明の第3実施形態を説明する。第3実施形態は、ホスト計算機3とストレージシステム2との間のデータ通信ネットワーク(上記例においてSAN)のトポロジに基づいて、仮想マシンの配置先ポート(ホスト計算機3のポート)の優先度を決定する。

30

【0183】

ストレージシステム間の経路について推奨度(優先度)を算出する第1実施形態と異なり、第3実施形態は、ホスト計算機3とストレージシステム2との間の経路の構成及び使用状況に基づいて、仮想マシンの移動先の推奨度(優先度)を算出する。また、第3実施形態は、ホスト計算機3とストレージシステム2との間の経路の構成(接続装置構成及び帯域)及び使用状況に基づいて、ボリュームの移動先ポートの推奨度(優先度)を算出する。

【0184】

第3実施形態において、ホスト計算機3上で、1又は複数の仮想マシンが動作している。ホスト計算機3は、それぞれ仮想化ソフトウェアを実行し、その仮想化ソフトウェア上で仮想マシンが動作する。仮想マシンはソフトウェア(プログラム)であって、仮想マシンの構成情報の他、OS(オペレーティングシステム)を含む。例えば、仮想マシンは仮想サーバであり、OSの他、仮想サーバのアプリケーションプログラム及びプログラムが動作するためのデータを含む。

40

【0185】

図28は、管理計算機1の不揮発性記憶装置12が格納している情報を示している。第1実施形態と異なる点として、ホスト管理テーブル190、ホストトポロジ管理の閾値のテーブル195、仮想マシン配置先推奨度テーブル197が追加されている。ホスト管理テーブル190は、計算機システムにおけるホスト計算機3を管理するためのテーブルである。仮想マシン配置先推奨度テーブル197は、仮想マシンの配置先ポートの推奨度を

50

管理するためのテーブルである。後述するように、移動範囲管理テーブル130は、仮想マシンの移動可能ポートを規定する情報をさらに有している。

【0186】

図29は、ホスト管理テーブル190の構成例を示している。ホスト管理テーブル190は、ホストIDのカラム1901、管理ポートIPアドレスのカラム1902、メモリ搭載量のカラム1903、メモリ使用量のカラム1904、CPUコア数のカラム1905、CPU周波数のカラム1906、CPU使用率のカラム1907、ポートIDのカラム1908を有する。ホスト管理テーブル190は、各ホスト計算機3の構成、性能及び使用状況の情報を有している。

【0187】

ホストIDは計算機システムにおいてホスト計算機3を一意に識別する識別子であり、管理計算機1は、ホスト計算機3の管理ポートIPアドレスにより、ホスト計算機3と通信を行うことができる。メモリ使用量及びCPU使用率以外の値は、例えば、管理者が、ホスト計算機3を計算機システムに組み込む際に登録する。メモリはホスト計算機3のメインメモリであり、ポートIDのカラム1908は、各ホスト計算機3内のポートの識別子を格納する。

【0188】

例えば、エージェント322は、それが動作しているホスト計算機3のメモリ使用量及びCPU使用率を定期的に取り得し、管理計算機1の要求に回答して、それらの値を管理計算機1に送信する。管理計算機1のホスト計算機管理プログラム123は、取得した1期間のメモリ使用量及びCPU使用率の値又は複数期間の平均値などを、ホスト管理テーブル190のメモリ使用量のカラム1904及びCPU使用率のカラム1907に格納する。

【0189】

図30は、ホストトポロジ管理の閾値のテーブル195の構成例を示す。ホストトポロジ管理の閾値のテーブル195は、最小帯域閾値のカラム1951、経路接続装置数閾値のカラム1952、最小帯域経路の平均帯域利用率閾値のカラム1953、最小帯域経路の最大帯域利用率閾値のカラム1954、遅延時間閾値のカラム1955を有する。

【0190】

ホストトポロジ管理の閾値のテーブル195の構成は、ストレージトポロジ管理の閾値のテーブル140の構成と同様である。ホストトポロジ管理の閾値のテーブル195は、ホスト計算機3のポートとストレージシステム2のポートとの間の経路について閾値を格納している点が、ストレージトポロジ管理の閾値のテーブル140と異なる。

【0191】

図31は、ストレージシステム管理情報格納部221が格納する情報を示している。第1実施形態と異なる点として、ホストトポロジ管理テーブル230が追加されている。ホストトポロジ管理テーブル230は、そのストレージシステム2のポートと、ホスト計算機3のポートとの経路の情報を管理するためのテーブルである。

【0192】

ストレージトポロジ管理テーブル200とホストトポロジ管理テーブル230が管理する経路が含まれるネットワークは同じであり、ホスト計算機3とストレージシステム2との間のデータ通信を行うネットワーク、典型的にSANである。ストレージトポロジ管理テーブル200は、ストレージシステム2のポート間の経路を管理し、ホストトポロジ管理テーブル230は、ホスト計算機3とストレージシステム2のポート間の経路を管理する。

【0193】

図32は、ホストトポロジ管理テーブル230の各カラムの構成例を示している。ホストトポロジ管理テーブル230は、ポートIDのカラム2301、対向ホスト計算機IDのカラム2302、対向ポートIDのカラム2303、最小帯域のカラム2304、経路接続装置数のカラム2305、最小帯域経路の平均帯域利用率のカラム2306、最小帯

10

20

30

40

50

域経路の最大帯域利用率のカラム 2 3 0 7、遅延時間のカラム 2 3 0 8 を有する。

【 0 1 9 4 】

ホストポロジ管理テーブル 2 3 0 の構成は、ストレージポロジ管理テーブル 2 0 0 と実質的に同一であり、ストレージシステム 2 のポートと接続する対向のポートが、ストレージシステム 2 のポートではなく、ホスト計算機 3 のポートである。それら以外のカラムの値の意味は、ストレージポロジ管理テーブル 2 0 0 において説明した通りである。

【 0 1 9 5 】

ホストポロジ管理テーブル 2 3 0 に格納する情報をストレージシステム 2 が取得する方法は、ストレージポロジ管理テーブル 2 0 0 と同様である。接続装置群 4 の各接続装置 4 0 は、その接続装置管理テーブル 4 0 0 において、それ自身のポートの対向装置及び接続ポート ID にホスト計算機 ID 及びそのポート ID を格納するエントリを有する。接続装置 4 0 は、ストレージシステム 2 からの要求に回答して、接続装置管理テーブル 4 0 0 から取得した必要な情報を送信する。

10

【 0 1 9 6 】

図 3 3 のフローチャートを参照して、仮想マシンの配置先（ホスト計算機 3 のポート）の推奨度を決定する方法を説明する。仮想マシンの配置先の決定は、例えば、稼働中の仮想マシンを現在のホスト計算機 3 から他のホスト計算機 3 に移動する、又は、新たに仮想マシンをホスト計算機 3 に配置するためなどに行われる。

【 0 1 9 7 】

管理計算機 1 の運用管理プログラム 1 2 4 又は管理者は、例えば、ホスト管理テーブル 1 9 0 におけるホスト計算機 3 の使用状況を参照して負荷が高いホスト計算機 3 を特定し、そのホスト計算機 3 で稼働している仮想マシンにおいて、他のホスト計算機 3 に移動する仮想マシンを選択する。仮想マシンの配置先の決定においては、基準となるボリュームのポートが決まっている。例えば、稼働中の仮想マシンが現在アクセスしているボリュームのポートであり、又は、仮想マシンの OS が格納されているボリュームのポートである。

20

【 0 1 9 8 】

運用管理プログラム 1 2 4 又は管理者は、移動範囲管理テーブル 1 3 0 を参照し、配置先ポート候補として、ボリュームのポートと接続可能なホスト計算機 3 のポートを選択する。移動範囲管理テーブル 1 3 0 は、図 1 0 に示す情報に加え、ストレージシステム 2 の各ポートに接続することができるホスト計算機 3 のポートを規定する情報を含む。

30

【 0 1 9 9 】

運用管理プログラム 1 2 4 又は管理者は、配置先ポート候補の選択において、ホスト管理テーブル 1 9 0 におけるホスト計算機 3 の性能及び使用状況を参照してもよい。これらの値が予め設定されている範囲のホスト計算機 3 のポートを配置先候補として選択する。運用管理プログラム 1 2 4 は、選択した各配置先候補ポートについて、図 3 3 のフローを実行する。

【 0 2 0 0 】

まず、管理計算機 1 の運用管理プログラム 1 2 4 は、推奨度項目数カウントを、項目数に合わせる（S 6 0 1）。ここでの項目数は 5 である。運用管理プログラム 1 2 4 は、選択したホスト計算機 3 のポートとボリュームのポートとの間の経路について、ステップ S 6 0 3 からステップ S 6 1 3 を実行する。

40

【 0 2 0 1 】

ステップ S 6 0 3 からステップ S 6 1 3 は、第 1 実施形態の図 2 3 のフローチャートにおけるステップ S 2 0 3 からステップ S 2 1 3 までと同様である。第 1 実施形態はストレージシステム 2 のポート間経路についてボリュームの配置先ポートの推奨度を算出するが、第 3 実施形態における管理計算機 1 は、ホスト計算機 3 のポートとストレージシステム 2 のポートとの間の経路について、仮想マシンの配置先候補ポートの推奨度を算出する。

【 0 2 0 2 】

運用管理プログラム 1 2 4 は、一つの仮想計算機の配置先候補ポートを選択して、その

50

ホストポートとボリュームのストレージシステムポートとの間の経路情報を、ボリュームが属するストレージシステム 2 から取得する (S 6 0 2)。ストレージシステム 2 は、ホストポロジ管理テーブル 2 3 0 から、ストレージシステムポートとホストポートとにより決まるエントリの情報を管理計算機 1 に送信する。

【 0 2 0 3 】

運用管理プログラム 1 2 4 は、ストレージシステム 2 から取得した経路情報を参照して、配置先候補ポート (により決まる経路) について、ステップ S 6 0 3 からステップ S 6 1 3 を実行する。運用管理プログラム 1 2 4 は、算出した推奨度を配置先推奨度テーブル 1 9 7 に格納する。仮想マシン配置先推奨度テーブル 1 9 7 は、例えば、ホスト計算機 I D、ポート I D、配置推奨度のカラムを含む。

10

【 0 2 0 4 】

運用管理プログラム 1 2 4 は、配置先推奨度テーブル 1 9 7 から、配置先候補のホスト計算機 3 及びポートの識別子並びに推奨度の値を取得して、画像出力装置 1 8 に出力する。管理者は、表示された推奨度の値を参考として、仮想マシンの配置先ポートを決定することができる。第 1 実施形態と同様に、管理計算機 1 は、管理者に配置先ポート候補及びその推奨度を提示することなく、算出した優先度 (本例で推奨度) に従って自動で仮想マシンの配置先ポートを決定してもよい。

【 0 2 0 5 】

次に、図 3 4 を参照して、ホスト計算機 3 の接続ポートが決まっている仮想マシンのボリュームの接続ポートの推奨度を算出する例を説明する。本例は、ボリュームを現在のストレージシステム 2 から異なるストレージシステム 2 に移動することをせず、同一ストレージシステム 2 における複数ポートのそれぞれの推奨度を算出する。

20

【 0 2 0 6 】

ストレージシステム 2 の異なるポートは、ホスト計算機 3 の一つのポートに対してそれぞれ異なる経路で接続されるため、それぞれの経路構成 (接続装置構成及び帯域を含む) が異なりうる。そのため、ストレージシステム 2 内の適切なポートを選択することで、ホスト計算機 3 のポートに対して、経路の転送性能が優れたポートを選択することができる。

【 0 2 0 7 】

図 3 4 のフローチャートにおいて、ステップ S 7 0 1、ステップ S 7 0 3 からステップ S 7 1 3 は、図 3 3 のフローチャートにおけるステップ S 6 0 1、ステップ S 6 0 3 からステップ S 6 1 3 と同様である。図 3 3 のフローと同様に、管理計算機 1 の運用管理プログラム 1 2 4 は、ホスト計算機 3 とストレージシステム 2 のポートの間の経路について、推奨度を算出する。

30

【 0 2 0 8 】

ステップ S 7 0 2 において、運用管理プログラム 1 2 4 は、ボリュームが存在するストレージシステム 2 のポートを一つ選択する。運用管理プログラム 1 2 4 は、ボリュームが存在するストレージシステム 2 のポートを順次選択して、ステップ S 7 0 3 からステップ S 7 1 3 を実行する。

【 0 2 0 9 】

運用管理プログラム 1 2 4 は、ストレージシステムのポートと算出した推奨度とを推奨度テーブル (不図示) に格納し、運用管理プログラム 1 2 4 が、それらを管理者に提示する。これは、他の推奨度計算のフローと同様である。他の処理と同様に、管理計算機 1 は、推奨度を提示することなく、自動でボリュームの接続ポートを決定してもよい。

40

【 0 2 1 0 】

本例では、図 3 4 のフローチャートを参照して、一つのストレージシステム 2 内でのポートの推奨度を算出する例を説明したが、管理計算機 1 は、図 3 4 のフローチャートに従って、他のストレージシステム 2 のポートを含めて、ボリュームが接続するストレージシステムポートの推奨度を算出してもよい。第 1 実施形態は、ストレージシステムポート間の経路についてポートの推奨度を算出するが、第 3 実施形態は、ホスト計算機 3 のポートと

50

ストレージシステム 2 のポートとの間経路に基づいて、ボリュームの接続ポートの推奨度を算出する。

【 0 2 1 1 】

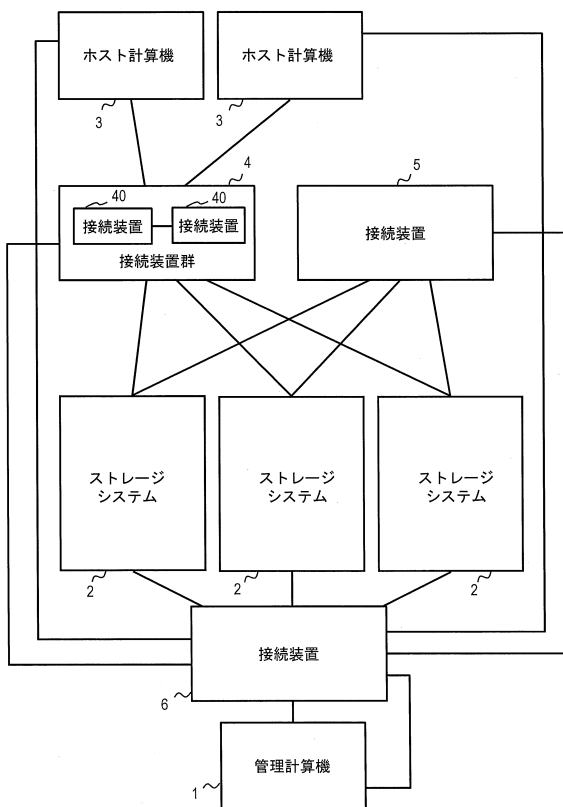
以上、本発明を添付の図面を参照して詳細に説明したが、本発明はこのような具体的構成に限定されるものではなく、添付した請求の範囲の趣旨内における様々な変更及び同等の構成を含むものである。本実施形態の構成の一部について、他の構成の追加・削除・置換をすることが可能である。システムは上記全ての機能を含むことが好ましいが、これらの一部のみを有していてもよい。

【 0 2 1 2 】

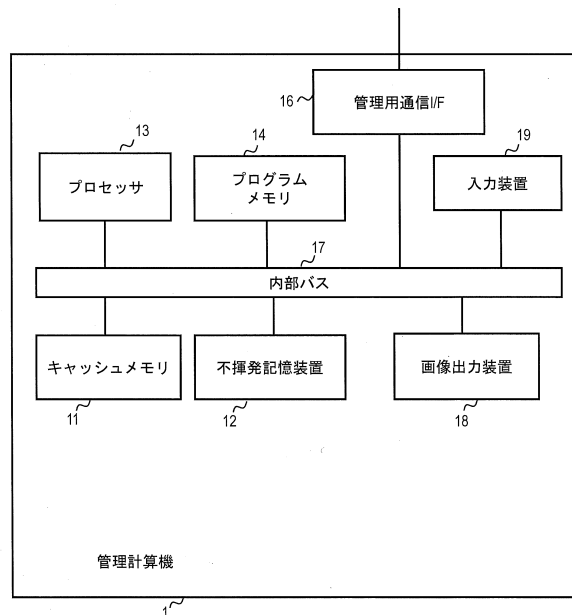
上記の各構成及び機能は、それらの一部又は全部を、例えば集積回路で設計して得られるハードウェアで実現してもよい。各機能を実現するプログラム、テーブル、ファイル等の情報は、不揮発性半導体メモリ、ハードディスクドライブ、SSD等の記憶デバイス、または、ICカード、SDカード、DVD等の計算機読み取り可能な非一時的データ記憶媒体に格納することができる。

10

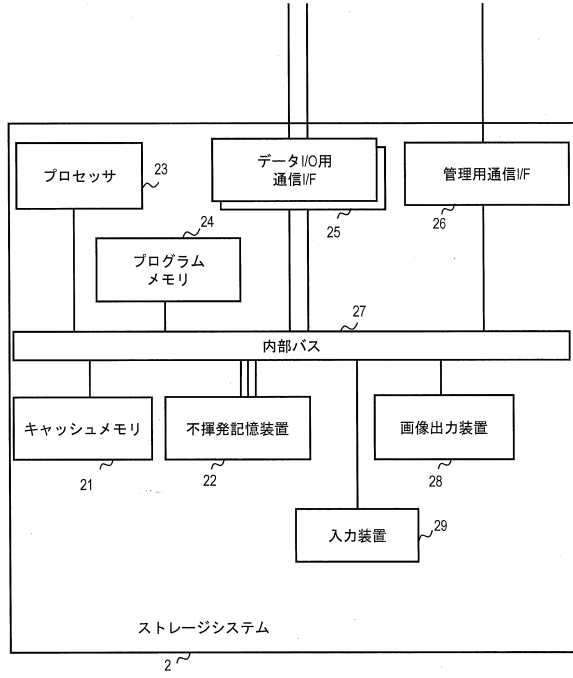
【 図 1 】



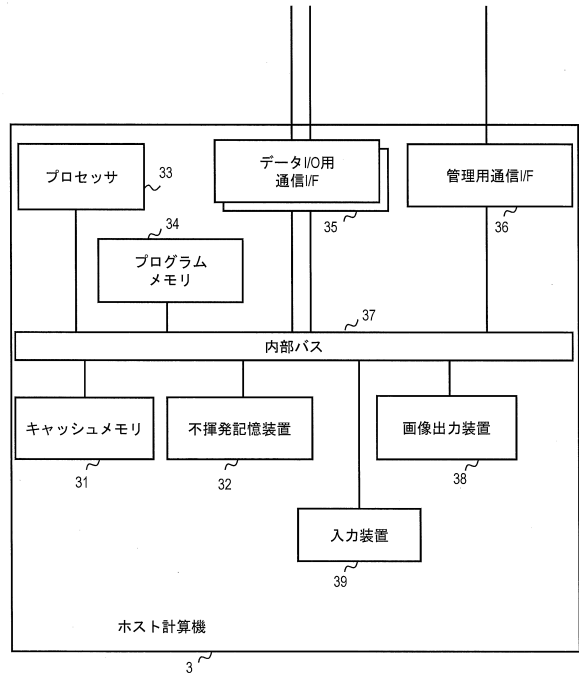
【 図 2 】



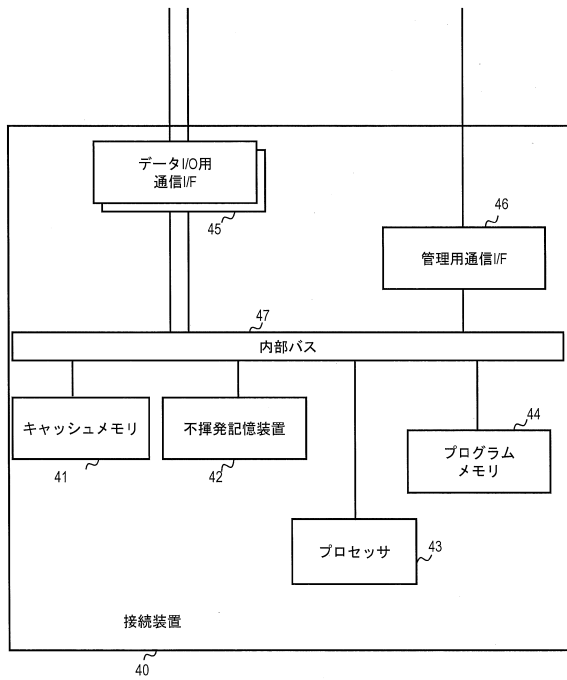
【図3】



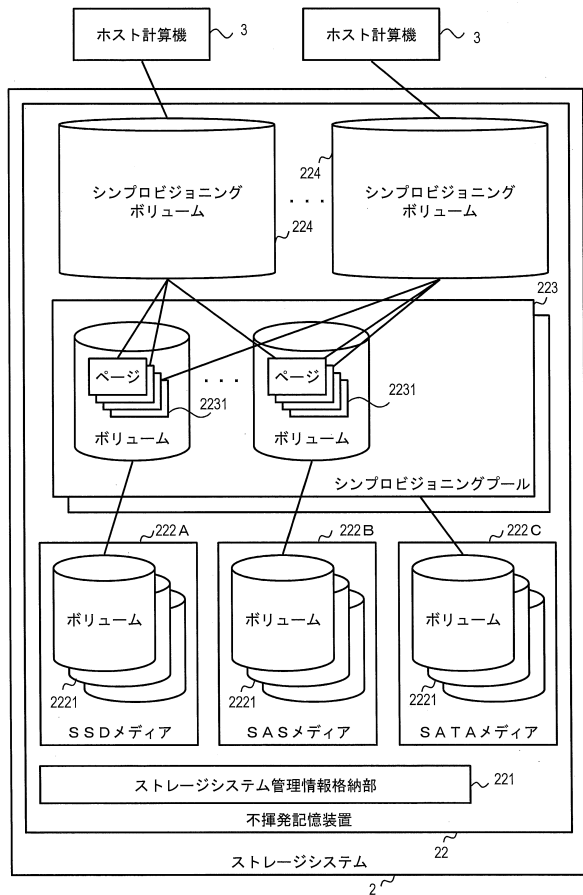
【図4】



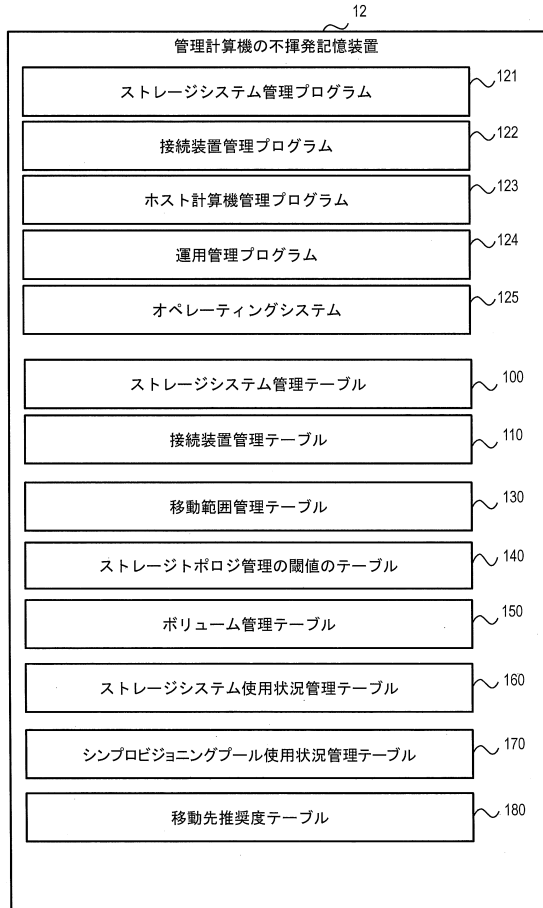
【図5】



【図6】



【図7】



【図8】

ストレージシステムID (1001)	管理ポートIPアドレス (1002)
ST1	192.168.1.101
ST2	192.168.1.102
ST3	192.168.1.103
ST4	192.168.1.104
...	...

【図9】

接続装置ID (1101)	管理ポートIPアドレス (1102)
SW1	192.168.1.201
SW2	192.168.1.202
SW3	192.168.1.203
SW4	192.168.1.204
...	...

【図10】

シングルストレージプールID (1301)	ストレージシステムID (1302)	ポートID (1303)
SPool1	ST1	PT1
		PT2
		PT3
		PT4
	ST2	PT1
		PT2
		PT3
		PT4
	ST3	PT1
		PT2
		PT3
		PT4
SPool2	ST4	PT1
		PT2
		PT3
		PT4
	ST5	PT1
		PT2
		PT3
...

【図11】

最小帯域閾値 (1401)	経由接続装置数閾値 (1402)	最小帯域経路の平均帯域利用率閾値 (1403)	最小帯域経路の最大帯域利用率閾値 (1404)	遅延時間閾値 (1405)
8Gbps	2	50%	80%	3ms

【図12】

150

ボリューム管理テーブル						
1501	1502	1503	1504	1505	1506	1507
ボリューム番号	容量	使用量	ボリュームタイプ	ストレージシステムID	ボリュームID	ポートID
1	10GB	5GB	通常	ST1	Vol1	PT1
2	10GB	8GB	通常	ST1	Vol2	PT1
3	10GB	3GB	通常	ST2	Vol1	PT1
4	40GB	3GB	シンプロビジョニング	ST1	Vol3	PT1
5	40GB	4GB	シンプロビジョニング	ST2	Vol2	PT1
6	10GB	1GB	シンプロビジョニング	ST1	Vol4	PT1
7	10GB	1GB	プール	ST1	Vol1	PT2
8	10GB	1GB	プール	ST1	Vol2	PT2
9	10GB	2GB	プール	ST2	Vol1	PT2
10	10GB	2GB	プール	ST2	Vol2	PT2

【図13】

160

ストレージシステム使用状況管理テーブル			
1601	1602	1603	1604
ストレージシステムID	総容量	容量閾値	使用量
ST1	14100GB	60%	6130GB
ST2	200000GB	50%	8300GB
ST3	100000GB	60%	9200GB
...

【図14】

170

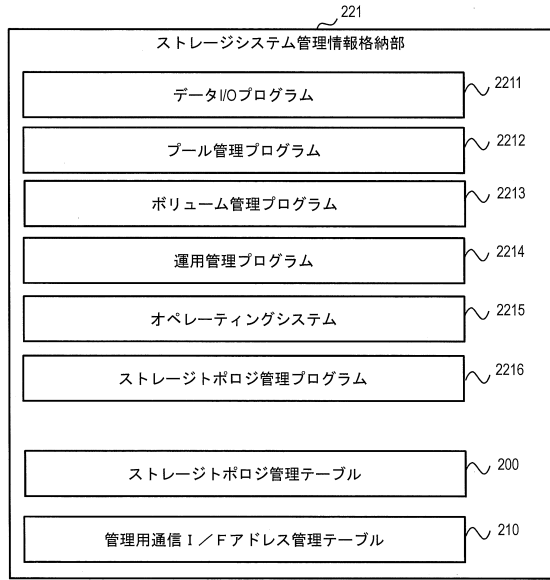
シンプロビジョニングプールの使用状況管理テーブル								
1701	1702	1703	1704	1705	1706	1707	1708	1709
プールID	全体容量	使用量	ストレージシステムID	容量	使用量	容量閾値	IOPS	性能閾値
TPool1	100GB	15GB	ST1	30GB	5GB	60%	100 IOPS	1000 IOPS
			ST2	70GB	10GB	50%	2000 IOPS	1500 IOPS

【図15】

180

移動先推奨度テーブル (ST1, PT1)			
1801	1802	1803	1804
シングルストレージプールID	ストレージシステムID	移動先ポートID	移動先推奨度
SPool1	ST1	PT2	★★★★★
		PT3	★★★★★
		PT4	★★★★★
	ST2	PT1	★★★★★
		PT2	★★★★★
		PT3	★★★★
	ST3	PT4	★★
		PT1	★
		PT2	★
		PT3	★
		PT4	★
		PT1	★

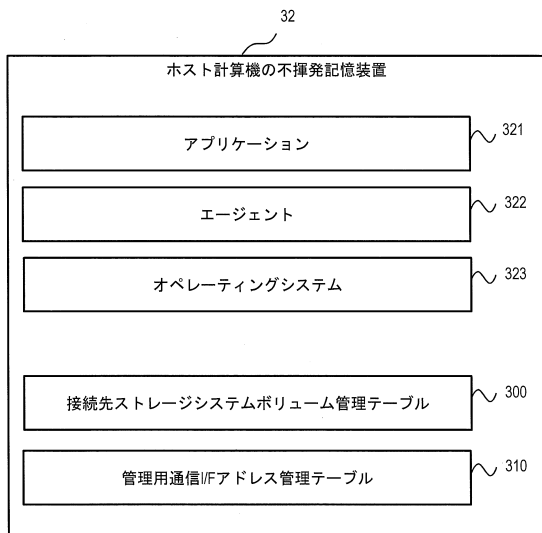
【図16】



【図17】

ポート ID	2002		2003		2004		2005		2006		2007		2008		
	対向ストレージシステムID	対向ポート ID	最小帯域	経由接続装置数	最小帯域接続の平均帯域利用率	最小帯域接続の最大帯域利用率	帯延時間	最小帯域接続の平均帯域利用率	最小帯域接続の最大帯域利用率	帯延時間	最小帯域接続の平均帯域利用率	最小帯域接続の最大帯域利用率	帯延時間	最小帯域接続の平均帯域利用率	最小帯域接続の最大帯域利用率
PT1	ST1	PT2	40Gbps	1	10%	50%	>1ms	10%	50%	>1ms	10%	50%	>1ms	10%	50%
		PT3	40Gbps	1	8%	40%	>1ms	8%	40%	>1ms	8%	40%	>1ms	8%	40%
		PT4	40Gbps	1	3%	25%	>1ms	3%	25%	>1ms	3%	25%	>1ms	3%	25%
		PT1	40Gbps	1	15%	20%	>1ms	15%	20%	>1ms	15%	20%	>1ms	15%	20%
PT2	ST2	PT2	2Gbps	2	80%	90%	2.4ms	80%	90%	2.4ms	80%	90%	2.4ms	80%	90%
		PT3	2Gbps	2	80%	90%	2.5ms	80%	90%	2.5ms	80%	90%	2.5ms	80%	90%
		PT4	2Gbps	2	80%	90%	2.0ms	80%	90%	2.0ms	80%	90%	2.0ms	80%	90%
		PT1	20Gbps	3	85%	90%	3.7ms	85%	90%	3.7ms	85%	90%	3.7ms	85%	90%
PT3	ST3	PT1	40Gbps	1	15%	55%	>1ms	15%	55%	>1ms	15%	55%	>1ms	15%	55%
		PT3	40Gbps	1	12%	40%	>1ms	12%	40%	>1ms	12%	40%	>1ms	12%	40%
		PT4	40Gbps	1	10%	25%	>1ms	10%	25%	>1ms	10%	25%	>1ms	10%	25%
		PT1	40Gbps	1	15%	20%	>1ms	15%	20%	>1ms	15%	20%	>1ms	15%	20%
PT4	ST1	PT1	40Gbps	1	15%	55%	>1ms	15%	55%	>1ms	15%	55%	>1ms	15%	55%
		PT3	40Gbps	1	12%	40%	>1ms	12%	40%	>1ms	12%	40%	>1ms	12%	40%
		PT4	40Gbps	1	10%	25%	>1ms	10%	25%	>1ms	10%	25%	>1ms	10%	25%
		PT1	40Gbps	1	15%	20%	>1ms	15%	20%	>1ms	15%	20%	>1ms	15%	20%
PT5	ST2	PT1	40Gbps	1	15%	55%	>1ms	15%	55%	>1ms	15%	55%	>1ms	15%	55%
		PT3	40Gbps	1	12%	40%	>1ms	12%	40%	>1ms	12%	40%	>1ms	12%	40%
		PT4	40Gbps	1	10%	25%	>1ms	10%	25%	>1ms	10%	25%	>1ms	10%	25%
		PT1	40Gbps	1	15%	20%	>1ms	15%	20%	>1ms	15%	20%	>1ms	15%	20%
PT6	ST3	PT1	40Gbps	1	15%	55%	>1ms	15%	55%	>1ms	15%	55%	>1ms	15%	55%
		PT3	40Gbps	1	12%	40%	>1ms	12%	40%	>1ms	12%	40%	>1ms	12%	40%
		PT4	40Gbps	1	10%	25%	>1ms	10%	25%	>1ms	10%	25%	>1ms	10%	25%
		PT1	40Gbps	1	15%	20%	>1ms	15%	20%	>1ms	15%	20%	>1ms	15%	20%

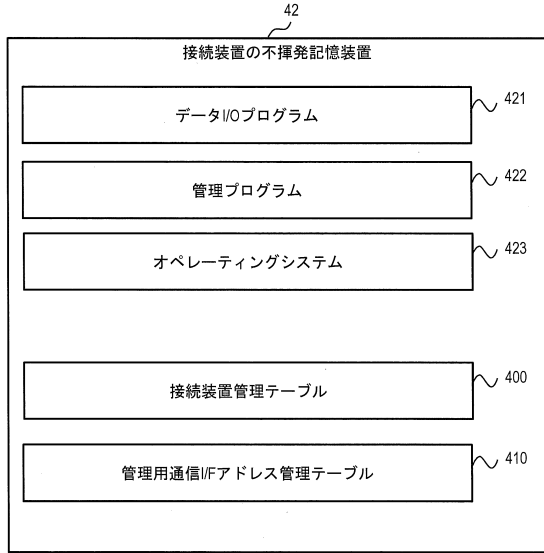
【図18】



【図19】

接続先ストレージシステムボリューム管理テーブル		
3001	3002	3003
ストレージシステム ID	ボリューム ID	ポート ID
ST1	Vol3	PT1
...

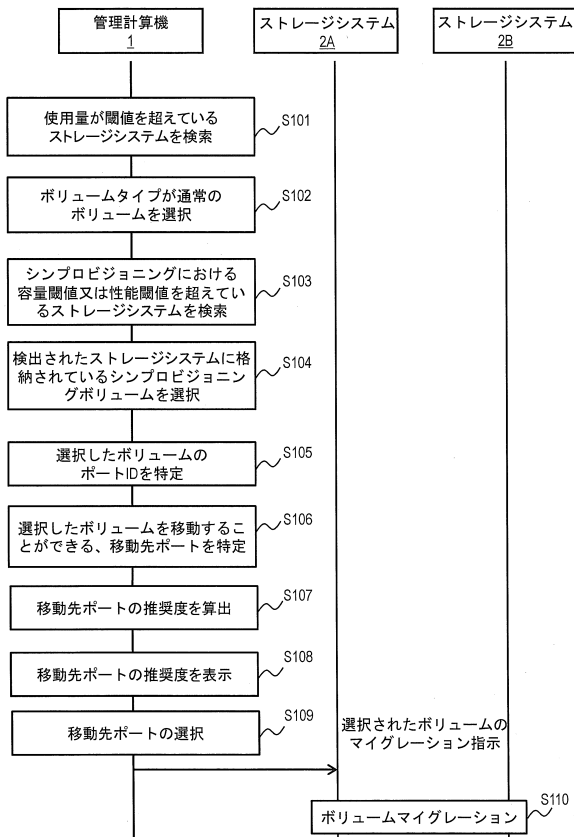
【図20】



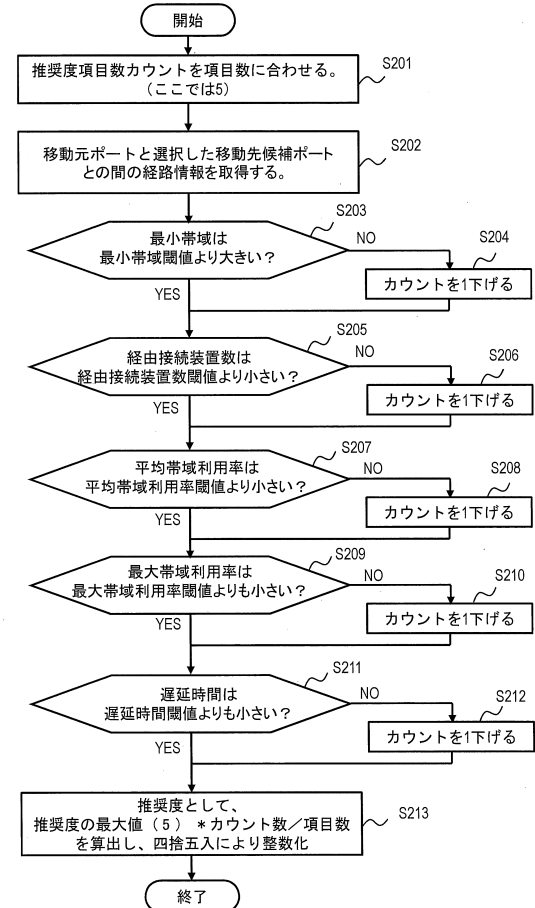
【図21】

接続装置管理テーブル					
ポートID	対向装置ID	接続ポートID	物理帯域	平均トラフィック量	最大トラフィック量
SPT1	ST1	PT1	40Gbps	6Gbps	22Gbps
SPT2	ST1	PT2	40Gbps	4Gbps	20Gbps
SPT3	ST1	PT3	40Gbps	3.2Gbps	16Gbps
SPT4	ST1	PT4	40Gbps	1.2Gbps	10Gbps
SPT5	ST2	PT1	40Gbps	6Gbps	8Gbps
SPT6	SW2	SPT1	2Gbps	1.6Gbps	1.8Gbps
SPT7	SW3	SPT1	8Gbps	4.2Gbps	6.0Gbps
SPT8	SW4	SPT1	8Gbps	5.0Gbps	6.4Gbps

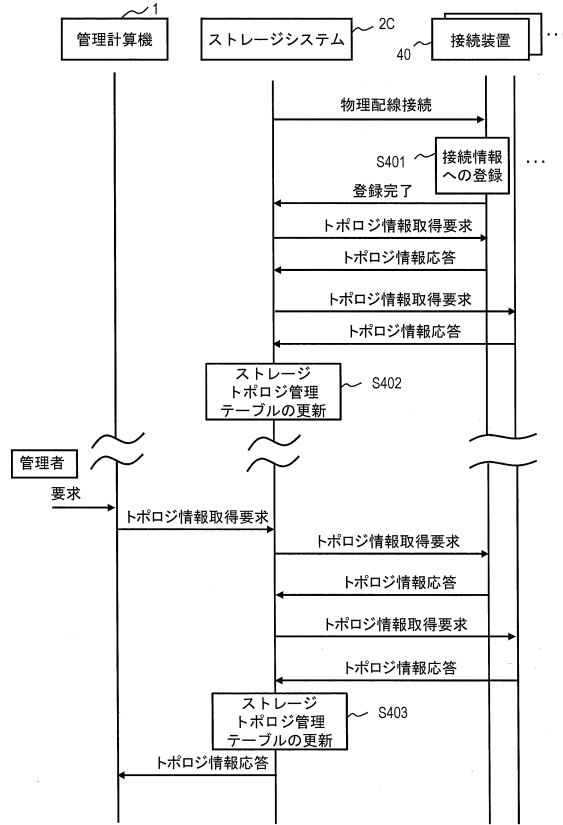
【図22】



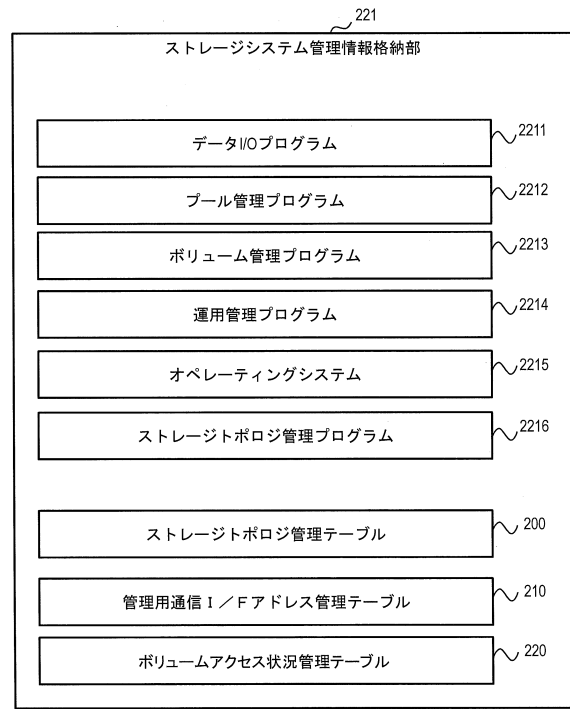
【図23】



【図 24】



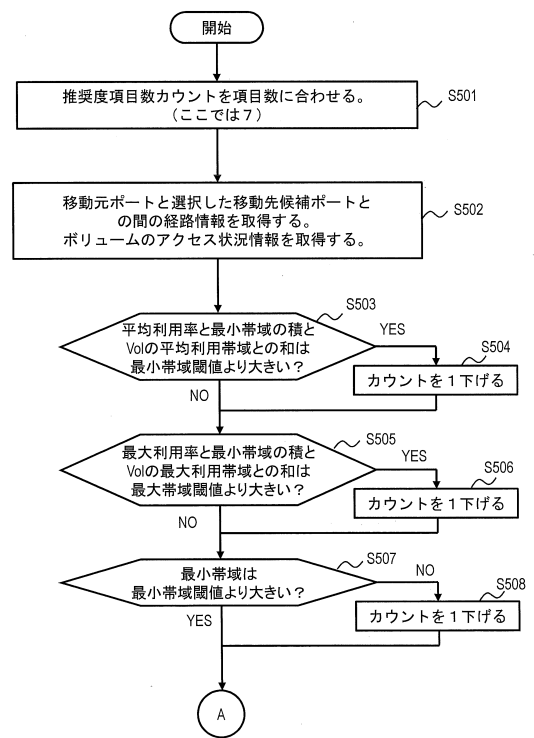
【図 25】



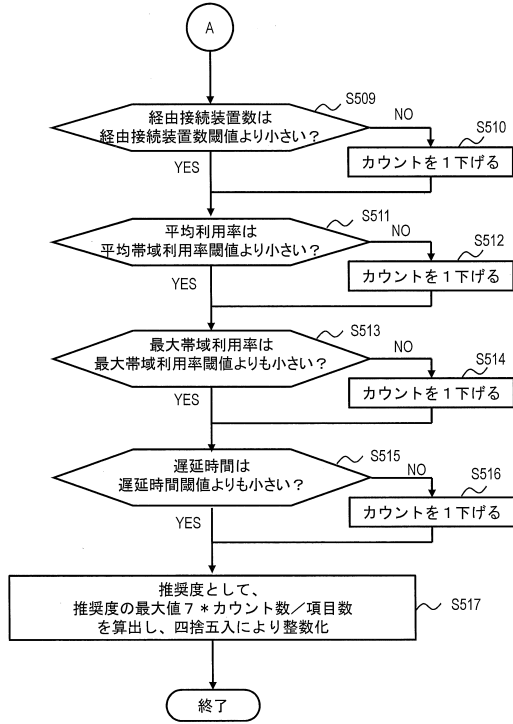
【図 26】

ボリュームアクセス状況管理テーブル (220)			
ポート ID (2201)	ボリューム ID (2202)	Volアクセスの平均利用帯域 (2203)	Volアクセスの最大利用帯域 (2204)
PT1	Vol1	300Mbps	800Mbps
	Vol2	200Mbps	300Mbps
	Vol3	50Mbps	200Mbps
	Vol4	100Mbps	150Mbps
PT2	Vol1	30Mbps	100Mbps
	Vol2	500Mbps	1Gbps
...

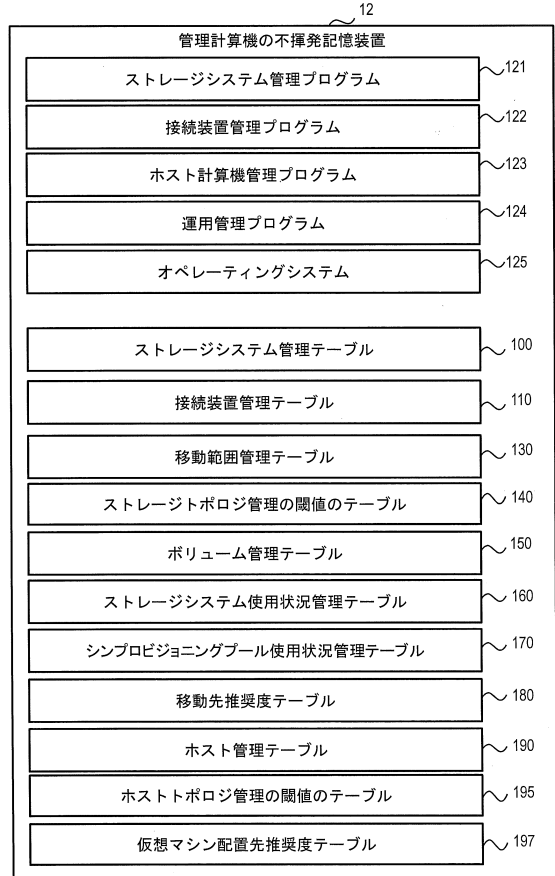
【図 27 A】



【図27B】



【図28】



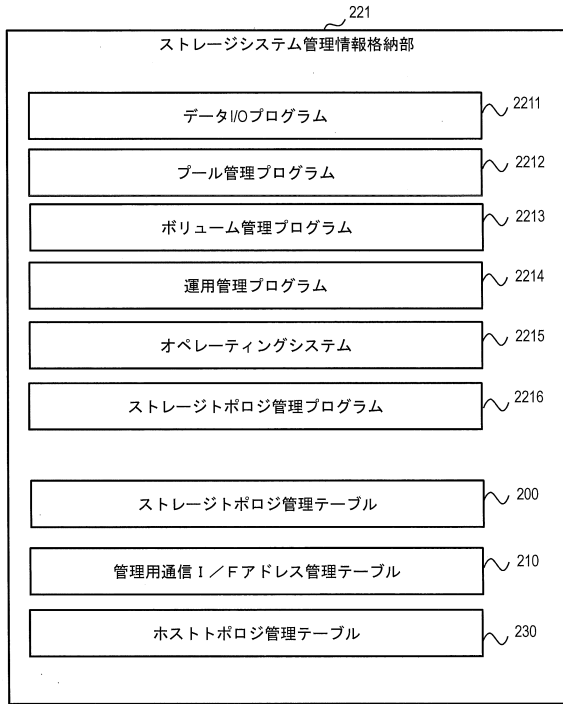
【図29】

1901	ホスト ID	HS1	HS2	HS3	HS4	...
	管理ポート IPアドレス	192.168.1.11	192.168.1.12	192.168.1.13	192.168.1.14	...
	メモリ搭載量	16GB	16GB	32GB	32GB	...
	メモリ使用量	10GB	12GB	20GB	28GB	...
	CPUコア数	4	4	8	8	...
1906	CPU周波数	2.0GHz	2.0GHz	2.3GHz	2.3GHz	...
	CPU利用率	20%	60%	30%	80%	...
1907	ポート ID	HBA1/HBA2	HBA1/HBA2	HBA1	HBA1	...

【図30】

ホストトポロジ管理の閾値のテーブル (195)				
1951	1952	1953	1954	1955
最小帯域閾値	経由スイッチ数閾値	最小帯域経路の平均帯域利用率閾値	最小帯域経路の最大帯域利用率閾値	遅延時間閾値
8Gbps	2	50%	80%	3ms

【図31】

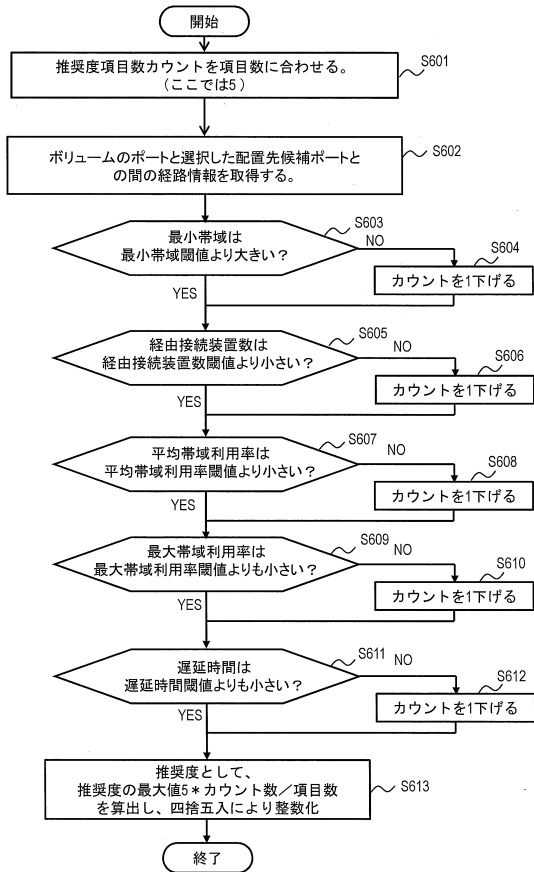


【図32】

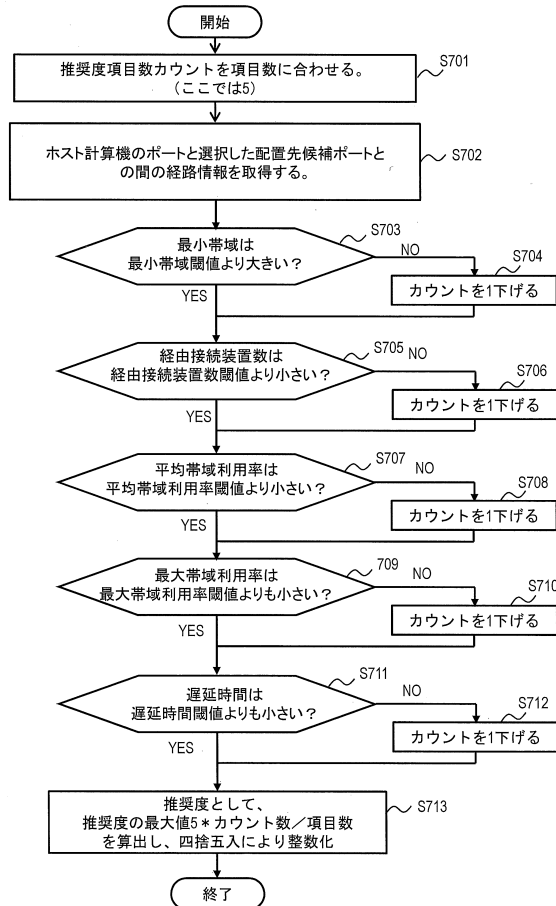
230
ホストトポロジ管理テーブル

2301 ポートID	2302 対向ホスト計算機ID	2303 対向ポートID	2304 最小帯域	2305 経由接続装置数	2306 最小帯域経路の平均帯域利用率	2307 最小帯域経路の最大帯域利用率	2308 遅延時間
PT1	HS1	HBA1	2Gbps	2	80%	90%	2.3ms
		HBA2	2Gbps	2	80%	90%	2.2ms
	HS2	HBA1	2Gbps	2	80%	90%	2.8ms
		HBA2	2Gbps	2	80%	90%	2.9ms
PT2	HS3	HBA1	40Gbps	1	15%	20%	<1ms
		HBA1	40Gbps	1	20%	25%	<1ms
	HS1	HBA1	2Gbps	2	80%	90%	2.3ms
		HBA2	2Gbps	2	80%	90%	2.2ms
HS2	HBA1	2Gbps	2	80%	90%	2.8ms	
	HBA2	2Gbps	2	80%	90%	2.9ms	
HS3	HBA1	40Gbps	1	15%	20%	<1ms	
	HBA1	40Gbps	1	20%	25%	<1ms	
...

【図33】



【図34】



フロントページの続き

- (56)参考文献 米国特許出願公開第2009/0172666 (US, A1)
特開2011-215794 (JP, A)
特開2008-047142 (JP, A)
特開2005-228278 (JP, A)
特開2010-086049 (JP, A)
特開2006-268093 (JP, A)
米国特許出願公開第2004/0024870 (US, A1)
特開2004-072135 (JP, A)
特開2005-242690 (JP, A)
特開2011-081580 (JP, A)
特開2006-344089 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 13/14
G06F 3/06
G06F 13/10