(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6231685号
(P6231685)

(45)発行日　平成29年11月15日(2017.11.15)　(24)登録日　平成29年10月27日(2017.10.27)

(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	301C
		G06F	13/14	310K

請求項の数 15　(全 35 頁)

(21)出願番号　特願2016-534025 (P2016-534025)
(86)(22)出願日　平成26年7月16日(2014.7.16)
(86)国際出願番号　PCT/JP2014/068873
(87)国際公開番号　WO2016/009504
(87)国際公開日　平成28年1月21日(2016.1.21)
審査請求日　平成28年7月22日(2016.7.22)

(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　110000279
特許業務法人ウィルフォート国際特許事務所
(72)発明者　末次　通夫
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　川口　智大
東京都千代田区丸の内一丁目６番６号　株式会社日立製作所内
(72)発明者　斎藤　秀雄
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及び通知制御方法

(57)【特許請求の範囲】
【請求項１】
ホストシステムに接続されるストレージシステムであって、
複数の論理ボリュームを含む複数種類の複数のリソースを管理し、前記複数の論理ボリュームが１つに仮想化された論理ボリュームである仮想ボリュームを前記ホストシステムに提供する複数のストレージ装置を有し、
前記複数のストレージ装置のうちのいずれかのストレージ装置である第１ストレージ装置が、第１コマンドを受信した場合、第１リソースの状態を変更し、前記第１リソースは、前記仮想ボリュームの基である前記複数の論理ボリュームのうち前記第１ストレージ装置が有する論理ボリュームである第１論理ボリュームと前記第１論理ボリュームに関連し前記第１ストレージ装置が管理するリソースとのうちの少なくとも１つであり、
前記第１ストレージ装置が、前記仮想ボリュームを指定したＲＴＰＧコマンド（Report TargetPortGroupsコマンド）及び前記仮想ボリュームを指定したＩ／Ｏコマンドのいずれかである第２コマンドを受信した場合、前記第１論理ボリュームに関わる状態変更の通知である状態変更通知を第２ストレージ装置に送信し、前記第２ストレージ装置は、第２論理ボリュームを有するストレージ装置であり、前記第２論理ボリュームは、前記仮想ボリュームの基であり前記第１論理ボリュームに関連付けられた論理ボリュームであり、
前記第２ストレージ装置は、前記状態変更通知を受信し、前記受信した状態変更通知に基づき状態変更を設定する、
ストレージシステム。

【請求項２】
前記複数のリソースは、複数のポートを含み、
前記第１リソースは、前記第１論理ボリュームへのパスであり前記第１論理ボリュームが関連付けられたパスを経由する第１パスであり、
前記複数のストレージ装置の各々は、そのストレージ装置が管理するパスの状態を表す情報を保持しているが、そのストレージ装置以外のストレージ装置が管理するパスの状態を表す情報を保持しておらず、
前記第２ストレージ装置において、前記状態変更通知に基づく状態変更は、前記第２論理ボリュームに関して設定される、
請求項１記載のストレージシステム。
【請求項３】
前記第１ストレージ装置は、前記第１コマンドを受信した場合、前記第１論理ボリュームに関して状態変更を設定し、
前記第２コマンドの受信は、前記仮想ボリュームを指定したＩ／Ｏコマンドの受信であり、
前記第２ストレージ装置は、前記受信した状態変更通知に基づき状態変更を設定した場合、完了応答を前記第１ストレージ装置に送信し、
前記第１ストレージ装置は、前記完了応答を受信した場合、前記第１論理ボリュームに関して設定されている状態変更を含んだ、前記Ｉ／Ｏコマンドの応答を、前記ホストシステムに送信する、
請求項２記載のストレージシステム。
【請求項４】
前記第１ストレージ装置は、
前記第１論理ボリュームに関して設定されている状態変更を含んだ前記応答を受信した前記ホストシステムから、前記仮想ボリュームを指定したＲＴＰＧコマンドを前記第２コマンドとして受信し、
前記第１論理ボリュームとペアを構成する前記第２論理ボリュームを有する前記第２ストレージ装置に、前記第２論理ボリュームに関連付けられたパスの状態の要求を送信し、前記要求を受信した前記第２ストレージ装置から、前記第２論理ボリュームに関連付けられた１以上のパスの状態を表す情報を含んだパス状態情報を受信し、
前記第１論理ボリュームに関連付けられ前記第１パスを含んだ１以上のパスの状態を表す第１パス状態と、前記第２論理ボリュームに関連付けられた１以上のパスの状態である第２パス状態とを含んだ、前記ＲＴＰＧコマンドの応答を、前記ホストシステムに送信する、
請求項３記載のストレージシステム。
【請求項５】
前記第１ストレージ装置は、前記第１論理ボリュームと前記第２論理ボリュームについてのペア状態及びＩ／Ｏ状態に応じて、前記ＲＴＰＧコマンドの応答に含まれる前記第１パス状態及び前記第２パス状態を制御する、
請求項４記載のストレージシステム。
【請求項６】
前記第１パス状態は、前記第１論理ボリュームに関連付けられた前記１以上のパスの優先度である第１パス優先度を表し、前記第２パス状態は、前記第２論理ボリュームに関連付けられた前記１以上のパスの優先度である第２パス優先度を表し、
前記第１ストレージ装置は、
前記ペア状態及び前記Ｉ／Ｏ状態が、前記第１論理ボリュームと前記第２論理ボリュームが同期状態であることを表していれば、前記第１パス優先度を、前記第１ストレージ装置が保持し前記第１論理ボリュームに関連付けられた前記１以上のパスの優先度を表す情報が表す優先度とし、前記第２パス優先度を、前記受信したパス状態情報が表す優先度とし、

前記ペア状態及び前記Ｉ／Ｏ状態が、前記第１論理ボリュームと前記第２論理ボリュームがサスペンド状態であり前記第１論理ボリューム内のデータよりも前記第２論理ボリューム内のデータの方が新しいことを表していれば、前記第１パス優先度よりも前記第２パス優先度を高くし、

前記ペア状態及び前記Ｉ／Ｏ状態が、前記第１論理ボリュームと前記第２論理ボリュームがサスペンド状態であり前記第２論理ボリューム内のデータよりも前記第１論理ボリューム内のデータの方が新しいことを表していれば、前記第２パス優先度よりも前記第１パス優先度を高くする、

請求項５記載のストレージシステム。

【請求項７】

前記第１ストレージ装置は、前記第２ストレージ装置との通信が不可能である場合、

前記ペア状態及び前記Ｉ／Ｏ状態が、前記第１論理ボリュームと前記第２論理ボリュームがサスペンド状態であり前記第１論理ボリューム内のデータよりも前記第２論理ボリューム内のデータの方が新しいことを表していれば、エラー応答を前記ホストシステムに送信し、

前記ペア状態及び前記Ｉ／Ｏ状態が、前記第１論理ボリュームと前記第２論理ボリュームが同期状態であること、又は、前記第１論理ボリュームと前記第２論理ボリュームがサスペンド状態であり前記第２論理ボリューム内のデータよりも前記第１論理ボリューム内のデータの方が新しいことを表していれば、前記第１パス優先度を前記第２パス優先度よりも高くすることを含んだ、前記ＲＴＰＧコマンドの応答を、前記ホストシステムに送信する、

請求項６記載のストレージシステム。

【請求項８】

前記状態変更は、ＡＬＵＡ（Asymmetric Logical Unit Access）に従うＵＡ（Unit Attention）であり、

前記複数のストレージ装置の各々に、前記仮想ボリュームに関連付けられた複数のターゲットポートグループが定義されており、

前記複数のターゲットポートグループは、１以上のポートの集合である、

請求項２記載のストレージシステム。

【請求項９】

前記第１コマンドは、前記第１ストレージ装置が有する複数の論理ボリュームのうち少なくとも前記第１論理ボリュームを含んだ２以上の論理ボリュームについてのコマンドであり、

前記第２コマンドは、前記第１論理ボリュームについてのコマンドであり、

前記第１ストレージ装置は、前記第１コマンドに応答して前記第１パスを含む複数のパスの各々の状態を変更しても、前記第２コマンドを受信した場合に送信する状態変更通知を、前記２以上の論理ボリュームのうちの前記第１論理ボリュームに関わる状態変更の通知とする、

請求項２記載のストレージシステム。

【請求項１０】

前記第１ストレージ装置は、前記第１コマンドを受信した場合、前記第１論理ボリュームに関して状態変更を設定し、

前記第２コマンドの受信は、前記仮想ボリュームを指定したＲＴＰＧコマンドの受信であり、

前記ＲＴＰＧコマンドを受信した前記第１ストレージ装置は、

前記第１論理ボリュームとペアを構成する前記第２論理ボリュームを有する前記第２ストレージ装置に、前記状態変更通知と、前記第２論理ボリュームに関連付けられたパスの状態の要求とを送信し、前記要求を受信した前記第２ストレージ装置から、前記第２論理ボリュームに関連付けられた１以上のパスの状態を表す情報を含んだパス状態情報を受信し、

前記第１論理ボリュームに関連付けられ前記第１パスを含んだ１以上のパスの状態を表す第１パス状態と、前記第２論理ボリュームに関連付けられた１以上のパスの状態である第２パス状態とを含んだ、前記ＲＴＰＧコマンドの応答を、前記ホストシステムに送信する、
請求項２記載のストレージシステム。
【請求項１１】
１以上のホストシステムが、複数のホスト装置で構成され、
前記複数のホスト装置の各々が、そのホスト装置が属するグループのＩＤを保持し、
前記複数のストレージ装置の各々が、そのストレージ装置が属するグループのＩＤを保持し、
各グループには、少なくとも１つのホスト装置と少なくとも１つのストレージ装置が属し、
前記第１ストレージ装置は、グループＩＤの問合せを、前記第１ストレージ装置に接続されたいずれかのホスト装置である第１ホスト装置から受信した場合、前記第１ストレージ装置が保持するグループＩＤを含んだ応答を前記第１ホスト装置に送信し、
前記第１ホスト装置は、
前記応答内のグループＩＤが、前記第１ホスト装置が保持するグループＩＤと同一であれば、前記第１ストレージ装置が有し前記第１論理ボリュームを含んだ１以上の論理ボリュームに関連付けられており前記第１パスを含んだ１以上のパスの各々の優先度を、第１の優先度とし、
前記応答内のグループＩＤが、前記第１ホスト装置が保持するグループＩＤと異なっていれば、前記１以上のパスの各々の優先度を、前記第１の優先度より低い第２の優先度とする、
請求項２記載のストレージシステム。
【請求項１２】
前記第２コマンドの受信は、前記仮想ボリュームを指定したＩ／Ｏコマンドの受信、及び、前記仮想ボリュームに関連した所定種類の問合せのうちの少なくとも１つであり、
前記第１ストレージ装置は、前記第１論理ボリュームに関して状態変更を管理している間に、前記仮想ボリュームを指定したＩ／Ｏコマンドを前記ホストシステムから受信した場合、状態変更を含んだ、前記Ｉ／Ｏコマンドの応答を、前記ホストシステムに送信する、
請求項１記載のストレージシステム。
【請求項１３】
前記複数のストレージ装置の各々は、そのストレージ装置が管理するリソースの状態を表す情報を保持しているが、そのストレージ装置以外のストレージ装置が管理するリソースの状態を表す情報を保持しておらず、
前記第１ストレージ装置は、前記仮想ボリュームを指定したＲＴＰＧコマンドを前記第２コマンドとして受信した場合、
前記第１論理ボリュームとペアを構成する前記第２論理ボリュームを有する前記第２ストレージ装置に、第２リソースの状態の要求を送信し、前記要求を受信した前記第２ストレージ装置から、前記第２リソースの状態を表す情報を含んだパス状態情報を受信し、
前記第２リソースは、前記第１リソースに関連付けられたリソースであり、前記第２論理ボリュームと前記第２論理ボリュームに関連し前記第２ストレージ装置が管理するリソースとのうちの少なくとも１つであり、
前記第１リソースの状態と、前記第２リソースの状態とを含んだ、前記ＲＴＰＧコマンドの応答を、前記ホストシステムに送信する、
請求項１記載のストレージシステム。
【請求項１４】
前記第１コマンドは、前記第１ストレージ装置が有する複数の論理ボリュームのうち少なくとも前記第１論理ボリュームを含んだ２以上の論理ボリュームについてのコマンドで

あり、
前記第２コマンドは、前記第１論理ボリュームについてのコマンドであり、
前記第１ストレージ装置は、前記第１コマンドに応答して前記第１リソースを含む複数のリソースの各々の状態を変更しても、前記第２コマンドを受信した場合に送信する状態変更通知を、前記２以上の論理ボリュームのうちの前記第１論理ボリュームに関わる状態変更の通知とする、
請求項１記載のストレージシステム。
【請求項１５】
複数の論理ボリュームを含む複数種類の複数のリソースを管理し前記複数の論理ボリュームが１つに仮想化された論理ボリュームである仮想ボリュームをホストシステムに提供する複数のストレージ装置を有したストレージシステムにおける通知制御方法であって、
前記複数のストレージ装置のうちのいずれかのストレージ装置である第１ストレージ装置により、第１コマンドを受信した場合、第１リソースの状態を変更し、前記第１リソースは、前記仮想ボリュームの基である前記複数の論理ボリュームのうち前記第１ストレージ装置が有する論理ボリュームである第１論理ボリュームと前記第１論理ボリュームに関連し前記第１ストレージ装置が管理するリソースとのうちの少なくとも１つであり、
前記第１ストレージ装置により、前記仮想ボリュームを指定したＲＴＰＧコマンド（ReportTargetPortGroupsコマンド）及び前記仮想ボリュームを指定したＩ／Ｏコマンドのいずれかである第２コマンドを受信した場合、前記第１論理ボリュームに関わる状態変更の通知である状態変更通知を第２ストレージ装置に送信し、前記第２ストレージ装置は、第２論理ボリュームを有するストレージ装置であり、前記第２論理ボリュームは、前記仮想ボリュームの基であり前記第１論理ボリュームに関連付けられた論理ボリュームであり、
前記第２ストレージ装置により、前記状態変更通知を受信し、前記受信した状態変更通知に基づき状態変更を設定する、
る通知制御方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、概して、ストレージシステムを構成する複数のストレージ装置間の情報の通知を制御する技術に関する。
【背景技術】
【０００２】
例えば、特許文献１によれば、１つのストレージ装置が２つのコントローラを有し、２つのコントローラのうちのどちらを利用するかが判断され、その判断結果を基にパス優先度が決定される。
【先行技術文献】
【特許文献】
【０００３】
【特許文献１】US2012/0254657
【発明の概要】
【発明が解決しようとする課題】
【０００４】
ところで、ホストシステムと論理ボリューム（ストレージ装置がホストシステムに提供する論理的な記憶デバイス）との間のパスの特定に関し、ＡＬＵＡ（Asymmetric Logical Unit Access）と呼ばれるプロトコルが知られている。ホストシステム（イニシエータ）は、ＡＬＵＡを利用して、論理ボリューム（ターゲット）についてパスの状態（例えば優先度）を照会できる。ＡＬＵＡが利用される環境では、ストレージ装置が有する複数のポートのうちの１以上のポートを「ターゲットポートグループ」（以下、ＴＰＧ）と定義することができる。
【０００５】

ＡＬＵＡが利用される環境では、１つのＴＰＧの状態が変更されると、そのＴＰＧ（以下、状態変更ＴＰＧ）に関連するＴＰＧ（以下、関連ＴＰＧ）に対して、状態変更が通知されなければならない。状態変更ＴＰＧと全ての関連ＴＰＧが１つのストレージ装置にあれば、複数のＴＰＧについてそのストレージ装置内で共有される構成情報を更新することで（例えば全ての関連ＴＰＧの各々に関する情報を更新することで）、状態変更の通知が可能である。

【０００６】

一方、複数のストレージ装置がそれぞれ有する複数の論理ボリュームを仮想的に１つの論理ボリュームとしてホスト装置に提供するストレージシステムでは、状態変更ＴＰＧと少なくとも１つの関連ＴＰＧが異なる２以上のストレージ装置に分かれていることがある。この場合、状態変更ＴＰＧを有するストレージ装置から関連ＴＰＧを有するストレージ装置へ状態変更が通知されなければならない。そのような状態変更の通知を実現する１つの仕組みとして、全てのストレージ装置が、それぞれ、各ストレージ装置の構成情報を保持し、全てのストレージ装置間で構成情報を同期することが考えられる。しかし、そのような仕組みでは、各ストレージ装置が保持する構成情報の総量が膨大となり、さらに、構成情報の同期をとるためにストレージシステム全体の性能が低下してしまう。

【０００７】

この種の問題は、論理ボリューム以外のオブジェクトが仮想化されるケースについてもあり得る。また、この種の問題は、ＴＰＧ及びパス以外のリソースの状態変更（例えば優先度変更）についてもあり得る。また、この種の問題は、ＡＬＵＡを利用するストレージシステムに限らず、複数のストレージ装置を有する他種のストレージシステム、例えば、いずれかのストレージ装置でのリソースの状態変更をホストシステムからのアクセスに応答してそのホストシステムに通知したり（気付かせたり）、状態変更に気付いたホストシステムから所定種類の問合せを受けた場合にリソースの変更後の状態をホストシステムに通知するようになっているストレージシステム、についてもあり得る。

【課題を解決するための手段】

【０００８】

ストレージシステムは、複数のストレージ装置を有する。複数のストレージ装置は、複数の論理ボリュームを含む複数種類の複数のリソースを管理し、複数の論理ボリュームが１つに仮想化された論理ボリュームである仮想ボリュームをホストシステムに提供する。第１ストレージ装置（いずれかのストレージ装置）が、第１イベントを検出した場合、第１リソースの状態を変更する。その後、第１リソースの状態をホストシステムに将来通知し得ることを意味する第２イベントを検出した場合、第１ストレージ装置は、第１論理ボリュームに関わる状態変更の通知である状態変更通知を、第１論理ボリュームに関連付けられ仮想ボリュームの基である第２論理ボリュームを有する第２ストレージ装置に送信する。第２ストレージ装置は、状態変更通知を受信し、受信した状態変更通知に基づき状態変更を設定する。なお、第１リソースは、仮想ボリューム（複数の論理ボリュームが１つに仮想化された論理ボリューム）の基である複数の論理ボリュームのうち第１ストレージ装置が有する論理ボリュームである第１論理ボリュームと、第１論理ボリュームに関連し第１ストレージ装置が管理するリソース（例えばパス）とのうちの少なくとも１つである。また、ホストシステムは、１以上のホスト装置（例えば、クラスタを構成する２以上のホスト装置）である。１以上のホスト装置のうちの少なくとも１つは、物理的なホスト装置でよく、１以上のホスト装置は、物理的なホスト装置で実行される仮想的なホスト装置を含んでもよい。

【発明の効果】

【０００９】

全てのストレージ装置が、それぞれ、各ストレージ装置の構成情報を保持し、且つ、全てのストレージ装置間で構成情報を同期すること無しに、第１論理ボリューム又はそれに関連したリソースであり第１ストレージ装置が管理する第１リソースの状態の変更を、ホストシステムにその状態変更を知られる前に、第１論理ボリュームに関連付けられた第２

論理ボリュームを有する第２ストレージ装置に通知することができる。
【図面の簡単な説明】
【００１０】
【図１】実施例１に係る計算機システムの構成を示す。
【図２】各ストレージ装置が有する管理情報の構成を示す。
【図３】ＡＬＵＡ管理テーブルの構成を示す。
【図４】ペア管理テーブルの構成を示す。
【図５】ＵＡ管理テーブルの構成を示す。
【図６】ＵＡ同期フラグビットリストの構成を示す。
【図７】パス管理テーブルの構成を示す。
【図８】ポート管理テーブルの構成を示す。
【図９】仮想ボックス管理テーブルの構成を示す。
【図１０】相対ポートＩＤ管理テーブルの構成を示す。
【図１１】有効ＩＤ範囲管理テーブルの構成を示す。
【図１２】ホスト装置がストレージ装置のパス優先度を把握するまでの流れを示す。
【図１３】異なる２つのストレージ装置のパス優先度をマージすることを模式的に示す。
【図１４】ＡＬＵＡ適用状態変更処理の流れを示す。
【図１５】ＡＬＵＡ適用状態確認処理の流れを示す。
【図１６】パス優先度変更処理の流れを示す。
【図１７】Ｉ／Ｏ処理の際のＵＡ伝播の流れを示す。
【図１８Ａ】ペア状態「ＰＡＩＲ」の場合の優先パスとアクセス経路の一例を示す。
【図１８Ｂ】ペア状態「ＳＵＳＰＥＮＤ」の場合の優先パスとアクセス経路の一例を示す。
【図１９】パス優先度変更処理の流れを示す。
【図２０】パス優先度レポート処理の流れの第１の部分を示す。
【図２１】パス優先度レポート処理の流れの第２の部分を示す。
【図２２】パス優先度レポート処理の流れの第３の部分を示す。
【図２３】パス優先度レポート処理の流れの第４の部分を示す。
【図２４】実施例２に係る計算機システムの構成、及び、各ストレージ装置が有する管理情報の構成を示す。
【図２５】データセンタＩＤ問合せコマンドの構成を示す。
【図２６】パス優先度管理テーブルの構成を示す。
【図２７Ａ】第１データセンタに属するホスト装置及びストレージ装置がそれぞれ有するデータセンタＩＤ管理テーブルを示す。
【図２７Ｂ】第２データセンタに属するホスト装置及びストレージ装置がそれぞれ有するデータセンタＩＤ管理テーブルを示す。
【図２８】パス優先度自動決定処理の流れを示す。
【図２９】Ｉ／Ｏ制御処理の流れを示す。
【図３０】実施例１に係る計算機システムの概要を示す。
【発明を実施するための形態】
【００１１】
以下の説明では「ａａａテーブル」等の表現にて情報を説明する場合があるが、これら情報は、テーブル等のデータ構造以外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に「ａａａテーブル」等について「ａａａ情報」と呼ぶことがある。さらに、各情報の内容を説明する際に、「ＩＤ」、「番号」という表現を用いるが、これらに代えて又は加えて他種の識別情報が使用されてよい。
【００１２】
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場合があるが、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で、定められた処理を、適宜に記憶資源（例えばメモリ）及び／又は通信インター

フェイスを用いながら行うため、処理の主語がプロセッサとされてもよい。プログラムを主語として説明された処理は、プロセッサを含むストレージコントローラ、ストレージ装置又はホスト装置が行う処理としてもよい。また、プロセッサが、処理の一部又は全部を行うハードウェア回路を含んでもよい。コンピュータプログラムは、プログラムソースからストレージ装置又はホスト装置にインストールされてもよい。プログラムソースは、例えば、プログラム配布サーバ、又は、計算機が読み取り可能な記憶メディアであってもよい。

【００１３】

また、以下の説明では、同種の要素を区別して説明する場合は、その要素の参照符号を使用し（例えばストレージ装置２０Ａ、ストレージ装置２０Ｂ）、同種の要素を区別しないで説明する場合は、その要素の参照符号のうちの共通符号のみ使用する（例えばストレージ装置２０）ことがある。

【００１４】

以下、図面を参照して、幾つかの実施例を説明する。なお、以下の実施例では、ＡＬＵＡ（Asymmetric Logical Unit Access）を例に取る。

【実施例１】

【００１５】

図３０は、実施例１に係る計算機システムの概要を示す。

【００１６】

ホスト装置１０と、ホスト装置１０が接続されたストレージシステム３０３とを有する。ホスト装置１０は、ホストシステムの一例である。ストレージシステム３０３は、ストレージ装置２０Ａ及び２０Ｂを有する。ストレージ装置２０Ａ及び２０Ｂは、複数のストレージ装置の一例であり、ストレージ装置２０の数は3以上でもよい。以下の説明では、ストレージ装置２０Ａを「第１ストレージ装置２０Ａ」と言い、ストレージ装置２０Ｂを「第２ストレージ装置２０Ｂ」と言うことがある。

【００１７】

ストレージ装置２０Ａ及び２０Ｂは、複数のＬＤＥＶ（論理記憶デバイス）３０１を含む複数種類の複数のリソースを管理する。ＬＤＥＶは、論理ボリュームである。リソースとしては、ＬＤＥＶの他に、ターゲットポートグループ（ＴＰＧ）３０７及びパス３０９等がある。ＴＰＧ３０７Ａ及び３０７Ｂの各々は、本実施例では１つのポートであるが、２以上のポートの集合でもよい。また、ＬＤＥＶ３０１は、論理的な記憶デバイスであり、実体的なＬＤＥＶでも仮想的なＬＤＥＶでもよい。実体的なＬＤＥＶは、物理的な記憶資源（例えば１以上の物理記憶デバイス）に基づくＬＤＥＶである。仮想的なＬＤＥＶは、外部のストレージ装置（図示せず）の記憶資源（例えばＬＤＥＶ）に基づいておりストレージ仮想化技術に従うＬＤＥＶである外部接続ＬＤＥＶであってもよいし、複数の仮想ページ（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ＬＤＥＶであってもよい。本実施例では、説明を分かり易くするため、各ＬＤＥＶは、いずれかのＶＤＥＶの基になるとする。

【００１８】

ストレージ装置２０Ａ及び２０Ｂは、ＬＤＥＶ３０１Ａ及び３０１Ｂが１つに仮想化されたＬＤＥＶであるＶＤＥＶ（仮想ボリューム）３０５をホスト装置１０に提供する。ホスト装置１０が、ＶＤＥＶ３０５を認識（例えばマウント）する。

【００１９】

ホスト装置１０は、物理的又は仮想的な計算機であり、パス管理プログラム３０２を実行する。パス管理プログラム３０２は、ホスト装置１０が認識したＶＤＥＶ３０５に関連付けられている複数のパス３０９を管理する。パス３０９は、ＴＰＧ３０７を経由し、そのＴＰＧ３０７に関連付けられているＬＤＥＶ３０１へと繋がる。図３０の例では、１つのＶＤＥＶ３０５の基になっているＬＤＥＶ３０１毎に１つのパス３０９であるが、１つのＬＤＥＶ３０１につきパス３０９は１つ以上でよく、故に、１つのＶＤＥＶ３０５につきパス３０９は２つ以上でよい。パス管理プログラム３０２は、各パスの優先度を管理し

ており、優先度の高いパス３０９を優先度の低いパス３０９よりも優先的に使用する。例えば、パス３０９Ｂの優先度よりパス３０９Ａの優先度が高ければ、パス管理プログラム３０２は、パス３０９Ａを優先的に使用してＩ／Ｏ（Input/Output）コマンドをストレージシステム３０３に送信する。パス優先度が、パス状態の一例である。

【００２０】

パス３０９Ａの優先度は、ホスト装置１０及び第１ストレージ装置２０Ａの各々において管理されており、パス３０９Ｂの優先度は、ホスト装置１０及び第２ストレージ装置２０Ｂの各々において管理されている。各ストレージ装置は、他のストレージ装置が管理するパスを管理しない。すなわち、第１ストレージ装置２０Ａは、第２ストレージ装置２０Ｂが管理するパス３０９Ｂについて優先度を管理しないし、第２ストレージ装置２０Ｂは、第１ストレージ装置２０Ａが管理するパス３０９Ａについて優先度を管理しない。従って、本実施例では、全てのストレージ装置２０が、それぞれ、各ストレージ装置のパス管理情報（パス優先度を含んだ情報）を保持し、且つ、全てのストレージ装置間でパス管理情報を同期する必要が無い。

【００２１】

この環境において、例えば、第１ストレージ装置２０Ａにおいて、パス３０９Ａの優先度が変更されたとする。第１ストレージ装置２０Ａは、パス３０９Ａの優先度が変更された場合、状態変更を、例えば、パス３０９Ａが関連付けられている第１ＬＤＥＶ３０１Ａについて設定する。状態変更の一例が、ＡＬＵＡではＵＡ（Unit Attention）である。

【００２２】

この段階では、未だ、ホスト装置１０は、パス３０９Ａの優先度が変更されたことを知らない。ホスト装置１０は、ストレージシステム３０３の詳細な構成を把握しておらず、ＡＬＵＡによれば、ストレージシステム３０３が、ホスト装置１０に、ホスト装置１０が使用するパスを決めさせる。具体的には、ストレージシステム３０３からホスト装置１０にパス優先度を通知することで、ホスト装置１０が、パス優先度を管理でき、以って、適切なパス３０９を選択できるようになる。より具体的には、ホスト装置１０が、パス優先度の問合せを第１ストレージ装置２０Ａに送信し、第１ストレージ装置２０Ａからその問合せの応答を受信し、その応答から、パス３０９Ａの変更後の優先度を知ることができる。ＡＬＵＡでは、そのような問合せは、ＳＣＳＩでサポートされているReportTargetPortGroupsコマンド（以下、ＲＴＰＧコマンド）である。

【００２３】

第１ストレージ装置２０Ａは、ＶＤＥＶ３０５に関連したＲＴＰＧコマンドをホスト装置１０から受信した場合（ステップ１）、第１ＬＤＥＶ３０１Ａとペアを構成する第２ＬＤＥＶ３０１Ｂを有する第２ストレージ装置２０Ｂに、状態変更通知と、第２ＬＤＥＶに関連付けられたパスの状態の要求であるパス状態要求とを送信する（ステップ２）。状態変更通知の一例が、後述のＵＡ同期フラグビットリストである。ＵＡ同期フラグビットリストからは、第１ストレージ装置２０Ａが有するどのＬＤＥＶについてＵＡが変更されたかがわかる。パス状態要求の一例が、第２ストレージ装置２０Ｂが管理するパスの優先度の要求であるパス優先度要求である。

【００２４】

第２ストレージ装置２０Ｂは、ＵＡ同期フラグビットリスト及びパス優先度要求を受信し、ＵＡ同期フラグビットリストを基に、ＵＡを、例えば第１ＬＤＥＶ３０１Ａとペアを構成する第２ＬＤＥＶ３０１Ｂについて設定し、且つ、第２ストレージ装置２０Ｂが保持するパス管理情報を基に、第２ＬＤＥＶ３０１Ｂに関連付いたパス３０９Ｂの優先度を表す情報を含んだパス優先度情報を第１ストレージ装置２０Ａに応答する（ステップ３）。

【００２５】

第１ストレージ装置２０Ａは、パス優先度情報を受信し、第１ＬＤＥＶ３０１Ａに関連付けられ第１パス３０９Ａを含んだ１以上のパスの優先度を表す第１パス優先度と、第２ＬＤＥＶ３０１Ｂに関連付けられ第２パス３０９Ｂを含んだ１以上のパスの状態である第２パス優先度とを含んだ応答（ＲＴＰＧコマンドの応答であるＲＴＰＧ応答）を、ホスト

装置１０に送信する（ステップ４）。つまり、ＲＴＰＧ応答において、２つのストレージ装置２０Ａ及び２０Ｂのパス優先度がマージされている。ホスト装置１０（パス管理プログラム３０２）は、ＲＴＰＧ応答に含まれる第１パス優先度及び第２パス優先度を、ホスト装置１０が保持するパス管理情報に設定し、以後、ＶＤＥＶ３０５に対するＩ／Ｏコマンドを送信する場合、そのパス管理情報が表す第１パス優先度及び第２パス優先度を基に、パス優先度が高い方のパス３０９Ａ又は３０９Ｂを選択する。

【００２６】

以上の処理において、第１ストレージ装置２０Ａから第２ストレージ装置２０Ｂへの状態変更通知（ＵＡ同期フラグビットリスト）が、第１ストレージ装置２０Ａでのパス優先度変更を第２ストレージ装置２０Ｂに通知することの一例である。以上の処理によれば、全てのストレージ装置２０Ａ及び２０Ｂが、それぞれ、各ストレージ装置のパス管理情報を保持し、且つ、全てのストレージ装置間でパス管理情報を同期すること無しに、第１ＬＤＥＶ３０１Ａでのパス優先度変更を第２ストレージ装置２０Ｂに通知することができる。

【００２７】

また、以上の処理によれば、パス優先度変更の通知は、第１ストレージ装置２０Ａでのパス優先度変更の都度に行われるのではなく、ＲＴＰＧコマンドを受信した場合に行われる。ＲＴＰＧコマンドの受信は、変更後のパス優先度をホスト装置１０に将来通知し得ることを意味する第２イベントの検出の一例である。パス優先度変更の都度にパス優先度変更を通知するとなると、第１ストレージ装置２０Ａの負荷が高くなり、結果として、ストレージシステム３０３の性能が低下し得る。以上の処理によれば、そのような性能低下を避けることができる。なお、第２イベントの検出の別の例として、後述するように、Ｉ／Ｏコマンドの受信（パス優先度変更を表すＵＡをホスト装置１０に送信するきっかけ）がある。

【００２８】

なお、本実施例では、ＲＴＰＧに設定される第１パス優先度及び第２パス優先度は、第１ＬＤＥＶ３０１Ａと第２ＬＤＥＶ３０１Ｂとのペアの状態であるペア状態と、第１ＬＤＥＶ３０１Ａと第２ＬＤＥＶ３０１ＢとについてのＩ／Ｏモードとに応じて、第１ストレージ装置２０Ａ（ＲＴＰＧコマンドを受信した方のストレージ装置）により制御される（Ｉ／Ｏモードが、Ｉ／Ｏ状態の一例である）。これにより、第１ストレージ装置２０Ａは、ペア状態及びＩ／Ｏモードに応じた最適パスをホスト装置１０に選択させることができるようになる。

【００２９】

具体的には、ペア状態が「ＰＡＩＲ」でありＩ／Ｏモードが「ミラー」、すなわち、第１ＬＤＥＶ３０１Ａと第２ＬＤＥＶ３０１Ｂが同期状態であることが表されていれば、第１ストレージ装置２０Ａは、ＲＴＰＧ応答に含められる第１パス優先度を、第１ストレージ装置が保持するパス管理情報から特定されるパス優先度とし、ＲＴＰＧ応答に含められる第２パス優先度を、第２ストレージ装置２０Ｂから受信したパス優先度情報が表すパス優先度とする。ペア状態が「ＳＵＳＰＥＮＤ」でありＩ／Ｏモードが「リモート」（第１ＬＤＥＶ３０１Ａ内のデータよりも第２ＬＤＥＶ３０１Ｂ内のデータの方が新しい）場合、第１ストレージ装置２０Ａは、第１パス優先度よりも第２パス優先度を高くする（例えば、第１パス優先度＝Active/non-optimized、第２パス優先度＝Active/optimized）。ペア状態が「ＳＵＳＰＥＮＤ」でありＩ／Ｏモードが「ローカル」（第２ＬＤＥＶ３０１Ｂ内のデータよりも第１ＬＤＥＶ３０１Ａ内のデータの方が新しい）場合、第１ストレージ装置２０Ａは、第２パス優先度よりも第１パス優先度を高くする（例えば、第２パス優先度＝Active/non-optimized、第１パス優先度＝Active/optimized）。これにより、第１ストレージ装置２０Ａは、ペア状態及びＩ／Ｏモードに応じた最適パスをホスト装置１０に選択させることができるようになる。

【００３０】

また、第１ストレージ装置２０Ａは、第２ストレージ装置２０Ｂとの通信が不可能であ

る場合、ペア状態が「ＳＵＳＰＥＮＤ」でありＩ／Ｏモードが「リモート」であれば、ＲＴＰＧ応答としてエラー応答をホスト装置１０に送信し、一方、ペア状態が「ＰＡＩＲ」でありＩ／Ｏモードが「ミラー」、又は、ペア状態が「ＳＵＳＰＥＮＤ」でありＩ／Ｏモードが「ローカル」であれば、第２パス優先度を含まず第１パス優先度をActive/optimizedとしたＲＴＰＧ応答を、ホスト装置１０に送信する。これにより、第１ストレージ装置２０Ａは、ペア状態及びＩ／Ｏモードに加えて第２ストレージ装置２０Ｂとの通信状態に応じた制御をホスト装置１０に実行させることができるようになる。

【００３１】

ホスト装置１０は、定期的にＲＴＰＧコマンドを送信してもよいし、ストレージシステム３０３からＵＡ（Unit Attention）を受信した場合にＲＴＰＧコマンドをストレージシステム３０３に送信してもよい。第１ストレージ装置２０Ａは、ホスト装置１０からＶＤＥＶ３０５（又はＶＤＥＶ３０５の基のＬＤＥＶ３０１Ａ又は３０１Ｂ）を指定したＩ／Ｏコマンドを受信した場合、そのＩ／Ｏコマンドに従いＩ／Ｏを実行しＩ／Ｏ応答をホスト装置１０に送信するが、第１ストレージ装置２０Ａに設定されているＵＡが少なくともパス優先度変更を表す値であれば、Ｉ／Ｏ応答にそのＵＡを設定する。ホスト装置１０は、Ｉ／Ｏ応答内のＵＡがパス優先度変更を表している場合に、Ｉ／Ｏコマンドで指定したＶＤＥＶ３０５（又はＶＤＥＶ３０５に関連付けられているポート）を指定したＲＴＰＧコマンドを少なくとも第１ストレージ装置２０Ａに送信する。つまり、ＵＡを含んだＩ／Ｏ応答の送信は、パス優先度変更をホスト装置１０に気付かせるアクションの１つである。従って、第１ストレージ装置２０Ａは、Ｉ／Ｏコマンドを受信した場合に（ＵＡを含んだＩ／Ｏ応答をホスト装置１０に送信する前に）、パス優先度変更の通知を、ペア相手のストレージ装置２０Ｂ（受信したＩ／Ｏコマンドで指定されているＶＤＥＶの基になっている第１ＬＤＥＶ３０１Ａとペアを構成する第２ＬＤＥＶ３０１Ｂを有するストレージ装置２０Ｂ）に送信する。これにより、パス優先度変更がペア相手のストレージ装置２０Ｂに通知される。

【００３２】

また、本実施例では、第１ストレージ装置２０Ａにおいて、パス優先度変更は、ストレージシステム３０３の管理システム（図示せず）からのパス優先度変更要求（パス優先度の変更の要求）に応答して行われる。パス優先度変更要求は、第１イベントの一例である。パス優先度変更は、ＬＤＥＶ単位よりも大きな単位、例えば、ＴＰＧ単位、又は、ホストグループ単位で行うことができる（ホストグループについては後述）。すなわち、１つのパス優先度変更要求により、そのパス優先度変更要求で指定されたリソースに関連付けられている複数のＬＤＥＶの各々について、パス優先度の変更が可能である。一方、ＲＴＰＧコマンドもＩ／Ｏコマンドも、ＬＤＥＶ単位（ＬＵＮ単位）である。このため、第１ストレージ装置２０Ａは、パス優先度変更の通知（状態変更通知の送信）を、ＬＤＥＶ単位で行ってよい。つまり、第１ストレージ装置２０Ａは、パス優先度変更の単位よりも小さい単位でパス優先度変更の通知を行ってよい。これにより、第１ストレージ装置２０Ａの負荷軽減が期待でき、以って、ストレージシステム３０３の性能向上が期待できる。

【００３３】

以上が、実施例１の概要である。なお、以上の説明では、第１ストレージ装置２０Ａを代表的に例に取っているが、本実施例（及び実施例２）で説明する第１ストレージ装置２０Ａの処理は、他の各ストレージ装置２０も実行可能である。

【００３４】

以下、実施例１を詳細に説明する。

【００３５】

図１は、実施例１に係る計算機システムの構成を示す。

【００３６】

１以上のホスト装置１０と、ストレージ管理サーバ１１が、ストレージシステムを構成する複数のストレージ装置２０Ａ及び２０Ｂと、通信ネットワーク（例えばＳＡＮ（Storage Area Network）１２を介して通信可能に接続されている。ストレージ管理サーバ１

１が、管理システムの一例である。管理システムは、１以上の計算機で構成されてよい。具体的には、例えば、管理計算機が情報を表示する場合（具体的には、管理計算機が自分の表示デバイスに情報を表示する、或いは、管理計算機が表示用情報を遠隔の表示用計算機に送信する場合）、管理計算機が管理システムである。

【００３７】

ホスト装置１０は、通信インターフェイスデバイスと、記憶デバイスと、それらに接続されたプロセッサとを有する。通信インターフェイスデバイスを介してストレージ装置２０等と通信できる。記憶デバイスは、例えばメモリである。プロセッサは、例えばＣＰＵであり、記憶デバイスに記憶されたプログラム（例えば、パス管理プログラム３０２（図３０参照）、アプリケーションプログラム及びオペレーティングシステム）を実行することができる。

【００３８】

ストレージ装置２０Ａ及び２０Ｂのうちストレージ装置２０Ａを例に取り、ストレージ装置の構成を説明する。

【００３９】

ストレージ装置２０Ａは、１以上のＰＤＥＶ群２８Ａと、１以上のＰＤＥＶ群２８Ａに対するＩ／Ｏを制御するストレージコントローラ１２２Ａとを有する。

【００４０】

ＰＤＥＶ群２８Ａは、１以上のＰＤＥＶの集合、例えばＲＡＩＤ（Redundant Array of Independent (or Inexpensive) Disks）グループである。ＰＤＥＶは、不揮発性の物理記憶デバイスを意味し、例えばＨＤＤ（Hard Disk Drive）又はＳＳＤ（Solid State Drive）である。ＰＤＥＶ群２８Ａに基づき、実体的なＬＤＥＶを構築することができる。

【００４１】

ストレージコントローラ１２２Ａは、ホスト装置１０及びストレージ管理サーバ１１と通信するためのＦＥ　Ｉ／Ｆ部（フロントエンドのインターフェイス部）と、１以上のＰＤＥＶ群２８Ａと通信するためのＢＥ　Ｉ／Ｆ部（バックエンドのインターフェイス部）と、メモリ部と、それらに接続された制御部とを有する。ＦＥ　Ｉ／Ｆ部が、１以上のＣＨＡ（チャネルアダプタ）２２Ａにより実現されている。ＢＥ　Ｉ／Ｆ部が、１以上のＤＫＡ（ディスクアダプタ）２６Ａにより実現されている。メモリ部が、ＳＭ（共有メモリ）２１Ａ、ＣＭ（キャッシュメモリ）２４Ａ、１以上のＣＨＡ２２Ａ内の１以上のＬＭ（ローカルメモリ）２２２Ａ、及び、１以上のＤＫＡ２６Ａ内の１以上のＬＭ２６２Ａにより実現されている。制御部が、１以上のＣＨＡ内の１以上のＭＰ（マイクロプロセッサ）２２１Ａ、及び、１以上のＤＫＡ２６Ａ内の１以上のＭＰ２６１Ａにより実現されている。ＳＭ２１Ａ及びＣＭ２４Ａは、それぞれ同一のメモリ上に設けられた領域であってもよいし、異なるメモリであってもよい。ＣＨＡ２２Ａ、ＳＭ２１Ａ、ＣＭ２４Ａ及びＤＫＡ２６Ａが、接続部（例えばバス又はスイッチ）２５を介して通信可能である。

【００４２】

１以上のＣＨＡ２２Ａは、複数のポート２９Ａを有する。ＣＨＡ２２Ａは、ＭＰ２２１Ａ及びＬＭ（例えば揮発メモリ）２２２Ａを有する。ＣＨＡ２２Ａの動作はＭＰ２２１Ａにより制御される。ＣＨＡ２２Ａは、Ｉ／Ｏコマンドをホスト装置１０から受信する。受信したＩ／ＯコマンドがライトコマンドのＣＨＡ２２ＡによりＣＭ２４Ａに一時格納され、ＣＭ２４Ａ上のライト対象のデータが、ＤＫＡ２６ＡによりＬＤＥＶに書き込まれる。受信したＩ／Ｏコマンドがリードコマンドの場合、ＤＫＡ２６Ａにより、リードコマンドに従いＬＤＥＶから読み出されたリード対象のデータＣＭ２４Ａに一時格納され、ＣＨＡ２２Ａにより、ＣＭ２４Ａ上のリード対象のデータがホスト装置１０に送信される。Ｉ／Ｏコマンドは、Ｉ／Ｏ先のＬＤＥＶ（例えばＶＤＥＶ）の番号（例えばＬＵＮ（Logical Unit Number））と、そのＬＤＥＶにおける領域のアドレス（例えばＬＢＡ（Logical Block Address））とを含む。

【００４３】

ＤＫＡ２６Ａは、ＭＰ２６１Ａ及びＬＭ２６２Ａを有する。ＤＫＡ２６Ａの動作はＭＰ２６１Ａにより制御される。ＤＫＡ２６Ａは、ＬＤＥＶに対するＩ／Ｏ（例えば、ＰＤＥＶ群２８Ａに対するデータのＩ／Ｏ）を制御し、その際、ＣＭ２４Ａに対するデータのＩ／Ｏも制御する。

【００４４】

ＣＭ２４Ａは、揮発メモリ及び／又は不揮発メモリであり、ＬＤＥＶに対するＩ／Ｏ対象データが一時的に格納される。

【００４５】

ＳＭ２１Ａは、ポート管理情報等の情報が格納される。ＣＨＡ２２Ａ及びＤＫＡ２６Ａの各々は、適宜、ＳＭ２１Ａ内の情報を参照するか、或いは、ＳＭ２１Ａ内の情報の少なくとも一部を、自分のＬＭ（２２２Ａ又は２６２Ａ）に格納しそのＬＭに格納された情報を参照する。

【００４６】

ストレージ装置２０Ａが行う処理は、ストレージ装置２０Ａのストレージコントローラ１２２Ａにより行なわれる。

【００４７】

図２は、各ストレージ装置２０が有する管理情報の構成を示す。なお、ＰＤＥＶ群２８は、図示の通り、１以上のＰＤＥＶ２８１により構成されている。以下、第１ストレージ装置２０Ａを例に取るが、他の各ストレージ装置についても同様である。

【００４８】

第１ストレージ装置２０Ａが保持する管理情報は、ＡＬＵＡ管理テーブル２１１Ａ、ペア管理テーブル２１２Ａ、ＵＡ管理テーブル２１３Ａ、パス管理テーブル２１４Ａ、ポート管理テーブル２１５Ａ、仮想ボックス管理テーブル２１６Ａ、相対ポートＩＤ管理テーブル２１７Ａ、有効ＩＤ範囲管理テーブル２１８Ａ及びＵＡ同期フラグビットリスト２２３Ａを含む。これらのテーブル２１１Ａ～２１８Ａ及びリスト２２３Ａのうちの２以上が１つにマージされてもよいし、これらのテーブル２１１Ａ～２１８Ａ及びリスト２２３Ａのうちの少なくとも１つの各々が２以上に分割されていてもよい。また、本実施例では、テーブル２１１Ａ～２１８Ａは、ＳＭ２１Ａに格納され、ＵＡ同期フラグビットリスト２２３Ａは、ＣＨＡ２２ＡのＬＭ２２２Ａに格納されるが、リスト２２３ＡがＬＭ２２２Ａに代えて又は加えて、ＳＭ２１Ａに格納されてもよいし、テーブル２１１Ａ～２１８Ａのうちの少なくとも１つが、ＳＭ２１Ａに代えて又は加えてＬＭ２２２Ａに格納されてもよい。以下、図３～図１１を参照して、テーブル２１１Ａ～２１８Ａ及びリスト２２３Ａを説明する。

【００４９】

図３は、ＡＬＵＡ管理テーブル２１１Ａの構成を示す。

【００５０】

ＡＬＵＡ管理テーブル２１１Ａは、第１ストレージ装置２０Ａが有するＬＤＥＶ毎にＡＬＵＡが適用されているか否かを表す情報を有する。具体的には、ＡＬＵＡ管理テーブル２１１Ａは、第１ストレージ装置２０Ａが有するＬＤＥＶ毎に、以下の情報、すなわち、ＬＤＥＶ番号（ＬＤＥＶのＬＤＥＶ番号）４０１と、ＡＬＵＡ適用状態４０２（ＡＬＵＡが適用されているか否か）とを有する。例えば、ＡＬＵＡ適用状態「０」は、ＡＬＵＡ適用対象外を意味し、ＡＬＵＡ適用状態「１」は、ＡＬＵＡ適用対象を意味する。

【００５１】

なお、「ＬＤＥＶ番号」とは、ＬＤＥＶの識別番号でありホスト装置１０には認識されずストレージ装置２０において使用される番号である。ＬＤＥＶには、ＬＤＥＶ番号に加えて、後述のＬＵＮ（Logical Unit Number）が関連付けられる。ＬＵＮは、ホスト装置１０に認識され、Ｉ／Ｏコマンドに含まれる（Ｉ／Ｏコマンドで指定される）。本実施例では、同一のＬＵＮが、ＶＤＥＶの基であるＬＤＥＶ３０１Ａ及び３０１Ｂに関連付けられ、それ故、ホスト装置１０は、ＶＤＥＶの番号としてＬＵＮを認識できる。各ストレージ装置２０は、Ｉ／Ｏコマンドに含まれているＬＵＮに対応したＬＤＥＶ番号を特定す

ることで、Ｉ／Ｏ先のＬＤＥＶを特定できる。

【００５２】

図４は、ペア管理テーブル２１２Ａの構成を示す。

【００５３】

ペア管理テーブル２１２Ａは、第１ストレージ装置２０Ａが有しＶＤＥＶの基になっているＬＤＥＶ毎に、ペア相手のＬＤＥＶ、ペア状態、Ｉ／Ｏモード、及び、ＡＬＵＡ適用状態をペア相手のストレージ装置に通知するか否かを表す情報、を有する。具体的には、ペア管理テーブル２１２Ａは、第１ストレージ装置２０Ａが有しＶＤＥＶの基になっているＬＤＥＶ毎に、以下の情報、なわち、ＬＤＥＶ番号（自ストレージ）５０１と、ＬＤＥＶ番号（ペア相手ストレージ）５０２と、ペア状態５０３と、Ｉ／Ｏモード５０４と、ＡＬＵＡ同期ビット５０５とを有する。

【００５４】

ＬＤＥＶ番号（自ストレージ）５０１は、第１ストレージ装置２０Ａが有するＬＤＥＶ、例えば第１ＬＤＥＶ３０１ＡのＬＤＥＶ番号である。ＬＤＥＶ番号（ペア相手ストレージ）５０２は、第１ストレージ装置２０Ａが有するＬＤＥＶとペアを構成するＬＤＥＶ（他ストレージ装置が有するＬＤＥＶ）の番号、例えば第２ストレージ装置２０Ｂが有する第２ＬＤＥＶ３０１ＢのＬＤＥＶ番号である。

【００５５】

ペア状態５０３は、ＬＤＥＶペアの状態を表し、例えば「ＰＡＩＲ」（一方のＬＤＥＶにデータが格納されると他方のＬＤＥＶにそのデータがコピーされる状態）、「ＳＵＳＰＥＮＤ」（一方のＬＤＥＶにデータが格納されても他方のＬＤＥＶにそのデータがコピーされない状態）等がある。

【００５６】

Ｉ／Ｏモード５０４は、ＬＤＥＶペア（言い換えればＶＤＥＶ）に対するＩ／Ｏのモード（状態）を表し、例えば「Mirror」（ＬＤＥＶペアを構成するＬＤＥＶ間が同期）、「Local」（自ストレージ内のＬＤＥＶ内のデータがペア相手ストレージ内のＬＤＥＶよりも新しく故に自ストレージ内のＬＤＥＶへのＩ／Ｏを許可）、「Remote」（ペア相手ストレージ内のＬＤＥＶ内のデータが自ストレージ内のＬＤＥＶよりも新しく故に自ストレージ内のＬＤＥＶへのＩ／Ｏを禁止しペア相手ストレージ内のＬＤＥＶへのＩ／Ｏを許可）、及び「Block」（エラーをホスト装置へ応答）等がある。

【００５７】

ＡＬＵＡ同期ビット５０５は、ＡＬＵＡ適用状態をペア相手のストレージ装置に通知するか否かを表す。例えば、ＡＬＵＡ同期ビット「０」は、ペア相手ＬＤＥＶのＡＬＵＡ適用状態の変更が不要を意味し、ＡＬＵＡ同期ビット「１」は、ペア相手ＬＤＥＶのＡＬＵＡ適用状態の変更が必要を意味する。

【００５８】

図５は、ＵＡ管理テーブル２１３Ａの構成を示す。

【００５９】

ＵＡ管理テーブル２１３Ａは、ＵＡ管理単位毎にＵＡを有する。ＵＡ管理単位は、ＵＡが表す状態変更についてのリソース種類によって異なってよい。本実施例では、ＵＡはパス優先度変更の有無を表し、ＵＡ管理単位は、ポート及びホストグループ単位である。具体的には、ＵＡ管理テーブル２１３Ａは、ポート２９Ａ毎に、以下の情報、すなわち、ポート番号６０１、ホストグループ番号６０２及びＵＡ６０３を有する。

【００６０】

ポート番号６０１は、ポート２９Ａの識別番号である。ホストグループ番号６０２は、ポート２９Ａに関連付けられたホストグループの番号である。なお、「ホストグループ」とは、１以上のホスト装置１０の集合であり、ストレージ装置２０のポートに関連付けられるリソースの一種である。ポート２９Ａを経由したＩ／Ｏは、Ｉ／Ｏ元のホスト装置１０が、そのポート２９Ａに関連付けられたホストグループに含まれるホスト装置１０の場合に、第１ストレージ装置２０Ａにより許可される。

【００６１】
ＵＡ６０３は、ＵＡの値を表し、例えば、ＵＡ「０」は、ＵＡ設定無しを意味し、ＵＡ「１」は、パス優先度変更を意味し、ＵＡ「２」は、ＡＬＵＡ適用状態変更を意味する。
【００６２】
図６は、ＵＡ同期フラグビットリスト２２３Ａの構成を示す。
【００６３】
ＵＡ同期フラグビットリスト２２３Ａは、第１ストレージ装置２０Ａが有するＬＤＥＶ毎に、ペア相手ストレージにパス優先度変更のＵＡを通知するため自ストレージのＬＤＥＶについてパス優先度変更のＵＡ（例えばＵＡ「１」）が設定されたか否かを表す。具体的には、ＵＡ同期フラグビットリスト２２３Ａは、第１ストレージ装置２０Ａが有するＬＤＥＶ毎に、以下の情報、すなわち、ＬＤＥＶ番号７０１及びＵＡ同期フラグビット７０２を有する。ＬＤＥＶ番号７０１は、第１ストレージ装置２０Ａが有するＬＤＥＶの番号である。ＵＡ同期フラグビット７０２は、ＬＤＥＶについてパス優先度変更のＵＡが設定されたか否かを表す。例えば、ＵＡ同期フラグ「０」は、パス優先度変更無しを意味し、ＵＡ同期フラグ「１」は、パス優先度変更を意味する。
【００６４】
ＵＡ同期フラグの変更は、ホスト装置１０からのＩ／Ｏコマンドの受付を止めることなく行われ、且つ、状態変更通知の送信（ＵＡ同期フラグビットリストの送信）は、ＵＡ同期フラグの変更とは非同期に行われる。これにより、ストレージシステムの性能低下を回避できる。
【００６５】
図７は、パス管理テーブル２１４Ａの構成を示す。
【００６６】
パス管理テーブル２１４Ａは、パス管理情報の一例であり、パス管理単位毎のパス優先度を表す。パス管理単位は、本実施例ではホストグループ単位である。つまり、本実施例では、ホストグループ単位で、パス優先度が設定及び変更される。具体的には、パス管理テーブル２１４Ａは、ホストグループ毎に、以下の情報、すなわち、仮想ボックス番号８０１、ＨＧ番号８０２、相対ポートＩＤ８０３及びパス優先度８０４を有する。
【００６７】
仮想ボックス番号８０１は、仮想ボックスの番号であり、仮想ボックスは、仮想的なストレージシステムに相当する。ＨＧ番号８０２は、ホストグループの番号である。
【００６８】
相対ポートＩＤ８０３は、ポート２９Ａのポート番号と異なるＩＤであり、１つの仮想ボックス（仮想ストレージシステム）（異なるストレージ装置２０Ａ、２０Ｂ）においてポート番号が重複しないよう割り振られたＩＤである。相対ポートは、仮想ポートと呼ぶこともできる。
【００６９】
パス優先度８０４の具体的な値として、例えば、「Active/optimized」（パス優先度「高」でありＩ／Ｏ実行可能）、「Active/non-optimized」（パス優先度「低」であり待機状態）、「Unavailable」（Ｉ／Ｏ実行不可能）等がある。
【００７０】
なお、本実施例では、以下方針でパス優先度を管理することができる。すなわち、（１）自ストレージが管理している、自ストレージ内のホストグループのパス優先度は、常に正しい情報である。（２）ペア相手ストレージが管理している、自ストレージ内のホストグループのパス優先度は、常に正しい情報である必要はない。
【００７１】
また、本実施例において、同一ストレージ装置内で、ポート番号は重複できない。しかし、仮想ボックス番号が異なれば、同一ストレージ装置内で、相対ポートＩＤは重複してよい。また、同一仮想ボックス（仮想ストレージ）内で、ストレージ装置が異なれば、ポート番号は重複してよい。また、同一仮想ボックス（仮想ストレージ）内で、相対ポート

ＩＤは重複できない。
【００７２】
図８は、ポート管理テーブル２１５Ａの構成を示す。
【００７３】
ポート管理テーブル２１５Ａは、各ポート２９Ａに関連付けられている情報を表す。具体的には、ポート管理テーブル２１５Ａは、ポート２９Ａ毎に、以下の情報、すなわち、ポート番号１００１、インデックス番号１００２、ＬＵＮ１００３、ＬＤＥＶ番号１００４及びＨＧ番号１００５を有する。ポート番号１００１は、ポート２９Ａの番号である。インデックス番号１００２は、ポート２９Ａに関連付け可能なエントリ（ＬＤＥＶ）の通し番号である。ＬＵＮ１００３は、ポート２９Ａに関連付けられたＬＤＥＶに関連付けられているＬＵＮ（本実施例ではＶＤＥＶの番号に相当）である。ＬＤＥＶ番号１００４は、ポート２９Ａに関連付けられたＬＤＥＶの番号である。ＨＧ番号１００５は、ポート２９Ａに関連付けられたホストグループの番号である。
【００７４】
図９は、仮想ボックス管理テーブル２１６Ａの構成を示す。
【００７５】
仮想ボックス管理テーブル２１６Ａは、仮想ボックス毎に、第１ストレージ装置２０Ａが管理するリソースを表す。具体的には、仮想ボックス管理テーブル２１６Ａは、仮想ボックス毎に、以下の情報、すなわち、仮想ボックス番号１１０１、ＨＧ番号１１０２、ポート番号１００３及びＬＤＥＶ番号１１０４を有する。仮想ボックス番号１１０１は、仮想ボックスの番号である。ＨＧ番号１１０２は、仮想ボックス内のポート２９Ａに関連付けられたホストグループの番号である。ポート番号１００３は、仮想ボックス内のポート２９Ａの番号である。ＬＤＥＶ番号１１０４は、仮想ボックス内のＬＤＥＶ（仮想ボックス内のポート２９Ａに関連付けられたＬＤＥＶ）の番号である。
【００７６】
図１０は、相対ポートＩＤ管理テーブル２１７Ａの構成を示す。
【００７７】
相対ポートＩＤ管理テーブル２１７Ａは、ホストグループ毎に、ホストグループに関連付けられたポート２９Ａの相対ポートＩＤを表す。具体的には、相対ポートＩＤ管理テーブル２１７Ａは、ホストグループ毎に、以下の情報、すなわち、ＨＧ番号１２０１及び相対ポートＩＤ１２０２を有する。ＨＧ番号１２０１は、ホストグループの番号である。相対ポートＩＤ１２０２は、ホストグループが関連付けられているポート２９Ａに割り振られた相対ポートＩＤである。
【００７８】
図８～図１０に示したテーブル２１５Ａ～２１７Ａを参照することで、ポート番号、ＬＵＮ、ＬＤＥＶ番号及びＨＧ番号の対応関係を把握することができる。
【００７９】
図１１は、有効ＩＤ範囲管理テーブル２１８Ａの構成を示す。
【００８０】
有効ＩＤ範囲管理テーブル２１８Ａは、仮想ボックス毎に、付与可能な相対ポートＩＤの値（番号）の範囲を表す。具体的には、有効ＩＤ範囲管理テーブル２１８Ａは、仮想ボックス毎に、仮想ボックス番号１３０１、開始ＩＤ１３０２及び終了ＩＤ１３０３を有する。仮想ボックス番号１３０１は、仮想ボックスの番号である。開始ＩＤ１３０２は、付与可能な相対ポートＩＤ範囲のうちの最初のＩＤ（値）である。終了ＩＤ１３０３は、付与可能な相対ポートＩＤ範囲のうちの最後のＩＤ（値）である。
【００８１】
以上が、テーブル２１１Ａ～２１８Ａ及びリスト２２３Ａの説明である。なお、テーブル２１１Ａ～２１８Ａ及びリスト２２３Ａのうちの少なくとも１つにおいて、少なくとも１種類の情報に代えて又は加えて、他種の情報が登録されてもよい。例えば、ペア管理テーブル２１２Ａに、ペア相手ストレージ装置のＩＤ等が登録されてもよい。

【００８２】
図１２は、ホスト装置１０がストレージ装置２０のパス優先度を把握するまでの流れを示す。
【００８３】
ホスト装置１０が、第１ＬＤＥＶ３０１ＡのＬＵＮを含んだＩ／Ｏコマンドを、パス３０９Ａ経由で第１ストレージ装置２０Ａに送信する。第１ストレージ装置２０Ａは、そのＩ／Ｏコマンドを受信した場合（Ｓ１２０１）、そのＩ／ＯコマンドのＬＵＮから特定される第１ＬＤＥＶ３０１AについてのＵＡ管理テーブル２１３Ａ（図５参照）が「０」であれば、そのＩ／Ｏコマンドに従うＩ／Ｏを行い、Ｇｏｏｄ応答をホスト装置１０に送信する（Ｓ１２０２）。なぜなら、第１ＬＤＥＶ３０１AについてＵＡ設定無しのためである。
【００８４】
その後、第１ＬＤＥＶ３０１Ａが関連付いているホストグループについてパス優先度が変更され、それに伴い、第１ストレージ装置２０Ａにより、そのホストグループについてＵＡ「１」がＵＡ管理テーブル２１３Ａ（図５参照）に設定され、且つ、第１ＬＤＥＶ３０１AについてＵＡ同期フラグ「１」がＵＡ同期フラグビットリスト２２３Ａ（図６参照）に設定されたとする（Ｓ１２０３）。
【００８５】
その後、再び、第１ストレージ装置２０Ａが、第１ＬＤＥＶ３０１ＡのＬＵＮを含んだＩ／Ｏコマンドを、パス３０９Ａ経由で受信したとする（Ｓ１２０４）。この場合、第１ＬＤＥＶ３０１Ａが関連付いているホストグループのＵＡが「１」なので（正確にはＵＡが「０」以外なので）、第１ストレージ装置２０Ａが、そのＩ／Ｏコマンドに従うＩ／Ｏを行った後、そのＵＡ「１」を含んだ応答をホスト装置１０に送信する（Ｓ１２０５）。
【００８６】
ホスト装置１０は、ＵＡ「１」を含んだ応答を受信した場合に、Ｓ１２０４のＩ／Ｏコマンド内のＬＵＮについてパス優先度の変更があったことに気付く。ホスト装置１０は、Ｓ１２０４のＩ／Ｏコマンド内のＬＵＮと同一のＬＵＮを含んだＡＬＵＡ適用状態問合せを第１ストレージ装置２０Ａに送信する。第１ストレージ装置２０Ａは、ＬＵＮを含んだＡＬＵＡ適用状態問合せを受信した場合（Ｓ１２０６）、そのＬＵＮに対応した第１ＬＤＥＶ３０１AについてのＡＬＵＡ適用状態４０２（図３参照）が表すＡＬＵＡ適用状態を含んだ応答をホスト装置１０に送信する（Ｓ１２０７）。
【００８７】
ホスト装置１０は、その応答からＡＬＵＡ適用であることを特定した場合、Ｓ１２０４のＩ／Ｏコマンド内のＬＵＮと同一のＬＵＮを含んだポート番号問合せを第１ストレージ装置２０Ａに送信する。第１ストレージ装置２０Ａは、ＬＵＮを含んだポート番号問合せを受信した場合（Ｓ１２０８）、そのＬＵＮに対応したポート番号をポート管理テーブル２１５Ａから特定し、特定したポート番号を含んだ応答をホスト装置１０に送信する（Ｓ１２０９）。
【００８８】
ホスト装置１０は、Ｓ１２０４のＩ／Ｏコマンド内のＬＵＮと同一のＬＵＮを含んだＲＴＰＧ（ReportTargetPortGroups）コマンドを送信する。第１ストレージ装置２０Ａは、ＲＴＰＧコマンドを受信した場合（Ｓ１２１０）、図３０を参照して説明した処理を行うことで、例えば図１３に示すように、異なるストレージ装置２０Ａ及び２０Ｂの異なるパス管理テーブル２１４Ａ及び２１４Ｂの該当部分（図１３の太枠内の情報の全て）がＲＴＰＧ応答にマージされる。具体的には、第１ストレージ装置２０Ａは、パス管理テーブル２１４Ａ内の相対ポートＩＤ及びパス優先度の組と、パス管理テーブル２１４Ｂ内の相対ポートＩＤ及びパス優先度の組とを、ＲＴＰＧ応答に設定する。第１ストレージ装置２０Ａは、ＲＴＰＧコマンド内のＬＵＮに対応したポート番号をポート管理テーブル２１５Ａ、相対ポートID管理テーブル、仮想ボックス管理テーブルから特定し、特定したポート番号を、ＲＰＧ応答に設定してもよい。第１ストレージ装置２０Ａは、そのＲＴＰＧ応答を

ホスト装置１０に送信する（Ｓ１２１１）。これにより、ホスト装置１０は、変更後のパス優先度を知ることができる。

【００８９】

以下、本実施例で行われる処理を説明する。なお、以下、説明を分かり易くするため、コマンド受信ストレージ装置を、第１ストレージ装置２０Ａとし、ペア相手ストレージを、第２ストレージ装置２０Ｂとする。コマンド受信ストレージ装置とは、ストレージ管理サーバ１１又はホスト装置１０からコマンド（Ｉ／Ｏコマンド又は何らかの問合せ）を受信するストレージ装置である。ペア相手ストレージ装置とは、コマンド受信ストレージ装置が受信したコマンドに関わるＬＤＥＶとペアを構成するＬＤＥＶを有するストレージ装置である。

【００９０】

図１４は、ＡＬＵＡ適用状態変更処理の流れを示す。

【００９１】

第１ストレージ装置２０Ａは、ストレージ管理サーバ１１からＡＬＵＡ適用状態変更コマンドを受信する（Ｓ１４０１）。ＡＬＵＡ適用状態変更コマンドは、ＨＧ番号を含む。以下、そのＨＧ番号を、図１４の説明において「対象ＨＧ番号」と言う。

【００９２】

第１ストレージ装置２０Ａは、対象ＨＧ番号に対応するＬＤＥＶを、ポート管理テーブル２１５Ａから特定する（Ｓ１４０２）。第１ストレージ装置２０Ａは、特定したＬＤＥＶのＡＬＵＡ適用状態４０２（図３参照）を、Ｓ１４０１で受信したコマンドに従い変更する（Ｓ１４０３）。また、第１ストレージ装置２０Ａは、対象ＨＧ番号に対応するＵＡ６０３（図５参照）を、「２」（ＡＬＵＡ適用状態の変更を意味する値）に変更する（Ｓ１４０４）。また、第１ストレージ装置２０Ａは、Ｓ１４０２で特定したＬＤＥＶに対応したＡＬＵＡ同期ビット５０５（図４参照）を「１」（ＡＬＵＡ適用状態の変更を意味する値）に変更する（Ｓ１４０５）。そして、第１ストレージ装置２０Ａは、Ｓ１４０１で受信したコマンドの応答として完了応答をストレージ管理サーバ１１に送信する（Ｓ１４０６）。

【００９３】

Ｓ１４０５と非同期に、第１ストレージ装置２０Ａは、ＡＬＵＡ同期ビット５０５が「１」のＬＤＥＶの番号を含んだ変更通知を、第２ストレージ装置２０Ｂに送信する。その変更通知は、第１ストレージ装置２０ＡのＬＤＥＶの番号（又はＬＵＮ）、又は、そのＬＤＥＶのペア相手のＬＤＥＶの番号を含む。第２ストレージ装置２０Ｂは、その変更通知が表すＬＤＥＶに対応したＡＬＵＡ適用状態（ＡＬＵＡ管理テーブル２１１ＡのＡＬＵＡ適用状態）を「１」に変更する（Ｓ１４０７）。また、第２ストレージ装置２０Ｂは、その変更通知内の対象ＨＧ番号に対応したＵＡ（ＵＡ管理テーブル２１３ＢのＵＡ６０３）を「２」に変更する（Ｓ１４０８）。第２ストレージ装置２０Ｂは、Ｓ１４０８と非同期に、その変更通知に対して完了通知をＳ１４０８の完了後に第１ストレージ装置１０Ａに送信する。完了通知は、ＡＬＵＡ適用状態の変更が完了したＬＤＥＶの番号（第２ストレージ装置２０ＢのＬＤＥＶの番号）、又は、そのＬＤＥＶとペア相手のＬＤＥＶの番号（第１ストレージ装置２０ＡのＬＤＥＶの番号）を含んでよい。

【００９４】

第１ストレージ装置２０Ａは、完了通知を受信した場合、すなわち、第２ストレージ装置２０ＢにおいてＡＬＵＡ適用状態の変更が完了した場合、変更完了したＬＤＥＶに対応したＡＬＵＡ同期ビット５０５（図４参照）を「０」に変更する（Ｓ１４０９）。

【００９５】

なお、ＶＤＥＶの基になるＬＤＥＶ間でペアが構成されていない場合、そのＶＤＥＶの基になる各ＬＤＥＶについてＡＬＵＡ同期ビットを「１」に変更する必要は無い。

【００９６】

図１５は、ＡＬＵＡ適用状態確認処理の流れを示す。

【００９７】

第１ストレージ装置２０Ａは、ストレージ管理サーバ１１からＡＬＵＡ適用状態問合せを受信する（Ｓ１５０１）。ＡＬＵＡ適用状態問合せは、ＨＧ番号を含む。以下、そのＨＧ番号を、図１５の説明において「対象ＨＧ番号」と言う。

【００９８】

第１ストレージ装置２０Ａは、対象ＨＧ番号に対応するＬＤＥＶを、ポート管理テーブル２１５Ａから特定する（Ｓ１５０２）。第１ストレージ装置２０Ａは、特定したＬＤＥＶのＡＬＵＡ適用状態４０２（図３参照）を、Ｓ１５０１で受信した問合せに従い特定する（Ｓ１５０３）。

【００９９】

第１ストレージ装置２０Ａは、特定したＡＬＵＡ適用状態を含んだ応答をストレージ管理サーバ１１に送信する（Ｓ１５０４）。

【０１００】

なお、Ｓ１５０４で送信された応答に含まれているＡＬＵＡ適用状態は、Ｓ１５０２で特定されたＬＤＥＶとペアを構成するＬＤＥＶのＡＬＵＡ適用状態でもある。なぜなら、図１４の処理により、ペア相手ＬＤＥＶのＡＬＵＡ適用状態も同じ状態に変更されているからである。従って、管理者（ストレージ管理サーバ１１を使用するユーザ）は、ＬＤＥＶペアを構成する２つのＬＤＥＶのうちの一方のＬＤＥＶのＡＬＵＡ適用状態を確認すればよい。

【０１０１】

図１６は、パス優先度変更処理の流れを示す。

【０１０２】

第１ストレージ装置２０Ａは、ストレージ管理サーバ１１からパス優先度変更コマンドを受信する（Ｓ１６０１）。パス優先度変更コマンドは、ＨＧ番号を含む。以下、そのＨＧ番号を、図１６の説明において「対象ＨＧ番号」と言う。

【０１０３】

第１ストレージ装置２０Ａは、対象ＨＧ番号に対応するＬＤＥＶを、ポート管理テーブル２１５Ａから特定する（Ｓ１６０２）。第１ストレージ装置２０Ａは、特定したＬＤＥＶのＵＡ同期フラグビット７０２（図６参照）を、「１」（パス優先度変更を意味する値）に変更する（Ｓ１６０３）。第１ストレージ装置２０Ａは、対象ＨＧ番号に対応するパス優先度８０４（図７参照）を、Ｓ１６０１で受信したコマンドに従い変更する（Ｓ１６０４）。第１ストレージ装置２０Ａは、対象ＨＧ番号に対応したＵＡ６０３（図５参照）を「１」（パス優先度変更を意味する値）に変更する（Ｓ１６０５）。

【０１０４】

第１ストレージ装置２０Ａは、完了を表す応答をストレージ管理サーバ１１に送信する（Ｓ１６０６）。

【０１０５】

なお、上述したように、各ストレージ装置２０は、パス優先度変更を意味する値のＵＡが存在する場合にそのＵＡに対応したＬＤＥＶをＩ／Ｏ先としたＩ／Ｏコマンドを受信すると、ＵＡを含んだＩ／Ｏ応答（Ｉ／Ｏコマンドの応答）を送信することになる。しかし、パス優先度変更処理の最中に、第２ストレージ装置２０Ｂ（ペア相手ストレージ装置）が、そのようなＩ／Ｏコマンドを受信しても、ＵＡを含んだＩ／Ｏ応答を送信しない。なぜなら、第２ストレージ装置２０Ｂは未だＵＡ変更を知らないからである。しかし、それは、以下の理由により問題ではない。

（１）パス優先度変更後にＵＡ変更が第２ストレージ装置２０Ｂに通知されていないということは、パス優先度変更後に第１ストレージ装置２０Ａが、ＵＡ変更に関わるＬＤＥＶを指定したＩ／Ｏコマンドを受信していないということ。従って、その時点では、ＵＡ変更に関わるＬＤＥＶペアを構成する２つのＬＤＥＶ内のデータは同じである。このため、どちらのＬＤＥＶからデータが読み出されてもよい。

（２）パス優先度変更後に第１ストレージ装置２０Ａは、Ｉ／Ｏコマンドを受信すると、ＵＡを含んだＩ／Ｏ応答がホスト装置１０に送信する。それにより、ホスト装置１０は、

パス優先度の変更があったことを知ることでき、どこのパス優先度が変更されたかをＲＴＰＧコマンドで確認できる。
（なお、ＵＡ変更通知（伝播）前に第２ストレージ装置２０Ｂがライトコマンドを受信しそれにより第２ストレージ装置２０Ｂ内のＬＤＥＶ（ペア相手のＬＤＥＶ）内のデータが新しくなったとしても、それは、以下の理由により問題ではない。すなわち、同期コピーのため、ＵＡ設定無しのＬＤＥＶを指定したライトコマンドを第２ストレージ装置２０Ｂが受信した場合、そのＵＡ設定無しのＬＤＥＶとペアのＬＤＥＶ（ＵＡ設定有のＬＤＥＶ）を有する第１ストレージ装置２０Ａに対して、ストレージ装置20Bからライトデータが伝播される。ストレージ装置20Aはストレージ装置20Bからのライトデータ伝播を契機に、図１７のＵＡ伝播を実施する。それにより、ＵＡ設定無のＬＤＥＶ（第２ストレージ装置２０Ｂ）にもＵＡが伝播され、第２ストレージ装置２０ＢがホストにＵＡを応答することができる。この場合、図１７におけるＳ１７０１（ホスト装置からのＩ／Ｏコマンド受信）を、「第２ストレージ装置からのコピーコマンド受信」に読み替えることができる。）
【０１０６】
図１７は、Ｉ／Ｏ処理の際のＵＡ伝播の流れを示す。
【０１０７】
第１ストレージ装置２０Ａは、ホスト装置１０からＩ／Ｏコマンドを受信する（Ｓ１７０１）。Ｉ／ＯコマンドはＬＵＮを含んでいる。
【０１０８】
第１ストレージ装置２０Ａは、ＵＡ同期フラグビット７０２が「１」であるＬＤＥＶが存在するか否かを、ＵＡ同期フラグビットリスト２２３Ａ（図６参照）を基に判断する（Ｓ１７０２）。
【０１０９】
Ｓ１７０２の判断結果が否定の場合（Ｓ１７０２：Ｎｏ）、第１ストレージ装置２０Ａは、ホスト装置１０にＧｏｏｄ応答（Ｇｏｏｄを表すＩ／Ｏ応答）を送信する（Ｓ１７０３）。
【０１１０】
一方、Ｓ１７０２の判断結果が肯定の場合（Ｓ１７０２：Ｙｅｓ）、第１ストレージ装置２０Ａは、ペア管理テーブル２１２Ａを基に、ＵＡ同期フラグビット「１」のＬＤＥＶのペア相手ＬＤＥＶを特定する（Ｓ１７０４）。なお、Ｉ／Ｏ先ＬＤＥＶ（Ｉ／Ｏコマンド内のＬＵＮに対応したＬＤＥＶ）のＵＡ同期フラグビットが「０」であっても、ＵＡ同期フラグビットリスト２２３Ａ上の少なくとも１つのＵＡ同期フラグビットが「１」であれば、Ｓ１７０２の判断結果は肯定である。また、Ｓ１７０４では、ＵＡ同期フラグが「１」である１以上のＬＤＥＶについて１以上のペア相手ＬＤＥＶが特定され、特定された１以上のペア相手ＬＤＥＶを有する１以上のストレージ装置の各々について、後述のＳ１７０５～Ｓ１７０７が行われる。ここでは、説明を分かり易くするため、ＵＡ同期フラグが「１」であるＬＤＥＶは１つであり、故に、ペア相手ＬＤＥＶも１つであり、ペア相手ストレージ装置は第２ストレージ装置２０Ｂであるとする。
【０１１１】
第１ストレージ装置２０Ａは、Ｓ１７０４で特定したペア相手ＬＤＥＶを有する第２ストレージ装置２０Ｂに、ＵＡ同期フラグビットリスト２２３Ａを送信する（Ｓ１７０５）。つまり、第１ストレージ装置２０Ａは、第１ストレージ装置２０ＡのＵＡ変更を第２ストレージ装置２０Ｂに通知する。なお、送信されるＵＡ同期フラグビットリスト２２３Ａは、Ｉ／Ｏコマンドを受信したＣＨＡ２２Ａ内のＬＭ２２２Ａ上のリスト２２３Ａである。リスト２２３Ａの中身は、全てのＣＨＡ２２Ａで同じでもよい。或いは、アクセス可能なＬＤＥＶ（受信するＩ／ＯコマンドでＩ／Ｏ先として指定され得るＬＤＥＶ）がＣＨＡ２２Ａによって異なっていれば、ＣＨＡ２２Ａによって、リスト２２３Ａの中身が違っていてもよい。
【０１１２】
第２ストレージ装置２０Ｂは、ＵＡ同期フラグビットリスト２２３Ａを受信し、受信し

たリスト２２３ＡからＵＡ同期フラグビットが「１」のＬＤＥＶを特定し、特定したＬＤＥＶをペア相手ＬＤＥＶとするＬＤＥＶをペア管理テーブル２１２Ｂを基に特定し、特定したＬＤＥＶに関連付いているＨＧ番号をポート管理テーブル２１５Ｂを基に特定し、特定したＨＧ番号に対応するＵＡ（ＵＡ管理テーブル２１３ＢにおけるＵＡ）を「１」（パス優先度変更を意味する値）に変更する（Ｓ１７０６）。これにより、第１ストレージ装置２０ＡのＵＡ変更が第２ストレージ装置２０Ｂに通知された（通知完了）となる。第２ストレージ装置２０Ｂは、完了を第１ストレージ装置２０Ａに通知する。

【０１１３】

第１ストレージ装置２０Ａは、完了の通知を第２ストレージ装置２０Ｂから受信した場合、ＵＡ同期フラグビットリスト２２３Ａにおける全てのＵＡ同期フラグビット「１」を「０」に変更する（Ｓ１７０７）。そして、第１ストレージ装置２０Ａは、ＵＡ応答（ＵＡ＝１を含んだＩ／Ｏ応答）をホスト装置１０に送信する（Ｓ１７０８）。

【０１１４】

図１７のＩ／Ｏ処理によれば、Ｉ／Ｏ先ＬＤＥＶのＵＡ同期フラグビットが「０」であっても、少なくとも１つの他のＬＤＥＶのＵＡ同期フラグビットが「１」でありさえすれば、ホスト装置１０にＵＡ応答が送信される。しかし、必ずしもそうである必要は無く、例えば、下記変形例が考えられる。

（変形例１）Ｓ１７０２において、第１ストレージ装置１０は、Ｉ／Ｏ先ＬＤＥＶのＵＡ同期フラグビットが「１」であるか否かを判断する。Ｉ／Ｏ先ＬＤＥＶのＵＡ同期フラグビットが「０」の場合、少なくとも１つの他のＬＤＥＶのＵＡ同期フラグビットが「１」であっても、Ｓ１７０２の判断結果が否定であり、Ｉ／Ｏ先ＬＤＥＶのＵＡ同期フラグビットが「１」の場合、Ｓ１７０２の判断結果が肯定である。この変形例１では、Ｓ１７０２の判断結果が肯定の場合、Ｓ１７０５で送信されるＵＡ同期フラグビットは、Ｉ／Ｏ先ＬＤＥＶに対応したＵＡ同期フラグビットのみでよい。Ｓ１７０７で「０」に変更されるＵＡ同期フラグビットは、Ｉ／Ｏ先ＬＤＥＶに対応したビットのみでよい。

（変形例２）Ｓ１７０２の判断結果が肯定の場合、Ｓ１７０５で送信されるＵＡ同期フラグビットは、ＵＡ同期フラグビットリスト２２３Ａのうち、Ｉ／Ｏコマンドの送信元ホスト装置を含んだＨＧに関連付いたＬＤＥＶに対応したＵＡ同期フラグビットのみである。

【０１１５】

図１８Ａ及び図１８Ｂは、ペア状態とパス優先度に関連性があることの一例を示す。具体的には、図１８Ａは、ペア状態「ＰＡＩＲ」の場合の適切パスの一例を示し、図１８Ｂは、ペア状態「ＳＵＳＰＥＮＤ」の場合の適切パスの一例を示す。

【０１１６】

図１８Ａに示すように、ペア状態「ＰＡＩＲ」（且つＩ／Ｏモード「ミラー」）の場合、ホスト装置１０は、パス管理テーブルのパス優先度に従い、パス３０９Ａを使用する、すなわち、第１ストレージ装置２０ＡのＬＤＥＶに対してＩ／Ｏを行うのが望ましい。

【０１１７】

しかし、ペア状態が、例えば図１８Ｂに示すように、ペア状態「ＳＵＳＰＥＮＤ」（且つＩ／Ｏモード「リモート」）に変更された場合、パス管理テーブルのパス優先度に関わらず、パス３０９Ｂを使用する、すなわち、第２ストレージ装置２０ＢのＬＤＥＶに対してＩ／Ｏを行うのが望ましい。従って、このケースでは、パス優先度が、パス３０９Ａについてパス優先度「Active/non-optimized」に変更され、パス３０９Ｂについてパス優先度「Active/optimized」に変更されることが望ましい。本実施例では、次の流れにより、ホスト装置は、パス優先度を知ることができる。

（１）ペア状態により、ストレージ装置で管理されているパス優先度が変更される。

（２）ペア状態変更によるパス優先度変更の際もＵＡが設定される。具体的には、図１９の処理が行われた後に、Ｉ／Ｏコマンドを受信する等のＵＡ変更イベント（図１７のＳ１７０１）が行われると、図１７の処理により、ＵＡが伝播する。

（３）ＵＡ応答により、ホスト装置は、パス優先度変更を知る。

（４）ホスト装置は、ＲＴＰＧコマンドを送信することにより、パス優先度を知る。

（なお、図１８Ｂに破線矢印で示すように、パス３０９Ａ経由で第２ストレージ装置２０ＢのＬＤＥＶに対してＩ／Ｏを行うことも可能ではあるが、パス３０９Ｂを使用することに比べて性能が低い。）

【０１１８】

そこで、本実施例では、ペア状態の変更を行ったストレージ装置（例えば２０Ａ）は、ペア状態の変更に伴い、ＵＡの値を「１」（パス優先度変更を意味する値）に変更する。ＵＡ「１」の場合、後述するように、やがてホスト装置１０からＲＴＰＧコマンドを第１ストレージ装置２０Ａが受信することになるが、ＲＴＰＧコマンドの受信を含むパス優先度レポート処理では、ペア状態及びＩ／Ｏモードに応じたパス優先度を表すレポートを含んだＲＴＰＧ応答が作成されホスト装置１０に対して送信される。これにより、ホスト装置１０が保持するパス管理情報が表すパス優先度が、ペア状態及びＩ／Ｏモードに応じたパス優先度に変更され、以後、ホスト装置１０により、その変更後のパス管理情報を基に、パスが使用されることになる。

【０１１９】

図１９は、パス優先度変更処理の流れを示す。

【０１２０】

第１ストレージ装置２０Ａは、ストレージ管理サーバ１１からペア状態変更コマンドを受信する（Ｓ１９０１）。ペア状態変更コマンドは、ペア状態の変更対象のＬＤＥＶペアのうち第１ストレージ装置２０Ａが有するＬＤＥＶの番号を含む。以下、そのＬＤＥＶの番号を、図１９の説明において「対象ＬＤＥＶ番号」と言う。

【０１２１】

第１ストレージ装置２０Ａは、対象ＬＤＥＶ番号に対応したＨＧ番号をポート管理テーブル２１５Ａから特定し、特定したＨＧ番号に対応したＵＡ６０３（図５参照）を「１」（パス優先度変更を意味する値）に変更する（Ｓ１９０２）。また、第１ストレージ装置２０Ａは、対象ＬＤＥＶ番号に対応したＵＡ同期フラグビット７０２（図５参照）を「１」に変更する（Ｓ１９０３）。

【０１２２】

第１ストレージ装置２０Ａは、完了を表す応答をストレージ管理サーバ１１に送信する（Ｓ１９０４）。

【０１２３】

なお、Ｓ１９０２でＵＡが「１」に変更されたが、パス優先度変更処理においては、そのＵＡ変更は、ペア相手ストレージ装置（２０Ｂ）に通知されない。第１ストレージ装置２０ＡがＩ／Ｏコマンドを受信等のＵＡ伝搬イベント発生した場合に（図１７参照）、ＵＡ変更がペア相手ストレージ装置に通知される。

【０１２４】

図２０～図２３は、パス優先度レポート処理の流れを示す。

【０１２５】

図２０に示すように、第１ストレージ装置２０Ａが、ＲＴＰＧコマンドをホスト装置１０から受信する（Ｓ２００１）。ＲＴＰＧコマンドは、ＬＵＮを含む。このＬＵＮは、例えば、ＵＡ応答に対応したＩ／Ｏコマンド内のＬＵＮと同一のＬＵＮである。すなわち、ホスト装置１０は、Ｉ／Ｏコマンドの応答としてＵＡ応答を受信した場合、そのＩ／Ｏコマンド内のＬＵＮと同一ＬＵＮを含んだＲＴＰＧコマンドを、そのＩ／Ｏコマンドの送信先ストレージ装置と同一のストレージ装置に送信できる。なお、ホスト装置１０は、ＲＴＰＧコマンドを、ＵＡ応答を受信したというイベントとは別のイベントに応答して（例えば定期的に）送信してよい。以下、図２０～図２３の説明において、ＲＴＰＧコマンド内のＬＵＮに対応したＬＤＥＶ（第１ストレージ装置２０Ａが有するＬＤＥＶ）を「対象ＬＤＥＶ」と言い、対象ＬＤＥＶの番号を、「対象ＬＤＥＶ番号」と言う。

【０１２６】

第１ストレージ装置２０Ａは、対象ＬＤＥＶがＶＤＥＶの基になるＬＤＥＶペアを構成しているか否かを、ペア管理テーブル２１２Ａを基に判断する（Ｓ２００２）。対象ＬＤ

ＥＶ番号がペア管理テーブル２１２Ａにあれば、Ｓ２００２の判断結果が肯定であり、対象ＬＤＥＶ番号がペア管理テーブル２１２Ａに無ければ、Ｓ２００２の判断結果が否定である。

【０１２７】

Ｓ２００２の判断結果が否定の場合（Ｓ２００２：Ｎｏ）、第１ストレージ装置２０Ａは、自ストレージ装置内のパス優先度（例えば、図示しないテーブルから特定されたパス優先度）を含んだ応答をホスト装置１０に送信する（Ｓ２００５）。

【０１２８】

Ｓ２００２の判断結果が肯定の場合（Ｓ２００２：Ｙｅｓ）、第１ストレージ装置２０Ａは、対象ＬＤＥＶ番号に対応したＩ／Ｏモードを、ペア管理テーブル２１２Ａから特定する（Ｓ２００３）。

【０１２９】

Ｓ２００４で特定されたＩ／Ｏモードが「Ｂｌｏｃｋの場合、第１ストレージ装置２０Ａは、ホスト装置１０にエラー応答を送信する（Ｓ２００６）。Ｓ２００４で特定されたＩ／Ｏモードが「Ｒｅｍｏｔｅ」の場合、第１ストレージ装置２０Ａは、図２１に示す処理を行う。Ｓ２００４で特定されたＩ／Ｏモードが「Ｍｉｒｒｏｒ」の場合、第１ストレージ装置２０Ａは、図２２に示す処理を行う。Ｓ２００４で特定されたＩ／Ｏモードが「Ｌｏｃａｌ」の場合、第１ストレージ装置２０Ａは、図２３に示す処理を行う。

【０１３０】

図２１に示すように、Ｉ／Ｏモードが「リモート」の場合、第１ストレージ装置２０Ａは、対象ＬＤＥＶに関連付けられている全てのポートの情報（例えばポート番号及び相対ポートＩＤ）と、対象ＬＤＥＶのペア相手ＬＤＥＶ（ペア相手ストレージ）の情報とを、対象ＬＤＥＶ番号を用いて、ペア管理テーブル２１２Ａ、相対ポートＩＤ管理テーブル２１７Ａ、仮想ボックス管理テーブル２１６Ａ及びポート管理テーブル２１５Ａから特定する（Ｓ２１０１）。特定される情報（ポートの情報及びペア相手ＬＤＥＶの情報）は、例えば、対象ＬＤＥＶを含んだ仮想ボックスの範囲とされる。これは、図２２（Ｓ２２０３）及び図２３（Ｓ２３０３）でも同様である。以下、図２１～図２３の説明において、対象ＬＤＥＶのペア相手ストレージ装置は、第２ストレージ装置２０Ｂとする。

【０１３１】

第１ストレージ装置２０Ａは、特定した第２ストレージ装置２０Ｂ（ペア相手ストレージ装置）と通信可能か否かを判断する（Ｓ２１０２）。例えば、第１ストレージ装置２０Ａは、所定のコマンドを第２ストレージ装置２０Ｂに送信して所定時間内に応答が返ってくるか否かを判断する。

【０１３２】

Ｓ２１０２の判断結果が否定の場合（Ｓ２１０２：Ｎｏ）、第１ストレージ装置２０Ａは、ホスト装置１０にエラー応答を送信する（Ｓ２１０３）。

【０１３３】

Ｓ２１０２の判断結果が肯定の場合（Ｓ２１０２：Ｙｅｓ）、第１ストレージ装置２０Ａは、ＵＡ同期フラグビットリスト２２３Ａと、パス優先度要求とを、第２ストレージ装置２０Ｂに送信する（Ｓ２１０４）。

【０１３４】

第２ストレージ装置２０Ｂは、対象ＬＤＥＶ番号と、ＵＡ同期フラグビットリスト２２３Ａと、パス優先度要求とを受信する。第２ストレージ装置２０Ｂは、受信したＵＡ同期フラグビットリスト２２３Ａを基に、ＵＡ同期フラグビットが「１」の全てのＬＤＥＶの各々について、ＨＧ番号を特定し、特定したＨＧ番号に対応したＵＡを「１」に変更する（Ｓ２１０５）。また、第２ストレージ装置２０Ｂは、パス優先度要求に応答して、対象ＬＤＥＶに対応したペア相手ＬＤＥＶに関連付いたＨＧ番号のパス優先度（パス管理テーブル２１４Ｂから特定されたパス優先度）を表すパス優先度情報、第１ストレージ装置２０Ａに通知する（Ｓ２１０６）。

【０１３５】

第１ストレージ装置２０Ａは、パス優先度情報を第２ストレージ装置２０Ｂから受信した場合、ＵＡ同期フラグビットリスト２２３Ａにおける全てのＵＡ同期フラグビット「１」を「０」に変更し（Ｓ２１０７）、第１パス優先度（対象ＬＤＥＶを含む１以上のＬＤＥＶ（第１ストレージ装置２０Ａ内のＬＤＥＶ）についてのパス優先度）と第２パス優先度（ペア相手ＬＤＥＶを含む１以上のＬＤＥＶ（第２ストレージ装置２０Ｂ内のＬＤＥＶ）についてのパス優先度）とをマージしたレポートを含むＲＴＰＧ応答をホスト装置１０に送信する（Ｓ２１０８）。但し、そのレポートでは、第１ストレージ装置２０Ａが有する第１パス優先度情報（対象ＬＤＥＶのＨＧ番号に対応したパス優先度を表す情報）が表すパス優先度に関わらず、第１パス優先度は「Active/non-optimized」とされ、第２ストレージ装置２０Ｂから受信した第２パス優先度情報（ペア相手ＬＤＥＶのＨＧ番号に対応したパス優先度を表す情報）が表すパス優先度に関わらず、第２パス優先度は「Active/optimized」とされる。つまり、以後、ペア相手ＬＤＥＶへのパスが優先的に使用されるようにする。なぜなら、Ｉ／Ｏモードが「リモート」の場合、ペア相手ＬＤＥＶ内のデータの方が対象ＬＤＥＶ内のデータより新しい可能性があり信頼できるデータであるためである。

【０１３６】

なお、Ｓ２１０４では、ＵＡ同期フラグビットリスト２２３Ａ全体ではなく、ＵＡ同期フラグビットリスト２２３Ａのうち対象ＬＤＥＶのＵＡ同期フラグビットが送信されてもよい。このため、Ｓ２１０５において、ＵＡ変更は、対象ＬＤＥＶのペア相手ＬＤＥＶに関連付いたＨＧ番号についてのみ行われてよい。また、Ｓ２１０４において、ＵＡ同期フラグビットリスト２２３Ａに少なくとも１つのＵＡ同期フラグビット「１」があるか否か（或いは、うち対象ＬＤＥＶのＵＡ同期フラグビットが「１」であるか否か）が判断され、その判断の結果が肯定の場合、ＵＡ同期フラグビットリスト（又は対象ＬＤＥＶのＵＡ同期フラグビット）が第２ストレージ装置２０Ｂに送信され、その判断の結果が否定の場合、ＵＡ同期フラグビットリスト（又は対象ＬＤＥＶのＵＡ同期フラグビット）が第２ストレージ装置２０Ｂに送信されず且つＳ２１０５がスキップされてよい。これは、図２２（Ｓ２２０６及びＳ２２０７）及び図２３（Ｓ２３０６及びＳ２３０７）についても同様である。

【０１３７】

図２２に示すように、Ｉ／Ｏモードが「ミラー」の場合、第１ストレージ装置２０Ａは、対象ＬＤＥＶが閉塞しているか否かを判断する（Ｓ２２０１）。

【０１３８】

Ｓ２２０１の判断結果が肯定の場合（Ｓ２１０１：Ｙｅｓ）、第１ストレージ装置２０Ａは、ホスト装置１０にエラー応答を送信する（Ｓ２２０２）。

【０１３９】

Ｓ２２０１の判断結果が否定の場合（Ｓ２１０１：Ｎｏ）、第１ストレージ装置２０Ａは、図２１のＳ２１０１と同じ処理を行う（Ｓ２２０３）。

【０１４０】

第１ストレージ装置２０Ａは、特定した第２ストレージ装置２０Ｂ（ペア相手ストレージ装置）と通信可能か否かを判断する（Ｓ２２０４）。

【０１４１】

Ｓ２２０４の判断結果が否定の場合（Ｓ２１０２：Ｎｏ）、第１ストレージ装置２０Ａは、第１パス優先度情報が表すパス優先度に関わらず第１パス優先度を「Active/optimized」としたレポートを含むＲＴＰＧ応答をホスト装置１０に送信する（Ｓ２２０５）。なぜなら、第２ストレージ装置２０Ｂと通信できないので、以後、対象ＬＤＥＶへのパスが優先的に使用されるようにするためである。

【０１４２】

Ｓ２２０４の判断結果が肯定の場合（Ｓ２２０４：Ｙｅｓ）、図２１の２１０４～Ｓ２１０７と同じ処理が行われる（Ｓ２２０６～Ｓ２２０９）。第１ストレージ装置２０Ａは、第１パス優先度及び第２パス優先度をマージしたレポートを含むＲＴＰＧ応答をホスト

装置１０に送信する（Ｓ２２１０）。そのレポートでは、第１パス優先度は、第１ストレージ装置２０Ａが有する第１パス優先度情報が表すパス優先度通りであり、第２パス優先度は、第２ストレージ装置２０Ｂから受信した第２パス優先度情報が表すパス優先度通りである。

【０１４３】

図２３に示すように、Ｉ／Ｏモードが「ローカル」の場合、図２２のＳ２２０１～Ｓ２２０９と同じ処理が行われる（Ｓ２３０１～Ｓ２３０９）。第１ストレージ装置２０Ａは、第１パス優先度と第２パス優先度とをマージしたレポートを含むＲＴＰＧ応答をホスト装置１０に送信する（Ｓ２３１０）。但し、そのレポートでは、第１ストレージ装置２０Ａが有する第１パス優先度情報が表すパス優先度に関わらず、第１パス優先度は「Active/optimized」とされ、第２ストレージ装置２０Ｂから受信した第２パス優先度情報が表すパス優先度に関わらず、第２パス優先度は「Active/non-optimized」とされる。つまり、以後、対象ＬＤＥＶへのパスが優先的に使用されるようにする。なぜなら、Ｉ／Ｏモードが「ローカル」の場合、対象ＬＤＥＶ内のデータの方がペア相手ＬＤＥＶ内のデータより新しい可能性があり信頼できるデータであるためである。

【実施例２】

【０１４４】

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

【０１４５】

まず、実施例２の概要を述べる。

【０１４６】

ホスト装置１０の数が多いと、ホストグループが増える可能性があり、そうすると、ホストグループ単位にパス優先度を設定するには手間がかかる。

【０１４７】

そこで、実施例２では、パス優先度を設定する手間が、実施例１より軽減される。具体的には、実施例２では、管理者はパス優先度を設定する必要は無い。１以上のホスト装置と１以上のストレージ装置でグループが定義される。本実施例では、そのグループを「データセンタ」と呼ぶ。ホスト装置と同一のデータセンタに属するストレージ装置との間の物理的な距離は、そのホスト装置と異なるデータセンタに属するストレージ装置との間の物理的な距離よりも短いことが望ましい。

【０１４８】

データセンタ毎にＩＤが割り振られる。データセンタＩＤは、管理者又はユーザ（仮想ボックスのユーザ）等により割り振られてもよい。データセンタＩＤは、そのデータセンタＩＤが割り振られたデータセンタに属するホスト装置及びストレージ装置がそれぞれ保持する。ホスト装置は、同一データセンタ内のストレージ装置（そのホスト装置が保持するデータセンタＩＤと同一のデータセンタＩＤを保持するストレージ装置）とそのストレージ装置内のＬＤＥＶとの間のパスの優先度を、異なるデータセンタ内のストレージ装置（そのホスト装置が保持するデータセンタＩＤと異なるデータセンタＩＤを保持するストレージ装置）とそのストレージ装置内のＬＤＥＶとの間のパスの優先度よりも高くする。具体的には、例えば、ホスト装置２４１０は、同一データセンタに属するストレージ装置２４２０内のＬＤＥＶとの間のパスの優先度を「Active/optimized」に決定し、異なるデータセンタに属するストレージ装置２４２０内のＬＤＥＶとの間のパスの優先度を「Active/non-optimized」に決定する。以上のことを、例えば下記のように表現できる。

少なくとも１つのホスト装置と、
前記複数のホスト装置に接続された複数のストレージ装置と
を有し、
前記複数のホスト装置の各々が、そのホスト装置が属するグループのＩＤを保持し、
前記複数のストレージ装置の各々が、そのストレージ装置が属するグループのＩＤを保持し、

各グループには、少なくとも１つのホスト装置と少なくとも１つのストレージ装置が属し、
前記少なくとも１つのホスト装置の各々は、
グループＩＤの問合せをいずれかのストレージ装置に送信し、
前記問合せの応答として、前記問合せの送信先のストレージ装置が保持するグループＩＤを含んだ応答を受信し、
前記応答内のグループＩＤが、当該ホスト装置が保持するグループＩＤと同一であれば、前記応答の送信元ストレージ装置が有する全ての論理ボリュームへの１以上のパスの各々の優先度を、第１の優先度とし、
前記応答内のグループＩＤが、当該ホスト装置が保持するグループＩＤと異なっていれば、前記応答の送信元ストレージ装置が有する全ての論理ボリュームへの１以上のパスの各々の優先度を、第１の優先度よりも低い第２の優先度とする、
計算機システム。

【０１４９】

図２４は、実施例２に係る計算機システムの構成、及び、各ストレージ装置が有する管理情報の構成を示す。

【０１５０】

複数のホスト装置２４１０と複数のストレージ装置２４２０がある。図２４の例では、第１及び第２ホスト装置２４１０Ａ及び２４１０Ｂと、第１及び第２ストレージ装置２４２０Ａ及び２４２０がある。以下、説明を分かり易くするために、第１ホスト装置２４１０Ａと第１ストレージ装置２４２０Ａが第１データセンタに属し、第２ホスト装置２４１０Ｂと第２ストレージ装置２４２０Ｂが第２データセンタに属しているとする。

【０１５１】

ホスト装置２４１０は、データセンタＩＤ問合せをストレージ装置２４２０に送信することで、そのストレージ装置２４２０が保持するデータセンタＩＤを含んだ応答をストレージ装置２４２０から受信する。データセンタＩＤ問合せは、例えば図２５に示す構成でよい。すなわち、データセンタＩＤ問合せは、オペレーションコードフィールドとデータセンタＩＤフィールドとを有してよい。オペレーションコードフィールドには、オペレーションコードが記述される。オペレーションコードは、この問合せを受信したストレージ装置に実施させたい命令の種類をあらわすコードである。また、データセンタＩＤ問合せにおけるデータセンタＩＤフィールドは、ブランクでよく、応答の際に、この問合せを受信したストレージ装置２４２０が保持するデータセンタＩＤがそのストレージ装置２４２０に記述され、図示の問合せの構造体を含んだ応答（データセンタＩＤが記述された構造体を含んだ応答）が、ストレージ装置２４２０からホスト装置２４１０に返される。

【０１５２】

ホスト装置２４１０は、記憶デバイスの一例としてメモリ１０１を有する。メモリ１０１は、パス優先度管理テーブル１０２を記憶する。パス優先度管理テーブル１０２は、例えば図２６に示すように、ストレージ装置毎に、ストレージ製番（ストレージの製造番号）２６０１、パス状態２６０２及びパス優先度２６０３を有する。つまり、本実施例では、パス優先度の設定は、ストレージ装置単位で行われる。すなわち、ホスト装置２４１０がアクセス可能な１つのストレージ装置内の複数のＬＤＥＶに対応した複数のパスについて、まとめて、パス優先度が設定される。なお、ストレージ製番２６０１は、ストレージ装置のＩＤの一例である。パス状態２６０２の値としては、「Normal」（ストレージ装置と通信可能）と、「Blocked」（ストレージ装置と通信不可能）がある。ホスト装置２４１０は、パス状態「Blocked」のストレージ装置についてのパスの優先度が「Active/optimized」でも、そのパスを使用せず、パス状態「Normal」の他のストレージ装置についてのパスを使用する。

【０１５３】

また、ホスト装置２４１０のメモリ１０１は、データセンタＩＤ管理テーブル１０３を記憶し、ストレージ装置２４２０（例えばＳＭ２１）も、データセンタＩＤ管理テーブル

２１９を記憶する。テーブル１０３及び２１９はそれぞれ構成が同じであり、また、同一データセンタ内のテーブル１０３及び２１９は保持する値も同じである。具体的には、第１ホスト装置２４１０Ａと第１ストレージ装置２４２０Ａが第１データセンタに属するが、テーブル１０３Ａ及び２１９Ａは、図２７Ａに示すように、同一のデータセンタＩＤ「0x0001」を保持する。同様に、第２ホスト装置２４１０Ｂと第２ストレージ装置２４２０Ｂが第２データセンタに属するが、テーブル１０３Ｂ及び２１９Ｂは、図２７Ｂに示すように、同一のデータセンタＩＤ「0x0002」を保持する。

【０１５４】

以下、実施例２で行われる処理を説明する。

【０１５５】

図２８は、パス優先度自動決定処理の流れを示す。なお、この処理は、ホスト装置２４１０内のパス管理プログラム（図３０参照）により行うことができる。また、この処理の主体として、ホスト装置２４１０Ａを例に取る。

【０１５６】

ホスト装置２４１０Ａは、データセンタＩＤ問合せを、ストレージ装置２４２０に送信し（Ｓ２８０１）、そのストレージ装置２４２０から、そのストレージ装置２４２０が保持するデータセンタＩＤを含んだ応答を受信する（Ｓ２８０２）。

【０１５７】

ホスト装置２４１０Ａは、受信した応答内のデータセンタＩＤと、データセンタＩＤ管理テーブル１０３Ａ内のデータセンタＩＤとが同一か否かを判断する（Ｓ２８０３）。

【０１５８】

Ｓ２８０３の判断結果が否定の場合（Ｓ２８０３：Ｎｏ）、Ｓ２８０２で受信した応答の送信元ストレージ装置がホスト装置２４１０Ａと異なるデータセンタに属しているということである。この場合、ホスト装置２４１０Ａは、Ｓ２８０２で受信した応答の送信元ストレージ装置について、パス優先度２６０３として「Active/non-optimized」を設定する（Ｓ２８０４）。これは、応答の送信元ストレージ装置がストレージ装置２４２０Ｂの場合に行われる処理である。

【０１５９】

一方、Ｓ２８０３の判断結果が肯定の場合（Ｓ２８０３：Ｙｅｓ）、Ｓ２８０２で受信した応答の送信元ストレージ装置がホスト装置２４１０Ａと同一データセンタに属しているということである。この場合、ホスト装置２４１０Ａは、Ｓ２８０２で受信した応答の送信元ストレージ装置について、パス優先度２６０３として「Active/optimized」を設定する（Ｓ２８０５）。これは、応答の送信元ストレージ装置がストレージ装置２４２０Ａの場合に行われる処理である。

【０１６０】

図２９は、Ｉ／Ｏ制御処理の流れを示す。この処理は、ホスト装置２４１０内のパス管理プログラムにより行うことができる。また、この処理の主体として、ホスト装置２４１０Ａを例に取る。この処理は、ＶＤＥＶに対するＩ／Ｏを、例えばパス管理プログラムがアプリケーションプログラムから指示された場合に行われる。

【０１６１】

ホスト装置２４１０Ａは、パス優先度２６０３が「Active/optimized」のストレージ装置をパス優先度管理テーブル１０２Ａから検索する（Ｓ２９０１）。ホスト装置２４１０Ａは、その検索により見つかったストレージ装置のパス状態２６０２が「Blocked」か否かを判断する（Ｓ２９０２）。

【０１６２】

Ｓ２９０２の判断結果が否定の場合（Ｓ２９０２：Ｎｏ）、ホスト装置２４１０Ａは、Ｉ／Ｏ先ＶＤＥＶの基であり見つかったストレージ装置のＬＤＥＶへのパスを選択し、選択したパスを使用してＩ／Ｏコマンドを送信する（Ｓ２９０３）。

【０１６３】

Ｓ２９０２の判断結果が肯定の場合（Ｓ２９０２：Ｙｅｓ）、ホスト装置２４１０Ａは

、パス状態２６０２が「Blocked」ではないストレージ装置があるか否かを、パス優先度管理テーブル１０２Ａを基に判断する（Ｓ２９０４）。

【０１６４】

Ｓ２９０４の判断結果が肯定の場合（Ｓ２９０４：Ｙｅｓ）、ホスト装置２４１０Ａは、Ｉ／Ｏ先ＶＤＥＶの基でありパス状態２６０２が「Blocked」ではないいずれかのストレージ装置のＬＤＥＶへのパスを選択し、選択したパスを使用してＩ／Ｏコマンドを送信する（Ｓ２９０５）。パス状態２６０２が「Blocked」ではないストレージ装置として、パス優先度２６０３が「Active/optimized」のストレージ装置とパス優先度２６０３が「Active/non-optimized」のストレージ装置の両方があれば、パス優先度２６０３が「Active/optimized」のストレージ装置が選択される。

【０１６５】

Ｓ２９０４の判断結果が否定の場合（Ｓ２９０４：Ｎｏ）、ホスト装置２４１０Ａは、通信不可で異常終了（例えばエラーを発行）する（Ｓ２９０６）。

【０１６６】

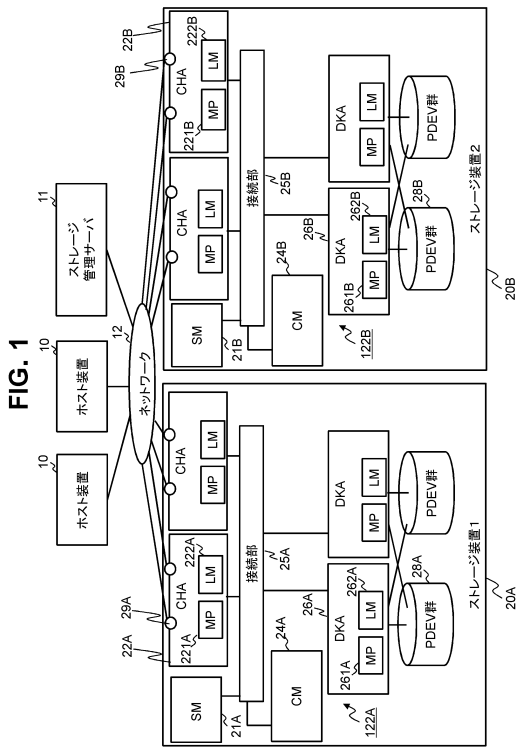
以上、幾つかの実施例及び変形例を説明したが、本発明は、これらの実施例及び変形例に限定されるものでなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。例えば、本発明は、ＡＬＵＡに限らず、ストレージ装置が状態変更をホストシステムに通知することでホストシステムが状態変更を知ることができるようになっている環境下のシステムにも適用できる。

【符号の説明】

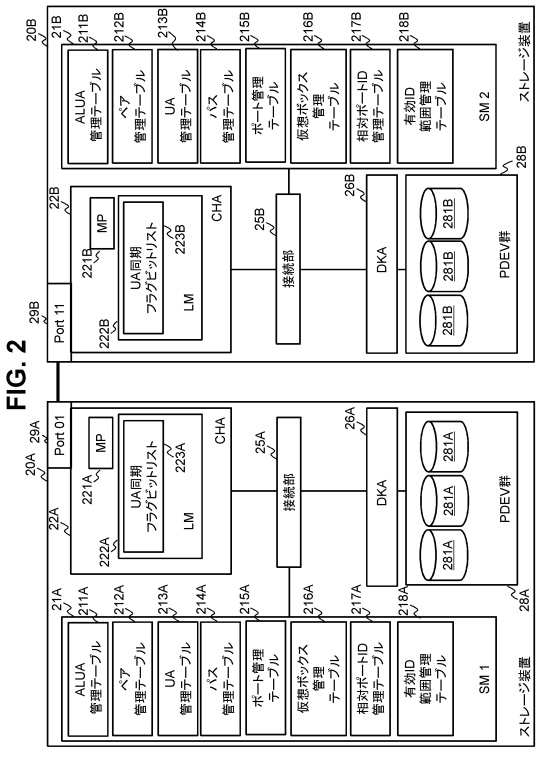
【０１６７】

１０：ホスト装置　２０Ａ：第１ストレージ装置　２０Ｂ：第２ストレージ装置

【図１】



【図２】



【図３】

FIG. 3

ALUA管理テーブル
211A

401 LDEV番号	402 ALUA適用状態
01:01	0
01:02	1
....	....

【図４】

FIG. 4

ペア管理テーブル
212A

501 LDEV番号（自ストレージ）	502 LDEV番号（ペア相手ストレージ）	503 ペア状態	504 I/Oモード	505 ALUA同期ビット
01:01	02:01	PAIR	Mirror	0
01:02	02:02	SUSPEND	Local	0
01:03	02:03	SUSPEND	Remote	0
01:04	02:04	SUSPEND	Block	1
....	....	....	....	....

【図５】

FIG. 5

UA管理テーブル
213A

601 ポート番号	602 ホストグループ番号	603 UA
01	HG01	0
02	HG02	1
....	....	....

【図６】

FIG. 6

UA同期フラグビットリスト
262A

701 LDEV番号	702 UA同期フラグビット
01:01	0
01:02	1
....	....

【図７】

FIG. 7

パス管理テーブル
214A

801 仮想ボックス番号	802 HG番号	803 相対ポートID	804 パス優先度
01	HG01	0x0000	Active/optimized
01	HG02	0x0001	Active/non-optimized
...	...	...	...
02	HG10	0x2000	Active/optimized
02	HG11	0x2001	Active/non-optimized
...	...	...	...

【図８】

FIG. 8

ポート管理テーブル
215A

1001 ポート番号	1002 インデックス番号	1003 LUN	1004 LDEV番号	1005 HG番号
01	001	00	00:01	HG01
01	002	01	00:02	HG02
...	...	...	...	...
02	001	11	11:11	HG11
02	002	12	11:12	HG12
...	...	...	...	...

【図９】

FIG. 9

仮想ボックス管理テーブル
216A

1101 仮想ボックス番号	1102 HG番号	1103 ポート番号	1104 LDEV番号
01	HG01	01	00:01
01	HG02	02	00:02
...	...	...	...

【図１０】

FIG. 10

相対ポートID管理テーブル
217A

1201 HG番号	1202 相対ポートID
HG01	0x0001
HG02	0x0002
...	...

【図１１】

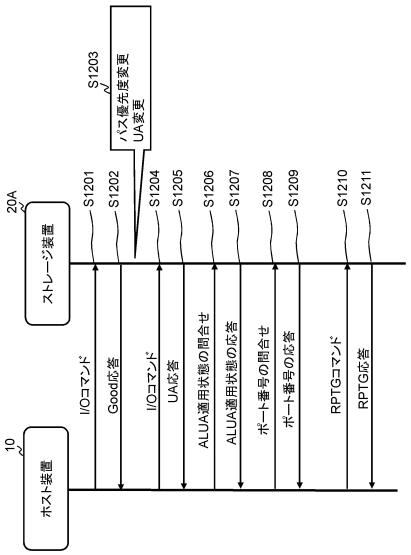
FIG. 11

有効ID範囲管理テーブル
218A

1301 仮想ボックス番号	1302 開始ID	1303 終了ID
01	0x0001	0x1999
02	0x2000	0x2999
...	...	...

【図１２】

FIG. 12



【図１３】

FIG. 13

214B

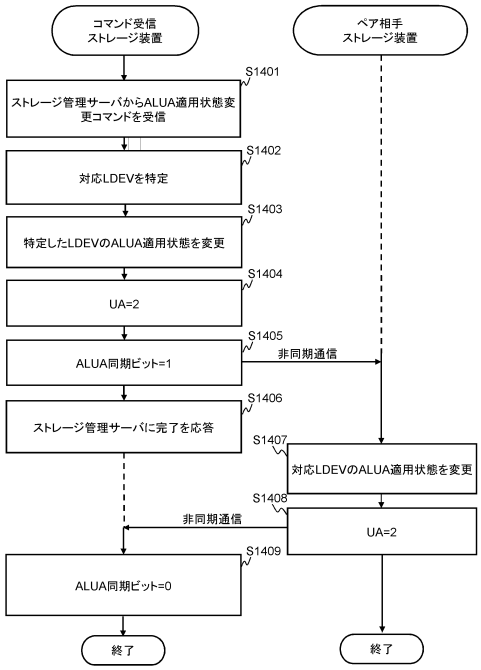
仮想ボックス番号	HG番号	相対ポートID	パス優先度
01	HG01	0x1000	Active/non-optimized
01	HG02	0x1001	Active/non-optimized
01	…	…	…
01	HG09	0x1999	Active/non-optimized
02	HG10	0x3000	Active/non-optimized
02	HG11	0x3001	Active/non-optimized
…	…	…	…

214A

仮想ボックス番号	HG番号	相対ポートID	パス優先度
01	HG01	0x0000	Active/optimized
01	HG02	0x0001	Active/optimized
01	…	…	…
01	HG09	0x0999	Active/optimized
02	HG10	0x2000	Active/optimized
02	HG11	0x2001	Active/optimized
…	…	…	…

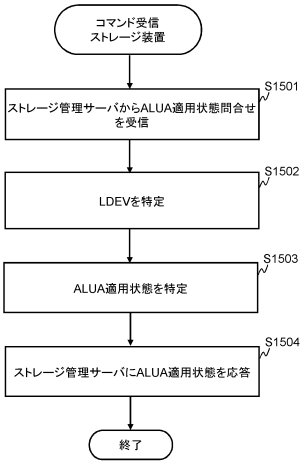
【図１４】

FIG. 14



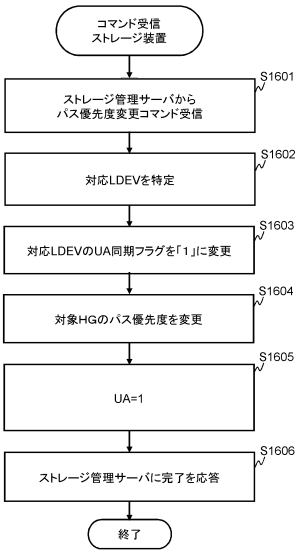
【図１５】

FIG. 15



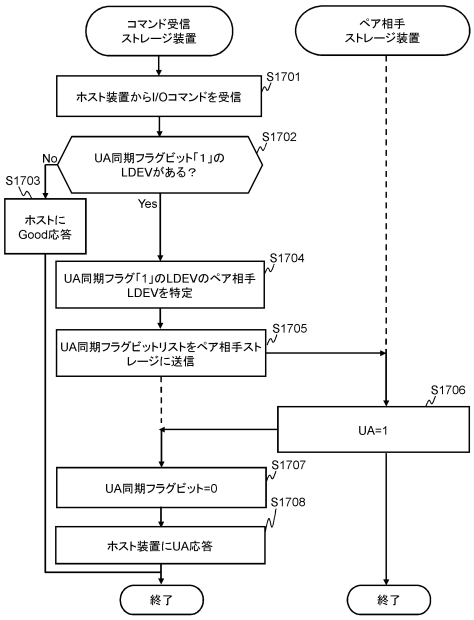
【図１６】

FIG. 16



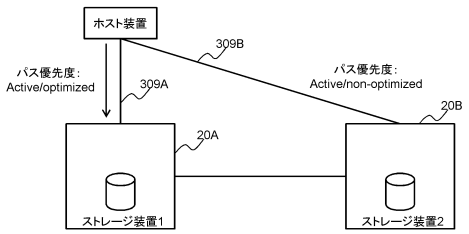
【図１７】

FIG. 17



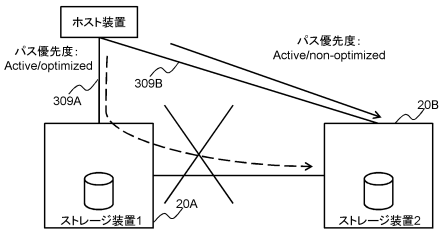
【図１８Ａ】

FIG. 18A



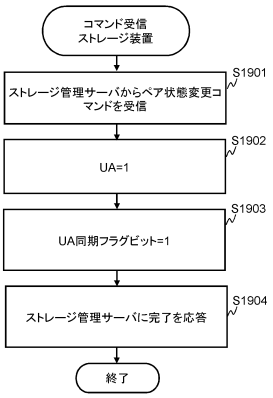
【図１８Ｂ】

FIG. 18B



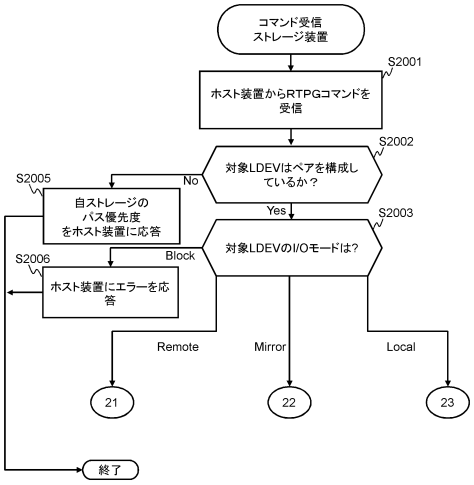
【図１９】

FIG. 19



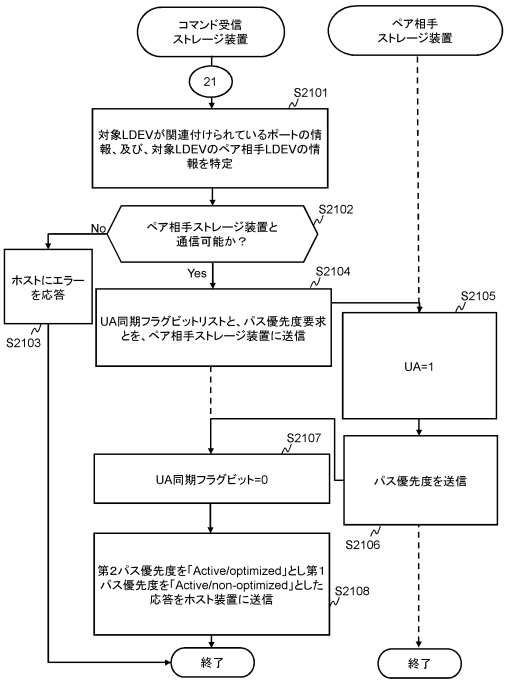
【図２０】

FIG. 20



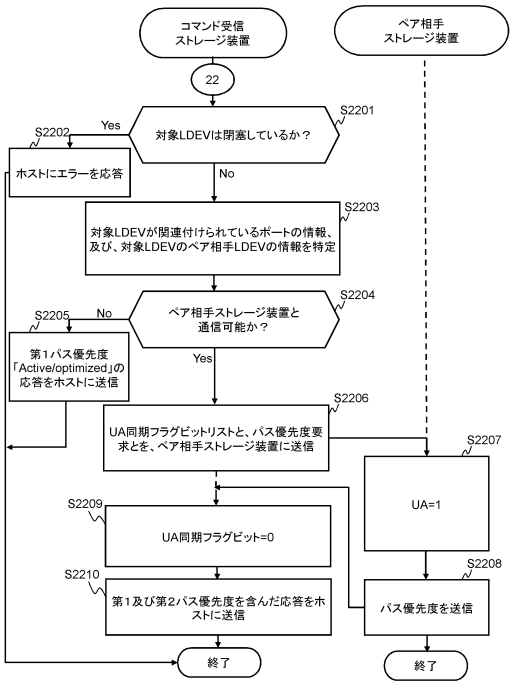
【図２１】

FIG. 21



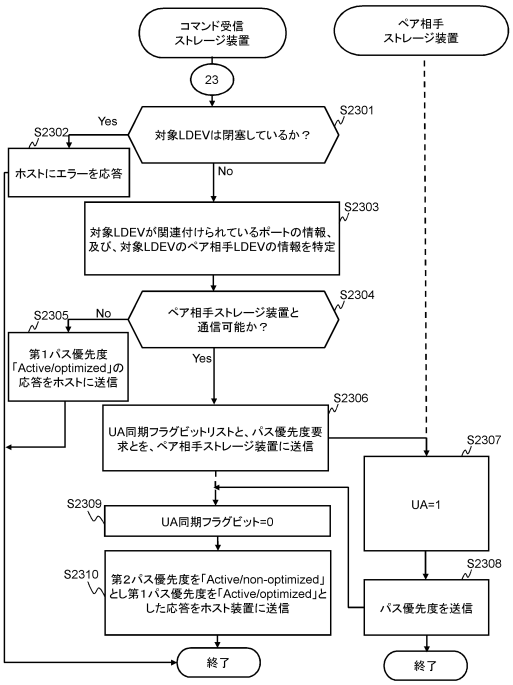
【図２２】

FIG. 22



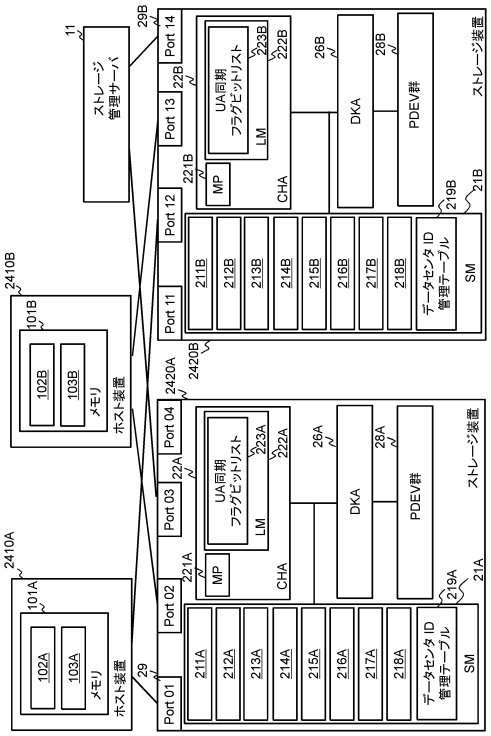
【図２３】

FIG. 23



【図２４】

FIG. 24



【図２５】

FIG. 25

データセンタID問合せ

オペレーションコード	データセンタID

【図２６】

FIG. 26

パス優先度管理テーブル
102

2601 ストレージ製番	2602 パス状態	2603 パス優先度
11111	Normal	Active/optimized
22222	Normal	Active/non-optimized
....	....	....

【図２７Ａ】

FIG. 27A

データセンタID管理テーブル
103A, 219A

データセンタID	0x0001

【図２７Ｂ】

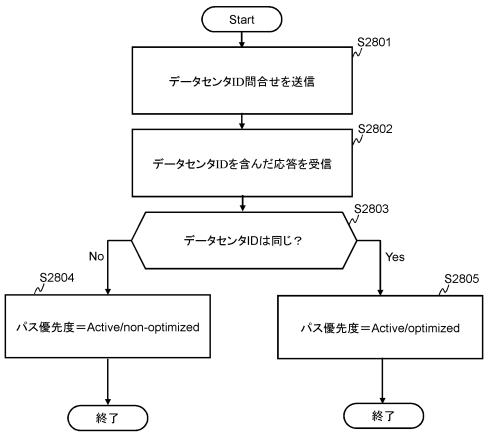
FIG. 27B

データセンタID管理テーブル
103B, 219B

データセンタID	0x0002

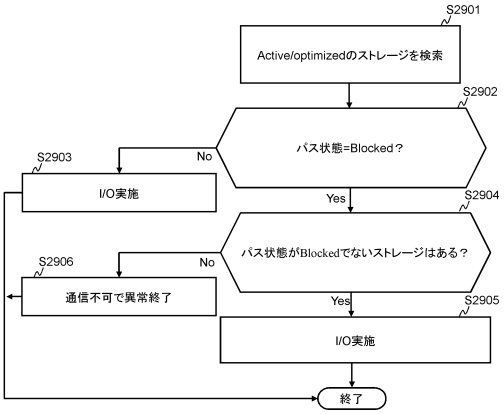
【図２８】

FIG. 28



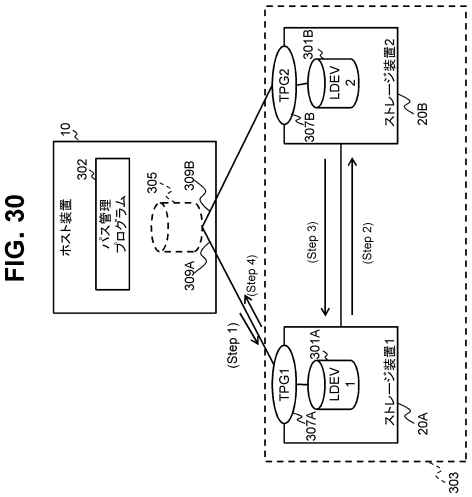
【図２９】

FIG. 29



【図３０】

FIG. 30



フロントページの続き

(72)発明者　南波　優
東京都品川区東品川四丁目１２番７号　株式会社日立ソリューションズ内

審査官　田中　啓介

(56)参考文献　特表２０１２－５０４７９３（ＪＰ，Ａ）
特開２０１４－０４８７１０（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ３／０６－３／０８
Ｇ０６Ｆ１３／００－１３／１４
Ｇ０６Ｆ１３／２０－１３／４２
Ｈ０４Ｌ１２／００－１２／２８
Ｈ０４Ｌ１２／４４－１２／９５５
Ｈ０４Ｗ８／２６、２４／００、２８／０２
Ｈ０４Ｗ７２／０４、７４／０４、７４／０８
Ｈ０４Ｗ８４／１２、８８／０８