

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4733399号
(P4733399)

(45) 発行日 平成23年7月27日 (2011.7.27)

(24) 登録日 平成23年4月28日 (2011.4.28)

(51) Int.Cl. F I
G O 6 F 13/10 (2006.01) G O 6 F 13/10 3 4 O A
G O 6 F 9/46 (2006.01) G O 6 F 9/46 3 5 O

請求項の数 10 (全 46 頁)

(21) 出願番号	特願2005-20908 (P2005-20908)	(73) 特許権者	000005108
(22) 出願日	平成17年1月28日 (2005.1.28)		株式会社日立製作所
(65) 公開番号	特開2006-209487 (P2006-209487A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成18年8月10日 (2006.8.10)	(74) 代理人	100075513
審査請求日	平成19年10月11日 (2007.10.11)		弁理士 後藤 政喜
		(74) 代理人	100084537
			弁理士 松田 嘉夫
		(74) 代理人	100114236
			弁理士 藤井 正弘
		(72) 発明者	橋本 顕義
			神奈川県川崎市麻生区王禅寺1069番地
			株式会社日立製作所 システム開発研究 所内

最終頁に続く

(54) 【発明の名称】 計算機システム、計算機、ストレージ装置及び管理端末

(57) 【特許請求の範囲】

【請求項 1】

サーバ装置と、前記サーバ装置に接続されるストレージ装置と、前記サーバ装置及び前記ストレージ装置と接続され、計算機システムを管理する管理端末と、を備える計算機システムであって、

前記サーバ装置は、第1のCPU、第1のメモリ及び複数のI/Oアダプタを含む計算機資源を備え、

前記第1のメモリは、

論理的に分割された前記計算機資源が割り当てられる仮想CPU、仮想メモリ及び仮想I/Oアダプタを備え、アプリケーションプログラムが動作する仮想計算機を管理する第1のハイパーバイザと、

前記I/Oアダプタ及び前記仮想I/Oアダプタの対応関係を示す仮想I/Oアダプタ管理情報と、を格納し、

前記ストレージ装置は、制御CPUと制御メモリとを備える制御部、ディスクキャッシュ、複数のチャネルアダプタ及び物理ディスクを含むストレージ資源を備え、

前記制御メモリは、

論理的に分割された前記ストレージ資源が割り当てられる仮想制御CPU、仮想ディスク、及び仮想チャネルアダプタを備え、前記仮想計算機に割り当てられる仮想ストレージ装置を管理する第2のハイパーバイザと、

前記チャネルアダプタ及び前記仮想チャネルアダプタの対応関係を示す仮想チャネルア

10

20

ダブタ管理情報と、を格納し、

前記管理端末は、第2のCPUと、第2のメモリ及び入出力部を備え、

前記第2のメモリは、前記I/Oアダプタと前記仮想I/Oアダプタとの対応関係、及び、前記チャンネルアダプタと前記仮想チャンネルアダプタとの対応関係を含む、前記仮想計算機と前記仮想ストレージ装置とを接続するパスの情報である仮想バス管理情報を格納し

、
前記第1のハイパーバイザによって生成される仮想計算機は、前記複数のI/Oアダプタに含まれる第1のI/Oアダプタと対応づけられる第1の仮想I/Oアダプタを備える第1の仮想計算機を含み、

前記第2のハイパーバイザによって生成される仮想ストレージ装置は、前記複数のチャンネルアダプタに含まれる第1のチャンネルアダプタと対応づけられた第1の仮想チャンネルアダプタを備える第1の仮想ストレージ装置を含み、

前記第1の仮想計算機及び前記第1の仮想ストレージ装置は、前記第1のI/Oアダプタ、前記第1の仮想I/Oアダプタ、前記第1のチャンネルアダプタ及び前記第1の仮想チャンネルアダプタが関連づけられた第1のパスによって接続され、

前記第1のハイパーバイザは、

前記第1のパスの障害発生を検知した場合に、前記第1のパスにおける前記第1の仮想I/Oアダプタが前記複数のI/Oアダプタに含まれる第2のI/Oアダプタと対応づけられるように前記仮想I/Oアダプタ管理情報を更新することによって、前記第1の仮想I/Oアダプタと対応づけられた前記第1のI/Oアダプタを、前記第2のI/Oアダプタに変更し、

前記第1の仮想I/Oアダプタと前記第2のI/Oアダプタとの対応づけを含む第1の変更情報を前記管理端末に送信し、

前記第2のハイパーバイザは、

前記第1のパスの障害発生を検知した場合に、前記第1のパスにおける前記第1の仮想チャンネルアダプタが前記複数のチャンネルアダプタに含まれる第2のチャンネルアダプタと対応づけられるように前記仮想チャンネルアダプタ管理情報を更新することによって、前記第1の仮想チャンネルアダプタと対応づけられた前記第1のチャンネルアダプタを、前記第2のチャンネルアダプタに変更し、

前記第1の仮想チャンネルアダプタと前記第2のチャンネルアダプタとの対応づけを含む第2の変更情報を前記管理端末に送信し、

前記管理端末は、前記第1の変更情報、及び、前記第2の変更情報の少なくとも一方を受信した場合に、前記仮想バス管理情報を更新することを特徴とする計算機システム。

【請求項2】

前記管理端末は、

前記サーバ装置から前記仮想I/Oアダプタ管理情報を取得し、

前記ストレージ装置から前記仮想チャンネルアダプタ管理情報を取得し、

前記取得された仮想I/Oアダプタ管理情報及び前記取得された仮想チャンネルアダプタ管理情報に基づいて、前記仮想バス管理情報を生成することを特徴とする請求項1に記載の計算機システム。

【請求項3】

前記仮想バス管理情報は、前記第1の仮想計算機が備える前記第1の仮想I/Oアダプタ、前記第1の仮想I/Oアダプタと対応づけられる前記第1のI/Oアダプタ、前記第1の仮想計算機と接続される前記第1の仮想ストレージ装置が備える前記第1の仮想チャンネルアダプタ、前記第1の仮想チャンネルアダプタと対応づけられた前記第1のチャンネルアダプタが関連づけられた前記第1のパスに関する第1のパス情報を含み、

前記管理端末は、前記第1の変更情報を受信した場合に、前記第1の仮想I/Oアダプタと対応づけられる前記第1のI/Oアダプタを、前記第2のI/Oアダプタに変更することによって前記仮想バス管理情報を更新することを特徴とする請求項1に記載の計算機システム。

10

20

30

40

50

【請求項 4】

前記仮想バス管理情報は、前記第 1 の仮想計算機が備える前記第 1 の仮想 I / O アダプタ、前記第 1 の仮想 I / O アダプタと対応づけられる前記第 1 の I / O アダプタ、前記第 1 の仮想計算機と接続される前記第 1 の仮想ストレージ装置が備える前記第 1 の仮想チャネルアダプタ、前記第 1 の仮想チャネルアダプタと対応づけられた前記第 1 のチャネルアダプタが関連づけられた前記第 1 のバスに関する第 1 のバス情報を含み、

前記管理端末は、前記第 2 の変更情報を受信した場合に、前記第 1 の仮想チャネルアダプタと対応づけられた前記第 1 のチャネルアダプタを、前記第 2 のチャネルアダプタに変更することによって前記仮想バス管理情報を更新することを特徴とする請求項 1 に記載の計算機システム。

10

【請求項 5】

前記第 1 の仮想計算機は、さらに、前記複数の I / O アダプタに含まれる第 3 の I / O アダプタと対応づけられた第 2 の仮想 I / O アダプタを備え、

前記第 1 の仮想ストレージ装置は、さらに、前記複数のチャネルアダプタに含まれる第 3 のチャネルアダプタと対応づけられた第 2 の仮想チャネルアダプタを備え、

前記第 1 の仮想計算機及び前記第 1 の仮想ストレージ装置は、さらに、前記第 3 の I / O アダプタ、前記第 2 の仮想 I / O アダプタ、前記第 3 のチャネルアダプタ及び前記第 2 の仮想チャネルアダプタが関連づけられた第 2 のバスによって接続され、

前記管理端末は、

前記仮想バス管理情報に基づいて、前記第 1 のバス及び前記第 2 のバスに関連する情報を前記入出力部に表示することを特徴とする請求項 1 に記載の計算機システム。

20

【請求項 6】

サーバ装置と、前記サーバ装置に接続されるストレージ装置と、前記サーバ装置及び前記ストレージ装置と接続され、計算機システムを管理する管理端末と、を備える計算機システムにおける仮想接続管理方法であって、

前記サーバ装置は、第 1 の CPU、第 1 のメモリ及び複数の I / O アダプタを含む計算機資源を備え、

前記第 1 のメモリは、

論理的に分割された前記計算機資源が割り当てられる仮想 CPU、仮想メモリ及び仮想 I / O アダプタを備え、アプリケーションプログラムが動作する仮想計算機を管理する第 1 のハイパーバイザと、

30

前記 I / O アダプタ及び前記仮想 I / O アダプタの対応関係を示す仮想 I / O アダプタ管理情報と、を格納し、

前記ストレージ装置は、制御 CPU と制御メモリとを備える制御部、ディスクキャッシュ、複数のチャネルアダプタ及び物理ディスクを含むストレージ資源を備え、

前記制御メモリは、

論理的に分割された前記ストレージ資源が割り当てられる仮想制御 CPU、仮想ディスク、及び仮想チャネルアダプタを備え、前記仮想計算機に割り当てられる仮想ストレージ装置を管理する第 2 のハイパーバイザと、

前記チャネルアダプタ及び前記仮想チャネルアダプタの対応関係を示す仮想チャネルアダプタ管理情報と、を格納し、

40

前記管理端末は、第 2 の CPU と、第 2 のメモリ及び入出力部を備え、

前記第 2 のメモリは、前記 I / O アダプタと前記仮想 I / O アダプタとの対応関係、及び、前記チャネルアダプタと前記仮想チャネルアダプタとの対応関係を含む、前記仮想計算機と前記仮想ストレージ装置とを接続するバスの情報である仮想バス管理情報を格納し、

前記第 1 のハイパーバイザによって生成される仮想計算機は、前記複数の I / O アダプタに含まれる第 1 の I / O アダプタと対応づけられる第 1 の仮想 I / O アダプタを備える第 1 の仮想計算機を含み、

前記第 2 のハイパーバイザによって生成される仮想ストレージ装置は、前記複数のチャ

50

ネルアダプタに含まれる第 1 のチャンネルアダプタと対応づけられた第 1 の仮想チャンネルアダプタを備える第 1 の仮想ストレージ装置を含み、

前記第 1 の仮想計算機及び前記第 1 の仮想ストレージ装置は、前記第 1 の I / O アダプタ、前記第 1 の仮想 I / O アダプタ、前記第 1 のチャンネルアダプタ及び前記第 1 の仮想チャンネルアダプタが対応づけられた第 1 のバスによって接続され、

前記仮想接続管理方法は、

前記第 1 のバスの障害発生が検知された場合に、前記第 1 のハイパーバイザが、前記第 1 のバスにおける前記第 1 の仮想 I / O アダプタが前記複数の I / O アダプタに含まれる第 2 の I / O アダプタに対応づけられるように前記仮想 I / O アダプタ管理情報を更新することによって、前記第 1 の仮想 I / O アダプタと対応づけられた前記第 1 の I / O アダプタを、前記第 2 の I / O アダプタに変更する第 1 のステップと、

前記第 1 のハイパーバイザが、前記第 1 の仮想 I / O アダプタと前記第 2 の I / O アダプタとの対応づけを含む第 1 の変更情報を前記管理端末に送信する第 2 のステップと、

前記第 1 のバスの障害発生が検知された場合に、前記第 2 のハイパーバイザが、前記第 1 のバスにおける前記第 1 の仮想チャンネルアダプタが前記複数のチャンネルアダプタに含まれる第 2 のチャンネルアダプタに対応づけられるように前記仮想チャンネルアダプタ管理情報を更新することによって、前記第 1 の仮想チャンネルアダプタと対応づけられた前記第 1 のチャンネルアダプタを、前記第 2 のチャンネルアダプタに変更する第 3 のステップと、

前記第 2 のハイパーバイザが、前記第 1 の仮想チャンネルアダプタと前記第 2 のチャンネルアダプタとの対応付けを含む第 2 の変更情報を前記管理端末に送信する第 4 のステップと

前記管理端末が、前記第 1 の変更情報、又は、前記第 2 の変更情報のいずれかを受信した場合に、前記仮想バス管理情報を更新する第 5 のステップと、を含むことを特徴とする仮想接続管理方法。

【請求項 7】

前記管理端末が、

前記サーバ装置から前記仮想 I / O アダプタ管理情報を取得するステップと、

前記ストレージ装置から前記仮想チャンネルアダプタ管理情報を取得するステップと、

前記取得された仮想 I / O アダプタ管理情報及び前記取得された仮想チャンネルアダプタ管理情報に基づいて、前記仮想バス管理情報を生成するステップと、を含むことを特徴とする請求項 6 に記載の仮想接続管理方法。

【請求項 8】

前記仮想バス管理情報は、前記第 1 の仮想計算機が備える前記第 1 の仮想 I / O アダプタ、前記第 1 の仮想 I / O アダプタと対応づけられる前記第 1 の I / O アダプタ、前記第 1 の仮想計算機と接続される前記第 1 の仮想ストレージ装置が備える前記第 1 の仮想チャンネルアダプタ、前記第 1 の仮想チャンネルアダプタと対応づけられた前記第 1 のチャンネルアダプタが対応づけられた前記第 1 のバスに関する第 1 のバス情報を含み、

前記第 5 のステップは、前記管理端末が、前記第 1 の変更情報を受信した場合に、前記第 1 のバス情報に含まれる前記第 1 の仮想 I / O アダプタと対応づけられる前記第 1 の I / O アダプタを、前記第 2 の I / O アダプタに変更することによって前記仮想バス管理情報を更新することを特徴とする請求項 6 に記載の仮想接続管理方法。

【請求項 9】

前記仮想バス管理情報は、前記第 1 の仮想計算機が備える前記第 1 の仮想 I / O アダプタ、前記第 1 の仮想 I / O アダプタと対応づけられる前記第 1 の I / O アダプタ、前記第 1 の仮想計算機と接続される前記第 1 の仮想ストレージ装置が備える前記第 1 の仮想チャンネルアダプタ、前記第 1 の仮想チャンネルアダプタと対応づけられた前記第 1 のチャンネルアダプタが対応づけられた前記第 1 のバスに関する第 1 のバス情報を含み、

前記第 5 のステップでは、前記管理端末が、前記第 2 の変更情報を受信した場合に、前記第 1 の仮想チャンネルアダプタと対応づけられた前記第 1 のチャンネルアダプタを、前記第 2 のチャンネルアダプタに変更することによって前記仮想バス管理情報を更新することを特

徴とする請求項 6 に記載の仮想接続管理方法。

【請求項 10】

前記第 1 の仮想計算機は、さらに、前記複数の I / O アダプタに含まれる第 3 の I / O アダプタと対応づけられた第 2 の仮想 I / O アダプタを備え、

前記第 1 の仮想ストレージ装置は、さらに、前記複数のチャンネルアダプタに含まれる第 3 のチャンネルアダプタと対応づけられた第 2 の仮想チャンネルアダプタを備え、

前記第 1 の仮想計算機及び前記第 1 の仮想ストレージ装置は、さらに、前記第 3 の I / O アダプタ、前記第 2 の仮想 I / O アダプタ、前記第 3 のチャンネルアダプタ及び前記第 2 の仮想チャンネルアダプタが対応づけられた第 2 のバスによって接続され、

前記仮想接続管理方法は、

前記管理端末が、前記仮想バス管理情報に基づいて、前記第 1 のバス及び前記第 2 のバスを前記入出力部に表示するステップを含むことを特徴とする請求項 6 に記載の仮想接続管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システムに関し、特に、ストレージ装置が含まれる計算機システムの論理分割技術に関する。

【背景技術】

【0002】

情報処理システムの能力を向上させるための方法として、情報処理システムが有する計算機の台数を増やす方法がある。しかし、計算機を多数設置すると個々の計算機の管理に手間がかかり、またこれらの複数の計算機の設置面積や電力等の合計も大きくなる問題があった。

【0003】

これを解決するために、処理能力の大きい計算機を設置し、その計算機の資源を論理的に複数の分割 (L P A R : Logical Partitioning) し、論理分割された仮想的な計算機の各々を独立に使用する技術が提案されている。このような論理分割によって、1 台の計算機を仮想的に複数の計算機に見せることができ、各論理区画にプロセッサ、メモリ等の資源の割り当てを制御することによって、各仮想計算機毎の性能が保証される。

【0004】

また、各仮想的計算機に種々のオペレーティングシステムを自由に搭載し、運用・停止や障害処理も仮想計算機毎に独立して行えるなど、柔軟な運用が可能となる。

【0005】

また、物理的な装置台数が少なくなるので、装置管理、設置面積、消費電力等で有利である。このような論理分割技術は、例えば、特許文献 1 に開示されている。

【0006】

また、特許文献 2 には、論理分割された仮想計算機からストレージ装置をアクセスする技術が記載されている。

【0007】

さらに、非特許文献 1 には、計算機の論理分割技術を適用して、論理的に分割するストレージ装置が提案されている。

【0008】

さらに、計算機システムの信頼性を向上させるために、システムダウンを起こりにくくする技術が提案されている。このために、クラスタリング、多重化、自動バックアップ、ホットスワップなどの技術が使用され、計算機システムの高い可用性 (ハイアベイラビリティ) を実現している。

【特許文献 1】特開 2003 - 157177 号公報

【特許文献 2】米国特許出願公開第 2003 / 0055933 号明細書

【非特許文献 1】IBM 社、" The IBM TotalStorage DS8000 Series: Concepts and cep

10

20

30

40

50

ts Architecture”、[online]、インターネット<URL : <http://www.redbooks.ibm.com/redpieces/pdfs/sg246452.pdf>>

【発明の開示】

【発明が解決しようとする課題】

【0009】

従来の計算機における論理分割技術では、計算機内のプロセッサ、メモリ等の資源が論理分割されて、各仮想計算機に割り当てられていた。同様に、従来のストレージ装置における論理分割技術では、ストレージ装置内のキャッシュメモリ、ディスクドライブ等の資源が論理分割されて、各仮想ストレージ装置に割り当てられていた。

【0010】

しかし、前述した従来技術では、計算機側の仮想化エンジンとストレージ装置側の仮想化エンジンとは連携していなかった。

【0011】

また、仮想計算機と仮想ストレージ装置との間に形成される仮想バスは、物理的なバスと必ずしも一致していなかった。よって、管理者は、仮想バスと物理バスの対応関係を常に確認しながら、サーバ及びストレージを設定しなければならなかった。特に、高い可用性を実現する計算機システムでは、サーバ及びストレージの関係が複雑であるため、サーバとストレージの対応関係を確認しながら、これらを設定することは高度な技術を要した。

【0012】

すなわち、従来技術では、サーバ側の仮想化エンジンとストレージ装置側の仮想化エンジンとが連携しないため、仮想計算機と仮想ストレージとの間の経路情報が存在しない。よって、高い可用性を実現するために、サーバ側及びストレージ側の構成情報を各々参照しなければならない問題があった。

【0013】

本発明は、サーバとストレージとの間の物理的バス、及び仮想計算機と仮想ストレージ装置との間の経路情報を一元的に管理することを目的とする。

【課題を解決するための手段】

【0014】

本発明は、サーバ装置と、前記サーバ装置に接続されるストレージ装置と、前記サーバ装置及び前記ストレージ装置と接続され、計算機システムを管理する管理端末と、を備える計算機システムであって、前記サーバ装置は、第1のCPU、第1のメモリ及び複数のI/Oアダプタを含む計算機資源を備え、前記第1のメモリは、論理的に分割された前記計算機資源が割り当てられる仮想CPU、仮想メモリ及び仮想I/Oアダプタを備え、アプリケーションプログラムが動作する仮想計算機を管理する第1のハイパーバイザと、前記I/Oアダプタ及び前記仮想I/Oアダプタの対応関係を示す仮想I/Oアダプタ管理情報と、を格納し、前記ストレージ装置は、制御CPUと制御メモリとを備える制御部、ディスクキャッシュ、複数のチャネルアダプタ及び物理ディスクを含むストレージ資源を備え、前記制御メモリは、論理的に分割された前記ストレージ資源が割り当てられる仮想制御CPU、仮想ディスク、及び仮想チャネルアダプタを備え、前記仮想計算機に割り当てられる仮想ストレージ装置を管理する第2のハイパーバイザと、前記チャネルアダプタ及び前記仮想チャネルアダプタの対応関係を示す仮想チャネルアダプタ管理情報と、を格納し、前記管理端末は、第2のCPUと、第2のメモリ及び入出力部を備え、前記第2のメモリは、前記I/Oアダプタと前記仮想I/Oアダプタとの対応関係、及び、前記チャネルアダプタと前記仮想チャネルアダプタとの対応関係を含む、前記仮想計算機と前記仮想ストレージ装置とを接続するバスの情報である仮想バス管理情報を格納し、前記第1のハイパーバイザによって生成される仮想計算機は、前記複数のI/Oアダプタに含まれる第1のI/Oアダプタと対応づけられる第1の仮想I/Oアダプタを備える第1の仮想計算機を含み、前記第2のハイパーバイザによって生成される仮想ストレージ装置は、前記複数のチャネルアダプタに含まれる第1のチャネルアダプタと対応づけられた第1の仮想

10

20

30

40

50

チャンネルアダプタを備える第１の仮想ストレージ装置を含み、前記第１の仮想計算機及び前記第１の仮想ストレージ装置は、前記第１のＩ／Ｏアダプタ、前記第１の仮想Ｉ／Ｏアダプタ、前記第１のチャンネルアダプタ及び前記第１の仮想チャンネルアダプタが関連づけられた第１のバスによって接続され、前記第１のハイパーバイザは、前記第１のバスの障害発生を検知した場合に、前記第１のバスにおける前記第１の仮想Ｉ／Ｏアダプタが前記複数のＩ／Ｏアダプタに含まれる第２のＩ／Ｏアダプタと対応づけられるように前記仮想Ｉ／Ｏアダプタ管理情報を更新することによって、前記第１の仮想Ｉ／Ｏアダプタと対応づけられた前記第１のＩ／Ｏアダプタを、前記第２のＩ／Ｏアダプタに変更し、前記第１の仮想Ｉ／Ｏアダプタと前記第２のＩ／Ｏアダプタとの対応づけを含む第１の変更情報を前記管理端末に送信し、前記第２のハイパーバイザは、前記第１のバスの障害発生を検知した場合に、前記第１のバスにおける前記第１の仮想チャンネルアダプタが前記複数のチャンネルアダプタに含まれる第２のチャンネルアダプタと対応づけられるように前記仮想チャンネルアダプタ管理情報を更新することによって、前記第１の仮想チャンネルアダプタと対応づけられた前記第１のチャンネルアダプタを、前記第２のチャンネルアダプタに変更し、前記第１の仮想チャンネルアダプタと前記第２のチャンネルアダプタとの対応づけを含む第２の変更情報を前記管理端末に送信し、前記管理端末は、前記第１の変更情報、及び、前記第２の変更情報の少なくとも一方を受信した場合に、前記仮想バス管理情報を更新することを特徴とする。

10

【発明の効果】

【００１５】

20

本発明によると、仮想計算機と仮想ストレージ装置間を備える計算機システムのハイアベイラビリティ構成を容易に設定することができる。

【発明を実施するための最良の形態】

【００１６】

以下、本発明の実施の形態を図面を参照して説明する。

【００１７】

（第１実施形態）

図１は、第１の実施の形態の計算機システムのハードウェア構成を示すブロック図である。

【００１８】

30

第１の実施の形態の計算機システムは、サーバ装置（０）１００、サーバ装置（１）１５０、ストレージ装置２００、及び管理端末３００によって構成されている。サーバ装置（０）１００及びサーバ装置（１）１５０では、アプリケーションプログラムが動作している。ストレージ装置２００は、サーバ装置１００、１５０の動作に必要なデータを記憶する。管理端末３００は、計算機システムの管理情報を管理し、計算機システム全体の動作を管理する。なお、物理資源の後に付された括弧書きの数字は、各物理資源の識別子である。

【００１９】

サーバ装置（０）１００は、ＣＰＵ（０）１０１、不揮発性メモリ（０）１０２、主記憶メモリ（０）１０４、ＬＡＮアダプタ（０）１０５、Ｉ／Ｏアダプタ（０）１０６、Ｉ／Ｏアダプタ（１）１０７及びシステムバス１０８等の物理資源を備える計算機である。

40

【００２０】

ＣＰＵ（０）１０１は、サーバ装置（０）１００で実行されるオペレーティングシステム（ＯＳ）及びアプリケーションプログラムに関する演算処理を行う。

【００２１】

主記憶メモリ（０）１０４は、ＣＰＵ（０）１０１の動作に必要なプログラムやデータを記憶する。

【００２２】

システムバス（０）１０８は、ＣＰＵ（０）１０１とＬＡＮアダプタ（０）１０５、Ｉ／Ｏアダプタ１０６及びＩ／Ｏアダプタ１０７の間を接続して、データや制御信号を転送

50

する。

【 0 0 2 3 】

I/Oアダプタ(0)106及びI/Oアダプタ(1)107は、I/Oチャンネル(例えば、ファイバチャンネル)400を介してストレージ装置200と接続されている。これらのI/Oアダプタ106、107は、ストレージ装置200に対してデータ入出力要求を送信し、ストレージ装置200に格納されたデータを受信する。I/Oアダプタは二つ設けられている。なお、さらに多くのI/Oアダプタを設けてもよい。この二つのI/Oアダプタ106、107は別個に動作する。よって、一方のI/Oアダプタに障害が発生しても、サーバ装置(0)100からストレージ装置200へのアクセスが停止しないように、処理系を2重化している。

10

【 0 0 2 4 】

LANアダプタ(0)105は、ネットワーク410を介して、他のサーバ装置150、ストレージ装置200及び管理端末300と接続されている。LANアダプタ(0)105は、ネットワーク410を介して接続された装置との間で制御信号や管理情報(各種管理テーブルの内容)を送受信する。

【 0 0 2 5 】

不揮発性メモリ(0)102には、ハイパーバイザ103が格納されている。ハイパーバイザ103は、CPU(0)101が実行する処理によって実現され、サーバ装置(0)100の物理資源の論理区画を実現する一手段である。

【 0 0 2 6 】

ハイパーバイザ103は、サーバ装置(0)100の電源投入時に不揮発性メモリ102から専用のプログラムを実行することにより、読み出される。そして、このプログラムの実行によってハイパーバイザ103が起動され、サーバ装置(0)100に備わる資源を管理する。すなわち、ハイパーバイザ103は、サーバ装置(0)100内に論理区画を構成し、独立して動作する仮想計算機を生成するための管理プログラムである。ハイパーバイザ103は、CPU101において所定の処理が実行されることによって実現される。

20

【 0 0 2 7 】

なお、ハイパーバイザ103は、サーバ装置(0)100の電源投入時に起動されるのではなく、サーバ装置(0)100のOS起動後に仮想化エンジンを起動し、OSと仮想化エンジンによってハイパーバイザを構成してもよい。この仮想化エンジンは不揮発メモリ(0)102に格納しても、ストレージ装置200に格納してもよい。この場合、サーバ装置(0)100の電源投入時に起動したOSが仮想化エンジンを読み出して実行する。

30

【 0 0 2 8 】

なお、本明細書では、多くの場合、ソフトウェアを主体として動作を説明するが、実際には、CPUがソフトウェアを実行することにより、ハイパーバイザ103等やストレージハイパーバイザ214等が動作する。

【 0 0 2 9 】

また、ハイパーバイザ103をソフトウェアで構成せず、ハードウェアによって構成してもよい。例えば、ハイパーバイザ用の専用チップを設けたり、CPU(0)101に、仮想計算機を管理するハイパーバイザ部を設けてもよい。

40

【 0 0 3 0 】

サーバ装置(1)150は、前述したサーバ装置(0)100と同じ構成を有するので、その説明は省略する。

【 0 0 3 1 】

ストレージ装置200は、CPU(2)211、主記憶メモリ(2)212及び不揮発性メモリ(2)213を含むストレージ制御部210を備える。また、CPU(3)221、主記憶メモリ(3)222及び不揮発性メモリ(3)223を含むストレージ制御部220を備える。この複数のストレージ制御部210、220は別個に動作する。よって

50

、一方のストレージ制御部に障害が発生しても、ストレージ装置 200 が停止しないように、処理系を 2 重化している。

【0032】

ストレージ制御部 210、220 は、物理ディスクドライブ 236 に対するデータの入出力を制御し、ストレージ装置 200 の動作を管理する。なお、ストレージ装置 200 が N A S (Network Attached Storage) である場合には、ストレージ制御部 210、220 においてファイルシステムが動作する。

【0033】

C P U (2) 211 は、ストレージ装置 200 で実行される各種管理プログラムに関する演算処理を行う。

【0034】

主記憶メモリ (2) 212 は、C P U (2) 211 の動作に必要なプログラムやデータを記憶する。

【0035】

不揮発性メモリ (2) 213 には、ストレージハイパーバイザ 214 が格納されている。ストレージハイパーバイザ 214 は、C P U 211、221 が実行する処理によって実現され、ストレージ装置 200 の物理資源の論理区画を実現する一手段である。

【0036】

ストレージハイパーバイザ 214 は、ストレージ装置 200 内に論理区画を構成し、独立して動作する仮想ストレージ装置を生成する管理プログラムによって実現されている。ストレージハイパーバイザ 214 の実現には、前述したサーバ装置 (0) 100 のハイパーバイザ 103 と同様な多様な方法を採用できる。

【0037】

なお、ストレージ制御部 220 の C P U (3) 221、主記憶メモリ (3) 222、不揮発性メモリ (3) 223 及びストレージハイパーバイザ 224 は、前述したストレージ制御部 210 の各部と同様に動作する。

【0038】

また、ストレージ装置 200 は、L A N アダプタ (2) 230 及びチャネルアダプタ (0) 231 ~ チャネルアダプタ (3) 234 を備える。

【0039】

チャネルアダプタ (0) 231 及びチャネルアダプタ (1) 232 は、I / O チャンネル 400 を介してサーバ装置 (0) 100 と接続されている。チャネルアダプタ 231、232 は、サーバ装置 (0) 100 からのデータ入出力要求を受信し、ストレージ装置 200 に格納されたデータを送信する。

【0040】

同様に、チャネルアダプタ (2) 233 及びチャネルアダプタ (3) 234 は、I / O チャンネル 400 を介してサーバ装置 (1) 150 と接続されている。チャネルアダプタ 233、234 は、サーバ装置 (1) 150 からのデータ入出力要求を受信し、ストレージ装置 200 に格納されたデータを送信する。

【0041】

ストレージ装置 200 には複数のチャネルアダプタが設けられている。この複数のチャネルアダプタ 231 ~ 234 は別個に動作する。よって、一方のチャネルアダプタに障害が発生しても、サーバ装置 100、150 からのストレージ装置 200 へのアクセスが停止しないように、処理系を 2 重化している。

【0042】

L A N アダプタ (2) 230 は、ネットワーク 410 を介して、サーバ装置 100、150 及び管理端末 300 と接続されている。L A N アダプタ (2) 230 は、ネットワーク 410 を介して接続された装置との間で制御信号や管理情報 (各種管理テーブルの内容) を送受信する。

【0043】

10

20

30

40

50

また、ストレージ装置 200 は、ディスクアダプタ 237、ディスクキャッシュ 238 及び物理ディスクドライブ 236 を備える。

【0044】

ディスクアダプタ 237 は、物理ディスクドライブ 236 に対するインターフェースである。ディスクアダプタ 237 は、例えば、ATA、SAS (Serial Attached SCSI)、ファイバチャネル等のプロトコルによって、物理ディスクドライブ 236 との間でデータや制御信号を送受信する。

【0045】

ディスクアダプタ 237 は二つ設けられている。なお、さらに多くのディスクアダプタ 237 を設けてもよい。各ディスクアダプタ 237 と物理ディスクドライブ 236 は異なる配線で接続されている。この二つのディスクアダプタ 237 は別個に動作する。よって、一方のディスクアダプタに障害が発生しても、物理ディスクドライブ 236 へのアクセスが停止しないように、処理系を 2 重化している。

【0046】

ディスクキャッシュ 238 は、物理ディスクドライブ 236 に読み書きされるデータを一時的に格納するメモリであり、サーバ装置 100、150 からのストレージ装置 200 へのアクセス性能を向上させる。

【0047】

物理ディスクドライブ 236 は、データを記憶する記憶媒体である。一般に、この記憶媒体は磁気ディスクであるが、光ディスク等、他の媒体を用いたものであってもよい。なお、複数の物理ディスクドライブ 236 によって RAID (Redundant Array of Independent Disks) が構成されており、記憶されるデータに冗長性を持たせている。このため、物理ディスクドライブ 236 の一部に障害が生じて、記憶されたデータが消失しないようになっている。

【0048】

また、ストレージ装置 200 は、ストレージ制御部 210、220、LAN アダプタ 230、チャネルアダプタ 231 ~ 234、ディスクアダプタ 237 及びディスクキャッシュ 238 を接続する相互結合網 239 を備える。相互結合網 239 は、例えば、クロスバスイッチによって構成される。

【0049】

管理端末 300 は、計算機システムを統括して管理するコンピュータ装置で、CPU、メモリ、入出力装置及び LAN インターフェースを備える。管理端末 300 の CPU では、仮想計算機管理プログラム 301 が動作している。

【0050】

管理端末 300 は、仮想バス管理テーブル 310 (図 5)、計算機資源管理テーブル 320 及びストレージ資源管理テーブル 330 を有する。仮想計算機管理プログラム 301 は、仮想バス管理テーブル 310、計算機資源管理テーブル 320 及びストレージ資源管理テーブル 330 を使用して、計算機システムの仮想的な構成を管理する。これらのテーブルの内容は、後述する。第 1 の実施の形態では、管理端末 300 に、仮想計算機と仮想ストレージ装置との接続関係を定める接続情報 (仮想バス管理テーブル 310) を有する管理部 (図示省略) が備わっている。

【0051】

管理端末 300 の LAN インターフェースは、ネットワーク 410 を介して、サーバ装置 100、サーバ装置 150 及びストレージ装置 200 と接続されている。当該 LAN インターフェースは、ネットワーク 410 を介して接続された装置との間で制御信号や管理情報 (各種管理テーブルの内容) を送受信する。

【0052】

I/O チャネル 400 は、例えばファイバチャネルプロトコルのような、データの転送に適するプロトコルで通信可能な伝送媒体である。なお、第 1 の実施の形態では、サーバ装置 100、150 とストレージ装置 200 との間は 1 対 1 で接続されているが、後述す

10

20

30

40

50

る第6の実施の形態のように、ネットワーク(SAN)によって接続されてもよい。

【0053】

ネットワーク410は、サーバ装置(0)100、サーバ装置(1)150、ストレージ装置200及び管理端末300を接続する。ネットワーク410は、例えば、TCP/IPプロトコルによって、コンピュータ間で制御信号や管理情報が通信可能に構成されており、例えば、イーサネット(登録商標、以下同じ)が用いられる。

【0054】

なお、図1には、二つのサーバ装置100、150と一つのストレージ装置200とが接続されている例を説明したが、三つ以上のサーバ装置を設けてもよく、二つ以上のストレージ装置を設けてもよい。

10

【0055】

図2は、第1の実施の形態の計算機システムの機能ブロック図である。

【0056】

サーバ装置(0)100は、機能的に、物理層、ハイパーバイザ層及び仮想計算機層に大別できる。

【0057】

物理層は、LANアダプタ及びI/Oアダプタ等の計算機資源を備えた物理計算機(0)120である。計算機資源とは、CPU、メモリ及びI/Oアダプタを含む、サーバ装置が備える物理資源を示す。なお、物理計算機(0)120は、図示した以外の計算機資源(CPU、メモリ等)を備えるが、本図における説明には不要なため、図示を省略する。

20

【0058】

ハイパーバイザ層は、前述したハイパーバイザ103によって実現される。物理計算機(0)120に備わる計算機資源は、ハイパーバイザ103によって管理されている。ハイパーバイザ103は、仮想I/Oアダプタ管理テーブル110及び計算機資源管理テーブル115を有している。

【0059】

なお、物理資源の後に付された括弧書きの数字は、各物理資源の識別子である。また、仮想資源の後に付された括弧書きの数字は、各仮想資源の識別子である。

【0060】

仮想I/Oアダプタ管理テーブル110(図3A)には、物理的なI/Oアダプタと仮想I/Oアダプタとの対応関係が記載されている。

30

【0061】

計算機資源管理テーブル115(図6)には、サーバ装置(0)100の資源と仮想計算機との対応関係が規定されている。計算機資源管理テーブル115によって、サーバ装置(0)100の計算機資源が管理される。

【0062】

仮想I/Oアダプタ管理テーブル110及び計算機資源管理テーブル115は、主記憶メモリ104に格納されている。なお、仮想I/Oアダプタ管理テーブル110及び計算機資源管理テーブル115を、不揮発メモリ102に格納してもよい。これらの管理テーブル110、115を不揮発メモリ102に格納することによって、サーバ装置(0)100の電源遮断時にも仮想計算機の管理情報が保持される。

40

【0063】

また、仮想I/Oアダプタ管理テーブル110及び計算機資源管理テーブル115をストレージ装置200に格納し、サーバ装置(0)100の起動後に、これらの管理テーブル110、115をストレージ装置200から読み出して、主記憶メモリ104に格納してもよい。

【0064】

また、仮想I/Oアダプタ管理テーブル110及び計算機資源管理テーブル115を管理端末300に格納し、サーバ装置(0)100の起動後に、これらの管理テーブル11

50

0、115を管理端末300から読み出して、主記憶メモリ104に格納してもよい。

【0065】

なお、ハイパーバイザ103は、仮想計算機を構成する他の計算機資源に関する情報も有している。

【0066】

仮想計算機層には、ハイパーバイザ103によって、物理計算機(0)120の計算機資源が論理区画によって分割された仮想計算機130、140が構成されている。仮想計算機(0)130は、仮想I/Oアダプタ(0)131、仮想I/Oアダプタ(1)132、CPU資源133及びメモリ資源134を含む。同様に、仮想計算機(1)140は、仮想I/Oアダプタ(2)141、仮想I/Oアダプタ(3)142、CPU資源143及びメモリ資源144を含む。なお、仮想計算機130、140には、サーバ装置(0)100に備わる他の計算機資源も含まれるが、本図における説明には不要なため、図示を省略する。

10

【0067】

そして、仮想計算機(0)130上では、OS(0)135が動作している。また、仮想計算機(1)140上では、OS(1)145が動作している。すなわち、OS(0)135は、仮想計算機(0)130に割り当てられた計算機資源を使用して演算処理をする。同様に、OS(1)145は、仮想計算機(1)140に割り当てられた計算機資源を使用して演算処理をする。

【0068】

20

また、各OS135、145上では、アプリケーションプログラムが動作している。このアプリケーションプログラムによって、サーバ装置(0)100に接続されたクライアント端末(図示省略)にデータベースサービスやウェブサービス等が提供される。

【0069】

サーバ装置(1)150は、前述したサーバ装置(0)100と同じ機能を有する。例えば、ハイパーバイザ層は、ハイパーバイザ153によって実現される。ハイパーバイザ153は、仮想I/Oアダプタ管理テーブル160(図3B)及び計算機資源管理テーブル165(図6)を有している。よって、サーバ装置(1)150の機能の詳細な説明は省略する。

【0070】

30

ストレージ装置200は、物理層、ハイパーバイザ層及び仮想ストレージ層に、機能的に大別できる。

【0071】

物理層は、LANアダプタ及びチャネルアダプタ等のストレージ資源を備えた物理ストレージ240である。ストレージ資源とは、CPU、ディスクキャッシュ、チャネルアダプタ及び物理ディスクドライブを含む、ストレージ装置が備える物理資源を示す。なお、物理ストレージ240は、図示した以外の資源(物理ディスクドライブ、ディスクキャッシュ等)を備えるが、本図における説明には不要なため、図示を省略する。

【0072】

ハイパーバイザ層は、前述したストレージハイパーバイザ214、224によって実現される。物理ストレージ240に備わるストレージ資源は、ハイパーバイザ214、224によって管理されている。ストレージハイパーバイザ214、224は、仮想チャネルアダプタ管理テーブル260及びストレージ資源管理テーブル265を有している。

40

【0073】

仮想チャネルアダプタ管理テーブル260(図4)には、物理的なチャネルアダプタと仮想チャネルアダプタとの対応関係が記載されている。

【0074】

ストレージ資源管理テーブル265(図7)には、ストレージ装置200の資源と仮想計算機との対応関係が規定されている。ストレージ資源管理テーブル223によって、ストレージ資源の配分が管理される。

50

【 0 0 7 5 】

仮想チャネルアダプタ管理テーブル 2 6 0 及びストレージ資源管理テーブル 2 6 5 は、主記憶メモリ 2 1 2 及び主記憶メモリ 2 2 2 に格納されている。なお、これらの仮想チャネルアダプタ管理テーブル 2 6 0 及びストレージ資源管理テーブル 2 6 5 を、不揮発メモリ 2 1 3 及び不揮発メモリ 2 2 3 に格納してもよい。これらの管理テーブル 2 6 0、2 6 5 を不揮発メモリに格納することによって、ストレージ装置 2 0 0 の電源遮断時にも仮想ストレージ装置の管理情報が保持される。

【 0 0 7 6 】

また、仮想チャネルアダプタ管理テーブル 2 6 0 及びストレージ資源管理テーブル 2 6 5 を物理ディスクドライブ 2 3 6 に格納し、ストレージ装置 2 0 0 の起動後に、これらの管理テーブル 2 6 0、2 6 5 を物理ディスクドライブ 2 3 6 から読み出して、主記憶メモリ 2 1 2 及び 2 2 2 に格納してもよい。

【 0 0 7 7 】

また、仮想チャネルアダプタ管理テーブル 2 6 0 及びストレージ資源管理テーブル 2 6 5 を管理端末 3 0 0 に格納し、ストレージ装置 2 0 0 の起動後に、これらの管理テーブル 2 6 0、2 6 5 を管理端末 3 0 0 から読み出して、主記憶メモリ 2 1 2 及び 2 2 2 に格納してもよい。

【 0 0 7 8 】

なお、ストレージハイパーバイザ 2 1 4、2 2 4 は、仮想ストレージ装置を構成する他のストレージ資源に関する情報を有している。

【 0 0 7 9 】

仮想ストレージ層は、ハイパーバイザ 2 1 4、2 2 4 によって、物理ストレージ 2 4 0 のストレージ資源が論理区画によって分割された仮想ストレージ装置 2 4 0、2 5 0 が構成されている。仮想ストレージ (0) 2 4 0 には、仮想チャネルアダプタ (0) 2 4 1、仮想チャネルアダプタ (1) 2 4 2、ディスクキャッシュ資源 2 4 3 及び仮想ディスク 2 4 4 が含まれる。同様に、仮想ストレージ (1) 2 5 0 には、仮想チャネルアダプタ (2) 2 5 1、仮想チャネルアダプタ (3) 2 5 2、ディスクキャッシュ資源 2 5 3 及び仮想ディスク 2 5 4 が含まれる。なお、仮想ストレージ装置 2 4 0、2 5 0 には、ストレージ装置 2 0 0 に備わる他のストレージ資源も含まれるが、本発明の説明に必要がないため、図示を省略する。

【 0 0 8 0 】

すなわち、仮想ストレージ装置 (0) 2 4 0 及び仮想ストレージ装置 (1) 2 5 0 は、ストレージハイパーバイザ 2 1 4、2 2 4 によって、物理ディスクドライブ 2 3 6 の領域が分割され複数の仮想ディスク 2 4 4、2 5 4 となったり、複数の物理ディスクドライブ 2 3 6 が統合され単一の仮想ディスク 2 4 4、2 5 4 となる。

【 0 0 8 1 】

そして、ストレージ装置 2 0 0 は、仮想ディスク 2 4 4、2 5 4 のうち、一つ又は複数の仮想ディスクを選択し、仮想計算機 1 0 0、1 5 0 に記憶領域として提供する。この選択された仮想ディスクを論理ユニット (L U : Logical Unit) という。論理ユニットとは、OS が一つのディスクとして認識できる単位を指す。

【 0 0 8 2 】

管理端末 3 0 0 は、仮想パス管理テーブル 3 1 0 (図 5) を有している。仮想パス管理テーブル 3 1 0 には、仮想計算機と仮想ストレージとの対応関係が記載されている。管理端末 3 0 0 は、サーバ装置 (0) 1 0 0 の仮想 I / O アダプタ管理テーブル 1 1 0 に含まれる管理情報、サーバ装置 (1) 1 5 0 の仮想 I / O アダプタ管理テーブル 1 6 0 に含まれる管理情報及びストレージ装置 2 0 0 の仮想チャネルアダプタ管理テーブル 2 6 0 に含まれる管理情報を収集することによって、仮想パス管理テーブル 3 1 0 を生成する。

【 0 0 8 3 】

図 3 A、図 3 B は、第 1 の実施の形態の仮想 I / O アダプタ管理テーブルの説明図である。前述したように、仮想 I / O アダプタ管理テーブルは、サーバ装置毎に設けられてい

10

20

30

40

50

る。仮想 I / O アダプタ管理テーブルには、物理的な I / O アダプタと仮想 I / O アダプタとの対応関係が記載されている。

【 0 0 8 4 】

図 3 A は、サーバ装置 (0) 1 0 0 の仮想 I / O アダプタ管理テーブル 1 1 0 の説明図である。

【 0 0 8 5 】

仮想 I / O アダプタ管理テーブル 1 1 0 には、I / O アダプタ番号 1 1 1、仮想 I / O アダプタ番号 1 1 2 及び仮想計算機番号 1 1 3 が含まれており、これらの対応が記憶されている。

【 0 0 8 6 】

I / O アダプタ番号 1 1 1 は物理的な I / O アダプタの識別子である。仮想 I / O アダプタ番号 1 1 2 は、仮想 I / O アダプタの識別子である。仮想計算機番号 1 1 3 は、サーバ装置 (0) 1 0 0 に設けられた仮想計算機の識別子である。

【 0 0 8 7 】

図 3 B は、サーバ装置 (1) 1 5 0 の仮想 I / O アダプタ管理テーブル 1 6 0 の説明図である。

【 0 0 8 8 】

仮想 I / O アダプタ管理テーブル 1 6 0 には、I / O アダプタ番号 1 6 1、仮想 I / O アダプタ番号 1 6 2 及び仮想計算機番号 1 6 3 が含まれており、これらの対応が記憶されている。

【 0 0 8 9 】

I / O アダプタ番号 1 6 1 は物理的な I / O アダプタの識別子である。仮想 I / O アダプタ番号 1 6 2 は、仮想 I / O アダプタの識別子である。仮想計算機番号 1 6 3 は、サーバ装置 (1) 1 5 0 に設けられた仮想計算機の識別子である。

【 0 0 9 0 】

この仮想 I / O アダプタ管理テーブル 1 1 0、1 6 0 によって、どの仮想計算機にどの仮想 I / O アダプタが含まれており、各仮想 I / O アダプタがどの物理的 I / O アダプタによって実現されるかが分かる。

【 0 0 9 1 】

例えば、仮想 I / O アダプタ管理テーブル 1 1 0 (図 3 A) の第 1 エントリによると、仮想計算機 (0) の仮想 I / O アダプタ (0) は、物理的 I / O アダプタ (0) の計算機資源によって実現される。また、仮想 I / O アダプタ管理テーブル 1 6 0 (図 3 B) の第 1 エントリによると、仮想計算機 (2) の仮想 I / O アダプタ (4) は、物理的 I / O アダプタ (2) の計算機資源によって実現される。

【 0 0 9 2 】

図 4 は、第 1 の実施の形態の仮想チャネルアダプタ管理テーブル 2 6 0 の説明図である。

【 0 0 9 3 】

仮想チャネルアダプタ管理テーブル 2 6 0 には、物理的なチャネルアダプタと仮想チャネルアダプタとの対応関係が記載されている。具体的には、仮想チャネルアダプタ管理テーブル 2 6 0 には、仮想ストレージ装置番号 2 6 1、仮想チャネルアダプタ番号 2 6 2 及びチャネルアダプタ番号 2 6 3 が含まれており、これらの対応が記憶されている。

【 0 0 9 4 】

仮想ストレージ装置番号 2 6 1 は、ストレージ装置 2 0 0 に設けられた仮想ストレージ装置の識別子である。仮想チャネルアダプタ番号 2 6 2 は、仮想チャネルアダプタの識別子である。チャネルアダプタ番号 2 6 3 は、物理的なチャネルアダプタの識別子である。

【 0 0 9 5 】

この仮想チャネルアダプタ管理テーブル 2 6 0 によって、どの仮想ストレージ装置に仮想チャネルアダプタが含まれており、各仮想チャネルアダプタがどの物理的チャネルアダプタによって実現されるかが分かる。

10

20

30

40

50

【 0 0 9 6 】

例えば、仮想チャネルアダプタ管理テーブル 2 6 0 の第 1 エントリによると、仮想ストレージ装置 (0) の仮想チャネルアダプタ (0) は、物理的チャネルアダプタ (0) のストレージ資源によって実現される。また、第 3 エントリによると、仮想ストレージ装置 (1) の仮想チャネルアダプタ (2) は、物理的チャネルアダプタ (0) のストレージ資源によって実現される。

【 0 0 9 7 】

図 5 は、第 1 の実施の形態の仮想パス管理テーブル 3 1 0 の説明図である。

【 0 0 9 8 】

仮想パス管理テーブル 3 1 0 には、仮想計算機と仮想ストレージとの対応関係が記載されている。特に、物理的な I / O アダプタと物理的なチャネルアダプタとの対応関係によって、サーバ装置とストレージ装置との間の物理パスが特定される。すなわち、仮想パス管理テーブル 3 1 0 には、仮想計算機と仮想ストレージと間の経路情報が含まれている。

10

【 0 0 9 9 】

管理端末 3 0 0 は、サーバ装置 (0) 1 0 0 の仮想 I / O アダプタ管理テーブル 1 1 0 に含まれる情報、サーバ装置 (1) 1 5 0 の仮想 I / O アダプタ管理テーブル 1 6 0 に含まれる情報及びストレージ装置 2 0 0 の仮想チャネルアダプタ管理テーブル 2 6 0 に含まれる情報を収集することによって、仮想パス管理テーブル 3 1 0 を生成する。すなわち、仮想パス管理テーブル 3 1 0 は、仮想 I / O アダプタ管理テーブル 1 1 0、仮想 I / O アダプタ管理テーブル 1 6 0 及び仮想チャネルアダプタ管理テーブル 2 6 0 が統合されたものであり、これらのテーブルの管理情報が含まれる。

20

【 0 1 0 0 】

具体的には、仮想パス管理テーブル 3 1 0 には、仮想ストレージ装置番号 3 1 1、仮想チャネルアダプタ番号 3 1 2、チャネルアダプタ番号 3 1 3、I / O アダプタ番号 3 1 4、仮想 I / O アダプタ番号 3 1 5 及び仮想計算機番号 3 1 6 が含まれており、これらの対応が記憶されている。

【 0 1 0 1 】

仮想ストレージ装置番号 3 1 1 は、仮想チャネルアダプタ管理テーブル 2 6 0 (図 4) の仮想ストレージ装置番号 2 6 1 に対応しており、ストレージ装置 2 0 0 に設けられた仮想ストレージ装置の識別子である。

30

【 0 1 0 2 】

仮想チャネルアダプタ番号 3 1 2 は、仮想チャネルアダプタ管理テーブル 2 6 0 (図 4) の仮想チャネルアダプタ番号 2 6 2 に対応しており、仮想チャネルアダプタの識別子である。

【 0 1 0 3 】

チャネルアダプタ番号 3 1 3 は、仮想チャネルアダプタ管理テーブル 2 6 0 (図 4) のチャネルアダプタ番号 2 6 3 に対応しており、物理的なチャネルアダプタの識別子である。

【 0 1 0 4 】

I / O アダプタ番号 3 1 4 は、仮想 I / O アダプタ管理テーブル 1 1 0、1 6 0 (図 3 A、図 3 B) の I / O アダプタ番号 1 1 1、1 6 1 に対応しており、物理的な I / O アダプタの識別子である。

40

【 0 1 0 5 】

仮想 I / O アダプタ番号 3 1 5 は、仮想 I / O アダプタ管理テーブル 1 1 0、1 6 0 (図 3 A、図 3 B) の仮想 I / O アダプタ番号 1 1 2、1 6 2 に対応しており、仮想 I / O アダプタの識別子である。

【 0 1 0 6 】

仮想計算機番号 3 1 6 は、仮想 I / O アダプタ管理テーブル 1 1 0、1 6 0 (図 3 A、図 3 B) の仮想計算機番号 1 1 3、1 6 3 に対応しており、仮想 I / O アダプタの識別子である。

50

【 0 1 0 7 】

この仮想パス管理テーブル 3 1 0 によって、どの仮想ストレージ装置の仮想チャンネルアダプタ（仮想チャンネルアダプタを実現する物理的チャンネルアダプタ）が、どの仮想計算機の仮想 I / O アダプタ（仮想 I / O アダプタを実現する物理的 I / O アダプタ）と接続されているかが分かる。

【 0 1 0 8 】

すなわち、チャンネルアダプタ番号 3 1 3 と I / O アダプタ番号 3 1 4 との対応によって、物理的な I / O アダプタと物理的なチャンネルアダプタとの接続関係が規定される。この接続関係によって、サーバ装置とストレージ装置との間の物理パスの経路情報が分かる。

【 0 1 0 9 】

例えば、仮想パス管理テーブル 3 1 0 の第 1 エントリによると、仮想ストレージ装置（0）の仮想チャンネルアダプタ（0）は、物理的チャンネルアダプタ（0）のストレージ資源によって実現される。また、仮想計算機（0）の仮想 I / O アダプタ（0）は、物理的 I / O アダプタ（0）の計算機資源によって実現される。また、この物理的チャンネルアダプタ（0）で実現される仮想チャンネルアダプタ（0）は、物理的 I / O アダプタ（0）で実現される仮想 I / O アダプタ（0）と対応している。

【 0 1 1 0 】

同様に、仮想パス管理テーブル 3 1 0 の第 2 エントリによると、物理的チャンネルアダプタ（1）で実現される仮想チャンネルアダプタ（1）は、物理的 I / O アダプタ（1）で実現される仮想 I / O アダプタ（1）と対応している。

【 0 1 1 1 】

すなわち、仮想ストレージ装置（0）と仮想計算機（0）とは、物理的チャンネルアダプタ（0）と物理的 I / O アダプタ（0）とによって接続され、さらに、物理的チャンネルアダプタ（1）と物理的 I / O アダプタ（1）とによっても接続される。

【 0 1 1 2 】

図 6 は、第 1 の実施の形態の計算機資源管理テーブル 1 1 5 の説明図である。

【 0 1 1 3 】

前述したように、計算機資源管理テーブル 1 1 5 は、ハイパーバイザ 1 0 3（サーバ装置（0）1 0 0 内）に設けられる。

【 0 1 1 4 】

計算機資源管理テーブル 1 1 5 には、仮想計算機番号 7 0 1、CPU 配分率 7 0 2、メモリ容量 7 0 3 及び I / O アダプタ番号 7 0 4 が対応するように格納されている。計算機資源管理テーブル 1 1 5 は、サーバ装置（0）1 0 0 に備わる資源（CPU（0）1 0 1、主記憶メモリ（0）1 0 4、I / O アダプタ 1 0 6、1 0 7）の対応関係を記憶している。

【 0 1 1 5 】

仮想計算機番号 7 0 1 は、サーバ装置（0）1 0 0 に設けられた仮想計算機を示す。CPU 配分率 7 0 2 は、サーバ装置（0）1 0 0 に設けられた CPU（0）1 0 1 のうち当該仮想計算機に割り当てられる割合を示す。メモリ容量 7 0 3 は、当該仮想計算機に割り当てられる主記憶メモリ（0）1 0 4 の容量である。I / O アダプタ番号 7 0 4 は、当該仮想計算機からストレージ装置 2 0 0 へのアクセスを受け持つ I / O アダプタを示す。

【 0 1 1 6 】

計算機資源管理テーブル 1 1 5 は、管理者が管理端末 3 0 0 から、計算機資源の割り当てを設定することによって生成される。

【 0 1 1 7 】

なお、サーバ装置（1）1 5 0（ハイパーバイザ 1 5 3）に設けられる計算機資源管理テーブル 1 6 5 は、前述した計算機資源管理テーブル 1 1 5 と同じ項目を有し、サーバ装置（1）1 5 0 に備わる資源の対応関係を記憶している。

【 0 1 1 8 】

また、管理端末 3 0 0 は、計算機資源管理テーブル 3 2 0 を有する。計算機資源管理テ

10

20

30

40

50

ーブル 3 2 0 は、前述した計算機資源管理テーブル 1 1 5 と同じ項目を有し、サーバ装置 (0) 1 0 0 に備わる資源の対応関係と、サーバ装置 (1) 1 5 0 に備わる資源の対応関係を記憶している。

【 0 1 1 9 】

図 7 は、第 1 の実施の形態のストレージ資源管理テーブル 2 6 5 の説明図である。

【 0 1 2 0 】

前述したように、ストレージ資源管理テーブル 2 6 5 は、ストレージハイパーバイザ 2 1 4、2 2 4 (ストレージ装置 2 0 0 内) に設けられる。

【 0 1 2 1 】

ストレージ資源管理テーブル 2 6 5 には、仮想計算機番号 6 0 1、仮想ディスク番号 6 0 2、ディスクキャッシュ容量 6 0 3、CPU 番号 6 0 4 及び I / O アダプタ番号 6 0 5 が対応するように格納されている。ストレージ資源管理テーブル 2 6 5 は、ストレージ装置 2 0 0 の資源 (物理ディスクドライブ 2 3 6、CPU 2 1 1、2 2 1、チャンネルアダプタ 2 3 1 ~ 2 3 4、ディスクキャッシュ 2 3 8) と仮想計算機との対応関係を記憶している。

【 0 1 2 2 】

仮想計算機番号 6 0 1 は、サーバ装置 1 0 0、1 5 0 に設けられた仮想計算機に対応する。仮想ディスク番号 6 0 2 は、ストレージハイパーバイザ 2 1 4、2 2 4 によって構成された仮想ディスク 2 4 4、2 5 4 の番号であり、仮想計算機番号 6 0 1 に規定された仮想計算機に割り当てられた仮想ディスクを示す。すなわち、仮想ディスク 2 4 4 は、仮想ディスク番号 1 2 1 及び 1 2 2 に対応する領域を備え、仮想ディスク 2 5 4 は、仮想ディスク番号 1 6 及び 1 7 に対応する領域を備えている。なお、仮想ディスクは、論理ユニットであってもよい。

【 0 1 2 3 】

ディスクキャッシュ容量 6 0 3 は、仮想計算機番号 6 0 1 に規定された仮想計算機に割り当てられたディスクキャッシュ 2 4 3、2 5 3 の容量である。CPU 番号 6 0 4 は、仮想計算機番号 6 0 1 に規定された仮想計算機からのアクセスを制御する制御用 CPU 2 1 1、2 2 1 を示す。

【 0 1 2 4 】

チャンネルアダプタ番号 6 0 5 は、仮想計算機番号 6 0 1 に規定された仮想計算機からのアクセス (仮想ディスク番号 6 0 2 に規定された仮想ディスクへのアクセス) を受け持つ チャンネルアダプタ 2 3 1 ~ 2 3 4 を示す。

【 0 1 2 5 】

すなわち、仮想計算機 (0) 1 3 0 には、1 2 1 番 ~ 1 2 2 番の 2 個の仮想ディスクが割り当てられている。この 1 2 1 番 ~ 1 2 2 番の仮想ディスクに対してアクセスするために仮想計算機 (0) 1 3 0 は 5 1 2 M バイトのディスクキャッシュを使用することができる。仮想計算機 (0) 1 3 0 から 1 2 1 番 ~ 1 2 2 番の仮想ディスクに対して 0 番 ~ 1 番の 2 個の I / O アダプタを介してアクセスされる。仮想計算機 (0) 1 3 0 から 1 2 1 番 ~ 1 2 2 番の仮想ディスクに対するアクセスを処理するために 2 番 ~ 3 番の 2 個の CPU が動作することを意味している。

【 0 1 2 6 】

ストレージ資源管理テーブル 2 6 5 は、管理者が管理端末 3 0 0 から、ストレージ資源の割り当てを設定することによって生成される。

【 0 1 2 7 】

また、管理端末 3 0 0 は、ストレージ資源管理テーブル 3 3 0 を有する。ストレージ資源管理テーブル 3 3 0 は、前述したストレージ資源管理テーブル 2 6 5 と同じ項目を有し、ストレージ装置 2 0 0 に備わる資源の対応関係を記憶している。

【 0 1 2 8 】

なお、計算機資源管理テーブル 1 1 5 (図 6) に CPU 配分率を記載しているが、CPU 番号を用いてもよい。また、ストレージ資源管理テーブル 2 6 5 (図 7) に CPU 番号

10

20

30

40

50

を記載しているが、CPU配分率を用いてもよい。

【0129】

図8は、第1の実施の形態の計算機システムの接続状態の説明図である。図8は、計算機システムの機能をハードウェア的に示した図2と異なり、ソフトウェア的観点からの（仮想世界からみた）接続状態を示す。よって、図8には、計算機システムのハードウェア構成（例えば、I/Oアダプタ106等、チャンネルアダプタ231等）は図示されない。

【0130】

仮想計算機(0)130は、二つの仮想I/Oアダプタ(0)131及び仮想I/Oアダプタ(1)132を備える。仮想ストレージ装置(0)240は、二つの仮想チャンネルアダプタ(0)241及び仮想チャンネルアダプタ(1)242を備える。仮想I/Oアダプタ(0)131は仮想チャンネルアダプタ(0)241と接続され、仮想I/Oアダプタ(1)132は仮想チャンネルアダプタ(1)242と接続されている。

10

【0131】

同様に、仮想計算機(1)140は、二つの仮想I/Oアダプタ(2)141及び仮想I/Oアダプタ(3)142を備える。仮想ストレージ装置(1)250は、二つの仮想チャンネルアダプタ(2)251及び仮想チャンネルアダプタ(3)252を備える。仮想I/Oアダプタ(2)141は仮想チャンネルアダプタ(2)251と接続され、仮想I/Oアダプタ(3)142は仮想チャンネルアダプタ(3)252と接続されている。

【0132】

また、仮想計算機(2)180は、二つの仮想I/Oアダプタ(4)181及び仮想I/Oアダプタ(5)182を備える。仮想I/Oアダプタ(4)181は仮想チャンネルアダプタ(0)241と接続されており、仮想I/Oアダプタ(5)182は仮想チャンネルアダプタ(1)242と接続されている。

20

【0133】

同様に、仮想計算機(3)190は、二つの仮想I/Oアダプタ(6)191及び仮想I/Oアダプタ(7)192を備える。仮想I/Oアダプタ(6)191は仮想チャンネルアダプタ(2)251と接続されており、仮想I/Oアダプタ(7)192は仮想チャンネルアダプタ(3)252と接続されている。

【0134】

換言すれば、図8には、仮想バス管理テーブル310（図5）の仮想計算機番号316、仮想I/Oアダプタ番号315、仮想チャンネルアダプタ番号312及び仮想ストレージ装置番号311の対応関係が示されている。

30

【0135】

図9は、第1の実施の形態の計算機システムのケーブル接続時処理のフローチャートである。このケーブル接続時処理はサーバ装置(0)100で実行される。なお、サーバ装置(1)150及びストレージ装置200においても同じ処理が実行される。

【0136】

まず、サーバ装置(0)100のI/Oアダプタ（ストレージ装置200のチャンネルアダプタ）が、ケーブルが接続されたことを検出する（S101）。

【0137】

次に、サーバ装置(0)100は、通信可能な装置と物理アドレスを交換する（S102）。例えば、サーバ装置(0)100がストレージ装置200に接続されている場合、サーバ装置(0)100が接続相手に物理アドレスを問い合わせることによって、ストレージ装置200のチャンネルアダプタの物理アドレスを取得する。ステップS102で交換される物理アドレスは、ケーブル（I/Oチャンネル）が接続されるポートが一意に特定されるものであればよい。

40

【0138】

例えば、ファイバチャンネルプロトコルではワールド・ワイド・ネーム（WWN）を交換する。特に、本発明では、ポートを特定するために、ワールド・ワイド・ポート・ネームを用いるとよい。また、iSCSIプロトコルでは、アドレス解決手法（例えば、ARP

50

）によって求めたMACアドレスを交換するとよい。

【0139】

そして、取得した物理アドレスをハイパーバイザ103に通知する。

【0140】

次に、ハイパーバイザ103は、取得したケーブル接続状況（I/Oアダプタの物理アドレスとチャンネルアダプタの物理アドレス）を、ネットワーク410経由で管理端末300に報告する（S103）。

【0141】

管理端末300は、I/Oアダプタの物理アドレスとチャンネルアダプタの物理アドレスを受信すると、接続相手であるチャンネルアダプタを備えるストレージ装置からの物理アドレスの報告を待つ。

10

【0142】

そして、接続された双方の装置から接続相手の物理アドレスを受信すると、I/Oアダプタとチャンネルアダプタとの対応関係を仮想バス管理テーブル310に登録する（S104）。すなわち、管理端末300は、検出されたケーブル数分のチャンネルアダプタ番号313及びI/Oアダプタ番号314を、仮想バス管理テーブル310に登録する。

【0143】

図10は、第1の実施の形態の計算機システムの仮想計算機生成処理のフローチャートである。この仮想計算機生成処理はサーバ装置（0）100で実行される。なお、サーバ装置（1）150においても同じ処理が実行される。

20

【0144】

まず、管理者が管理端末300を操作して、サーバ装置（0）100に備わる計算機資源の割り当てを設定して、仮想計算機の生成を指示する（S111）。管理端末300は、設定された資源割り当てをサーバ装置（0）100に送信する。

【0145】

サーバ装置（0）100は、管理端末300から仮想計算機生成指示を受信すると、ハイパーバイザ103は、受信した仮想計算機生成指示に従って、サーバ装置（0）100内のI/Oアダプタ以外の計算機資源を割り当て、仮想計算機を生成する（S112）。

【0146】

そして、ハイパーバイザ103は、仮想計算機の生成が終了すると、管理端末300に仮想計算機の生成が完了した旨を報告する。

30

【0147】

管理端末300は、サーバ装置（0）100から仮想計算機の生成が完了した旨の報告を受信すると、ストレージ装置200に仮想ストレージ装置を生成する必要があるか否かを判定する（S113）。

【0148】

その結果、仮想ストレージ装置を生成する必要があると判定されると、ステップS114に進む。一方、仮想ストレージ装置を生成する必要があると判定されず、ステップS115に進む。

【0149】

例えば、サーバ装置（0）100に初めて仮想計算機が生成されるときは、仮想計算機がアクセスする仮想ストレージ装置を生成する必要がある。一方、生成される仮想計算機が、既に生成されている仮想ストレージ装置を使用するのであれば、新たに仮想ストレージ装置を生成する必要がある。

40

【0150】

ステップS114では、管理端末300は、設定されたストレージ装置200に備わるストレージ資源の割り当てを、ストレージ装置200に送信して、仮想ストレージ装置の生成を指示する。ストレージ装置200が管理端末300から仮想ストレージ装置生成指示を受信すると、ストレージハイパーバイザ214、224は、受信した仮想ストレージ装置生成指示に従って、ストレージ装置200内のチャンネルアダプタ以外のストレージ資

50

源を割り当て、仮想ストレージ装置を生成する (S 1 1 4)。

【 0 1 5 1 】

そして、ストレージハイパーバイザ 2 1 4、2 2 4 は、仮想ストレージ装置の生成が終了すると、管理端末 3 0 0 に仮想ストレージ装置の生成が完了した旨を報告する。

【 0 1 5 2 】

管理端末 3 0 0 は、ストレージ装置 2 0 0 から仮想ストレージ装置の生成が完了した旨の報告を受信すると、仮想アダプタ (仮想チャネルアダプタ及び仮想 I / O アダプタ) を生成する必要があるか否かを判定する (S 1 1 5)。

【 0 1 5 3 】

その結果、仮想アダプタを生成する必要があると判定されると、ステップ S 1 1 6 に進む。一方、仮想アダプタを生成する必要があると判定されないと、ステップ S 1 1 7 に進む。

10

【 0 1 5 4 】

例えば、サーバ装置 (0) 1 0 0 に初めて仮想計算機が生成されるときは、仮想計算機が使用する仮想アダプタを生成する必要がある。一方、既に生成された仮想アダプタを使用する場合、又は、冗長ポートを使用する場合には、仮想アダプタを生成する必要がある。

【 0 1 5 5 】

ステップ S 1 1 6 では、I / O アダプタ (又は、チャネルアダプタ) の資源を割り当て、仮想アダプタを生成する (S 1 1 6)。仮想 I / O アダプタはハイパーバイザ 1 0 3、1 5 3 が生成し、仮想チャネルアダプタはストレージハイパーバイザ 2 1 4、2 2 4 が生成する。

20

【 0 1 5 6 】

ハイパーバイザ 1 0 3 は、仮想 I / O アダプタを生成すると、仮想 I / O アダプタ管理テーブル 1 1 0 にエントリを追加する。このとき、仮想 I / O アダプタ管理テーブル 1 1 0 には、仮想 I / O アダプタ番号 1 1 2 及び仮想計算機番号 1 1 3 が登録される。なお、ステップ S 1 1 2 では I / O アダプタの資源が割り当てられていないので、I / O アダプタ番号 1 1 1 は空白である。

【 0 1 5 7 】

また、ストレージハイパーバイザ 2 1 4 は、仮想チャネルアダプタを生成すると、仮想チャネルアダプタ管理テーブル 2 6 0 にエントリを追加する。このとき、仮想チャネルアダプタ管理テーブル 2 6 0 には、仮想ストレージ装置番号 2 6 1 及び仮想チャネルアダプタ番号 2 6 2 が登録される。なお、ステップ S 1 1 4 ではチャネルアダプタの資源が割り当てられていないので、チャネルアダプタ番号 2 6 3 は空白である。

30

【 0 1 5 8 】

その後、チャネルアダプタと仮想チャネルアダプタとの間のパスを生成する (S 1 1 7)。すなわち、仮想チャネルアダプタ番号 2 6 2 に対応するチャネルアダプタ番号 2 6 3 が、仮想チャネルアダプタ管理テーブル 2 6 0 に登録される。これによって、仮想チャネルアダプタ管理テーブル 2 6 0 が完成し、サーバ装置 (0) 1 0 0 内のパスが形成される。

40

【 0 1 5 9 】

次に、ストレージハイパーバイザ 2 1 4、2 2 4 は、仮想チャネルアダプタ管理テーブル 2 6 0 の仮想ストレージ装置番号 2 6 1、仮想チャネルアダプタ番号 2 6 2 及びチャネルアダプタ番号 2 6 3 を、管理端末 3 0 0 に送る。管理端末 3 0 0 は、仮想チャネルアダプタ管理テーブル 2 6 0 に含まれる情報を受信すると、仮想パス管理テーブル 3 1 0 のチャネルアダプタ番号 3 1 3 に対応するエントリを検索する。そして、検索されたエントリの仮想ストレージ装置番号 3 1 1 及び仮想チャネルアダプタ番号 3 1 2 に、該受信した情報を登録する (S 1 1 8)。これによって、管理端末 3 0 0 は、仮想ストレージ装置からチャネルアダプタまでの関係を管理可能になる。

【 0 1 6 0 】

50

次に、管理者が、管理端末300を用いて、目的とするチャンネルアダプタと接続されているI/Oアダプタを選択する(S119)。このとき、仮想計算機管理プログラム301が、目的とするチャンネルアダプタに物理的に接続されているI/Oアダプタ(一つ又は複数)を表示して、管理者がその中から選択することもできる。管理端末300は、選択されたI/Oアダプタの情報をハイパーバイザ103に送る。

【0161】

ハイパーバイザ103は、選択されたI/Oアダプタの情報を受信すると、I/Oアダプタと仮想I/Oアダプタとの間のパスを生成する(S120)。すなわち、仮想I/Oアダプタ番号112に対応するI/Oアダプタ番号111が、仮想I/Oアダプタ管理テーブル110に登録される。これによって、仮想I/Oアダプタ管理テーブル110が完成し、サーバ装置内のパスが形成される。

10

【0162】

次に、ハイパーバイザ103は、仮想I/Oアダプタ管理テーブル110のI/Oアダプタ番号111仮想I/Oアダプタ番号112及び仮想計算機番号113を、管理端末300に送る。管理端末300は、仮想I/Oアダプタ管理テーブル110に含まれる情報を受信すると、I/Oアダプタ番号に対応するエントリを検索する。そして、検索されたエントリの仮想I/Oアダプタ番号315及び仮想計算機番号316に、該受信した情報を登録する(S121)。これによって、仮想ストレージ装置から仮想計算機までの経路が仮想パス管理テーブル310に登録される。そして、管理端末300は、仮想ストレージ装置から仮想計算機までの関係を管理可能になる。

20

【0163】

その後、冗長パスを生成する必要があるか否かを判定する(S122)。冗長パスを生成する必要があると判定されると、ステップS115に戻り、冗長パスにおいて使用する仮想アダプタを生成する必要があるか否かを判定する。その後、ステップS116～S122の処理によって冗長パスが生成される。

【0164】

なお、冗長パスを生成する際に新たにアダプタを生成する(冗長アダプタが存在しない)場合には、ステップS116に戻るようにしてもよい。また、冗長パスを生成する際に新たなアダプタを生成しない(冗長アダプタが存在する)場合には、ステップS117に戻るようにしてもよい。

30

【0165】

冗長パスの生成が終了すると、冗長構成を生成する必要があるか否かを判定する(S123)。冗長構成を生成する必要があると判定されると、ステップS112に戻り、ステップS112～S122の処理によって冗長構成において使用する仮想計算機が生成される。

【0166】

図11は、第1の実施の形態の計算機システムのI/Oチャンネル400の通信プロトコル層の構造の説明図である。

【0167】

仮想計算機(0)130上のOS(0)135は、ストレージ装置200内の論理ユニットにアクセスする場合、ディスクI/Oプロトコル(例えば、SCSI規格)に従って入出力を行う。本実施の形態では、ディスクI/Oプロトコル層を「ディスクI/O層」1100、1106と称す。

40

【0168】

OS(0)135が発行したディスクI/Oコマンドは、ハイパーバイザ103によって受け取られるが、ハイパーバイザ103とストレージハイパーバイザ214、224との間にも通信プロトコル層が存在する。これを「ハイパーバイザ通信層」1101、1105と称す。

【0169】

さらに、本実施の形態では、I/Oチャンネル400による汎用的な通信を行う層を「I

50

「Ｉ／Ｏチャンネルプロトコル層」１１０２、１１０４と称す。さらに、物理的な通信媒体などのハードウェアの層を「物理層」１１０３と称す。

【０１７０】

Ｉ／Ｏチャンネル４００の通信プロトコルがこのような層構造をなすことによって、Ｉ／Ｏチャンネル４００を構成する通信媒体が変わっても、ディスクＩ／Ｏ層１１００、１１０６やハイパーバイザ通信層１１０１、１１０５が影響を受けない。

【０１７１】

ＯＳ（０）１３５が発行したディスクＩ／Ｏコマンドは、仮想計算機（０）１３０に伝えられる。仮想計算機（０）１３０は仮想ストレージ装置（０）２４０に対して当該Ｉ／Ｏコマンドを発行する。この当該Ｉ／Ｏコマンドは、実際にはハイパーバイザ１０３が受け取る。

10

【０１７２】

ハイパーバイザ１０３は、当該ディスクＩ／Ｏコマンドに情報を付加して（図１２参照）、ストレージハイパーバイザ２１４、２２４に送信する。ストレージハイパーバイザ２１４、２２４はこれを受信し、当該ディスクＩ／Ｏコマンドを抽出して仮想ストレージ装置（０）２４０に送信する。

【０１７３】

このような層構造で通信をすることによって、ＯＳ（０）１３５は直接仮想ストレージ装置（０）と通信しているように認識する。

【０１７４】

20

図１２は、第１の実施の形態のサーバ装置（０）１００とストレージ装置２００との間で伝達される情報の説明図である。

【０１７５】

本実施の形態のＩ／Ｏチャンネル４００においては、フレーム１２００を単位として通信が行われる。これは一般的なファイバチャネル、イーサネットと同様である。フレーム１２００は、Ｉ／Ｏチャンネルプロトコルヘッダ１２０１及びＩ／Ｏチャンネルプロトコルペイロード１２０２からなる。

【０１７６】

Ｉ／Ｏチャンネルプロトコルヘッダ１２０１は、Ｉ／Ｏチャンネルプロトコル層１１０２、１１０４での通信に必要な制御情報を含む。具体的には、Ｉ／Ｏチャンネルプロトコルヘッダ１２０１は、送信元の識別子及び送信先の識別子を含む。Ｉ／Ｏチャンネルプロトコルペイロード１２０２は、Ｉ／Ｏチャンネルプロトコル層１１０２、１１０４において伝達される情報であり、Ｉ／Ｏチャンネルプロトコル層１１０２、１１０４ではその内容に関知しない。

30

【０１７７】

Ｉ／Ｏチャンネルプロトコルペイロード１２０２は、ハイパーバイザ通信ヘッダ１２０３及びハイパーバイザ通信ペイロード１２０４からなる。

【０１７８】

ハイパーバイザ通信ヘッダ１２０３には、ハイパーバイザ通信層１１０１、１１０５での通信に必要な制御情報を含んでいる。このハイパーバイザ通信ヘッダ１２０３の内容は図１３において詳述する。ハイパーバイザ通信ペイロード１２０４は、ハイパーバイザ通信層１１０１、１１０５において伝達されるデータであり、ハイパーバイザ通信層１１０１、１１０５ではその内容には関知しない。

40

【０１７９】

本実施の形態では、ハイパーバイザ通信ペイロード１２０４は、ディスクＩ／Ｏ層１１００、１１０６の通信に必要な情報からなる。具体的には、ディスクＩ／Ｏコマンドや、転送されるデータである。本実施の形態では、ハイパーバイザ通信層１１０１、１１０５の上位層がディスクＩ／Ｏ層であるため、ハイパーバイザ通信ペイロード１２０４にディスクＩ／Ｏ層１１００、１１０６の情報が含まれる。しかし、ハイパーバイザ、ストレージハイパーバイザ間の通信であれば、ディスクＩ／Ｏではない別の形式の情報が含まれる

50

。すなわち、ハイパーバイザ、ストレージハイパーバイザ間でディスク I / O 以外の目的の通信を行う場合、その通信にはディスク I / O プロトコル以外のプロトコルが規定されている。前述したディスク I / O ではない別の形式の情報は、そのプロトコルに従った情報である。

【 0 1 8 0 】

図 1 3 は、第 1 の実施の形態のハイパーバイザ通信ヘッダ 1 2 0 3 の内容の説明図である。

【 0 1 8 1 】

ハイパーバイザ通信ヘッダ 1 2 0 3 は、送信元ハイパーバイザ番号 1 3 0 0、送信元仮想計算機番号 1 3 0 1、送信元仮想アダプタ番号 1 3 0 2、送信先ハイパーバイザ番号 1 3 1 0、送信先仮想計算機番号 1 3 1 1 及び送信先仮想アダプタ番号 1 3 1 2 を含む。本実施の形態では、サーバ装置 1 0 0、1 5 0 及びストレージ装置 2 0 0 が、それぞれ複数存在するシステムにも対応できるように、ハイパーバイザ、ストレージハイパーバイザに一意的な識別子を付与する。

【 0 1 8 2 】

送信元ハイパーバイザ番号 1 3 0 0 は、本フレームを送信したハイパーバイザ又はストレージハイパーバイザの識別子である。

【 0 1 8 3 】

送信先ハイパーバイザ番号 1 3 1 0 は、本フレームを受信すべきハイパーバイザ又はストレージハイパーバイザの識別子である。

【 0 1 8 4 】

送信元仮想計算機番号 1 3 0 1 は、本フレームを送信した仮想計算機、又は仮想ストレージ装置の識別子である。

【 0 1 8 5 】

送信先仮想計算機番号 1 3 1 1 は、本フレームを受信すべき仮想計算機、又は仮想ストレージ装置の識別子である。

【 0 1 8 6 】

送信元仮想アダプタ番号 1 3 0 2 は、本フレームを送信した仮想 I / O アダプタ、又は仮想チャネルアダプタの識別子である。

【 0 1 8 7 】

送信先仮想アダプタ番号 1 3 1 2 は、本フレームを受信すべき仮想 I / O アダプタ、又は仮想チャネルアダプタの識別子である。

【 0 1 8 8 】

図 1 4 は、第 1 の実施の形態のディスク I / O コマンド送信処理のフローチャートである。なお、仮想計算機 (0) 1 3 0 が I / O アダプタ (0) 1 0 6 及びチャネルアダプタ (0) 2 3 1 を経由して、仮想ストレージ装置 (0) 2 4 0 にアクセスする場合を説明するが、他の仮想計算機 1 4 0、1 8 0、1 9 0 や他の仮想ストレージ装置 2 5 0 でも同じ処理を実行することができる。

【 0 1 8 9 】

O S (0) 1 3 5 は、ディスク I / O コマンドが発行されると、当該ディスク I / O コマンドを仮想計算機 (0) 1 3 0 (仮想 I / O アダプタ (0) 1 3 1) に伝える。(S 1 3 1)。仮想 I / O アダプタ (0) 1 3 1 に送られたディスク I / O コマンドは、実際にはハイパーバイザ 1 0 3 が受け取る。

【 0 1 9 0 】

ハイパーバイザ 1 0 3 は、当該ディスク I / O コマンドに情報を付加して (図 1 2 参照)、物理的な I / O アダプタ (0) 1 0 6 に転送する (S 1 3 2)。

【 0 1 9 1 】

I / O アダプタ (0) 1 0 6 は、接続されたチャネルアダプタ (0) 2 3 1 (仮想ストレージ装置 (0) 2 4 0) に当該ディスク I / O コマンドを送信する (S 1 3 3)。

【 0 1 9 2 】

そして、チャンネルアダプタ(0)231は、I/Oアダプタ(0)106が送信したディスクI/Oコマンドを受信する(S134)。チャンネルアダプタ(0)231が受信したディスクI/Oコマンドは、実際にはストレージハイパーバイザ214が受け取る。

【0193】

ストレージハイパーバイザ214は、受信したディスクI/Oコマンドのハイパーバイザ通信ヘッダ1203に含まれる送信先仮想アダプタ番号1312を抽出する。そして、ストレージハイパーバイザ214は抽出した仮想チャンネルアダプタ(0)241に、当該ディスクI/Oコマンドを転送する(S135)。

【0194】

図15は、第1の実施の形態のディスクI/Oコマンド完了報告処理のフローチャートである。なお、仮想ストレージ装置(0)240が、チャンネルアダプタ(0)231及びI/Oアダプタ(0)106を経由して仮想計算機(0)130にアクセスする場合を説明するが、他の仮想ストレージ装置や他の仮想計算機140、180、190でも同じ処理を実行することができる。

【0195】

仮想ストレージ装置(0)240は、仮想計算機(0)130からディスクI/Oコマンドを受信すると、当該ディスクI/Oコマンドによるアクセスの対象となる仮想ディスク番号を取得して、アクセス先の仮想ディスクを特定して、対応する仮想ディスク244にデータを読み書きする。

【0196】

このときストレージハイパーバイザ214は、アクセス対象の仮想ディスク224の仮想ブロックアドレスを、物理ディスクドライブ236の物理ブロックアドレスに変換する。そして、仮想ディスク244に対するアクセスを物理ディスクドライブ236に対するアクセスに変換する。そして、ストレージハイパーバイザ214が物理ディスクドライブ236にアクセスし、データの読み出し処理又は書き込み処理を行う。

【0197】

そして、仮想ストレージ装置(0)240は、物理ディスクに対するデータ読み書き処理が終了すると、ディスクI/O完了報告を発行する。すると、仮想ストレージ装置(0)240は、当該ディスクI/O完了報告を仮想チャンネルアダプタ(0)241に伝える(S141)。仮想チャンネルアダプタ(0)241に送られたディスクI/O完了報告は、実際にはストレージハイパーバイザ214が受け取る。

【0198】

ストレージハイパーバイザ214は、当該ディスクI/O完了報告に情報を付加して(図12参照)、物理的なチャンネルアダプタ(0)231に転送する(S142)。

【0199】

チャンネルアダプタ(0)231は、接続されたI/Oアダプタ(0)106(仮想計算機(0)130)に当該ディスクI/O完了報告を送信する(S143)。

【0200】

そして、I/Oアダプタ(0)106は、チャンネルアダプタ(0)231が送信したディスクI/O完了報告を受信する(S144)。I/Oアダプタ(0)106が受信したディスクI/O完了報告は、実際にはハイパーバイザ103が受け取る。

【0201】

ハイパーバイザ103は、受信したディスクI/O完了報告のハイパーバイザ通信ヘッダ1203に含まれる送信先仮想アダプタ番号1312を抽出する。そして、ハイパーバイザ103は抽出した仮想I/Oアダプタ(0)131に、当該ディスクI/O完了報告を転送する(S145)。

【0202】

図16は、第1の実施の形態の障害処理のフローチャートである。なお、ハイパーバイザ103がI/Oチャンネル400の障害を検出する場合を説明するが、ハイパーバイザ153でも同様に障害を検出することができる。

【 0 2 0 3 】

まず、ハイパーバイザ 1 0 3 が I / O チャンネル 4 0 0 の断線を検出する (S 1 5 1) 。
なお、I / O チャンネル 4 0 0 の断線が検出されなければ、この障害処理は実行されない。

【 0 2 0 4 】

I / O チャンネル 4 0 0 の断線は、例えば、ディスク I / O コマンドを発行したサーバ装置 (0) 1 0 0 の仮想計算機 1 3 0 が、ディスク I / O 完了報告を受信せず、タイムアウトとなったときに、I / O チャンネル 4 0 0 が断線したと判定する。また、サーバ装置 (0) 1 0 0 に I / O チャンネル 4 0 0 の断線を検出する断線検出回路を設けて断線を監視してもよい。

【 0 2 0 5 】

すなわち、この障害処理は、ディスク I / O が発生したタイミングで実行されてもよいし、ディスク I / O が発生したタイミング以外で実行されてもよい。

【 0 2 0 6 】

次に、ハイパーバイザ 1 0 3 は、仮想計算機 (0) 1 3 0 と仮想ストレージ装置 (0) 2 4 0 との接続を維持するため、仮想 I / O アダプタ (0) 1 3 1 と物理的な I / O アダプタ (0) 1 0 6 との接続パス (サーバ装置 (0) 1 0 0 内の接続パス) を変更する (S 1 5 2) 。例えば、仮想 I / O アダプタ (0) 1 3 1 と I / O アダプタ (1) 1 0 7 との接続パスを設定する。

【 0 2 0 7 】

これに伴い、仮想 I / O アダプタ管理テーブル 1 1 0 の内容が変更される。本実施の形態では、I / O アダプタ番号 1 1 1 が " 0 " 、仮想 I / O アダプタ番号 1 1 2 が " 0 " 、仮想計算機番号 3 1 3 が " 0 " のエントリにおいて、I / O アダプタ番号 1 1 1 が " 1 " に変更される。

【 0 2 0 8 】

なお、障害が検出されたエントリの内容を変更せず、新たに設定したパスのエントリを追加してもよい。

【 0 2 0 9 】

その後、ハイパーバイザ 1 0 3 は、仮想 I / O アダプタと物理的な I / O アダプタとの接続パス (サーバ装置 (0) 1 0 0 内の接続パス) の変更を管理端末 3 0 0 に通知する (S 1 5 3) 。

【 0 2 1 0 】

ハイパーバイザ 1 0 3 からのパス変更通知を受信した管理端末 3 0 0 は、仮想パス管理テーブル 3 1 0 を変更する (S 1 5 4) 。具体的には、I / O アダプタ番号 3 1 4 が " 0 " 、仮想 I / O アダプタ番号 3 1 5 が " 0 " 、仮想計算機番号 3 1 6 が " 0 " のエントリの、I / O アダプタ番号 3 1 4 が " 1 " に変更される。

【 0 2 1 1 】

その後、管理端末 3 0 0 は、仮想パス管理テーブル 3 1 0 を変更が終了した旨を、ハイパーバイザ 1 0 3 に通知する。

【 0 2 1 2 】

管理端末 3 0 0 から変更終了通知を受信したハイパーバイザ 1 0 3 は、未完了の送信情報 (例えば、完了報告を受信していない発行済のディスク I / O コマンド) を、新たなパスを構成する I / O アダプタに再送する (S 1 5 5) 。

【 0 2 1 3 】

図 1 7 は、第 1 の実施の形態の障害処理のフローチャートである。なお、ストレージハイパーバイザ 2 1 4 が、I / O チャンネルプロトコル層 1 1 0 4 で、I / O チャンネル 4 0 0 の障害を検出する場合を説明するが、ストレージハイパーバイザ 2 2 4 でも同様に障害を検出することができる。

【 0 2 1 4 】

まず、ストレージハイパーバイザ 2 1 4 が、I / O チャンネル 4 0 0 の断線を検出する (S 1 6 1) 。なお、I / O チャンネル 4 0 0 の断線が検出されなければ、この障害処理は実

10

20

30

40

50

行されない。

【0215】

I/Oチャネル400の断線は、例えば、ディスクI/O完了報告を発行したストレージ装置200の仮想ストレージ装置210が、当該ディスクI/O完了報告に対する確認信号を受信せず、タイムアウトとなったときに、I/Oチャネル400が断線したと判定する。また、ストレージ装置200にI/Oチャネル400の断線を検出する断線検出回路を設けて断線を監視してもよい。

【0216】

すなわち、この障害処理は、ディスクI/Oが発生したタイミングに実行されてもよいし、ディスクI/Oが発生したタイミング以外で実行されてもよい。

10

【0217】

次に、ストレージハイパーバイザ214は、仮想ストレージ装置(0)240と仮想計算機(0)130との接続を維持するため、仮想チャネルアダプタ(0)241と物理的なチャネルアダプタ(0)231との接続パス(ストレージ装置200内の接続パス)を変更する(S162)。例えば、仮想チャネルアダプタ(0)241とチャネルアダプタ(1)232との接続パスを設定する。

【0218】

これに伴い、仮想チャネルアダプタ管理テーブル260の内容が変更される。本実施の形態では、仮想ストレージ装置番号261が"0"、仮想チャネルアダプタ番号262が"0"、チャネルアダプタ番号263が"0"のエントリにおいて、チャネルアダプタ番号263が"1"に変更される。

20

【0219】

なお、障害が検出されたエントリの内容を変更せず、新たに設定したパスのエントリを追加してもよい。

【0220】

その後、ストレージハイパーバイザ214は、仮想I/Oアダプタと物理的なI/Oアダプタとの接続パス(サーバ装置(0)100内の接続パス)の変更を管理端末300に通知する(S163)。

【0221】

ストレージハイパーバイザ214からのパス変更通知を受信した管理端末300は、仮想パス管理テーブル310を変更する(S164)。具体的には、仮想ストレージ装置番号311が"0"、仮想チャネルアダプタ番号312が"0"、チャネルアダプタ番号313が"0"のエントリにおいて、チャネルアダプタ番号313が"1"に変更される。

30

【0222】

その後、管理端末300は、仮想パス管理テーブル310を変更が終了した旨を、ストレージハイパーバイザ214に通知する。

【0223】

管理端末300から変更終了通知を受信したストレージハイパーバイザ214は、未完了の送信情報を、新たなパスを構成するI/Oアダプタに再送する(S165)。

【0224】

このような送信情報には、ディスクI/Oコマンドに対する完了報告がある。I/Oチャネルとして複数の通信プロトコルがある。その一つのファイバチャネル規格では、I/Oアダプタ、チャネルアダプタがなんらかの情報を受信した場合に、送信元に確かに受信した旨(アクノレジメント)を通知する場合がある。ステップS165では、I/Oアダプタに送信した情報のうち、アクノレジメントを受信していない情報を再送する。

40

【0225】

以上に説明したように、図16に示す仮想計算機側の障害処理及び図17に示す仮想ストレージ装置側の障害処理によって、仮想パス管理テーブル310が変更され、仮想計算機側(0)130及び仮想ストレージ装置(0)230間のパスが復旧し、両者間で通信が可能になる。このパス変更処理は、ハイパーバイザ103、ストレージハイパーバイザ

50

214が実行するため、OS(0)135、OS(1)145、OS(2)185、OS(3)195が障害を検出することができない。そのため、OS側で障害処理をする必要がないという利点がある。

【0226】

図18は、第1の実施の形態の管理端末300に表示される画面の説明図である。

【0227】

管理端末300の表示画面320には、管理者が本計算機システムを管理する際に必要な管理情報が表示される。すなわち、管理端末300の表示画面320には、仮想バス管理テーブル310(図5)に記憶された情報(仮想計算機番号316、仮想I/Oアダプタ番号315、I/Oアダプタ番号314、チャンネルアダプタ番号313、仮想チャンネルアダプタ番号312及び仮想ストレージ装置番号311)の対応関係が視覚的に示されている。

10

【0228】

具体的には、この表示画面320には、仮想計算機表示領域321、I/Oアダプタ表示領域322、チャンネルアダプタ表示領域323及びストレージ装置表示領域324が設けられている。

【0229】

仮想計算機表示領域321には、本計算機システムに備わる仮想計算機(0)～仮想計算機(3)が表示される。各仮想計算機内には、仮想計算機に備わる仮想I/Oアダプタが表示される。

20

【0230】

I/Oアダプタ表示領域322には、本計算機システムのサーバ装置に備わる物理的なI/Oアダプタ(0)～I/Oアダプタ(3)が表示される。

【0231】

チャンネルアダプタ表示領域323には、本計算機システムのストレージ装置に備わる物理的なチャンネルアダプタ(0)～チャンネルアダプタ(3)が表示される。

【0232】

ストレージ装置表示領域324には、本計算機システムに備わる仮想ストレージ装置(0)及び仮想ストレージ装置(1)が表示される。各仮想ストレージ装置内には、仮想ストレージ装置に備わる仮想チャンネルアダプタが表示される。

30

【0233】

さらに、この表示画面320には、仮想計算機(仮想I/Oアダプタ)と仮想ストレージ装置(仮想チャンネルアダプタ)との間に設定されたパスが示される。各パスがどのI/Oアダプタ及びどのチャンネルアダプタを経由するかが視覚的に分かるように示されている。また、各パスは、異なる種類(又は、異なる色彩)の線で示されており、パスの経路が容易に分かるようになっている。

【0234】

このように、管理端末300の表示画面320に、パスの経路を視覚的に表示することによって、仮想計算機と仮想ストレージ装置とを接続しているパスが物理的に異なるパスに分けられており、冗長構成をとっていることが分かる。そして、管理者は、仮想計算機、仮想ストレージ装置間の物理的な接続と仮想的な接続の対応を確認することができる。

40

【0235】

すなわち、仮想計算機(0)と仮想ストレージ装置(0)との間に、仮想I/Oアダプタ及び仮想チャンネルアダプタが異なる二つのパスが設けられている。この二つのパスは、物理的にも異なるI/Oアダプタ及びチャンネルアダプタを経由するので、仮想的にも物理的にも、冗長構成を実現している。

【0236】

前述したように、本発明の第1の実施の形態では、仮想計算機を生成する場合には、管理コンソールプログラムが統合バス資源テーブルに従って、正しい物理資源(物理パス)と仮想資源(仮想パス)との対応関係を確認しながら、計算機システムの構成を設定する

50

。すなわち、サーバ及びストレージを連携して論理分割する際に、物理的資源と論理的資源の対応関係を確認することができる。よって、仮想計算機を用いた計算機システムのハイパバイリティ構成を容易に設定できる。

【0237】

(第2実施形態)

本発明の第2の実施の形態では、仮想計算機においてパス制御プログラムが動作している。このパス制御プログラムは、仮想計算機と仮想ストレージ装置との間に設定されたパスを管理し、I/Oチャンネルの障害処理をする。

【0238】

図19は、第2の実施の形態の計算機システムの接続状態の説明図である。図19は、図8と同様に、ソフトウェア的観点からの(仮想世界からみた)計算機システムの接続状態を示す。よって、図19には、計算機システムのハードウェア構成(例えば、I/Oアダプタ106等、チャンネルアダプタ231等)は図示されない。

【0239】

仮想計算機(0)130では、OS(0)135及びパス制御プログラム136が動作している。パス制御プログラム136は、ストレージ装置200に記憶されており、仮想計算機(0)130の起動時にストレージ装置200から読み出され、主記憶メモリ104にロードされて、CPU101で実行される。また、パス制御プログラム136は、サーバ装置内の不揮発性メモリに記憶され、当該不揮発性メモリからロードされてもよい。パス制御プログラム136は、仮想計算機(0)130に関する仮想I/Oアダプタと仮想ストレージ装置(0)240に関する仮想チャンネルアダプタとの間に設定されたパスを管理する。

【0240】

同様に、仮想計算機140では、OS145及びパス制御プログラム146が動作している。仮想計算機180では、OS185及びパス制御プログラム186が動作している。仮想計算機190では、OS195及びパス制御プログラム196が動作している。各パス制御プログラム156、186、196は、各仮想計算機に設定されたパスを管理する。

【0241】

図19には、前述した図8と同様に、仮想パス管理テーブル310(図5)の仮想計算機番号316、仮想I/Oアダプタ番号315、仮想チャンネルアダプタ番号312及び仮想ストレージ装置番号311の対応関係が示されている。なお、仮想計算機及び仮想ストレージ装置間の接続関係は、前述した図8と同じであるため、その説明は省略する。

【0242】

図20は、第2の実施の形態の障害処理のフローチャートである。なお、仮想計算機(0)130のパス制御プログラム136がI/Oチャンネル400の障害を検出する場合を説明するが、他のパス制御プログラム146、186、196でも同様に障害を検出することができる。

【0243】

まず、パス制御プログラム136がI/Oチャンネル400の断線を検出する(S171)。なお、I/Oチャンネル400の断線が検出されなければ、この障害処理は実行されない。

【0244】

I/Oチャンネル400の断線は、例えば、ディスクI/Oコマンドを発行したサーバ装置(0)100の仮想計算機130が、ディスクI/O完了報告を受信せず、タイムアウトとなったときに、I/Oチャンネル400が断線したと判定する。また、サーバ装置(0)100にI/Oチャンネル400の断線を検出する断線検出回路を設け、パス制御プログラム136が断線を監視してもよい。

【0245】

すなわち、この障害処理は、ディスクI/Oが発生したタイミングに実行されてもよい

し、ディスク I / O が発生したタイミング以外で実行されてもよい。

【 0 2 4 6 】

次に、バス制御プログラム 1 3 6 は、仮想計算機 (0) 1 3 0 が仮想ストレージ装置 (0) 2 4 0 へのアクセスに使用する仮想 I / O アダプタを切り替える (S 1 7 2)。具体的には、仮想 I / O アダプタ (0) 1 3 1 を使用する I / O を仮想 I / O アダプタ (1) 1 3 2 を使用するように切り替える。

【 0 2 4 7 】

その後、バス制御プログラム 1 3 6 は、未完了の送信情報 (例えば、完了報告を受信していない発行済のディスク I / O コマンド) を、該当する I / O アダプタに再送する (S 1 7 3)。

10

【 0 2 4 8 】

このように、第 2 の実施の形態では、I / O チャンネルに障害が発生したときに、バス制御プログラムが、仮想計算機が使用する仮想 I / O アダプタを切り替える。よって、管理端末 3 0 0 の仮想バス管理テーブル 3 1 0 を変更することなく、障害を復旧することができる。

【 0 2 4 9 】

(第 3 実施形態)

本発明の第 3 の実施の形態では、仮想計算機においてクラスタ制御プログラムが動作している。このクラスタ制御プログラムは、仮想計算機が対となり、仮想計算機のクラスタ構成を管理する。

20

【 0 2 5 0 】

図 2 1 は、第 3 の実施の形態の計算機システムの接続状態の説明図である。図 2 1 は、図 8 と同様に、ソフトウェア的観点からの (仮想世界からみた) 計算機システムの接続状態を示す。よって、図 2 1 には、計算機システムのハードウェア構成 (例えば、I / O アダプタ 1 0 6 等、チャンネルアダプタ 2 3 1 等) は図示されない。

【 0 2 5 1 】

仮想計算機 (2) 1 8 0 では、OS (2) 1 8 5 及びクラスタ制御プログラム 1 8 7 が動作している。クラスタ制御プログラム 1 8 7 は、ストレージ装置 2 0 0 に記憶されており、仮想計算機 (2) 1 8 0 の起動時にストレージ装置 2 0 0 から読み出され、主記憶メモリ (1) 1 5 4 にロードされて、CPU 1 5 1 で実行される。また、クラスタ制御プログラム 1 8 7 は、サーバ装置内の不揮発性メモリに記憶され、当該不揮発性メモリからロードされてもよい。

30

【 0 2 5 2 】

本実施の形態では、異なるハードウェア (サーバ装置) に構成される仮想計算機 (0) 1 3 0 と仮想計算機 (2) 1 8 0 とがクラスタを構成している。クラスタ対を構成する仮想計算機同士は通信バス 4 2 1 で接続されている。通信バス 4 2 1 は、サーバ装置 (0) 1 0 0 とサーバ装置 (1) 1 5 0 とを接続する専用線で構成してもよい。また、ネットワーク 4 1 0 を用いてもよい。

【 0 2 5 3 】

クラスタ対を構成する仮想計算機同士は、通信バス 4 2 1 を介して、障害検知のための制御用 I / O (例えば、ハートビート信号) を所定のタイミング (例えば、周期的に) で発行する。そして、相手方が発行する制御用 I / O が受信できなければ、クラスタ対を構成する仮想計算機に障害が発生したと判定する。

40

【 0 2 5 4 】

クラスタ対を構成する仮想計算機同士は、フェイルオーバー時に仮想ストレージ装置 2 4 0 等に格納された同じデータを使用してサービスを提供するので、同じ構成の仮想 I / O アダプタが設定されており、該クラスタ間で対になる仮想 I / O アダプタは、同じように仮想ストレージ装置と接続されている。

【 0 2 5 5 】

同様に、仮想計算機 (1) 1 4 0 と仮想計算機 (3) 1 9 0 とは、通信バス 4 2 2 で接

50

続され、クラスタを構成している。

【0256】

図21には、前述した図8と同様に、仮想バス管理テーブル310（図5）の仮想計算機番号316、仮想I/Oアダプタ番号315、仮想チャネルアダプタ番号312及び仮想ストレージ装置番号311の対応関係が示されている。なお、仮想計算機及び仮想ストレージ装置間の接続関係は、前述した図8と同じであるため、その説明は省略する。

【0257】

図22は、第3の実施の形態の障害処理のフローチャートである。なお、ハイパーバイザ103が仮想計算機(0)130を再起動する場合を説明するが、他のハイパーバイザや他の仮想計算機でも同じ再起動処理を行うことができる。

10

【0258】

まず、クラスタ制御プログラム137が、仮想計算機(0)130から送られるハートビート信号を所定時間以上受信できないと、ハートビート断と判定する(S181)。なお、ハートビート信号の遮断が検出されなければ、この障害処理は実行されない。

【0259】

具体的には、ハートビート信号の遮断は、例えば、周期的に送信されるハートビート信号が、所定時間以上受信できないことが、複数回生じたときに、ハートビート断と判定する。すなわち、仮想計算機(0)130に障害が発生したことが判定される。

【0260】

次に、クラスタ制御プログラム187は、障害が発生した仮想計算機(0)130が提供していたサービスのプログラムを仮想計算機(2)180で起動して、フェイルオーバーの準備をする(S182)。そして、仮想計算機(0)130を計算機システムから切り離す。これによってクライアント端末（図示省略）から仮想計算機(0)130へのアクセスが停止する。

20

【0261】

その後、クラスタ制御プログラム187は、仮想計算機(0)130に障害が発生したこと、及び業務の引継ぎの準備ができたことを、管理者に通知する(S183)。

【0262】

なお、管理者の指示を待って、仮想計算機(0)130から仮想計算機(2)180へ業務が引継がれる。

30

【0263】

図23は、第3の実施の形態の仮想計算機の再起動処理のフローチャートである。なお、ハイパーバイザ103が仮想計算機(0)130を再起動する場合を説明するが、他のハイパーバイザや他の仮想計算機でも同じ再起動処理を行うことができる。

【0264】

まず、管理者は、停止した仮想計算機(0)130の再起動を、管理端末300から指示する(S191)。

【0265】

仮想計算機(0)130では、ネットワーク410経由で、管理端末300からの再起動指示を受信すると、クラスタ制御プログラム137が当該再起動指示を受け取る。そして、クラスタ制御プログラム137は、ハイパーバイザ103に、仮想計算機(0)130の再起動を指示する(S192)。

40

【0266】

その後、ハイパーバイザ103によって、計算機資源が再度割り当てられ、仮想計算機(0)130が再起動する(S193)。

【0267】

その後、管理者が、管理端末300からフェイルバックを指示する(S194)。このフェイルバック指示によって、仮想計算機(2)180に引き継がれた業務を、仮想計算機(0)130で再開する。具体的には、再起動される仮想計算機が提供していたサービスのプログラムを仮想計算機(0)130で起動する。その後、仮想計算機(0)130

50

を計算機システムに接続する。これによってクライアント端末から仮想計算機（０）１３０へのアクセスが再開する。

【０２６８】

（第４実施形態）

本発明の第４の実施の形態は、管理端末３００が仮想パス管理テーブルを有さず、ストレージ装置２００が仮想パス管理テーブル２６６を有している点で、前述した第１の実施の形態と異なる。

【０２６９】

図２４は、第４の実施の形態の計算機システムの機能ブロック図である。

【０２７０】

第４の実施の形態の計算機システムは、サーバ装置（０）１００、サーバ装置（１）１５０、ストレージ装置２００、及び管理端末３００によって構成されている。なお、前述した第１の実施の形態と同じ構成には同じ符号を付し、その説明は省略する。

【０２７１】

サーバ装置（０）１００及びサーバ装置（１）１５０は、機能的に、物理層、ハイパーバイザ層及び仮想計算機層に大別できる。各層に備わる各部の構成及び機能は、前述した第１の実施の形態と同じである。

【０２７２】

ストレージ装置２００は、機能的に、物理層、ハイパーバイザ層及び仮想ストレージ層に大別できる。各層に備わる各部の構成及び機能は、前述した第１の実施の形態と同じである。

【０２７３】

本実施の形態のストレージ装置２００は、仮想パス管理テーブル２６６を有する。ストレージハイパーバイザ２１４、２２４は、仮想パス管理テーブル２６６を使用して、ストレージ装置２００の仮想的な構成を管理する。すなわち、第４の実施の形態では、仮想パス管理テーブル２６６を有するストレージ装置２００に、仮想計算機と仮想ストレージ装置との接続関係を定める接続情報（仮想パス管理テーブル２６６）を有する管理部（図示省略）が備わっている。

【０２７４】

仮想パス管理テーブル２６６には、仮想計算機と仮想ストレージとの対応関係が記載されている。具体的には、仮想パス管理テーブル２６６は、第１の実施の形態の仮想パス管理テーブル３１０（図５）と同じ構成で、同じ管理情報を含む。

【０２７５】

ストレージハイパーバイザ２１４、２２４は、サーバ装置（０）１００の仮想Ｉ／Ｏアダプタ管理テーブル１１０に含まれる管理情報、サーバ装置（１）１５０の仮想Ｉ／Ｏアダプタ管理テーブル１６０に含まれる管理情報及びストレージ装置２００の仮想チャネルアダプタ管理テーブル２６０に含まれる管理情報を収集することによって、仮想パス管理テーブル２６６を生成する。

【０２７６】

管理端末３００は、計算機システムを統括して管理するコンピュータ装置で、ＣＰＵ、メモリ、入出力装置及びインターフェースが備わっている。管理端末３００は、保守用端末装置（ＳＶＰ：サービスプロセッサ）である。管理者は、管理端末３００を操作することによって、本計算機システムを管理することができる。

【０２７７】

管理端末３００のインターフェースは、ストレージ装置２００に接続されている。なお、管理端末３００をストレージ装置２００に直接接続することなく、ネットワーク４１０経由でストレージ装置２００に接続してもよい。

【０２７８】

なお、図２４には、二つのサーバ装置１００、１５０と一つのストレージ装置２００とが接続されている例について説明したが、三つ以上のサーバ装置を設けてもよく、二つ以

10

20

30

40

50

上のストレージ装置を設けてもよい。この場合、管理端末 300 を、いずれのストレージ装置に接続してもよいが、仮想パス管理テーブル 266 を有するストレージ装置に接続することが望ましい。また、仮想パス管理テーブル 116 を、複数のストレージ装置が有していてもよい。このようにすれば、障害が発生した場合に備えて、当該テーブルを二重化しておくことができる。

【0279】

(第5実施形態)

本発明の第5の実施の形態は、管理端末 300 が仮想パス管理テーブルを有さず、サーバ装置(0)100 が仮想パス管理テーブル 116 を有している点で、前述した第1の実施の形態と異なる。

【0280】

図25は、第5の実施の形態の計算機システムの機能ブロック図である。

【0281】

第5の実施の形態の計算機システムは、サーバ装置(0)100、サーバ装置(1)150、ストレージ装置200、及び管理端末300によって構成されている。なお、前述した第1の実施の形態と同じ構成には同じ符号を付し、その説明は省略する。

【0282】

サーバ装置(0)100及びサーバ装置(1)150は、機能的に、物理層、ハイパーバイザ層及び仮想計算機層に大別できる。各層に備わる各部の構成及び機能は、前述した第1の実施の形態と同じである。

【0283】

本実施の形態のサーバ装置(0)100は、仮想パス管理テーブル 116 を有する。ハイパーバイザ103は、仮想パス管理テーブル 116 を使用して、計算機システムの仮想的な構成を管理する。すなわち、第5の実施の形態では、仮想パス管理テーブル 116 を有するサーバ装置(0)100に、仮想計算機と仮想ストレージ装置との接続関係を定める接続情報(仮想パス管理テーブル 116)を有する管理部(図示省略)が備わっている。

【0284】

仮想パス管理テーブル 116 には、仮想計算機と仮想ストレージとの対応関係が記載されている。具体的には、仮想パス管理テーブル 116 は、第1の実施の形態の仮想パス管理テーブル 310(図5)と同じ構成で、同じ管理情報を含む。

【0285】

ハイパーバイザ103は、サーバ装置(0)100の仮想I/Oアダプタ管理テーブル 110に含まれる管理情報、サーバ装置(1)150の仮想I/Oアダプタ管理テーブル 160に含まれる管理情報及びストレージ装置200の仮想チャネルアダプタ管理テーブル 260に含まれる管理情報を収集することによって、仮想パス管理テーブル 116 を生成する。

【0286】

ストレージ装置200は、機能的に、物理層、ハイパーバイザ層及び仮想ストレージ層に大別できる。各層に備わる各部の構成及び機能は、前述した第1の実施の形態と同じである。

【0287】

管理端末300は、計算機システムを統括して管理するコンピュータ装置で、CPU、メモリ、入出力装置及びインターフェースが備わっている。管理端末300は、保守用端末装置(SVP:サービスプロセッサ)である。管理者は、管理端末300を操作することによって、本計算機システムを管理することができる。

【0288】

管理端末300のインターフェースは、ネットワーク410を介して、サーバ装置100、150及びストレージ装置200と接続されている。当該インターフェースは、ネットワーク410を介して接続された装置との間で制御信号や管理情報(各種管理テーブル

10

20

30

40

50

の内容)を送受信する。なお、管理端末300をサーバ装置(0)100に直接接続してもよい。

【0289】

なお、図25には、二つのサーバ装置100、150と一つのストレージ装置200とが接続されている例について説明したが、三つ以上のサーバ装置を設けてもよく、二つ以上のストレージ装置を設けてもよい。この場合、管理端末300を、いずれのサーバ装置に接続してもよいが、仮想バス管理テーブル116を有するサーバ装置に接続することが望ましい。また、仮想バス管理テーブル116を、複数のサーバが有していてもよい。このようにすれば、障害が発生した場合に備えて、当該テーブルを二重化しておくことができる。

10

【0290】

(第6実施形態)

本発明の第6の実施の形態は、I/Oチャンネル400が、I/Oアダプタとチャンネルアダプタとを直接接続するのではなく、I/OアダプタとチャンネルアダプタとがI/Oチャンネルスイッチ401を介して接続されている点で、前述した第1の実施の形態と異なる。

【0291】

図26は、第6の実施の形態の計算機システムのハードウェア構成を示すブロック図である。

【0292】

第6の実施の形態の計算機システムは、サーバ装置(0)100、サーバ装置(1)150、ストレージ装置200、管理端末300、及びI/Oチャンネルスイッチ401によって構成されている。なお、前述した第1の実施の形態と同じ構成には同じ符号を付し、その説明は省略する。

20

【0293】

サーバ装置(0)100及びサーバ装置(1)150では、アプリケーションプログラムが動作している。サーバ装置(0)100及びサーバ装置(1)150の各部の構成及び機能は、前述した第1の実施の形態と同じである。

【0294】

I/Oアダプタ(0)106及びI/Oアダプタ(1)107は、I/Oチャンネル400を介してI/Oチャンネルスイッチ401と接続されている。

30

【0295】

I/Oアダプタ106、107がストレージ装置200に対してデータ入出力要求を送信すると、I/Oチャンネルスイッチ401は当該データ入出力要求を受信する。そして、I/Oチャンネルスイッチ401は、当該データ入出力要求の宛先を判定し、当該データ入出力要求を宛先のストレージ装置200に転送する。

【0296】

また、I/Oアダプタ106、107は、ストレージ装置200に格納されたデータを、I/Oチャンネルスイッチ401を経由して受信する。

【0297】

なお、第6の実施の形態でも、I/Oアダプタは二つ設けられている。なお、さらに多くのI/Oアダプタを設けてもよい。この二つのI/Oアダプタ106、107は別個に動作する。よって、一方のI/Oアダプタに障害が発生しても、サーバ装置(0)100からストレージ装置200へのアクセスが停止しないように、処理系を2重化している。

40

【0298】

ストレージ装置200は、サーバ装置100、150の動作に必要なデータを記憶する。ストレージ装置200の各部の構成及び機能は、前述した第1の実施の形態と同じである。

【0299】

チャンネルアダプタ(0)231~チャンネルアダプタ(3)234は、I/Oチャンネル400を介してI/Oチャンネルスイッチ401と接続されている。また、I/Oチャンネルス

50

イチ 401 は、サーバ装置 (0) 100 及びサーバ装置 (1) 150 と接続されている。

【0300】

チャンネルアダプタ (0) 231 及びチャンネルアダプタ (1) 232 は、I/Oチャンネル 400 及びI/Oチャンネルスイッチ 401 を介してサーバ装置 (0) 100 と接続されている。チャンネルアダプタ 231、232 は、サーバ装置 (0) 100 からのデータ入出力要求を受信し、ストレージ装置 200 に格納されたデータを送信する。

【0301】

同様に、チャンネルアダプタ (2) 233 及びチャンネルアダプタ (3) 234 は、I/Oチャンネル 400 及びI/Oチャンネルスイッチ 401 を介してサーバ装置 (1) 150 と接続されている。チャンネルアダプタ 233、234 は、サーバ装置 (1) 150 からのデータ入出力要求を受信し、ストレージ装置 200 に格納されたデータを送信する。

【0302】

ストレージ装置 200 には複数のチャンネルアダプタが設けられている。この複数のチャンネルアダプタ 231 ~ 234 は別個に動作する。よって、一方のチャンネルアダプタに障害が発生しても、サーバ装置 100、150 からのストレージ装置 200 へのアクセスが停止しないように、処理系を2重化している。

【0303】

管理端末 300 は、計算機システムを統括して管理するコンピュータ装置である。管理端末 300 は、前述した第1の実施の形態と異なる形式の仮想パス管理テーブル 317 (図 27) を有する。なぜなら、第6の実施の形態では、パスがI/Oチャンネルスイッチ 401 を経由するためである。

【0304】

管理端末 300 の他の各部の構成及び機能は、前述した第1の実施の形態と同じである。

【0305】

I/Oチャンネル 400 は、例えばファイバチャンネルプロトコルのような、データの転送に適するプロトコルで通信可能な伝送媒体である。

【0306】

第6の実施の形態では、I/Oチャンネル 400 は、サーバ装置 100、150 のI/OアダプタとI/Oチャンネルスイッチ 401 との間、及びストレージ装置 200 のチャンネルアダプタとI/Oチャンネルスイッチ 401 との間を接続する。

【0307】

I/Oチャンネルスイッチ 401 は、複数のポートを有するパケット転送装置である。例えば、I/Oチャンネル 400 にファイバチャンネルプロトコルを用いる場合には、I/Oチャンネルスイッチ 401 にはSANスイッチが用いられる。

【0308】

I/Oチャンネルスイッチ 401 の各ポートにはI/Oアダプタ又はチャンネルアダプタと接続するI/Oチャンネル 400 が接続されている。I/Oチャンネルスイッチ 401 は、経路情報を保持する。この経路情報はパケットの転送先を判定するために用いられるルーティングテーブル (図示省略) である。I/Oチャンネルスイッチ 401 は、ルーティングテーブルを参照して、各ポートに入力されたパケットの宛先を判定し、宛先となるI/Oアダプタ又はチャンネルアダプタに接続されているポートから当該パケットを送信する。

【0309】

I/Oチャンネルスイッチ 401 は、ゾーン設定テーブル (図 30 参照) を有する。ゾーン設定テーブルには、I/Oチャンネルスイッチ 401 のポート間の通信可否が規定されている。そして、I/Oチャンネルスイッチ 401 は、通信可能であると規定されたポート間のみパケットを転送する。よって、特定の装置間のみパケットの転送が実現する。なお、ゾーン設定テーブルとルーティングテーブルとを一つのテーブルで構成してもよい。

【0310】

また、I/Oチャンネルスイッチ401は、ネットワーク410にも接続されている。さらに、I/Oチャンネルスイッチ401間は、相互接続線402によって接続されている。I/Oチャンネルスイッチ401は、相互接続線402を介して、パケットの転送のための経路情報を交換する。これによって、I/Oチャンネルスイッチ401は、経路情報を共有する。

【0311】

なお、第6の実施の形態では、I/Oチャンネルスイッチ401は二つ設けられている。この二つのI/Oチャンネルスイッチ401も別個に動作する。よって、一方のI/Oチャンネルスイッチに障害が発生しても、サーバ装置(0)100からストレージ装置200へのアクセスが停止しないように、処理系を2重化している。

10

【0312】

また、I/Oチャンネルスイッチ401は相互接続線402を用いて互いの稼動状況を監視することもできる。稼動状況の監視と障害処理の方法は、前述した第3の実施形態におけるクラスタ制御プログラムと同様の方法を採用できる。

【0313】

ネットワーク410は、サーバ装置(0)100、サーバ装置(1)150、ストレージ装置200、管理端末300及びI/Oチャンネルスイッチ401を接続する。ネットワーク410は、例えば、TCP/IPプロトコルによって、コンピュータ間で制御信号や管理情報が通信可能に構成されており、例えば、イーサネットが用いられる。

【0314】

20

なお、図26には、二つのサーバ装置100、150と一つのストレージ装置200とが接続されている例について説明したが、三つ以上のサーバ装置を設けてもよく、二つ以上のストレージ装置を設けてもよい。

【0315】

図27は、第6の実施例の形態の仮想パス管理テーブル317の説明図である。

【0316】

仮想パス管理テーブル317は、前述した第1の実施の形態の仮想パス管理テーブル310に含まれる項目に加え、I/Oチャンネルスイッチのポート番号を記録する318、319。

【0317】

30

I/Oチャンネルポート番号(ストレージ側)318は、チャンネルアダプタ231~234と接続したI/Oチャンネルスイッチ401のポート番号である。

【0318】

I/Oチャンネルポート番号(サーバ側)319は、I/Oアダプタ106~157と接続したI/Oチャンネルスイッチ401のポート番号である。

【0319】

第6の実施の形態では、I/Oアダプタ106~157及びチャンネルアダプタ231~234はI/Oチャンネルスイッチ401としか接続されていないため、I/Oチャンネルスイッチ401のポート番号を含めて管理しなければ、仮想パスを管理できない。

【0320】

40

図28は、第6の実施の形態の計算機システムのケーブル接続時処理のフローチャートである。このケーブル接続時処理はサーバ装置(0)100で実行される。なお、サーバ装置(1)150及びストレージ装置200においても同じ処理が実行される。

【0321】

まず、サーバ装置(0)100のI/Oアダプタ(ストレージ装置200のチャンネルアダプタ)が、I/Oチャンネルスイッチ401との接続がされたことを検出する(S191)。

【0322】

次に、サーバ装置(0)100は、I/Oチャンネルスイッチ401と物理アドレスを交換する(S192)。なお、ステップS192で交換されるアドレスは、I/Oチャンネル

50

スイッチ 4 0 1 のポートを識別可能なものを用いる。

【 0 3 2 3 】

ステップ S 1 9 2 で交換される物理アドレスは、接続される I / O チャンネルスイッチ 4 0 1 のポートが一意に特定されればよい。例えば、ファイバチャネルプロトコルにおけるワールド・ワイド・ネーム (W W N) や、 i S C S I プロトコルにおける M A C アドレスを交換する。

【 0 3 2 4 】

次に、ハイパーバイザ 1 0 3 は、取得したケーブル接続状況 (I / O アダプタの物理アドレスと I / O チャンネルスイッチの物理アドレス) を、ネットワーク 4 1 0 経由で管理端末 3 0 0 に報告する (S 1 9 3)

10

その後、管理端末 3 0 0 は、仮想バス管理テーブル 3 1 7 に、 I / O アダプタと I / O チャンネルスイッチ 4 0 1 との対応関係を登録する。具体的には、チャンネルアダプタ番号 3 1 3、 I / O アダプタ番号 3 1 4、 I / O チャンネルポート番号 (ストレージ側) 3 1 8、 I / O チャンネルポート番号 (サーバ側) 3 1 9 に取得した物理アドレスからポートあるいはアダプタの識別子を登録する。

【 0 3 2 5 】

図 2 9 は、第 6 の実施の形態の計算機システムの仮想計算機生成処理のフローチャートである。

【 0 3 2 6 】

第 6 の実施の形態の仮想計算機生成処理は、前述した第 1 の実施の形態の仮想計算機生成処理 (図 1 0) と、仮想計算機生成指示 (S 1 1 1) から、 I / O アダプタと仮想 I / O アダプタとの間のバス生成処理 (S 1 2 1) までの処理は同じである。

20

【 0 3 2 7 】

管理端末 3 0 0 が、仮想 I / O アダプタ番号 3 1 5 及び仮想計算機番号 3 1 6 の仮想バス管理テーブル 3 1 7 への登録 (S 1 2 1) が終了すると、 I / O アダプタとチャンネルアダプタとの間の接続手続を行う (S 2 0 1)。

【 0 3 2 8 】

具体的には、仮想バス管理テーブル 3 1 7 のチャンネルアダプタ番号 3 1 3 と I / O アダプタ番号 3 1 4 を参照し、通信可能なチャンネルアダプタと I / O アダプタを特定する。そして、チャンネルアダプタに対応するポート及び I / O アダプタに対応するポートを特定して、ポート間の通信可否を特定する。その通信可否の結果を、 I / O チャンネルスイッチ 4 0 1 のゾーン設定テーブル (図 3 0 参照) に設定する。このゾーニングによって、 I / O チャンネルスイッチ 4 0 1 は、特定のポート間 (特定の装置間) のみパケットを転送する。

30

【 0 3 2 9 】

その後、冗長バスを生成する必要があるか否かを判定する (S 1 2 2)。さらに、冗長構成を生成する必要があるか否かを判定する (S 1 2 3)。このステップ S 1 2 2 及びステップ S 1 2 3 の処理は、前述した第 1 の実施の形態の仮想計算機生成処理 (図 1 0) と同じである。

【 0 3 3 0 】

図 3 0 は、第 6 の実施の形態のゾーン設定テーブル 1 3 0 0 の構成図である。

40

【 0 3 3 1 】

ゾーン設定テーブル 1 3 0 0 は、入力ポート番号 1 3 0 1、出力ポート番号 1 3 0 2、及び各ポート間の通信可否の情報が含まれている。

【 0 3 3 2 】

例えば、入力ポート (0) に入力されたパケットは、出力ポート (1) から出力できるが、出力ポート (n) からは出力できない。よって、入力ポート (0) に接続された装置から出力ポート (1) に接続された装置にはパケットを転送できるが、入力ポート (0) に接続された装置から出力ポート (n) に接続された装置にはパケットを転送できない。

【 0 3 3 3 】

すなわち、入力ポート (0) に入力されたパケットの宛先が、出力ポート (n) に接続

50

された装置であったとしても、当該装置にはパケットを転送できない。

【0334】

なお、第6の実施の形態においても、仮想バス管理テーブル317をサーバ装置100、150が有してもよいし、ストレージ装置200が有してもよい。更に、I/Oチャンネルスイッチ401が仮想バス管理テーブル317を有してもよい。サーバ装置100、150、ストレージ装置200のいずれにも接続されているI/Oチャンネルスイッチ401が仮想バス管理テーブル317を有し、計算機システムを管理すれば、管理端末300を計算機システムに設ける必要がない。

【0335】

このようにして、ハイパーバイザ103、ハイパーバイザ153、ストレージハイパーバイザ214、ストレージハイパーバイザ224及びI/Oチャンネルスイッチ401が連携することによって、I/Oチャンネルの冗長構成を実現することができる。また、ゾーニングによって、仮想計算機間、仮想ストレージ装置間のセキュリティがI/Oチャンネルにおいても確保できる。

【0336】

これまで、サーバ装置、ストレージ装置が複数設けられてもよいことを述べてきた。これまで述べてきた発明の実施の形態では、1つのサーバ装置を複数の仮想計算機に論理分割する例を説明したが、本発明は仮想計算機の実現方法によらないことはいうまでもない。

【0337】

例えば、本発明は、図31に示す仮想計算機の実現方法にも適用しうる。サーバ装置(0)100は、計算機資源3100、3101を備える。同様に、サーバ装置(1)150は、計算機資源3102、3103を備える。各計算機資源3100、3101、3102、3103は、CPU、メモリ等を含む。

【0338】

そして、サーバ装置(0)の資源3100とサーバ装置(1)の資源3102とを抽出して、仮想計算機(4)3104を構成することもできる。同様にサーバ装置(0)の資源3101とサーバ装置(1)の資源3103を抽出して、仮想計算機(5)3105を構成することもできる。

【0339】

同様の方法で、複数のストレージ装置の資源を抽出して1つの仮想ストレージ装置を構成することもできる。

【0340】

本発明は、このような仮想計算機、仮想ストレージ装置の構成にも適用することができる。

【図面の簡単な説明】

【0341】

【図1】第1の実施の形態の計算機システムのハードウェア構成を示すブロック図である。

【図2】第1の実施の形態の計算機システムの機能ブロック図である。

【図3A】第1の実施の形態のサーバ装置(0)の仮想I/Oアダプタ管理テーブルの説明図である。

【図3B】第1の実施の形態のサーバ装置(1)の仮想I/Oアダプタ管理テーブルの説明図である。

【図4】第1の実施の形態の仮想チャンネルアダプタ管理テーブルの説明図である。

【図5】第1の実施の形態の仮想バス管理テーブルの説明図である。

【図6】本発明の実施の形態の計算機資源管理テーブルの説明図である。

【図7】本発明の実施の形態のストレージ資源管理テーブルの説明図である。

【図8】第1の実施の形態の計算機システムの接続状態の説明図である。

【図9】第1の実施の形態のケーブル接続時処理のフローチャートである。

10

20

30

40

50

【図 1 0】第 1 の実施の形態の仮想計算機生成処理のフローチャートである。

【図 1 1】第 1 の実施の形態の I / O チャンネルの通信プロトコル層の構造の説明図である。

。

【図 1 2】第 1 の実施の形態のサーバ装置とストレージ装置との間で伝達される情報の説明図である。

【図 1 3】第 1 の実施の形態のハイパーバイザ通信ヘッダの内容の説明図である。

【図 1 4】第 1 の実施の形態のディスク I / O コマンド送信処理のフローチャートである。

。

【図 1 5】第 1 の実施の形態のディスク I / O コマンド完了報告処理のフローチャートである。

10

【図 1 6】第 1 の実施の形態の障害処理のフローチャートである。

【図 1 7】第 1 の実施の形態の障害処理のフローチャートである。

【図 1 8】第 1 の実施の形態の管理端末 3 0 0 に表示される画面の説明図である。

【図 1 9】第 2 の実施の形態の計算機システムの接続状態の説明図である。

【図 2 0】第 2 の実施の形態の障害処理のフローチャートである。

【図 2 1】第 3 の実施の形態の計算機システムの接続状態の説明図である。

【図 2 2】第 3 の実施の形態の障害処理のフローチャートである。

【図 2 3】第 3 の実施の形態の仮想計算機の再起動処理のフローチャートである。

【図 2 4】第 4 の実施の形態の計算機システムの機能ブロック図である。

【図 2 5】第 5 の実施の形態の計算機システムの機能ブロック図である。

20

【図 2 6】第 6 の実施の形態の計算機システムのハードウェア構成を示すブロック図である。

【図 2 7】第 6 の実施例の形態の仮想パス管理テーブルの説明図である。

【図 2 8】第 6 の実施の形態の計算機システムのケーブル接続時処理のフローチャートである。

【図 2 9】第 6 の実施の形態の計算機システムの仮想計算機生成処理のフローチャートである。

【図 3 0】第 6 の実施の形態のゾーン設定テーブルの構成図である。

【図 3 1】第 6 の実施の形態のゾーン設定テーブルの構成図である。

【符号の説明】

30

【 0 3 4 2 】

1 0 0 サーバ装置 (0)

1 0 3 ハイパーバイザ

1 1 0 仮想 I / O アダプタ管理テーブル

1 1 5 計算機資源管理テーブル

1 1 6 仮想パス管理テーブル

1 3 0 仮想計算機 (0)

1 4 0 仮想計算機 (1)

1 5 0 サーバ装置 (1)

1 5 3 ハイパーバイザ

40

1 6 0 仮想 I / O アダプタ管理テーブル

1 6 5 計算機資源管理テーブル

1 8 0 仮想計算機 (2)

1 9 0 仮想計算機 (3)

2 0 0 ストレージ装置

2 1 4、2 2 4 ストレージハイパーバイザ

2 6 0 仮想チャンネルアダプタ管理テーブル

2 6 5 ストレージ資源管理テーブル

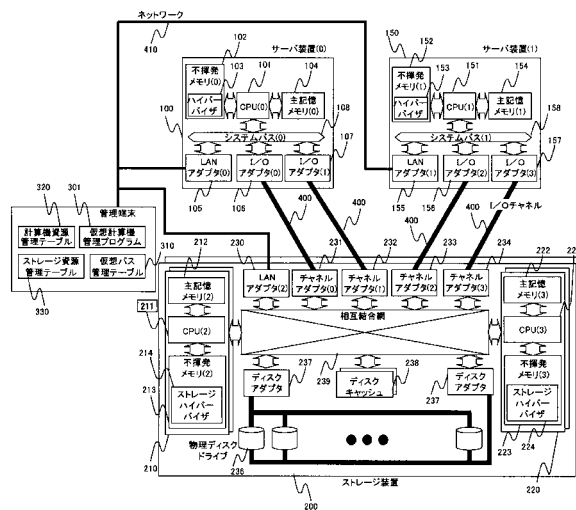
2 6 6 仮想パス管理テーブル

3 0 0 管理端末

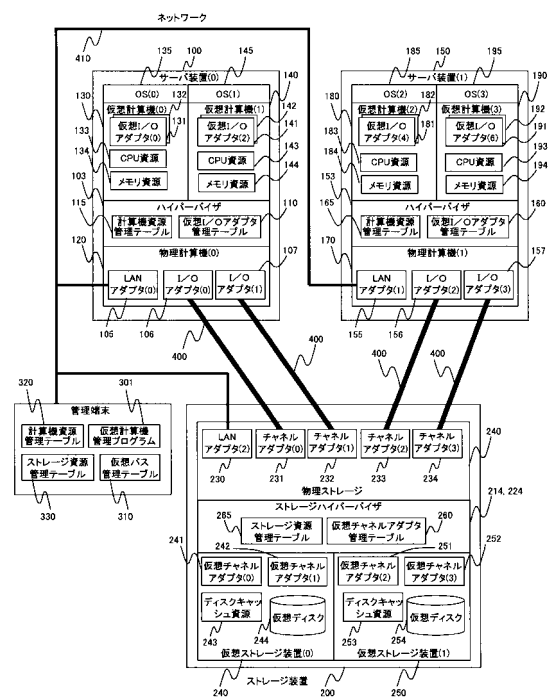
50

- 310、317 仮想バス管理テーブル
- 320 計算機資源管理テーブル
- 330 ストレージ資源管理テーブル
- 400 I/Oチャネル
- 401 I/Oチャネルスイッチ
- 410 ネットワーク

【図 1】



【図 2】



【図 3 A】

111 I/Oアダプタ番号	112 仮想I/Oアダプタ番号	113 仮想計算機番号
0	0	0
1	1	0
0	2	1
1	3	1

サーバ装置(0)の仮想I/Oアダプタ管理テーブル 110

【図 4】

261 仮想ストレージ 装置番号	262 仮想チャネル アダプタ番号	263 チャネル アダプタ番号
0	0	0
0	1	1
1	2	0
1	3	1
0	0	2
0	1	3
1	2	2
1	3	3

仮想チャネルアダプタ管理テーブル 260

【図 3 B】

161 I/Oアダプタ番号	162 仮想I/Oアダプタ番号	163 仮想計算機番号
2	4	2
3	5	2
2	6	3
3	3	3

サーバ装置(1)の仮想I/Oアダプタ管理テーブル 160

【図 5】

311 仮想ストレージ 装置番号	312 仮想チャネル アダプタ番号	313 チャネルアダプタ 番号	314 I/Oアダプタ 番号	315 仮想I/O アダプタ番号	316 仮想計算機番号
0	0	0	0	0	0
0	1	1	1	1	0
1	2	0	0	2	1
1	3	1	1	3	1
0	0	2	2	4	2
0	1	3	3	5	2
1	2	2	2	6	3
1	3	3	3	7	3

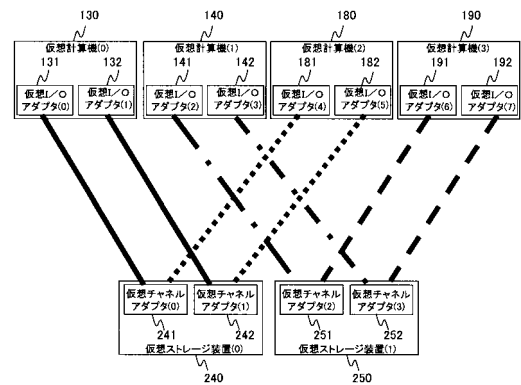
仮想バス管理テーブル 310

【図 6】

701 仮想計算機番号	702 CPU配分率	703 メモリ容量	704 I/Oアダプタ番号
0	20%	512MB	0 1
1	30%	128MB	0 1
⋮	⋮	⋮	⋮

計算機資源管理テーブル 115

【図 8】

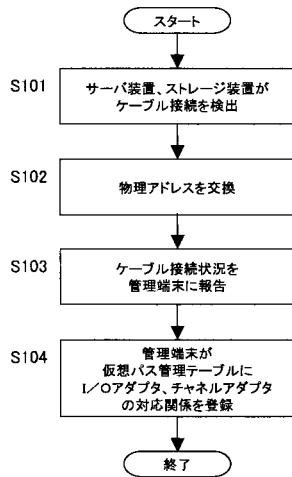


【図 7】

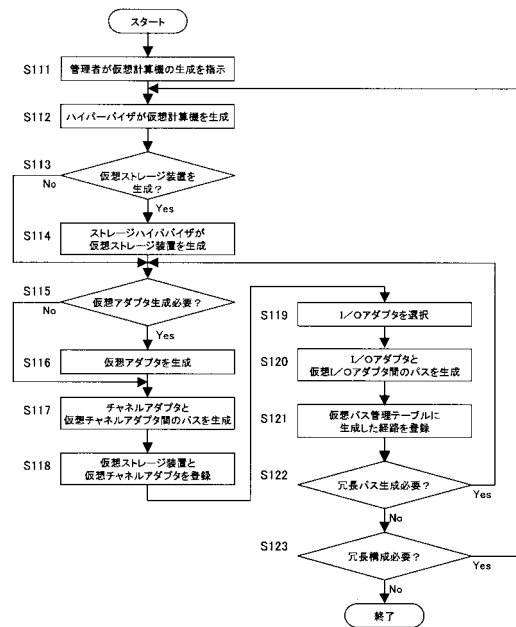
601 仮想計算機番号	602 仮想ディスク番号	603 ディスクキャッシュ容量	604 CPU番号	605 チャネルアダプタ 番号
0	121	512MB	2	0
	122		3	1
1	16	256MB	2	2
	17		3	3
⋮	⋮	⋮	⋮	⋮

ストレージ資源管理テーブル 265

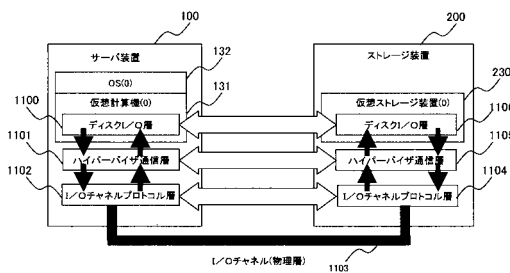
【図 9】



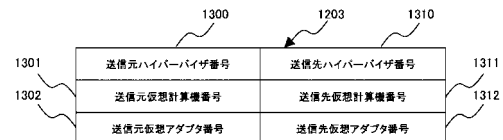
【図 10】



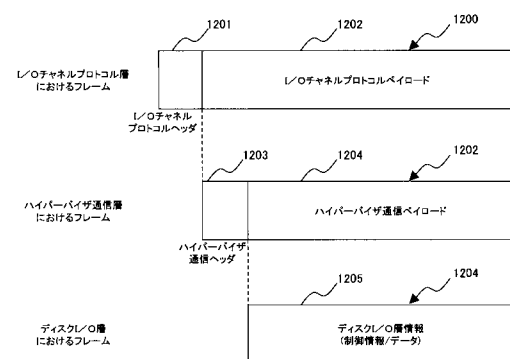
【図 11】



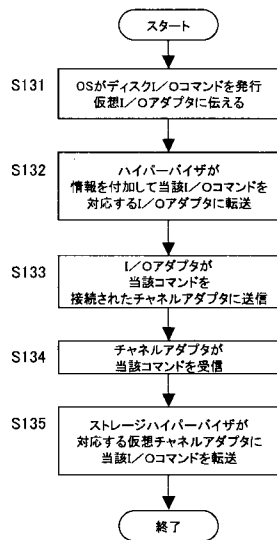
【図 13】



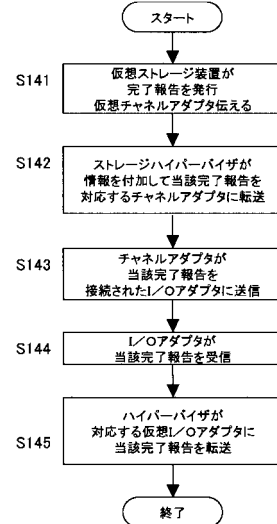
【図 12】



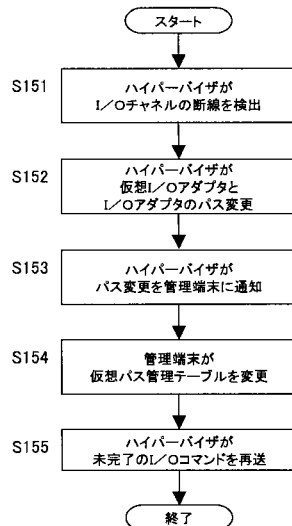
【図 14】



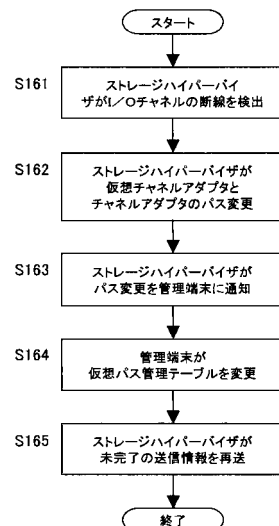
【図 15】



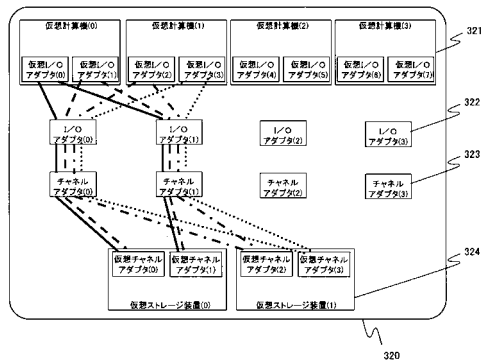
【図 16】



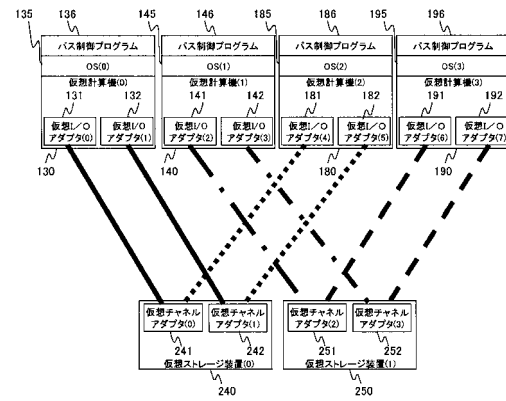
【図 17】



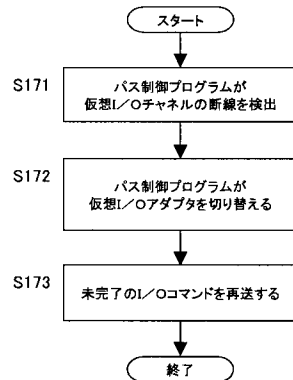
【図 18】



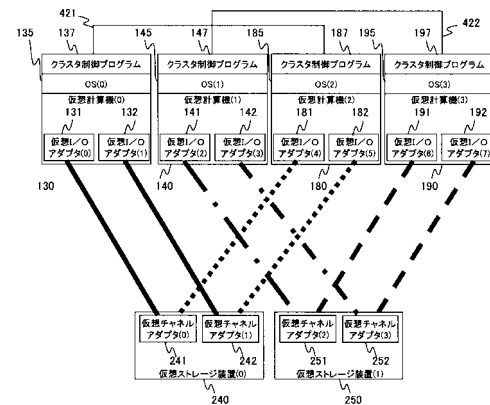
【図 19】



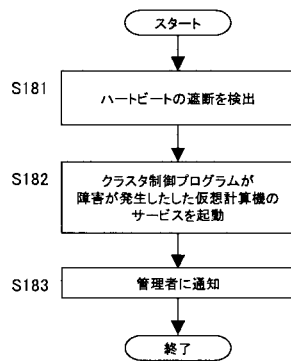
【図 20】



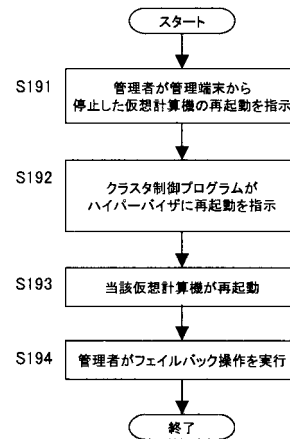
【図 21】



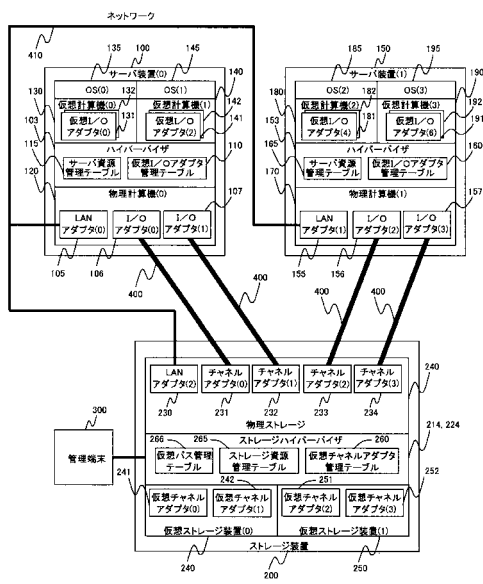
【図 22】



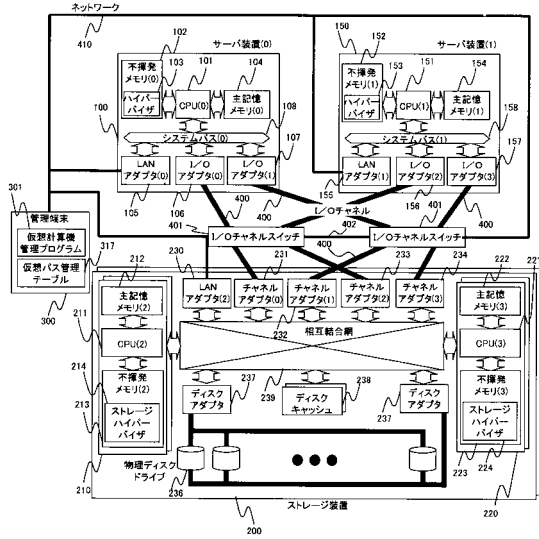
【図 23】



【図 24】



【図 26】

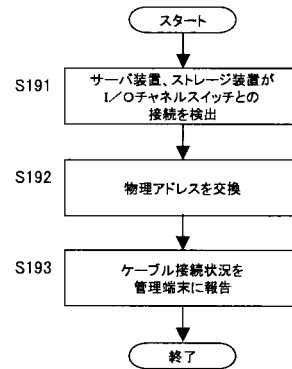


【図 27】

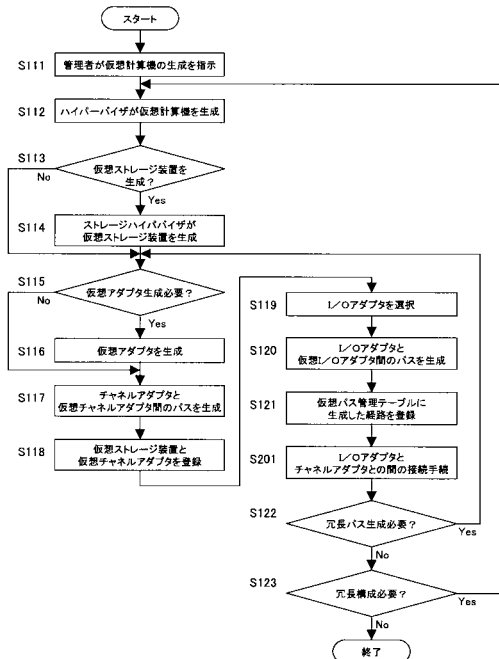
仮想ストレージ装置番号	仮想チャンネルアダプタ番号	チャンネルアダプタ番号	I/Oチャンネルポート番号 (ストレージ側)	I/Oチャンネルポート番号 (サーバ側)	レノアダプタ番号	仮想レノアダプタ番号	仮想計算機番号
0	0	0	1	5	0	0	0
0	1	1	2	6	1	1	0
1	2	0	1	5	0	2	1
1	3	1	2	6	1	3	1
0	0	2	3	7	2	4	2
0	1	3	4	8	3	5	2
1	2	2	3	7	2	6	3
1	3	3	4	8	3	7	3

仮想バス管理テーブル

【図 28】



【図 29】

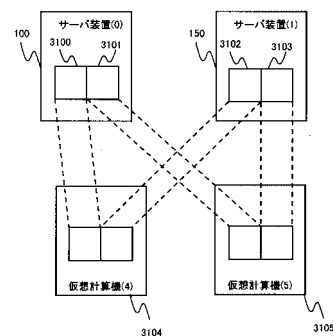


【図 30】

		出力ポート番号				
		0	1	...	n	
入力ポート番号	0	○				x
	1	○				○
	...					
	n-1	x	x			x
	n	x	○			

ゾーン設定テーブル

【図 31】



フロントページの続き

(72)発明者 岩寄 正明

神奈川県川崎市麻生区王禅寺 1 0 6 9 番地 株式会社日立製作所 システム開発研究所内

審査官 木村 貴俊

(56)参考文献 特開 2 0 0 2 - 2 2 2 1 1 0 (J P , A)

特開平 0 9 - 2 7 4 5 9 8 (J P , A)

特開 2 0 0 4 - 3 3 4 4 8 1 (J P , A)

岩山 知三郎, S A P、システム間連携の最新版「X I 3 . 0」リリース, C o m p u t o p i
a 2 0 0 4 - 1 2 , 日本, 株式会社コンピュータ・エージ社, 2 0 0 4 年 1 1 月 2 4 日, V o
l . 3 9 N o . 4 5 9 , P.89-P.93

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 F 3 / 0 6 - 3 / 0 8、 9 / 4 6

G 0 6 F 1 2 / 0 0、 1 3 / 1 0 - 1 3 / 4 2