

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2022-124054
(P2022-124054A)

(43)公開日 令和4年8月25日(2022.8.25)

(51)国際特許分類	F I	テーマコード(参考)
G 0 6 F 11/20 (2006.01)	G 0 6 F 11/20 6 6 9	5 B 0 3 4
G 0 6 F 3/06 (2006.01)	G 0 6 F 3/06 3 0 1 X	
	G 0 6 F 3/06 3 0 4 F	
	G 0 6 F 11/20 6 7 1	

審査請求 未請求 請求項の数 8 O L (全20頁)

(21)出願番号	特願2021-21588(P2021-21588)	(71)出願人	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
(22)出願日	令和3年2月15日(2021.2.15)	(74)代理人	110002365 特許業務法人サンネクスト国際特許事務所
		(72)発明者	弘中 和衛 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	出口 彰 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	柴山 司 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内

最終頁に続く

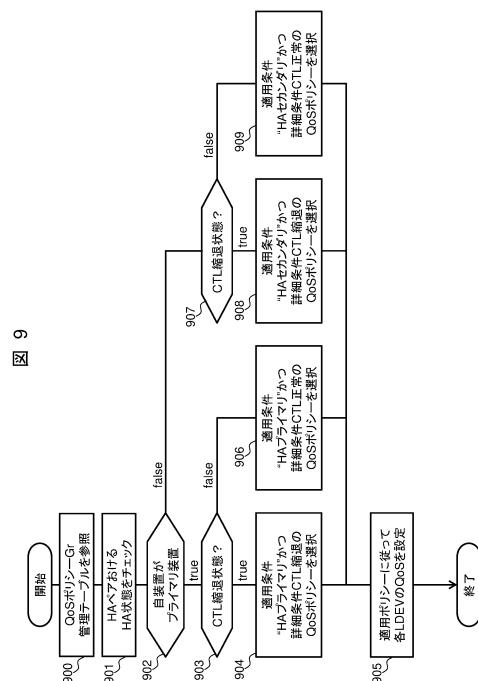
(54)【発明の名称】 ストレージシステム、ストレージ装置及びストレージ装置管理方法

(57)【要約】

【課題】ストレージ装置の状態に応じた適正な運用を実現すること。

【解決手段】ネットワーク経由で接続されたホストに第1のボリュームを提供する第1のストレージ装置と、前記第1のボリュームのリモートコピー先である第2のボリュームを含む複数のボリュームを提供する第2のストレージ装置とを有するストレージシステムであって、前記第2のストレージ装置の前記第2のボリューム及び他の前記複数のボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、前記第1のボリューム及び前記第2のストレージ装置の状態に応じて、複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部とを備えたことを特徴とする。かかる構成及び動作によれば、フェイルオーバーの前後や、縮退の有無などに応じたポリシーセットを選択して、ストレージ装置の運用を適正に行うことができる。

【選択図】図9



【特許請求の範囲】**【請求項 1】**

ネットワーク経由で接続されたホストに第 1 のボリュームを提供する第 1 のストレージ装置と、前記第 1 のボリュームのリモートコピー先である第 2 のボリュームを含む複数のボリュームを提供する第 2 のストレージ装置とを有するストレージシステムであって、

前記第 2 のストレージ装置の前記第 2 のボリューム及び他の前記複数のボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、

前記第 1 のボリューム及び前記第 2 のストレージ装置の状態に応じて、複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部と

を備えたことを特徴とするストレージシステム。

10

【請求項 2】

前記第 1 のストレージ装置に障害が発生した場合に前記第 1 のボリュームから前記第 2 のボリュームへのフェイルオーバーを実行する冗長構成を有し、

前記記憶部は、前記第 2 のストレージ装置が前記フェイルオーバーの実行前に用いるポリシーセットと、前記フェイルオーバーの実行後に用いるポリシーセットとを記憶し、

前記フェイルオーバーの実行後に用いるポリシーセットでは、前記フェイルオーバーの実行前に用いるポリシーセットに比べて、前記第 2 のボリュームに割り当てる性能が高く、前記他の複数のボリュームへに割り当てる性能が低い

ことを特徴とする請求項 1 に記載のストレージシステム。

20

【請求項 3】

前記第 2 のストレージ装置が前記フェイルオーバーの実行前に用いるポリシーセットは、前記ホストに係るデータ処理とは独立に行う処理を優先する規定群であり、前記第 2 のストレージ装置が前記フェイルオーバーの実行後に用いるポリシーセットは、前記ホストに係るデータ処理を優先する規定群であることを特徴とする請求項 2 に記載のストレージシステム。

【請求項 4】

前記ポリシー選択部は、前記第 2 のストレージ装置の処理能力と、前記第 2 のストレージ装置が前記フェイルオーバーの実行後に用いるポリシーセットとを比較し、前記第 2 のストレージ装置の処理能力に不足がある場合に警告を行うことを特徴とする請求項 2 に記載のストレージシステム。

30

【請求項 5】

前記第 1 のストレージ装置及び / 又は前記第 2 のストレージ装置は、コントローラの縮退運用が可能であり、

前記記憶部は、前記縮退運用が不要である場合に用いるポリシーセットと前記縮退運用中に用いるポリシーセットとを記憶する

ことを特徴とする請求項 1 に記載のストレージシステム。

【請求項 6】

前記記憶部は、時間に応じて異なる複数のポリシーセットを記憶し、

前記ポリシー選択部は、その時点に対応したポリシーセットを選択する

ことを特徴とする請求項 1 に記載のストレージシステム。

40

【請求項 7】

ネットワーク経由で接続されたホストにボリュームを提供するストレージ装置であって、

前記ボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、

自装置の状態に応じて複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部と

を備えたことを特徴とするストレージ装置。

【請求項 8】

ネットワーク経由で接続されたホストに第 1 のボリュームを提供する第 1 のストレージ

50

装置と、前記第 1 のボリュームのリモートコピー先である第 2 のボリュームを含む複数のボリュームを提供する第 2 のストレージ装置とを管理するストレージ装置管理方法であって、

前記第 2 のストレージ装置の前記第 2 のボリューム及び他の前記複数のボリュームのサービス品質に係る規定群であるポリシーセットを記憶部に複数格納するステップと、

前記第 1 のボリューム及び前記第 2 のストレージ装置の状態を判定するステップと、

前記第 1 のボリューム及び前記第 2 のストレージ装置の状態に応じて、複数の前記ポリシーセットから使用するポリシーセットを選択するステップと

を含むことを特徴とするストレージ装置管理方法。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、ストレージシステム、ストレージ装置及びストレージ装置管理方法に関する。

【背景技術】

【0002】

企業内で利用されるストレージシステムでは、一般的に 1 以上の装置を冗長に構成し、システムの可用性や信頼性を向上させて運用される。

【0003】

特許文献 1 には、可用性を向上させることができるストレージシステムが開示される。特許文献 1 によれば、第 1 のストレージ装置の第 1 のボリュームと、第 2 のストレージの第 2 のボリュームとは同一のボリュームとしてホストに提供される。この結果、ホストからは、前記第 1 のボリュームと第 2 のボリュームは違う経路を経由して接続される同一のボリュームとして認識することができる。

20

【0004】

特許文献 2 には、第 1 のストレージ装置の第 1 のボリュームと第 2 のストレージ装置の第 2 のボリュームとがそれぞれ同じ仮想 ID をホスト計算機に応答する High Availability (HA) ペア構成が開示されている。

【0005】

特許文献 2 によれば、第 1 ストレージ装置または第 2 ストレージ装置が、仮想 ID をコピー先として指定したデータコピー指示をホストから受信すると、コピー先ボリュームが存在するストレージ装置に応じて、第 1 ボリュームまたは第 2 ボリュームのいずれかをコピー元ボリュームとして選択し、コピー元ボリュームからコピー先ボリュームにデータをコピーする。

30

【0006】

また、ストレージシステムにおいては、ストレージ装置に接続される特定のホストやボリュームからのアクセスを他のホストやボリュームへのアクセスよりも優先的に処理することにより、一定の帯域や性能といったサービスレベル保証 (SLA) を実現する Quality of Service (QoS) 機能が広く実装されている。

【先行技術文献】

40

【特許文献】

【0007】

【特許文献 1】米国特許出願 第 2009 / 0271582 号明細書

【特許文献 2】国際公開 第 2015 / 189925 号

【発明の概要】

【発明が解決しようとする課題】

【0008】

第 1 のストレージ装置の第 1 のボリューム (以降、Primary Volume (PVOL) と呼ぶ) と、第 2 のストレージ装置の第 2 のボリューム (以降、Secondary Volume (SVOL) と呼ぶ) との間で HA ペアを構成することにより、

50

例えば、前記第1のストレージ装置で障害が発生し、前記P V O Lへのアクセスが不能となった場合でも、アクセス先ボリュームを前記P V O Lから前記S V O Lに自動的に切り替え（フェイルオーバー）されることで、ホストはその運用を継続できる。

【0009】

しかしながら、フェイルオーバーのみを目的として、前記第1のストレージ装置と同一の構成を持つ前記第2のストレージ装置として備えておくことは運用コストが高い。

【0010】

そこで企業では、前記第2のストレージ装置をより効率的に活用するため、前記第2のストレージ装置はH Aペアにおけるフェイルオーバー先として以外にも、前記第1のボリュームのデータバックアップソースとしての利用や、データの参照を目的としたバッチ処理、開発・テストなどの前記第1のストレージ装置や前記P V O Lを利用するホストへの影響なく実施されることが好ましい、2次的な目的で活用することで効率化を図る運用がなされる場合がある。

10

【0011】

このような運用において、例えば前記P V O Lについて、I O性能や帯域を保証するS L Aをする目的でQ o S設定することを考えた場合、前述の通り、前記第1のストレージ装置の障害発生時においても、前記第2のストレージ装置にフェイルオーバーされることでホストの運用が継続されるフェイルオーバー後の運用を考慮すると、前記P V O Lに適用したQ o S設定を前記S V O Lにおいても同様に適用したいと想定される。

【0012】

しかし、前記のQ o S設定は前記第1のストレージ装置の装置性能や利用状況に基づいて設計されるものであり、たとえ、前記第1のストレージ装置と前記第2のストレージ装置の装置構成が同一だとしても、その利用状況や負荷状況は異なるものであると考えられる。このため、前記第2のストレージ装置において前記のQ o S設定を同様に適用しても、フェイルオーバー後に第2のストレージ装置においてQ o S設定が意図通りのS L Aを満たせるとは限らない。

20

【0013】

このため、前記第1のストレージ装置から前記第2のストレージ装置へフェイルオーバーする場合において、前記第1のストレージ装置において実現されるS L Aを満たせるよう、前記第2のストレージ装置においてQ o Sを設定することが求められる。また、フェイルオーバーに限らず、ストレージ装置に部分的な障害が生じた縮退の発生によりストレージ装置の状態が変化した場合にも、かかる変化に応じて適切なQ o Sの設定を用いることが望ましい。時刻や曜日などでストレージ装置の処理量が異なる場合にも同様である。

30

そこで、本願発明は、ストレージ装置の状態に応じた適正な運用を実現する技術について開示する。

【課題を解決するための手段】

【0014】

代表的な本発明のストレージシステムの一つは、ネットワーク経由で接続されたホストに第1のボリュームを提供する第1のストレージ装置と、前記第1のボリュームのリモートコピー先である第2のボリュームを含む複数のボリュームを提供する第2のストレージ装置とを有するストレージシステムであって、前記第2のストレージ装置の前記第2のボリューム及び他の前記複数のボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、前記第1のボリューム及び前記第2のストレージ装置の状態に応じて、複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部とを備えたことを特徴とする。

40

また、代表的な本発明のストレージ装置の一つは、ネットワーク経由で接続されたホストにボリュームを提供するストレージ装置であって、前記ボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、自装置の状態に応じて複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部とを備えたことを特徴とする。

50

また、代表的な本発明のストレージ装置管理方法の一つは、ネットワーク経由で接続されたホストに第1のボリュームを提供する第1のストレージ装置と、前記第1のボリュームのリモートコピー先である第2のボリュームを含む複数のボリュームを提供する第2のストレージ装置とを管理するストレージ装置管理方法であって、前記第2のストレージ装置の前記第2のボリューム及び他の前記複数のボリュームのサービス品質に係る規定群であるポリシーセットを記憶部に複数格納するステップと、前記第1のボリューム及び前記第2のストレージ装置の状態を判定するステップと、前記第1のボリューム及び前記第2のストレージ装置の状態に応じて、複数の前記ポリシーセットから使用するポリシーセットを選択するステップとを含むことを特徴とする。

【発明の効果】

10

【0015】

本発明によれば、ストレージ装置の状態に応じた適正な運用を実現することができる。前記した以外の課題、構成および効果は、以下の実施例の説明により明らかにされる。

【図面の簡単な説明】

【0016】

【図1】ストレージ装置

【図2】HAペア構成におけるフェイルオーバー

【図3】管理情報

【図4】LDEV管理テーブル

【図5】HAペア管理テーブル

【図6】QoSポリシー管理テーブル

【図7】QoSポリシーGr管理テーブル

【図8】LDEV稼働情報テーブル

【図9】HAペアの状態に応じたQoSポリシー自動適用フロー

【図10】QoSポリシリスク検知フロー

【発明を実施するための形態】

【0017】

本発明の実施例について、図面を参照して説明する。

【0018】

なお、以下に説明する実施例は特許請求の範囲に係る発明を限定するものではなく、実施例の中で説明されている要素の組み合わせの全てが発明の解決手段に必須であるとは限らない。また、以下の説明では、「xxxテーブル」、「xxxリスト」、「xxxDB」、「xxxキュー」等の表現で各種の情報を説明することがあるが、各種情報はテーブル、リスト、DB、キュー、等以外のデータ構造で表現されていてもよい。そのため、データ構造に依存しないことを示すため、「xxxのテーブル」、「xxxリスト」、「xxxDB」、「xxxキュー」等を「xxx情報」と称することがある。

20

30

【0019】

さらに、以下の説明では各情報の内容を説明する際に、「識別情報」、「識別子」、「名」、「名前」、「ID」という表現を用いるが、これらについてはお互いに置換が可能である。

40

【0020】

また、以下の説明では、後述する本発明の実施例は、汎用コンピュータ上で稼動するソフトウェアで実装してもよいし、専用ハードウェア又はソフトウェアとハードウェアの組み合わせで実装してもよい。

【0021】

また、以下の説明では、「プログラム」を主語として処理を説明することがあるが、プログラムはプロセッサ（例えば、CPU: Central Processing Unit）によって実行されることによって、定められた処理を記憶資源（例えば、メモリ）、通信I/F、ポートを用いて処理を実行するため、プロセッサを主語として説明してもよい。

50

【 0 0 2 2 】

また、以下の説明では、プログラムを主語として説明された処理は、プロセッサを有する計算機（例えば、計算ホスト、ストレージ装置）が行う処理としてもよい。また、以下の説明では、「コントローラ」の表現で、プロセッサ又はプロセッサが行う処理の一部又は全部を行うハードウェア回路を指してもよい。プログラムは、プログラムソース（例えば、プログラム配布サーバや、計算機が読み取り可能な記憶メディア）から各計算機にインストールされてもよい。この場合、プログラム配布サーバはCPUと記憶資源を含み、記憶資源はさらに配布プログラムと配布対象であるプログラムを記憶している。そして、配布プログラムをCPUが実行することで、プログラム配布サーバのCPUは配布対象のプログラムを他の計算機に配布する。

10

【 0 0 2 3 】

また、以下の説明では、「PDEV」は、物理的な記憶デバイスを意味し、典型的には、不揮発性の記憶デバイス（例えば補助記憶デバイス）でよい。PDEVは、例えば、HDD（Hard Disk Drive）又はSSD（Solid State Drive）でよい。ストレージシステムに異なる種類のPDEVが混在していてもよい。

【 0 0 2 4 】

また、以下の説明では、「RAID」は、Redundant Array of Inexpensive Disksの略である。RAIDグループは、複数のPDEV（典型的には同種のPDEV）で構成され、そのRAIDグループに関連付けられたRAIDレベルに従いデータを記憶する。RAIDグループは、パリティグループ（Parity Group : PG）と呼ばれてもよい。パリティグループは、例えば、パリティを格納するRAIDグループのことでよい。

20

【 0 0 2 5 】

また、以下の説明では、「VOL」はボリュームの略であり、物理的な記憶デバイスまたは、論理的な記憶デバイスでもよい。VOLは、実体的なVOL（RVOL）であってもよいし、仮想的なVOL（VVOL）であってもよい。「RVOL」は、そのRVOLを有するストレージシステムが有する物理的な記憶資源（例えば、1以上のRAIDグループ）に基づくVOLでよい。「VVOL」は、外部接続VOL（EVOL）と、容量拡張VOL（TPVOL）と、スナップショットVOLとのうちのいずれでもよい。EVOLは、外部のストレージシステムの記憶空間（例えばVOL）に基づいており、ストレージ仮想化技術に従うVOLでよい。TPVOLは、複数の仮想領域（仮想的な記憶領域）で構成されており容量仮想化技術（典型的にはThin Provisioning）に従うVOLでよい。

30

【 0 0 2 6 】

また、以下の説明では、「プール（POOL）」は、論理的な記憶領域（例えば複数のプールVOLの集合）であり、用途ごとに用意されてよい。例えば、プールとして、TPプールと、スナップショットプールとのうちの少なくとも1種類があってもよい。TPプールは、複数のページ（実体的な記憶領域）で構成された記憶領域でよい。ストレージコントローラが、ホストシステム（以下、ホスト）から受信したライト要求が指定するアドレスが属する仮想領域（TPVOLの仮想領域）にページが割り当てられていない場合、その仮想領域（ライト先仮想領域）にTPプールからページを割り当てる（ライト先仮想領域にページが割り当て済みであってもページが新たにライト先仮想領域に割り当てられてもよい）。ストレージコントローラは、割り当てられたページに、そのライト要求に付随するライト対象データを書き込んでよい。スナップショットプールは、オリジナルのVOLから退避されたデータが格納される記憶領域でよい。1つのプールが、TPプールとしてもスナップショットプールとしても使用されてもよい。「プールVOL」は、プールの構成要素となるVOLでよい。プールVOLは、RVOLであってもよいしEVOLであってもよい。

40

【 0 0 2 7 】

また、以下の説明では、ホストから認識されるVOL（ホストに提供されるVOL）を

50

「LDEV」と言う。以下の説明では、LDEVは、TPVOL(又はRVOL)であり、プールは、TPプールである。しかし、本発明は、容量拡張技術(Thin Provisioning)が採用されていないストレージ装置にも適用できる。

【0028】

また、以下の説明では、「PVOL(Primary VOL)」はバックアップ、レプリケーション、スナップショットの元ボリュームとなるLDEVのことであり、「SVOL(Secondary VOL)」はバックアップ、レプリケーション、スナップショット先となるLDEVのことで良い。

【0029】

以下、本実施形態を詳細に説明する。

【0030】

図1は、本実施例に係るストレージ装置の構成例を示す。

【0031】

ストレージ装置100に、ネットワーク110を介して1以上のホスト107が接続されている。また、ストレージ装置100に管理システム108が接続されている。ネットワーク110は例えばFC(Fibre Channel)やiSCSI(Internet Small Computer System Interface)である。

【0032】

ホスト107は、ホストシステムの略であり、1以上のホストである。ホスト107は、H-I/Fホストインタフェース109を有しており、ホストインタフェース109経由で、アクセス要求(ライト要求又はリード要求)をストレージ装置100に送信したり、アクセス要求の応答(例えば、ライト完了を含んだライト応答、又は、リード対象チャンクを含んだリード応答)を受信したりする。ホストインタフェース109は、例えば、HBA(Host Bus Adapter)又はNIC(Network Interface Card)である。

【0033】

ストレージ装置100は、複数のドライブ101と、複数のドライブ101に接続されたストレージコントローラ106とを有する。複数のドライブ101を含んだ1以上のRAIDグループが構成されていてよい。

【0034】

管理システム108は、ストレージ装置100の構成及び状態を管理する。管理システム108、M-I/F管理インタフェースデバイス120を有しており、管理インタフェースデバイス120が接続される管理ネットワーク111経由で、ストレージ装置100に命令を送信したりその命令の応答を受信したりする。管理インタフェースデバイス120は、例えばNICである。

【0035】

また、管理システム108はストレージ装置100を管理するサーバやPC上で実行されるソフトウェアで良いし、ストレージ装置100に接続されるホスト107の管理を行うセキュリティアプライアンスやソフトウェアの機能として実装されても良い。

【0036】

統合管理システム121は、複数の拠点における複数のストレージ装置100の構成及び稼働状態の管理や監視を行う、例えばクラウド119上で実行される管理ソフトウェアである。統合管理システム121は管理ネットワーク111を経由して、管理対象のストレージ装置100と通信し、各々のストレージ装置100の構成情報や稼働情報を収集する。

【0037】

管理用コントローラ118は、ストレージ装置100を管理するための機能を提供するコントローラであり、管理システム108や統合管理システム121から命令を受信したり、命令へ応答を返信し、管理システム108や統合管理システム121にストレージ装置100の構成情報や稼働情報を送信したり、ストレージ装置100の設定や構成の変更

10

20

30

40

50

を実行する。管理用コントローラ 118 は、例えば Baseboard Micro Controller (BMC) である。

【0038】

ストレージコントローラ 106 は、F-I/F フロントエンドインタフェース 116 と、B-I/F バックエンドインタフェース 117 と、R-I/F レプリケーションインタフェース 122 と、CM キャッシュメモリ 105 と、NVRAM (Non-Volatile RAM) 104 と、プロセッサパッケージ 113A 及びプロセッサパッケージ 113B と、それらの要素間の通信を中継する中継器 112 とを有する。中継器 112 は、例えばバス又はスイッチである。

【0039】

フロントエンドインタフェース 116 は、ストレージ装置 100 に接続されるホスト 107 と通信するインタフェースであり、例えば、FC (Fibre Channel) や iSCSI (Internet Small Computer System Interface) のインタフェースである。

【0040】

バックエンドインタフェース 117 は、ドライブ 101 と通信するインタフェースである。バックエンドインタフェース 117 は、E/D 回路 (暗号化及び復号化のためのハードウェア回路) を含んでいてよい。具体的には、例えば、B-I/F バックエンドインタフェース 117 は、SAS (Serial Attached SCSI (Small Computer System Interface)) コントローラを含んでいてよく、SAS コントローラが、E/D 回路を含んでいてよい。

【0041】

レプリケーションインタフェース 122 は、複数のストレージ装置 100 の間において HA 構成を構成する際に用いるインタフェースであり、ネットワーク 110 を介して、HA を構成する他のストレージ装置 100 と接続される。レプリケーションインタフェース 122 は、例えば、FC や iSCSI である。

【0042】

CM キャッシュメモリ 105 (例えば DRAM (Dynamic Random Access Memory)) には、ドライブ 101 に書き込まれるデータ又はドライブ 101 から読み出されたデータがプロセッサパッケージ 113 により一時格納される。NVRAM 104 には、電断時にバッテリー (図示せず) から電力を受けたプロセッサパッケージ 113 によりキャッシュメモリ 105 内のデータ (例えばダーティデータ (ドライブ 101 に書き込まれていないデータ)) が退避される。

【0043】

ストレージコントローラ 106 は、プロセッサパッケージ 113A 及びプロセッサパッケージ 113B によりクラスタが構成されている。プロセッサパッケージ 113A (プロセッサパッケージ 113B) は、DRAM 115A (DRAM 115B) と、CPU 114A (CPU 114B) とを有する。DRAM 115A (DRAM 115B) には、CPU 114A (CPU 114B) に実行される制御プログラム 102A (制御プログラム 102B) と、CPU 114A (CPU 114B) により参照又は更新される管理情報 103A (管理情報 103B) とが格納される。CPU 114A (CPU 114B) に制御プログラム 102A (制御プログラム 102B) が実行されることにより、例えば、ストレージ装置 100 の I/O 処理、アドレスの変換処理が実行される。制御プログラム 102A (制御プログラム 102B) 及び管理情報 103 (管理情報 103B) のうちの少なくとも一方が、複数のプロセッサパッケージ 113A 及びプロセッサパッケージ 113B に共有の記憶領域 (例えば CM 105) に格納されてもよい。

【0044】

< HA ペア構成におけるフェイルオーバーの例 >

【0045】

図 2 は、本実施例に係る HA ペア構成におけるフェイルオーバーの構成例を示す。

10

20

30

40

50

【 0 0 4 6 】

ここでは、複数のストレージ装置 1 0 0 の間で構成した H A ペア構成において、フェイルオーバーを行う例を示し、本願技術が解決したい課題について具体例を用いて説明する。

【 0 0 4 7 】

図 2 は、ストレージ装置 1 0 0 A とストレージ装置 1 0 0 B の間において、H A ペア構成が構成されており、ストレージ装置 1 0 0 A の P V O L 2 0 1 A 及び P V O L 2 0 1 B が、ペア装置のストレージ装置 1 0 0 B の S V O L 2 0 0 A 及び S V O L 2 0 0 B に各々ペアとして関連付けられている。

【 0 0 4 8 】

また、P V O L 2 0 1 A に関連付けられた S V O L 2 0 0 A、P V O L 2 0 1 B に関連付けられた S V O L 2 0 0 B 間は、ネットワーク 1 1 0 を介したりリモートコピー 2 0 4 により内容が同期される。この際の同期方式は P V O L 2 0 1 A、P V O L 2 0 1 B へのデータ更新と同期する方式でも良いし、非同期で P V O L 2 0 1 A、P V O L 2 0 1 B への更新差分データを S V O L 2 0 0 A、S V O L 2 0 0 B に反映する方式でも良い。

【 0 0 4 9 】

ここで、一般的な構成例としては、ストレージ装置 1 0 0 A とストレージ装置 1 0 0 B は、システムの可用性を考慮して各々の装置が物理的に距離のある拠点にされていても良いし、近接距離、例えば、同じデータセンタ内に同床して設置されていても良い。

【 0 0 5 0 】

また、ストレージ装置 1 0 0 A では、ホスト 1 0 7 A、ホスト 1 0 7 C がストレージ装置 1 0 0 B では、ホスト 1 0 7 A、ホスト 1 0 7 D、ホスト 1 0 7 E がそれぞれ対応する V O L との間にパスを形成し接続され運用されている。

【 0 0 5 1 】

また、統合管理システム 1 2 1 は管理ネットワーク 1 1 1 を介してストレージ装置 1 0 0 A およびストレージ装置 1 0 0 B の構成情報や性能情報を定期的に監視しており、ストレージ管理者は統合管理システム 1 2 1 を介して各ストレージ装置の状態監視や、統合管理システム 1 2 1 への指示により各ストレージ装置の構成や設定変更を行うことができる。

【 0 0 5 2 】

また、ホスト 1 0 7 A では可用性が要求されるような重要なアプリケーションが実行されており、運用の可用性を担保するため、ネットワーク 1 1 0 を介して、ペアを構成するストレージ装置 1 0 0 A の P V O L 2 0 1 A と P V O L 2 0 1 B およびストレージ装置 1 0 0 B の S V O L 2 0 0 A と S V O L 2 0 0 B との間にアクセス可能なパスが定義される。

【 0 0 5 3 】

通常時ホスト 1 0 7 A は P V O L 2 0 1 A と P V O L 2 0 1 B にアクセスすることで運用されるが、例えば、ストレージ装置 1 0 0 A が障害などで P V O L 2 0 1 A や P V O L 2 0 1 B へのアクセスが不能な場合は、フェイルオーバー 2 0 5 により、ストレージ装置 1 0 0 B の S V O L 2 0 0 A や S V O L 2 0 0 B へアクセス先を切り替えることにより運用を継続することで可用性を担保した運用がなされる。

【 0 0 5 4 】

ここで、フェイルオーバー 2 0 5 は H A ペアを構成するストレージ装置 1 0 0 により障害が検知され、ストレージ装置 1 0 0 A がホスト 1 0 7 A に通知することでホスト 1 0 7 A によりフェイルオーバー 2 0 5 が実行されても良いし、ホスト 1 0 7 A による都合でフェイルオーバー 2 0 5 が実行されても良い。

【 0 0 5 5 】

以上のような構成において、例えば、可用性が要求されるような重要なアプリケーションが実行されているホスト 1 0 7 A に対しては、性能安定性の観点から、そのアクセスに対する I / O 帯域や I O P e r S e c o n d (I O P S) などの一定のサービス

10

20

30

40

50

レベルを保証したい場合が考えられる。そこで、P V O L 2 0 1 A、P V O L 2 0 1 B に対して Q o S の設定を行うことで、装置を共有する V O L 2 0 2 C へのアクセスにより、ホスト 1 0 7 A のアクセスが阻害されることを防ぎ、サービスレベルを一定に保つことを行う。

【 0 0 5 6 】

また、ホスト 1 0 7 A はフェイルオーバー 2 0 5 を実行することで、そのアクセス先を H A ペアである S V O L 2 0 0 A、S V O L 2 0 0 B に切り替えることがあるが、その場合においても、P V O L 2 0 1 A、P V O L 2 0 1 B と同様のサービスレベルを S V O L 2 0 0 A、S V O L 2 0 0 B において保証したい場合があると考えられる。

【 0 0 5 7 】

しかしながら既に説明した通り、ストレージ装置 1 0 0 A とストレージ装置 1 0 0 B は接続されているホストや定義される V O L 数や装置への負荷状況（ワークロード）が異なるため、P V O L 2 0 1 A、P V O L 2 0 1 B に設定した Q o S を同様に S V O L 2 0 0 A、S V O L 2 0 0 B に設定したとしても、期待するサービスレベルが必ずしも維持されるとは限らない。

【 0 0 5 8 】

また、H A ペア構成は正副最低 2 以上のストレージ装置 1 0 0 を必要とすることから、必然的にその運用コストが高い。このため、運用コストやストレージ装置 1 0 0 の稼働率を高めて運用効率を向上させる目的で、正常時の副側装置（例えば、ストレージ装置 1 0 0 B ）においては、実際にフェイルオーバー 2 0 5 が実行されるまでは他の 2 次的な用途

【 0 0 5 9 】

そこで、本発明ではフェイルオーバーの状況に応じた複数種類の Q o S ポリシーを予め設定しておくことで、H A ペア構成における副側装置（例えば、ストレージ装置 1 0 0 B ）においてフェイルオーバー 2 0 5 が実際に実行されるまでの間、ホスト 1 0 7 D やホスト 1 0 7 E により、2 次的な用途で利用されている V O L 2 0 2 D、V O L 2 0 2 E などのアクセスを優先し、一方で、フェイルオーバー 2 0 5 が実行され、ホスト 1 0 7 A のアクセスがストレージ装置 1 0 0 B 側に切り替った場合は、Q o S のポリシーを切り替えて適用し、S V O L 2 0 0 A、S V O L 2 0 0 B のアクセスを優先することで、例えば、ホスト 1 0 7 A が期待するサービスレベルがフェイルオーバー 2 0 5 の実行後においても維持できるように Q o S を制御する。

【 0 0 6 0 】

また、統合管理システム 1 2 1 により、予め設定された複数種類の Q o S ポリシーについて、例えば、各ストレージ装置の構成や性能情報などの稼働情報を利用した予測により、フェイルオーバー 2 0 5 実行後にサービスレベルの維持が困難であると判断する場合、ストレージ管理者にリスクを警告または実現可能な Q o S ポリシーの提案を行うことで、フェイルオーバー 2 0 5 実行後においても、Q o S ポリシーの不整合やサービスレベル破綻が生じないように制御する。

【 0 0 6 1 】

なお、どの Q o S のポリシーを各ストレージ装置 1 0 0 適用するのかを判断するのは、ストレージ装置 1 0 0 自体によって行われても良いし、統合管理システム 1 2 1 により判断し、各ストレージ装置 1 0 0 に対して Q o S の設定変更を命令する方式でも良い。

【 0 0 6 2 】

以上のようにして、H A ペア構成において、フェイルオーバー 2 0 5 前後の状態を加味して適切な Q o S を設定できるようにすることで、システムの安定稼働に必要なサービスレベルと可用性を維持しながら、フェイルオーバー 2 0 5 が実際に生じるまでの間は副側装置における運用効率を高めることができ、システム運用の効率化が図られる。

【 0 0 6 3 】

以下、本実施形態を詳細に説明する。

【 0 0 6 4 】

10

20

30

40

50

図 3 は、本願実施例のストレージ装置における管理情報の構成例を示す。

【 0 0 6 5 】

管理情報 1 0 3 は、複数の管理テーブルを含む。管理テーブルとして、例えば、L D E V に関する情報を保持する L D E V 管理テーブル 3 0 1、H A ペア構成を管理する情報を保持する H A ペア管理テーブル 3 0 0、Q o S ポリシーを管理する情報を保持する Q o S ポリシー管理テーブル 3 0 2、複数の Q o S ポリシーをグループとして管理するための情報を保持する Q o S ポリシー G r 管理テーブル 3 0 3、L D E V 毎の稼働情報を管理する L D E V 稼働情報テーブル 3 0 4、などで構成され、管理情報 1 0 3 A 及び管理情報 1 0 3 B 間、またはストレージ装置 1 0 0 A 及びストレージ装置 1 0 0 B 間で少なくとも一部の情報が同期してもよい。

10

【 0 0 6 6 】

図 4 は、本願実施例のストレージ装置の管理情報における L D E V 管理テーブルの構成例を示す。

【 0 0 6 7 】

L D E V 管理テーブル 3 0 1 は、「 P V O L 」や「 S V O L 」などの L D E V 毎にエン트리 (レコード) を有する。各エントリーが格納する情報は、L D E V 番号 4 0 0、L D E V 容量 4 0 1、V O L 種別 4 0 2、データ削減有効 4 0 4、プール番号 4 0 3、P G 番号 4 0 5 および { データ移行対象 } である。

【 0 0 6 8 】

L D E V 番号 4 0 0 は当該 L D E V の識別番号を表す。

20

【 0 0 6 9 】

L D E V 容量 4 0 1 は当該 L D E V の容量を表す。

【 0 0 7 0 】

V O L 種別 4 0 2 は当該 L D E V の種別を表し、(例えば、「 R V O L 」または「 T P V O L 」またはスナップショットの S V O L など) を表す。

【 0 0 7 1 】

データ削減有効 4 0 4 は当該 L D E V においてデータ圧縮や重複排除といったデータ削減機能が有効かどうかの属性を表す。

【 0 0 7 2 】

プール番号 4 0 3 は、例えば、「 E V O L 」、 「 T P V O L 」や「 S V O L 」といった論理的な容量を持つ L D E V において、それらの L D E V が関連付けられているプールの識別番号を表し、プール番号 4 0 3 が関連付けられているプール内の領域からデータ格納領域が割り当てられる。一方で、例えば「 R V O L 」など物理的な容量を持つ L D E V においては、L D E V がプールに関連付けられることはないため「 N / A 」となる。

30

【 0 0 7 3 】

P G 番号 4 0 5 は、例えば「 R V O L 」など、物理的な容量を持つ L D E V において、関連付けられたパリティグループ (P G) の識別番号を示し、一方で、例えば「 E V O L 」、 「 T P V O L 」や「 S V O L 」といった論理的な容量を持つ L D E V では直接関連付けられる P G は存在しないため、「 N / A 」となる。

【 0 0 7 4 】

図 5 は、本願実施例のストレージ装置の管理情報における H A ペア管理テーブルの構成例を示す。

40

【 0 0 7 5 】

H A ペア管理テーブル 3 0 0 は、H A ペア毎にエントリーを有する。各エントリーが格納する情報は、ペア番号 5 0 0、{ 利用可能 }、P V O L L D E V 番号 5 0 2、ペア装置 I D 5 0 3、S V O L L D E V 番号 5 0 4、同期方式 5 0 5、ペア状態 5 0 6、及び仮想 I D 5 0 7 といった情報を格納する。

【 0 0 7 6 】

ペア番号 5 0 0 は、ストレージ装置 1 0 0 が管理する H A ペアの番号である。P V O L L D E V 番号 5 0 2 は、ペアを構成する P V O L 2 0 1 の L D E V 番号 4 0 0 である。ペ

50

ア装置ID503は、PVOLLDEV番号502とペア関係となるSVOL200の属するストレージ装置100を識別するIDである。{SVOLHDEV番号}は、SVOLLの属するストレージ装置におけるLDEV番号400である。同期方式505は、ペアを構成するPVOL201とSVOL200のデータを同期するタイミングを指定する情報で、ペアを構成するPVOL201とSVOL200のデータをPVOL201へのデータライトと同期してPVOL201とSVOL200の内容を同期する「SYNC」方式と、PVOLのデータライトとは非同期にPVOL201とSVOL200の内容を同期する「ASYNC」方式などを指定する。ペア状態506は、PVOL201およびSVOL200で構成されるペアの状態を示し、例えば、データを同期したペア状態を構成していれば「PAIR」であり、データを同期していないペアを構成していれば「PSS」である。仮想ID507は、ペアを構成するPVOL201とSVOL200を同一のLDEVとホスト107が認識することが可能な仮想のLDEVのIDである。

10

【0077】

図6は、本願実施例のストレージ装置の管理情報におけるQoSポリシー管理テーブルの構成例を示す。

【0078】

QoSポリシー管理テーブル302は、QoSのポリシー毎にエントリを有する。各エントリーが格納する情報は、ポリシー番号600、LDEV番号400、IO種別601、最大IOPS602、最小IOPS603、最大帯域604、及び最小帯域605、といった情報を格納する。

20

【0079】

ポリシー番号600は、QoSポリシーの管理番号である。LDEV番号400は、ポリシー番号600のQoSポリシーにより制御される対象のLDEVのLDEV番号400を示す。IO種別601は、ポリシー番号600のQoSポリシーにより制御されるIOの種類を指定する情報で、例えばLDEVに対するリードアクセスが制御対象であれば「R」、ライトアクセスが制御対象であれば「W」となる。最大IOPS602は、QoS対象のLDEVに対して許容する最大IOPSを指定する情報で、例えば「5000」であれば毎秒5000回のIO処理を上限にストレージ装置100が処理を行い、超過する場合はIO処理を受け付けないことで、当該LDEVのIO処理が指定された負荷以上ならないように制御するために用いる。最小IOPS603は、QoS対象のLDEVに対して許容する最小IOPSを指定する情報で、例えば「2500」であれば毎秒2500回のIO処理を最小限ストレージ装置100が行うように制御することで、LDEVへのIOアクセスを保証する。最大帯域604は、QoS対象のLDEVに対して許容する最大のIO帯域を指定する情報で、例えば「200MB/s」であれば200MB/sのIO帯域を上限にストレージ装置100が処理を行い、超過する場合はIO処理を受け付けないことで、当該LDEVのIO処理帯域が指定された負荷以上ならないように制御するために用いる。最小帯域605は、QoS対象のLDEVに対して許容する最小のIO帯域を指定する情報で、例えば「50MB/s」であれば50MB/sのIO帯域を最小限ストレージ装置100が行うように制御することで、LDEVへのIOアクセスを保証する。

30

40

【0080】

図7は、本願実施例のストレージ装置の管理情報におけるQoSポリシーGr管理テーブルの構成例を示す。

【0081】

QoSポリシーGr管理テーブル303は、QoSの適用ポリシーをグループとして一括管理するためのテーブルであり、各エントリーが格納する情報は、ポリシーGr番号700適用条件701、詳細条件703及び適用ポリシー702といった情報を格納する。

【0082】

ポリシーGr番号700は、複数のQoSポリシーをグループとして一括管理するためのグループ番号である。適用条件701は、当該のQoSポリシーグループを適用する条

50

件を表し、例えば、HAペアにおける正常時の副側装置（ストレージ装置100B）では、正常時に適用すべきQoSポリシーを定義する場合であれば当該フィールドは「HAセカンダリ」であり、フェイルオーバー時には「HAプライマリ」となる。また、詳細条件703は適用条件701における詳細条件であり、例えば適用条件701が「HAプライマリ」であり、詳細条件703が「CTL正常」であれば、当該装置において、HAペア状態がプライマリであり、CTLが正常な状態で適用するQoSポリシーの定義となる。また、適用条件701および詳細条件703のフィールドは、当該エントリーのQoSポリシーを適用すべき条件を表すものであるため、HAペアの条件に限定される必要はなく、例えばストレージ装置100において生じる各種のイベントや状態、時刻などに応じて様々な条件を定義できて良い。適用ポリシー702は、当該のQoSポリシーグループにより一括適用されるポリシー番号600の集合を表す。

【0083】

図8は、本願実施例のストレージ装置の管理情報におけるLDEV稼働情報テーブルの構成例を示す。

【0084】

LDEV稼働情報テーブル304は、タイムスタンプ804が示す時刻毎に対応するLDEV毎の稼働情報を格納するテーブルが対応しており、ストレージコントローラ106により一定期間のLDEV毎にその稼働情報を各エントリーの値として格納する。各エントリーが格納する情報は、LDEV番号400リードIOPS800ライトIOPS801リード帯域802及びライト帯域803といった情報を格納する。

【0085】

LDEV番号400は対象のLDEVのLDEV番号400を表す。リードIOPS800は、監視対象のLDEVにおける毎秒のリードIOの処理回数を表す。ライトIOPS801は、監視対象のLDEVにおける毎秒のライトIOの処理回数を表す。リード帯域802は、監視対象のLDEVにおける毎秒のリードIOにより生じるデータ帯域を表す。ライト帯域803は、監視対象のLDEVにおける毎秒のライトIOにより生じるデータ帯域を表す。

< HAペアの状態に応じたQoSポリシー自動適用フロー >

【0086】

図9は、本願実施例のストレージ装置におけるHAペアの状態に応じたQoSポリシー自動適用フローの処理例を示す。

【0087】

この処理フローは概して、HAペア構成を構成するストレージ装置100において、HAペアの状態に応じて適切なQoSポリシーを選択して適用するための処理フローである。

【0088】

ステップ900で、QoSポリシーGr管理テーブル303を参照する。

【0089】

ステップ901で、ストレージコントローラ106によりHAペアのHA状態をチェックする。その後、ステップ902へ進む。

【0090】

ステップ902で、自装置にフェイルオーバー205されているかどうかをチェックし、自装置がプライマリ側となっている場合はステップ903へ進み、自装置がセカンダリ側となっている場合はステップ907へ進む。

【0091】

ステップ903で、ストレージコントローラ106が縮退状態になっているかどうかをチェックし、障害等が生じて縮退状態であればステップ904へ進み、正常であればステップ906へ進む。

【0092】

ステップ904でQoSポリシーGr管理テーブル303を参照し、適用条件701が

“ H A プライマリ ” で詳細条件 7 0 3 が “ C T L 縮退 ” となっている適用ポリシー 7 0 2 を選択する。その後、ステップ 9 0 5 へ進む。

【 0 0 9 3 】

また、ステップ 9 0 6 では Q o S ポリシー G r 管理テーブル 3 0 3 を参照し、適用条件 7 0 1 が “ H A プライマリ ” で詳細条件 7 0 3 が “ C T L 正常 ” となっている適用ポリシー 7 0 2 を選択する。その後、ステップ 9 0 5 へ進む。

【 0 0 9 4 】

一方、ステップ 9 0 2 で自装置がセカンダリ側であれば、ステップ 9 0 7 で、ストレージコントローラ 1 0 6 が縮退状態になっているかどうかをチェックし、障害等が生じて縮退状態であればステップ 9 0 8 へ進み、正常であればステップ 9 0 9 へ進む。

10

【 0 0 9 5 】

ステップ 9 0 8 で Q o S ポリシー G r 管理テーブル 3 0 3 を参照し、適用条件 7 0 1 が “ H A セカンダリ ” で詳細条件 7 0 3 が “ C T L 縮退 ” となっている適用ポリシー 7 0 2 を選択する。その後、ステップ 9 0 5 へ進む。

【 0 0 9 6 】

また、ステップ 9 0 9 では Q o S ポリシー G r 管理テーブル 3 0 3 を参照し、適用条件 7 0 1 が “ H A セカンダリ ” で詳細条件 7 0 3 が “ C T L 正常 ” となっている適用ポリシー 7 0 2 を選択する。その後、ステップ 9 0 5 へ進む。

【 0 0 9 7 】

ステップ 9 0 5 で適用ポリシー 7 0 2 のポリシー番号 6 0 0 に従って各 L D E V の Q o S を設定する。

20

< Q o S ポリシリスク検知フロー >

【 0 0 9 8 】

図 1 0 は、本願実施例のストレージ装置における Q o S ポリシリスク検知フローの処理例を示す。

【 0 0 9 9 】

この処理フローは概して、統合管理システム 1 2 1 により実行されるものであり、統合管理システム 1 2 1 は、監視対象のストレージ装置 1 0 0 において構成されている H A ペアについて、プライマリ側装置の P V O L において設定されている Q o S 設定が、セカンダリ側装置の S V O L においてフェイルオーバー後に適用される Q o S ポリシーとして設定されているかチェックする。チェックの結果、設定されていない場合は、統合管理システム 1 2 1 は、ストレージ管理者に警告をし、前記 P V O L において設定されている Q o S 設定を S V O L のフェイルオーバー後に適用される Q o S ポリシーにとして設定する。また、統合管理システム 1 2 1 は、前記フェイルオーバー後にセカンダリ側で適用される Q o S ポリシーにおける Q o S 設定値について、セカンダリ側のストレージ装置の処理能力が十分であるか否かチェックし、不足であれば警告や Q o S 設定値の制御を行う。例えば、Q o S を設定していない L D E V などの過去の稼働率を参照し、フェイルオーバー後の Q o S ポリシーに基づく処理がセカンダリ側のストレージ装置 1 0 0 で利用可能な I O 帯域や I O P S を超過する場合は、Q o S を設定していない高負荷な L D E V に対して Q o S を設定することにより、S V O L へのフェイルオーバー後における Q o S 設定値を実現可能とする。

30

40

【 0 1 0 0 】

ステップ 1 0 0 0 で、統合管理システム 1 2 1 により監視対象のストレージ装置 1 0 0 における H A ペア管理テーブル 3 0 0 を参照する。

【 0 1 0 1 】

ステップ 1 0 0 1 で、H A ペアにおける P V O L の L D E V 番号 4 0 0 を取得する。

【 0 1 0 2 】

ステップ 1 0 0 2 で、H A ペアにおける S V O L の L D E V 番号 4 0 0 を取得する。

【 0 1 0 3 】

ステップ 1 0 0 3 で、H A ペアにおけるペア装置 I D 5 0 3 を取得する。

50

【0104】

ステップ1004で、ペア装置のQoSポリシーGr管理テーブル303を参照し、フェイルオーバー発生後のQoSポリシーを取得する。具体的には、適用条件701が「H Aプライマリ」となっているエントリーの適用ポリシー702を取得する。

【0105】

ステップ1006で、当該SVOLにおけるQoSポリシーをQoSポリシー管理テーブル302から取得する。

【0106】

ステップ1007で、当該SVOLにおいてQoSポリシーが設定されているかをチェックし、QoSポリシーが未設定であればステップ1005へ進み、設定されていればステップ1008へ進む。

【0107】

ステップ1008で、ペア装置のLDEV稼働情報テーブル304を参照し、装置における稼働情報を取得する。

【0108】

ステップ1009で、ペア装置における稼働情報と現在設定されているQoSポリシーGr設定から、各LDEVに対してQoSポリシーで設定されている設定項目（例えば最大/最小IOPSまたはIO帯域）が装置の稼働情報を鑑みて装置能力を超過していないかを判断し、QoSポリシーで設定されている項目が実現可能かを判断する。

【0109】

ステップ1010で、QoSポリシーで設定されている項目が実現不能、リスクが高いと判断されるである場合はステップ1012へ進み、実現可能と判断される場合は終了する。

【0110】

ステップ1012で、ペア装置においてQoSが設定されていないLDEVのうち、LDEV稼働情報から稼働率が高い高負荷なLDEVに対してQoSポリシーを設定する。このときのQoSポリシーの決定方法は、例えば当該LDEVの稼働情報における平均値または前記QoSが設定されていないLDEV全体の平均値などに基づき、QoSポリシーGrにより設定されるQoSポリシーが装置能力を超過しない範囲の値を求めればよい。

【0111】

ステップ1005で、ストレージ管理者にSVOLにおいてQoSが設定されていない旨を警告し、ペア装置にフェイルオーバー後にSVOLにおいてSLAが保証されない旨を警告する。

【0112】

ステップ1011で、PVOLに設定されているQoSポリシーを同様にSVOLに適用し、ステップ1008～ステップ1010で当該ポリシーが実現可能かをチェックする。

【0113】

上述してきたように、本実施例に係るストレージシステムは、ネットワーク経由で接続されたホストにボリュームを提供するストレージ装置100を有し、前記ボリュームのサービス品質に係る規定群であるポリシーセットを複数記憶する記憶部と、前記ストレージ装置の状態に応じて複数の前記ポリシーセットから使用するポリシーセットを選択するポリシー選択部とを備える。

かかる構成及び動作により、ストレージ装置の状態に応じた適正な運用を実現することができる。

ここで、ポリシーセットは、QoSポリシーグループに対応する。

また、記憶部は、ストレージ装置100のDRAM115であってもよいし、統合管理システム121が有する記憶領域でもよい。

また、ポリシー選択部は、統合管理システム121であってもよいし、ストレージ装置

10

20

30

40

50

100のCPU114であってもよい。

【0114】

また、ストレージシステムは、第1のストレージ装置100Aに障害が発生した場合にストレージ装置100AのPVOL201から第2のストレージ装置100BのSVOL200へのフェイルオーバーを実行する冗長構成を有し、前記記憶部は、前記第2のストレージ装置100Bが前記フェイルオーバーの実行前に用いるポリシーセットと、前記フェイルオーバーの実行後に用いるポリシーセットとを記憶し、前記フェイルオーバーの実行後に用いるポリシーセットでは、前記フェイルオーバーの実行前に用いるポリシーセットに比べて、第2のボリュームに割り当てる性能が高く、他の複数のボリュームへに割り当てる性能が低い。

10

かかる構成及び動作により、前記第1のストレージ装置と前記第2のストレージ装置との間でHAペアを構成し、前記PVOLにおいてQoSを用いたSLA保証を行う場合において、前記SVOLに対してフェイルオーバーを実施する場合でも、前記PVOLと同等のQoSを用いたSLA保証を前記SVOLにおいて実現することができる。

【0115】

また、前記第2のストレージ装置が前記フェイルオーバーの実行前に用いるポリシーセットは、前記ホストに係るデータ処理とは独立に行う処理を優先する規定群であり、前記第2のストレージ装置が前記フェイルオーバーの実行後に用いるポリシーセットは、前記ホストに係るデータ処理を優先する規定群である。

かかる構成及び動作により、例えば、前記第2のストレージ装置ではフェイルオーバーが発生するまでの間、二次的な目的で利用するホストやアプリケーションからのIOを優先するようにQoSを設定し、フェイルオーバー後においては、前記SVOLへのIOが優先されるようにQoSを設定することができる。

20

【0116】

また、前記ポリシー選択部は、前記第2のストレージ装置の処理能力と、前記第2のストレージ装置が前記フェイルオーバーの実行後に用いるポリシーセットとを比較し、前記第2のストレージ装置の処理能力に不足がある場合に警告を行う。

かかる構成及び動作により、ストレージ管理者はフェイルオーバー時のSLAについてリスクがある場合は、事前に知ることができるため、フェイルオーバー時におけるSLAに起因する障害発生を予防できる。

30

【0117】

また、前記ストレージ装置は、コントローラの縮退運用が可能であり、前記記憶部は、前記縮退運用が不要である場合に用いるポリシーセットと前記縮退運用中に用いるポリシーセットとを記憶する。

かかる構成及び動作により縮退の発生によりストレージ装置の状態が変化した場合にも、その変化に応じて適切なQoSの設定を用いることができる。

【0118】

また、前記記憶部は、時間に応じて異なる複数のポリシーセットを記憶し、前記ポリシー選択部は、その時点に対応したポリシーセットを選択する構成とすれば、時刻や曜日による処理量の変化に応じた適切なQoSの設定を用いることができる。

40

【0119】

以上、一実施形態を説明したが、これは本発明の説明のための例示であって、本発明の範囲をこの実施形態のみに限定する趣旨ではない。本発明は他の形態でも実施することが可能である。

【符号の説明】

【0120】

100 ストレージ装置、101 ドライブ、102 制御プログラム、103 管理情報、104 NVRAM、105 キャッシュメモリ、106 ストレージコントローラ、107 ホスト、108 管理システム、109 ホストインタフェース、110 ネットワーク、111 管理ネットワーク、112 中継器、113 プロセッサパッケー

50

ジ、114 CPU、115 DRAM、116 フロントエンドインタフェース、117 バックエンドインタフェース、118 管理用コントローラ、119 クラウド、120 管理インタフェースデバイス、121 統合管理システム、122 レプリケーションインタフェース、200 SVOL、201 PVOL、202 VOL、203 リモートペア、204 リモートコピー、205 フェイルオーバー、300 HAペア管理テーブル、301 LDEV管理テーブル、302 QoSポリシー管理テーブル、303 QoSポリシーGr管理テーブル、304 LDEV稼働情報テーブル、400 LDEV番号、401 LDEV容量、402 VOL種別、403 プール番号、404 データ削減有効、405 PG番号、500 ペア番号、501 利用可能号、502 PVOLLDEV番号、503 ペア装置ID、504 SVOLLDEV番号、505 同期方式、506 ペア状態、507 仮想ID、600 ポリシ番号、601 IO種別、602 最大IOPS、603 最小IOPS、604 最大帯域、605 最小帯域、700 ポリシーGr番号、701 適用条件、702 適用ポリシー703 詳細条件、800 リードIOPS、801 ライトIOPS、802 リード帯域、803 ライト帯域、804 タイムスタンプ

10

20

30

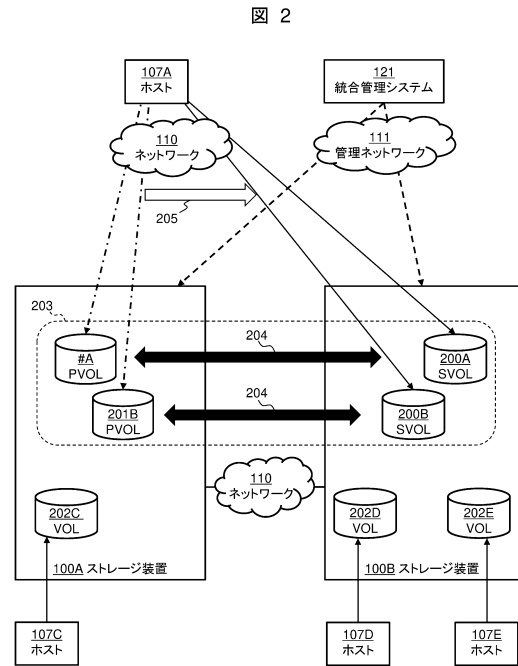
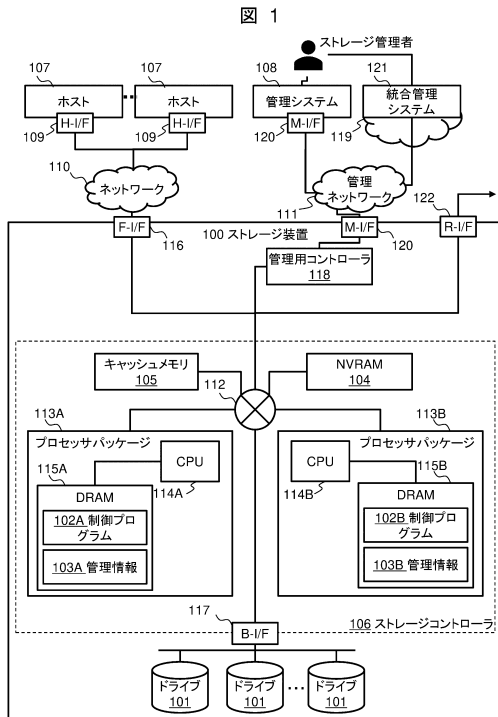
40

50

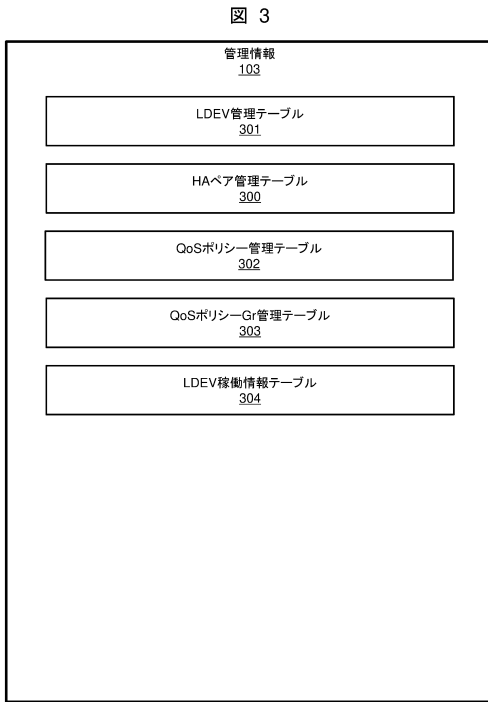
【図面】

【図1】

【図2】



【 図 3 】



【 図 4 】

図 4
LDEV管理テーブル
301

400	401	402	403	404	405
LDEV番号	LDEV容量	VOL種別	データ削減種別	プール番号	PG番号
1	1TB	EVOL	無効	1	N/A
2	2TB	TPVOL	有効	1	N/A
3	320GB	SVOL	無効	1	N/A
4	320GB	PVOL	無効	1	N/A
5	32TB	RVOL	無効	N/A	2
...

10

20

【 図 5 】

図 5
HAペア管理テーブル
300

500	501	502	503	504	505	506	507
ペア番号	利用可能	PVOLLDEV番号	ペア装置ID	SVOLLDEV番号	同期方式	ペア状態	装置ID
1	可能	1	10001	1	SYNC	PAIR	001
2	可能	2	10000	2	ASYNC	PSUS	002
3	不能	3	-	-	-	-	-
...

30

【 図 6 】

図 6
QoSレバレッジ管理テーブル
302

600	601	602	603	604	605
ホリゾン番号	LDEV番号	I/O種別	最大IOPS	最小IOPS	最大帯域
1	1	W	-	-	200MB/s
2	1	W	-	-	100MB/s
3	2	R	5000	2500	-
4	2	R	10000	5000	-
5	3	R	10000	5000	-
6	3	R	5000	2500	-
...

40

50

【 図 7 】

図 7
QoSポリシーGr管理テーブル
303

700 ポリシーGr番号	701 適用条件	703 詳細条件	702 適用ポリシー
1	HAプライマリ	CTR正常	{1,3,5}
2	HAプライマリ	CTR縮退	{2,4,6}
3	HAセカンダリ	CTR正常	...
4	HAセカンダリ	CTR縮退	...
...

【 図 8 】

図 8

LDEV稼働情報テーブル
304

400 LDEV番号	800 リードIOPS	801 ライトIOPS	802 リード帯域	803 ライト帯域
1	10	200	172MB/s	5MB/s
2	1000	10	10MB/s	183MB/s
3	10	5	3MB/s	1MB/s
...

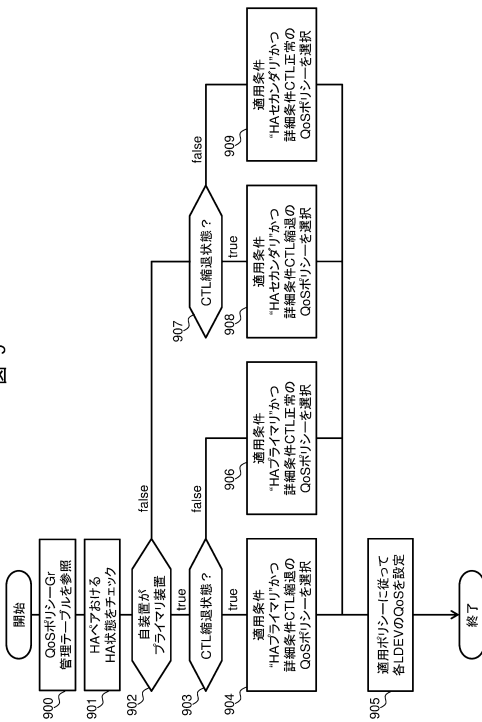
804 タイムスタンプ
2020.10.01 00:00:00
2020.10.01 00:00:30
2020.10.01 00:01:30
...

10

20

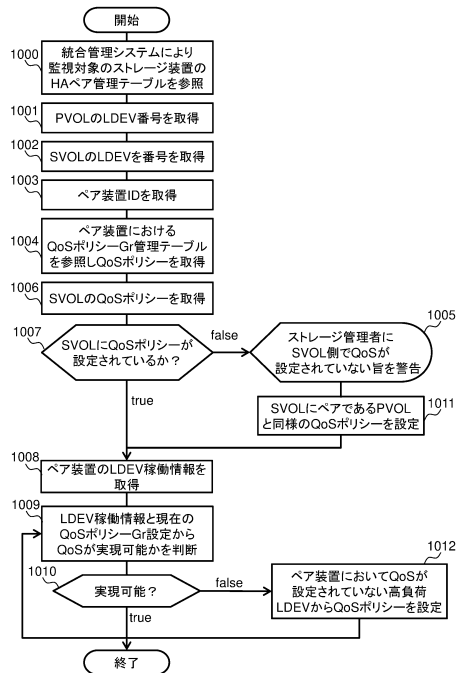
【 図 9 】

図 9



【 図 10 】

図 10



30

40

50

フロントページの続き

Fターム(参考) 5B034 BB01 CC04