

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2011-191854
(P2011-191854A)

(43) 公開日 平成23年9月29日(2011.9.29)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 11/20 (2006.01)	G06F 11/20 310E	5B014
G06F 13/14 (2006.01)	G06F 13/14 310F	5B034

審査請求 未請求 請求項の数 13 O L (全 20 頁)

(21) 出願番号 特願2010-55544 (P2010-55544)
(22) 出願日 平成22年3月12日 (2010.3.12)

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区丸の内一丁目6番6号
(74) 代理人 100075513
弁理士 後藤 政喜
(74) 代理人 100114236
弁理士 藤井 正弘
(74) 代理人 100120260
弁理士 飯田 雅昭
(72) 発明者 若松 峻彦
神奈川県横浜市戸塚区戸塚町5030番地
株式会社日立製作所ソフトウェア事業部
内

最終頁に続く

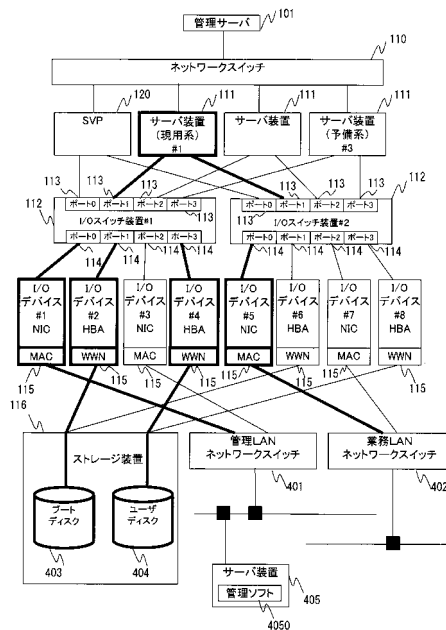
(54) 【発明の名称】 計算機システム、計算機システムの制御方法及びプログラム

(57) 【要約】

【課題】 I/OスイッチによりI/Oデバイスを共用する状態で現用系サーバから予備系サーバへの切り替えが行われた場合でも、管理サーバから各サーバの物理位置を把握する。

【解決手段】 複数の計算機と複数のI/Oデバイスを接続するI/Oスイッチを管理する構成管理情報を有する管理サーバが、計算機に対するI/Oデバイスの割り当てを制御する計算機システムの制御方法であって、管理サーバが、第1の計算機と第1の計算機に割り当てられたI/Oデバイスの識別子を取得して構成管理情報に格納し、第1の計算機から第2の計算機への切り換えを受け付け、第1の計算機を停止させ、第1の計算機に割り当てられた前記I/Oデバイスを第2の計算機へ割り当て、第2の計算機を起動させ、第2の計算機へ切り替えられたI/Oデバイスのうち、特定のI/Oデバイスの識別子を予め設定した仮想識別子に書き換える。

【選択図】 図4



【特許請求の範囲】

【請求項 1】

プロセッサとメモリとI/Oインターフェースとを有する複数の計算機と、
前記複数の計算機を前記I/Oインターフェース経由で接続する1以上のI/Oスイッチと、

前記I/Oスイッチに接続する複数のI/Oデバイスと、

前記I/Oスイッチを介して前記計算機に接続するI/Oデバイスを管理する構成管理情報を有し、前記計算機に対する前記I/Oデバイスの割り当てを制御する管理サーバと、を備えた計算機システムにおいて、

前記管理サーバは、

前記第1の計算機から前記複数の計算機のうちの第2の計算機への切り換えを受け付けて、前記第1の計算機に割り当てられた前記I/Oデバイスを前記第2の計算機へ割り当てる構成管理部を有し、

前記構成管理部は、

前記複数の計算機のうちの第1の計算機と、前記第1の計算機に割り当てられたI/Oデバイスの識別子を取得して、前記構成管理情報に格納する識別子検出部と、

前記第1の計算機に割り当てられた前記I/Oデバイスを前記第2の計算機へ切り替える指令を前記I/Oスイッチに送信するI/Oスイッチ切り替え部と、

前記構成管理情報のうち特定のI/Oデバイスの識別子を予め設定した仮想識別子に書き換えるデバイス識別子書き換え部と、を備え、

前記第1の計算機を停止させた後に、前記I/Oスイッチ切り替え部が前記第1の計算機に割り当てられた前記I/Oデバイスを前記第2の計算機へ切り替える指令を前記I/Oスイッチに送信し、

前記第2の計算機を起動した後に、前記デバイス識別子書き換え部が前記第2の計算機へ切り替えられたI/Oデバイスのうち、特定のI/Oデバイスの識別子を予め設定した仮想識別子に書き換えることを特徴とする計算機システム。

【請求項 2】

請求項1に記載の計算機システムであって、

前記構成管理情報は、

前記I/Oスイッチの前記計算機とI/Oデバイスの接続関係を保持し、前記I/Oデバイスの情報として前記識別子と、前記特定のI/Oデバイスを示す情報を有し、

前記識別子検出部は、

前記第1の計算機に割り当てられたI/Oデバイスの識別子を取得し、当該I/Oデバイスが特定のI/Oデバイスであるときに前記特定のI/Oデバイスを示す情報を設定することを特徴とする計算機システム。

【請求項 3】

請求項1に記載の計算機システムであって、

前記構成管理部は、

前記第1の計算機を監視して障害の発生を検知する障害検知部を有し、前記障害検知部が前記第1の計算機の障害の発生を検知したときには前記第1の計算機を停止して前記第2の計算機へ前記I/Oデバイス引き継ぐことを特徴とする計算機システム。

【請求項 4】

請求項1に記載の計算機システムであって、

前記複数の計算機にそれぞれ接続されて前記計算機の稼働状態を管理する第3の計算機をさらに有し、

前記デバイス識別子書き換え部は、

前記第3の計算機に対して、前記第2の計算機へ切り替えられたI/Oデバイスのうち、特定のI/Oデバイスの識別子を前記仮想識別子に書き換える指令を送信することを特徴とする計算機システム。

【請求項 5】

10

20

30

40

50

請求項 1 に記載の計算機システムであって、

前記複数の計算機を管理する第 4 の計算機を接続した第 1 のネットワークと、前記複数の計算機のサービスを提供する第 2 のネットワークと、をさらに有し、

前記 I / O デバイスは、前記第 1 のネットワークに接続された第 1 の I / O デバイスと、前記第 2 のネットワークに接続された第 2 の I / O デバイスとを含み、

前記デバイス識別子書き換え部は、

前記 I / O デバイスのうち前記第 1 のネットワークに接続された第 1 の I / O デバイスを前記特定の I / O デバイスとし、当該第 1 の I / O デバイスの識別子を前記仮想識別子に書き換えることを特徴とする計算機システム。

【請求項 6】

10

請求項 1 に記載の計算機システムであって、

前記デバイス識別子書き換え部は、

前記複数の I / O デバイスの識別子に対応する仮想識別子を予めそれぞれ設定したことを特徴とする計算機システム。

【請求項 7】

プロセッサとメモリと I / O インターフェースとを有する複数の計算機を前記 I / O インターフェース経由で接続する 1 以上の I / O スイッチに複数の I / O デバイスを接続し、前記 I / O スイッチを介して前記計算機に接続する I / O デバイスを管理する構成管理情報を有する管理サーバが、前記計算機に対する前記 I / O デバイスの割り当てを制御する計算機システムの制御方法であって、

20

前記管理サーバが、前記複数の計算機のうち第 1 の計算機と、前記第 1 の計算機に割り当てられた I / O デバイスの識別子を取得して、前記構成管理情報に格納するステップと、

前記管理サーバが、前記第 1 の計算機から前記複数の計算機のうち第 2 の計算機への切り換えを受け付けるステップと、

前記管理サーバが、前記第 1 の計算機を停止させるステップと、

前記管理サーバが、前記第 1 の計算機に割り当てられた前記 I / O デバイスを前記第 2 の計算機へ割り当てる指令を前記 I / O スイッチに送信するステップと、

前記管理サーバが、前記第 2 の計算機を起動させるステップと、

前記管理サーバが、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換えるステップと、を含むことを特徴とする計算機システムの制御方法。

30

【請求項 8】

請求項 7 に記載の計算機システムの制御方法であって、

前記構成管理情報は、

前記 I / O スイッチの前記計算機と I / O デバイスの接続関係を保持し、前記 I / O デバイスの情報として前記識別子と、前記特定の I / O デバイスを示す情報を有し、

前記管理サーバが、前記第 1 の計算機に割り当てられた I / O デバイスの識別子を取得して、前記構成管理情報に格納するステップは、

前記第 1 の計算機に割り当てられた I / O デバイスの識別子を取得し、当該 I / O デバイスが特定の I / O デバイスであるときに前記特定の I / O デバイスを示す情報を設定することを特徴とする計算機システムの制御方法。

40

【請求項 9】

請求項 7 に記載の計算機システムの制御方法であって、

前記管理サーバが、前記第 1 の計算機から前記複数の計算機のうち第 2 の計算機への切り換えを受け付けるステップは、

前記管理サーバが、前記第 1 の計算機を監視して障害の発生を検知したときに、前記第 1 の計算機から前記第 2 の計算機への切り換えを受け付けることを特徴とする計算機システムの制御方法。

【請求項 10】

50

請求項 7 に記載の計算機システムの制御方法であって、
前記複数の計算機にそれぞれ接続されて前記計算機の稼働状態を管理する第 3 の計算機をさらに有し、

前記管理サーバが、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換えるステップは、

前記第 3 の計算機に対して、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を前記仮想識別子に書き換える指令を送信することを特徴とする計算機システムの制御方法。

【請求項 11】

請求項 7 に記載の計算機システムの制御方法であって、
前記複数の計算機を管理する第 4 の計算機を接続した第 1 のネットワークと、前記複数の計算機のサービスを提供する第 2 のネットワークと、をさらに有し、

前記 I / O デバイスは、前記第 1 のネットワークに接続された第 1 の I / O デバイスと、前記第 2 のネットワークに接続された第 2 の I / O デバイスとを含み、

前記管理サーバが、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換えるステップは、

前記 I / O デバイスのうち前記第 1 のネットワークに接続された第 1 の I / O デバイスを前記特定の I / O デバイスとし、当該第 1 の I / O デバイスの識別子を前記仮想識別子に書き換えることを特徴とする計算機システムの制御方法。

【請求項 12】

請求項 7 に記載の計算機システムの制御方法であって、
前記管理サーバが、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換えるステップは、

前記複数の I / O デバイスの識別子に対応する仮想識別子を予めそれぞれ設定したことを特徴とする計算機システムの制御方法。

【請求項 13】

プロセッサとメモリと I / O インターフェースとを有する複数の計算機を前記 I / O インターフェース経由で接続する 1 以上の I / O スイッチに複数の I / O デバイスを接続し、前記 I / O スイッチを介して前記計算機に接続する I / O デバイスを管理する構成管理情報を有する管理サーバが、前記計算機に対する前記 I / O デバイスの割り当てを制御するプログラムであって、

前記複数の計算機の中の第 1 の計算機と、前記第 1 の計算機に割り当てられた I / O デバイスの識別子を取得して、前記構成管理情報に格納する手順と、

前記第 1 の計算機から前記複数の計算機の中の第 2 の計算機への切り換えを受け付ける手順と、

前記第 1 の計算機を停止させる手順と、

前記第 1 の計算機に割り当てられた前記 I / O デバイスを前記第 2 の計算機へ割り当てる指令を前記 I / O スイッチに送信する手順と、

前記第 2 の計算機を起動させる手順と、

前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換える手順と、

を前記管理サーバに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、PCI-Express Switch に接続された計算機の管理に関する。

【背景技術】

【0002】

従来、PCI デバイスは計算機内に搭載されていたが、PCI-Express Switch が実用化されたことにより、計算機外で取り扱うことが可能になった。これにより、P C I バスの切り

10

20

30

40

50

替えが容易となり、柔軟に I / O 構成を変更することが可能となる（例えば、特許文献 1）。

【 0 0 0 3 】

計算機システムの信頼性を向上させるために、現用系サーバと予備系サーバを用意することで、障害時に現用系サーバから予備系サーバへの切り替えを行うリカバリ方法がある。この現用系サーバと予備系サーバを PCI-Express Switch を接続して I / O デバイスを共用することで、計算機システムの信頼性を維持しつつ、柔軟な I / O 構成を組みたいというニーズが高まっている。

【 先行技術文献 】

【 特許文献 】

10

【 0 0 0 4 】

【 特許文献 1 】 特開 2 0 0 5 - 3 0 1 4 8 8

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 5 】

サーバ管理ソフトウェアの中には、管理対象サーバの NIC (Network Interface Card) に関連付けられる MAC アドレス (Media Access Control address) から管理対象サーバの物理位置を判断しているものがある。しかし、前記従来例のように、PCI-Express Switch に接続されている現用系サーバから予備系サーバへの切り替えが発生した場合、PCI-Express Switch により現用系サーバと予備系サーバが同じ PCI デバイスの NIC に接続されるため、NIC に関連付けられる MAC は同じものになる。このため、管理ソフトウェアが、管理対象サーバの物理位置の変更を検知することができず、管理者はサーバの運用及び管理を引き続き継続することができない、という問題があった。

20

【 0 0 0 6 】

そこで本発明は、上記問題点に鑑みてなされたもので、現用系サーバと予備系サーバを PCI-Express Switch に接続して I / O デバイスを共用する状態で、現用系サーバから予備系サーバへの切り替えが行われた場合でも、管理サーバから各サーバの物理位置を把握することを目的とする。

【 課題を解決するための手段 】

【 0 0 0 7 】

30

本発明は、プロセッサとメモリと I / O インターフェースとを有する複数の計算機を前記 I / O インターフェース経由で接続する 1 以上の I / O スイッチに複数の I / O デバイスを接続し、前記 I / O スイッチを介して前記計算機に接続する I / O デバイスを管理する構成管理情報を有する管理サーバが、前記計算機に対する前記 I / O デバイスの割り当てを制御する計算機システムの制御方法であって、前記管理サーバが、前記複数の計算機のうちの第 1 の計算機と、前記第 1 の計算機に割り当てられた I / O デバイスの識別子を取得して、前記構成管理情報に格納し、前記管理サーバが、前記第 1 の計算機から前記複数の計算機のうちの第 2 の計算機への切り換えを受け付け、前記管理サーバが、前記第 1 の計算機を停止させ、前記管理サーバが、前記第 1 の計算機に割り当てられた前記 I / O デバイスを前記第 2 の計算機へ割り当てる指令を前記 I / O スイッチに送信し、前記管理サーバが、前記第 2 の計算機を起動させ、前記管理サーバが、前記第 2 の計算機へ切り替えられた I / O デバイスのうち、特定の I / O デバイスの識別子を予め設定した仮想識別子に書き換える。

40

【 発明の効果 】

【 0 0 0 8 】

したがって、本発明により、管理者は I / O スイッチ (PCI-Express Switch) に接続された計算機に現用系と待機系の切り替えが発生した場合においても、計算機の物理位置が I / O デバイスに固有の識別子から仮想識別子に変わったことを判定することが可能となる。

【 図面の簡単な説明 】

50

【 0 0 0 9 】

【 図 1 】 本発明の実施形態を示し、計算機システムの全体を示すブロック図である。

【 図 2 】 本発明の実施形態を示し、管理サーバ 1 0 1 の構成を示すブロック図である。

【 図 3 】 本発明の実施形態を示し、サーバ装置 1 1 1 の構成を示すブロック図である。

【 図 4 】 本発明の実施形態を示し、動作概要の一つを示す計算機システムのブロック図である。

【 図 5 】 本発明の実施形態を示し、動作概要の一つを示す計算機システムのブロック図で、フェイルオーバの様子を示す。

【 図 6 】 本発明の実施形態を示し、サーバ管理テーブル 1 0 8 を示す説明図である。

【 図 7 】 本発明の実施形態を示し、サーバ I / O 構成情報テーブル 1 0 9 を示す説明図である。

10

【 図 8 】 本発明の実施形態を示し、仮想識別子テーブル 1 2 3 を示す説明図である。

【 図 9 】 本発明の実施形態を示し、管理サーバ 1 0 1 のデバイス識別子検出部 1 0 3 で行われる処理の一例を示すフローチャートである。

【 図 1 0 】 本発明の実施形態を示し、サーバ障害回復部 1 0 4 で行われる処理の一例を示すフローチャートである。

【 図 1 1 】 本発明の実施形態を示し、I / O スイッチ切り替え部 1 0 5 で行われる処理の一例を示すフローチャートである。

【 図 1 2 】 本発明の実施形態を示し、デバイス識別子取得選択部 1 0 6 で行われる処理の一例を示すフローチャートである。

20

【 図 1 3 】 本発明の実施形態を示し、デバイス識別子書き換え部 1 0 7 で行われる処理の一例を示すフローチャートである。

【 発明を実施するための形態 】

【 0 0 1 0 】

以下、本発明の一実施形態を添付図面に基づいて説明する。

【 0 0 1 1 】

図 1 は、本発明の実施形態を示し、計算機システムの全体を示すブロック図である。図 1 の計算機システムは、複数のサーバ装置 1 1 1 で現用系のサーバ装置 1 1 1 と予備系（または待機系）のサーバ装置 1 1 1 を構成し、現用系と予備系で I / O デバイス 1 1 5 を切り換え可能な I / O スイッチ装置 1 1 2 を共用し、管理サーバ 1 0 1 からの指示に応じて現用系と予備系を切り替えるものである。

30

【 0 0 1 2 】

管理サーバ 1 0 1 は本実施形態における制御の中心である。管理サーバ 1 0 1 は、I / O 構成管理部 1 0 2 と各種テーブル（1 0 8 ~ 1 0 9、1 2 3）、デバイス識別子取得プログラム 1 2 1、およびデバイス識別子書き換えプログラム 1 2 2 を実行する。I / O 構成管理部 1 0 2 は、デバイス識別子検出部 1 0 3、サーバ障害回復部 1 0 4、I / O スイッチ切り替え部 1 0 5、デバイス識別子取得選択部 1 0 6、デバイス識別子書き換え部 1 0 7 から構成される。

【 0 0 1 3 】

管理サーバ 1 0 1 は、ネットワークスイッチ 1 1 0 を介して、複数のサーバ装置 1 1 1、複数の I / O スイッチ装置 1 1 2、ファームウェア層の Service Processor（以後、S V P と表記する）1 2 0 に接続される。I / O スイッチ装置 1 1 2 は、サーバ装置 1 1 1 及び S V P 1 2 0 と接続する複数の上流ポート 1 1 3 と、複数の I / O デバイス 1 1 5 と接続する複数の下流ポート 1 1 4 を備えて、サーバ装置 1 1 1 及び S V P 1 2 0 と I / O デバイス 1 1 5 とを接続する。複数の I / O デバイス 1 1 5 のうちのいくつかは、ストレージ装置 1 1 6 と接続される H B A（Host Bus Adaptor）で構成され、サーバ装置 1 1 1 からストレージ装置 1 1 6 をアクセスすることができる。

40

【 0 0 1 4 】

また、複数の I / O デバイス 1 1 5 のうちのいくつかは、管理 LAN スイッチ 4 0 1、業務 LAN スイッチ 4 0 2 に接続される N I C（Network interface Card）で構成され

50

、サーバ装置 1 1 1 から管理 LAN スイッチ 4 0 1 と業務 LAN スイッチ 4 0 2 にアクセスすることができる。

【 0 0 1 5 】

なお、複数のサーバ装置 1 1 1 は添え字 # 1 ~ # 3 で個々のサーバ装置 1 1 1 を識別し、また、複数の I / O スイッチ装置 1 1 2 は、添え字 # 1、# 2 で識別し、上流ポート 1 1 3 と下流ポート 1 1 4 はそれぞれ 0 ~ 3 の添え字で識別し、I / O デバイス 1 1 5 は # 1 ~ # 8 で識別する。

【 0 0 1 6 】

管理 LAN スイッチ 4 0 1 は管理ソフトウェア 4 0 5 0 (図 4 参照) が稼動するサーバ装置 4 0 5 等が、サーバ装置 # 1 ~ # 3 を管理するための管理ネットワークを構成する。なお、サーバ装置 4 0 5 の管理ソフトウェア 4 0 5 0 は、前記従来例で述べたように、サーバ装置 # 1 ~ # 3 に接続された NIC の MAC アドレスで、サーバ装置 # 1 ~ # 3 を実行する。

10

【 0 0 1 7 】

業務 LAN スイッチ 4 0 2 はサーバ装置 # 1 ~ # 3 と外部などの計算機を接続し、サーバ装置 # 1 ~ # 3 のサービスを外部等の計算機に提供する業務ネットワークを構成する。

【 0 0 1 8 】

管理サーバ 1 0 1 は、サーバ装置 1 1 1 や I / O スイッチ装置 1 1 2 や I / O デバイス 1 1 5 の障害を検知し回復する機能を有する。デバイス識別子検出部 1 0 3 は、サーバ装置 1 1 1 に接続された I / O デバイス 1 1 5 のデバイス識別子を検出する機能を有する。I / O デバイス 1 1 5 のデバイス識別子としては、例えば、特定のネットワークに接続されている NIC の MAC、特定のストレージ装置に接続されている H B A (Host Bus Adapter) の WWN (World Wide Name) 等である。

20

【 0 0 1 9 】

サーバ障害回復部 1 0 4 は、サーバ装置 1 1 1 や I / O スイッチ装置 1 1 2 や I / O デバイス 1 1 5 の障害を検知し回復する機能を有する。I / O スイッチ切り替え部 1 0 5 は、サーバ管理テーブル 1 0 8、およびサーバ I / O 構成情報テーブル 1 0 9 の情報を取得し、I / O スイッチ装置 1 1 2 の切り替えを行う機能を有する。

【 0 0 2 0 】

デバイス識別子取得選択部 1 0 6 は、サーバ管理テーブル 1 0 8、およびサーバ I / O 構成情報テーブル 1 0 9 の情報を取得し、特定のデバイス識別子を選択する機能を有する。デバイス識別子書き換え部 1 0 7 は、デバイス識別子取得選択部 1 0 6 によって選択されたデバイス識別子を任意のデバイス識別子に書き換える機能を有する。

30

【 0 0 2 1 】

サーバ管理テーブル 1 0 8 は、サーバ装置 1 1 1 の構成、およびサーバ装置 1 1 1 に接続されている I / O スイッチ装置 1 1 2 の情報が格納される。サーバ I / O 構成情報テーブル 1 0 9 は、サーバ装置 1 1 1 に接続される一つまたは複数の I / O スイッチ装置 1 1 2 と、I / O デバイス 1 1 5 の I / O 構成定義情報や状態などが格納される。サーバ I / O 構成情報テーブル 1 0 9 はスイッチ管理テーブルとして機能する。デバイス識別子取得プログラム 1 2 1 は、I / O デバイス 1 1 5 が有する固有の識別子を取得する機能を有するプログラムが格納される。デバイス識別子書き換えプログラム 1 2 2 は、I / O デバイス 1 1 5 が有する固有の識別子を書き換える機能を有するプログラムが格納される。

40

【 0 0 2 2 】

本実施形態では、管理サーバ 1 0 1 が、複数のサーバ装置 1 1 1 の何れかに障害が発生した場合に、障害が発生したサーバ装置 1 1 1 を一旦停止し、I / O スイッチ装置 1 1 2 の切り替えを行い、障害が発生したサーバ装置 1 1 1 に接続された複数の I / O デバイス 1 1 5 の情報を書き換え、予備系のサーバ装置 1 1 1 を起動して障害が発生したサーバ装置 1 1 1 の I / O デバイス 1 1 5 を引き継ぐ一実施形態を示す。

【 0 0 2 3 】

図 2 は、管理サーバ 1 0 1 の構成を示すブロック図である。管理サーバ 1 0 1 は、メモ

50

り201とプロセッサ202とディスクインターフェース203とネットワークインターフェース204から構成される。メモリ201内には、サーバ管理テーブル108、サーバI/O構成情報テーブル109、デバイス識別子取得プログラム121、デバイス識別子書き換えプログラム122が格納される。

【0024】

I/O構成管理部102は、デバイス識別子検出部103、サーバ障害回復部104、I/Oスイッチ切り替え部105、デバイス識別子取得選択部106、デバイス識別子書き換え部107が含まれる。メモリ内のI/O構成管理部102やテーブル(108~109、123)、デバイス識別子取得プログラム121、およびデバイス識別子書き換えプログラム122は、プロセッサ202に読み込まれて実行される。ディスクインターフェース203は、管理サーバ101を起動するための上記各プログラムが格納された記憶媒体としてのディスク(図示省略)に接続される。ネットワークインターフェース204は、ネットワークスイッチ110等で構成されるネットワークに接続され各装置の障害情報などが転送され、また、管理サーバ101からの指令が転送される。なお、これらの機能はハードウェアで実装しても良い。

10

【0025】

図3は、サーバ装置111の構成を示すブロック図である。図1に示す複数のサーバ装置111(#1~#3)は同一の構成である。サーバ装置111は、メモリ301、プロセッサ302、I/Oスイッチインターフェース303、BMC(Base board Management Controller)304から構成される。メモリ301には、サーバ装置111で処理されるプログラムが格納され、このプログラムはプロセッサ302で実行される。I/Oスイッチインターフェース303は、I/Oスイッチ装置112に接続される。BMC304は、サーバ装置111内のハードウェアに障害が発生した場合に、ネットワークスイッチ110を介してSVP120に障害を通知する機能を有する。BMC304は障害の発生箇所とは独立に動作できるため、メモリ301やプロセッサ302に障害が発生したとしても障害通知を転送することができる。

20

【0026】

なお、本実施形態のI/Oスイッチ装置112、I/Oスイッチインターフェース303及びI/Oデバイス115はPCI-Expressの規格に準拠したものである。

【0027】

また、SVP120は、プロセッサとメモリとネットワークインターフェースを備えた計算機でありサーバ装置111の稼動状態を管理する。SVP120は、各サーバ装置111のBMC304を監視し、BMC304から障害の通知を受信すると、管理サーバ101に障害の発生したサーバ装置111を通知する。SVP120は管理サーバ101からサーバ装置111の起動やリセット等の指令を受信すると、対象となるサーバ装置111のBMC304に対して起動やリセット等を指令する。

30

【0028】

図4は、本発明における動作概要の一つを示している。サーバ装置111は複数のI/Oスイッチ装置112を介して複数のI/Oデバイス115と接続される。また、I/Oデバイス115はデバイスによって接続先が変わる。

40

【0029】

図4の例では、サーバ装置111(#1)が現用系を構成し、サーバ装置111(#3)が予備系を構成する。なお、以下では、各装置を上述の図1に示した添え字で識別する。図中I/Oデバイス#1、#3、#5、#7がNICで構成され、I/Oデバイス#2、#4、#6、#8がHBAで構成された例を示す。

【0030】

現用系のサーバ装置#1はI/Oスイッチインターフェース303を介してI/Oスイッチ装置#1の上流ポート1とI/Oスイッチ装置#2の上流ポート1に接続される。I/Oスイッチ装置#1では、上流ポート1と下流ポート0、1、4が接続される。そして、下流ポート0にはNICで構成されたI/Oデバイス#1が接続され、下流ポート2、

50

4にはHBAで構成されたI/Oデバイス#2、#4が接続される。I/Oスイッチ装置#2では、上流ポート1と下流ポート0が接続される。そして、I/Oスイッチ装置#2の下流ポート0にはNICで構成されたI/Oデバイス#5が接続される。

【0031】

I/Oデバイス#1のNICは、管理LANネットワークスイッチ401に接続され、I/Oデバイス#5のNICは、業務LANスイッチ402に接続される。I/Oデバイス#2のHBAはストレージ装置116のブートディスク403に接続され、I/Oデバイス#4のHBAはストレージ装置116のユーザディスク404に接続される。なお、ストレージ装置116のブートディスク403とユーザディスク404は、ロジカルユニットとして提供される。

10

【0032】

上記のように設定された現用系のサーバ装置#1は、I/Oスイッチ装置#1、#2を介してブートディスク403と、ユーザディスク404にアクセスし、管理LANスイッチ401を介してサーバ装置405と接続し、業務LANスイッチ402を介してサービスを提供する計算機に接続され。

【0033】

上記の構成において、現用系のサーバ装置#1はI/Oスイッチ装置#1、#2を介して接続されたI/Oデバイス#1、#2、#4、#5のうち、管理LANスイッチ401に接続された指定デバイス識別子のみを取得し、管理サーバ101に対して送信する。この指定デバイス識別子はユーザ(または管理者)によって任意に設定可能である。例えば、サーバ装置#1のI/Oデバイス#1、#5がNICの場合、サーバ装置#1は、I/Oスイッチインターフェース303に接続された複数のI/Oデバイス#1、#5のうち、管理LANスイッチ401に接続されたNIC(I/Oデバイス#1)の固有識別子(MAC)のみを指定デバイス識別子として管理サーバ101に送信する。

20

【0034】

すなわち、業務LANスイッチ401は、他の計算機と接続してサーバ装置#1~#3のサービスを提供するため、障害発生時に現用系のサーバ装置#1から予備系のサーバ装置3にフェイルオーバーを行った後も、現用系のサーバ装置#1から予備系のサーバ装置3が引き継いだNIC(I/Oデバイス#5)の識別子(MACアドレス)を変更してはならないネットワークを構成する。

30

【0035】

これに対して、管理LANスイッチ402は、サーバ装置405と接続して管理ソフトウェア4050によりサーバ装置#1~#3の管理を行うため、障害発生時に現用系のサーバ装置#1から予備系のサーバ装置3にフェイルオーバーを行った後は、現用系のサーバ装置#1から予備系のサーバ装置3が引き継いだNIC(I/Oデバイス#3)の識別子(MACアドレス)を変更するネットワークを構成する。

【0036】

図4の状態では、予備系のサーバ装置#3は、I/Oスイッチ装置#1の上流ポート3と、I/Oスイッチ装置#2の上流ポート3にそれぞれ接続されるが、各上流ポート3には下流ポートが接続されていない。

40

【0037】

図5は、本発明における動作概要の一つを示し、フェイルオーバーの例を示している。図5は、図4に示した環境で現用系のサーバ装置#1に障害が発生して、予備系のサーバ装置#3に処理を引き継ぐ例を示している。

【0038】

現用系のサーバ装置#1で障害が発生した場合に、管理サーバ101は現用系のサーバ装置#1を一旦停止する。そして、管理サーバ101はI/Oスイッチ装置112へ現用系のサーバ装置#1から予備系のサーバ装置#3への切り換えを指示し、I/Oスイッチ装置112は上流ポート113と下流ポート114の接続を切り替えることにより、現用系のサーバ装置#1に接続されている全てのI/Oデバイス115を、予備系のサーバ装

50

置 # 3 へ接続する。

【 0 0 3 9 】

つまり、サーバ装置 1 1 1 と I / O スイッチ装置 1 1 2 間の経路は、図 5 に示す経路 5 0 1 から経路 5 0 3 へ、経路 5 0 2 から経路 5 0 4 へと変更される。このとき、I / O スイッチ装置 1 1 2 と I / O デバイス 1 1 5 間の経路は変更されないことが重要である。

【 0 0 4 0 】

次に管理サーバ 1 0 1 は、予備系のサーバ装置 # 3 を起動し、管理 LAN スイッチ 4 0 1 に接続された NIC (I / O デバイス # 1) の特定のデバイス識別子 (MAC) のみを予め設定した仮想の識別子に書き換える。

【 0 0 4 1 】

このとき、管理サーバ 1 0 1 は、管理 LAN スイッチ 4 0 1 に接続されている I / O デバイス # 1 (NIC) のデバイス識別子 (MAC) のみの書き換えを指示し、業務 LAN スイッチ 4 0 2 に接続されている I / O デバイス # 5 (NIC) のデバイス識別子については書き換えは行わないところに特徴がある。また、このデバイス識別子の書き換えは I / O デバイス 1 1 5 が H B A の場合には、デバイス識別子 (WWN) などに対しても適用可能である。

【 0 0 4 2 】

図 6 は、サーバ管理テーブル 1 0 8 を示す。カラム 1 1 0 1 は、サーバ装置識別子を示す。カラム 1 1 0 2 は、サーバ装置 1 1 1 のプロセッサ構成、カラム 1 1 0 3 はメモリ容量が格納される。カラム 1 1 0 4 は、当該サーバ装置 1 1 1 が接続されている I / O スイッチ装置 1 1 2 の識別子が格納される。

【 0 0 4 3 】

カラム 1 1 0 5 は当該サーバ装置 1 1 1 が接続されている I / O スイッチ装置 1 1 2 の上流ポート 1 1 3 のポート番号が格納される。カラム 1 1 0 6 は当該サーバ装置 1 1 1 に割り当てられている I / O デバイス 1 1 5 が接続された下流ポート 1 1 4 のポート番号が格納される。

【 0 0 4 4 】

サーバ管理テーブル 1 0 8 によって、サーバ装置 # 1 ~ # 3 (図中 H O S T 1 ~ 3) に割り当てられた I / O デバイス 1 1 5 の I / O スイッチ装置 1 1 2 の識別子と、下流ポート 1 1 4 のポート番号と、上流ポート 1 1 3 のポート番号の対応関係が保持される。

【 0 0 4 5 】

図 7 は、サーバ I / O 構成情報テーブル 1 0 9 を示す。カラム 1 2 0 2 は I / O スイッチ装置 1 1 2 の識別子が格納される。カラム 1 2 0 2 は、I / O スイッチ装置 1 1 2 の下流ポート 1 1 4 のポート番号が格納される。カラム 1 2 0 3 は下流ポート 1 1 4 に接続された I / O デバイス 1 1 5 の種類が格納される。カラム 1 2 0 4 は、I / O デバイス 1 1 5 の固有の識別子がデバイス識別子として格納される。カラム 1 2 0 5 は、サーバ装置 1 1 1 から通知された指定デバイス識別子が格納される。また、指定デバイス識別子は、接続デバイス 1 2 0 3 に対して複数の指定デバイス識別子が格納される場合もある。

【 0 0 4 6 】

デバイス識別子は、管理対象の I / O デバイス 1 1 5 に固有の識別子で、例えば、MAC や WWN で構成される。指定デバイス識別子は、管理対象のサーバ装置 1 1 1 に接続されている I / O デバイス 1 1 5 のうち、管理ネットワークに接続されている I / O デバイス 1 1 5 のデバイス識別子を示す。なお、指定デバイス識別子はデバイス識別子に代わって、管理ネットワークに接続されていることを示すフラグを用いてもよい。

【 0 0 4 7 】

サーバ I / O 構成情報テーブル 1 0 9 を管理することによって、一つのサーバ装置 1 1 1 に対して複数の I / O 構成を管理することができる。

【 0 0 4 8 】

図 8 は、仮想識別子テーブル 1 2 3 を示す説明図である。仮想識別子テーブル 1 2 3 は、I / O スイッチ装置 1 1 2 に接続されている I / O デバイス 1 1 5 の固有の識別子を格

10

20

30

40

50

納するデバイス識別子として格納するカラム 1 2 3 1 と、管理サーバ 1 0 1 が設定した仮想デバイス識別子を格納するカラム 1 2 3 2 から構成される。

【 0 0 4 9 】

仮想デバイス識別子は、サーバ装置 1 1 1 がフェイルオーバなどで切り替わったことをサーバ装置 4 0 5 に通知するため、I / O デバイス 1 1 5 に固有のデバイス識別子に代わって I / O デバイス 1 1 5 に付与する識別子である。

【 0 0 5 0 】

図 9 は、管理サーバ 1 0 1 のデバイス識別子検出部 1 0 3 で行われる処理の一例を示すフローチャートである。本処理は、管理サーバ 1 0 1 がサーバ装置 1 1 1 を管理する場合に必ず行われる処理で、例えば、サーバ装置 1 1 1 の起動、停止や I / O デバイス 1 1 5 の変更などである。

10

【 0 0 5 1 】

ステップ 1 3 0 1 では、管理サーバ 1 0 1 のデバイス識別子検出部 1 0 3 がサーバ管理テーブル 1 0 8、およびサーバ I / O 構成情報テーブル 1 0 9 から、サーバ装置 1 1 1 の指定デバイス識別子を取得する。ステップ 1 3 0 2 では、デバイス識別子検出部 1 0 3 は、サーバ装置 1 1 1 の指定デバイス識別子情報の取得の有無を判定する。指定デバイス識別子が取得された場合にはステップ 1 3 0 3 へ進み、指定デバイス識別子がなければ処理を終了する。

【 0 0 5 2 】

ステップ 1 3 0 3 では、デバイス識別子検出部 1 0 3 がサーバ装置 1 1 1 に指定デバイス識別子の送信命令を発行する。例えば、サーバ装置 1 1 1 に I / O デバイス (N I C) 1 1 5 が接続されている場合には、M A C アドレスの送信命令を送信する。この送信命令は、複数のサーバ装置 1 1 1 に接続されている複数の I / O デバイス 1 1 5 に対し、複数の指定デバイス識別子の送信命令を行うことが可能である。

20

【 0 0 5 3 】

ステップ 1 3 0 4 では、デバイス識別子検出部 1 0 3 が指定デバイス識別子の送信命令に対する応答として受信した指定デバイス識別子をサーバ I / O 構成情報テーブル 1 0 9 に格納する。

【 0 0 5 4 】

上記処理により、デバイス識別子検出部 1 0 3 は、各サーバ装置 1 1 1 から管理ネットワークに接続されている I / O デバイス 1 1 5 のデバイス識別子を指定デバイス識別子として取得し、サーバ I / O 構成情報テーブル 1 0 9 の指定デバイス識別子 1 2 0 5 に格納する。なお、サーバ装置 1 1 1 は、デバイス識別子検出部 1 0 3 からの指定デバイス識別子の送信命令に対して、管理ネットワークに接続されていない I / O デバイス 1 1 5 については、デバイス識別子を通知しない。例えば、図 4 の構成では、サーバ装置 1 1 1 は、管理 L A N スイッチ 4 0 1 に接続されている I / O デバイス # 1 の M A C を管理サーバ 1 0 1 に応答するが、I / O デバイス # 2、# 4、# 5 のデバイス識別子については管理サーバ 1 0 1 に通知しない。また、サーバ装置 1 1 1 は、管理ネットワークの所定の装置 (例えば、サーバ装置 4 0 5) と通信可能な I / O デバイス 1 1 5 を、管理ネットワークに接続されている I / O デバイス 1 1 5 として判定することができる。

30

40

【 0 0 5 5 】

上記処理は、管理サーバ 1 0 1 が管理対象とするサーバ装置 1 1 1 の全てについて繰り返して行うことができる。

【 0 0 5 6 】

なお、管理サーバ 1 0 1 が管理ネットワークに接続されている場合は、管理サーバ 1 0 1 が管理ネットワークから I / O デバイス 1 1 5 のデバイス識別子を取得するようにしても良い。

【 0 0 5 7 】

図 1 0 は、サーバ障害回復部 1 0 4 で行われる処理の一例を示すフローチャートである。サーバ障害回復部 1 0 4 は、S V P 1 2 0 からサーバ装置 1 1 1 の障害の通知を受信す

50

ると図10の処理を実行する。なお、障害の検知は、SVP120からの通知に限らず、サーバ障害回復部104が各サーバ装置111のハートビートを検知するものであっても良く、公知ないし周知の手法を適用することができる。

【0058】

ステップ1041では、サーバ障害回復部104が、現用系のサーバ装置111(図4のサーバ装置#1)の障害を検知した場合にSVP120から通知された現用系のサーバ装置111の起動を停止する。ステップ1402では、サーバ障害回復部104がSVP120、およびI/Oスイッチ装置112からI/Oスイッチ情報を取得し、サーバ管理テーブル108、およびサーバI/O構成情報テーブル109の更新を行う。I/Oスイッチ情報は、全てのI/Oスイッチ装置112の上流ポート113と下流ポート114の接続関係を示す。ステップ1402では、サーバ障害回復部104が、障害発生により停止した現用系のサーバ装置111に接続されていた下流ポート114を特定し、停止した現用系のサーバ装置111が使用していたI/Oデバイス115を取得する。

10

【0059】

ステップ1403では、停止した現用系のサーバ装置111を予備系のサーバ装置111(図4のサーバ装置#3)に切り替えるため、I/Oスイッチ装置112の切り替えをI/Oスイッチ切り替え部105が実行する。すなわち、I/Oスイッチ切り替え部105は、サーバ障害回復部104が取得した各I/Oスイッチ装置112の上流ポート113と下流ポート114の接続関係から、障害により停止した現用系のサーバ装置111のI/Oデバイス115を、予備系のサーバ装置111に切り替えるよう指令する。この指令は、対象となるI/Oデバイス115の下流ポート114を、予備系のサーバ装置111が接続された上流ポート113に切り替える指令であり、I/Oスイッチ切り替え部105が各I/Oスイッチ装置112に指令する。なお、I/Oスイッチ切り替え部105で実行される処理の詳細については図11で後述する。

20

【0060】

ステップ1404では、I/Oスイッチ切り替え部105がステップ1403で指令したI/Oスイッチ装置112の切り替えの成功、失敗を判定する。この判定は、I/Oスイッチ切り替え部105の指令に対するI/Oスイッチ装置112の応答などに基づいて上流ポート113と下流ポート114の接続の切り替えが成功したか否かを判定することができる。

30

【0061】

ステップ1405では、I/Oスイッチ切り替え部105により障害の発生した現用系のサーバ装置111のI/Oデバイス115が予備系のサーバ装置111に接続された後、サーバ障害回復部104が、予備系のサーバ装置111を起動する。このとき、予備系のサーバ装置111に接続されるI/Oデバイス115が管理ネットワークに接続されるNIC(図4のI/Oデバイス#1)の場合、対象のNICにVLAN(Virtual LAN)を予め設定することで、NICを管理ネットワークから隔離しておいてもよい。これは、管理ネットワークに接続されたサーバ装置405の管理ソフトウェア4050がNICのMACアドレスでサーバ装置111の管理を行うため、I/Oデバイス115が管理ネットワークに接続されたNICで予備系のサーバ装置111をそのまま起動すると、管理ソフトウェア4050は、障害が発生したサーバ装置111が再起動したと誤認するのを防ぐため、このNICをVLANにより管理ネットワークから隔離しておく。

40

【0062】

ステップ1046では、予備系のサーバ装置111に接続されるI/Oデバイス115の指定デバイス識別子の取得、および選択をデバイス識別子取得選択部106が実行する。デバイス識別子取得選択部106は、図12で後述するように、管理ネットワークに接続されたI/Oデバイス115のうち、仮想デバイス識別子を付与するI/Oデバイス115を選択する。図4の例では、管理ネットワークに接続されたI/Oデバイス#1が仮想デバイス識別子の付与対象として選択される。

【0063】

50

ステップ1047では、予備系のサーバ装置111に接続されるI/Oデバイス115の指定デバイス識別子の書き換えをデバイス識別子書き換え部107が実行する。

【0064】

デバイス識別子書き換え部107は、図13で後述するように、上記ステップ1406で選択されたI/Oデバイス115（I/Oデバイス#1のNIC）のデバイス識別子（図8のMAC1）を、仮想識別子テーブル123の仮想デバイス識別子（図8のMAC11）で書き換えるように予備系のサーバ装置111に指令する。

【0065】

上記処理により、障害が発生した現用系のサーバ装置111のI/Oデバイス115を引き継いだ予備系のサーバ装置111は、I/Oデバイス115のうち管理ネットワークに接続されるNIC（I/Oデバイス#1）について、管理サーバ101から仮想デバイス識別子（MAC11）を受信し、NICのデバイス識別子（MAC1）を仮想デバイス識別子（MAC11）に書き換える。

10

【0066】

これにより、管理ネットワークに接続されたサーバ装置405の管理ソフトウェア4050は、新たな仮想デバイス識別子をデバイス識別子として認識し、停止したサーバ装置111を予備系のサーバ装置111が引き継いだことを認識することが可能となる。

【0067】

したがって、現用系サーバ装置111と予備系サーバ装置111をPCI-ExpressのI/Oスイッチ装置112にそれぞれ接続してI/Oデバイス115を共用する状態で、現用系サーバ装置111から予備系サーバ装置111への切り替えが行われた場合でも、管理ネットワークのサーバ装置405の管理ソフトウェア4050は各サーバ装置111の物理位置を把握することができる。

20

【0068】

一方、I/Oデバイス115のうち業務ネットワークに接続されるNICのデバイス識別子は、障害発生前と同一であるので、他の計算機などは障害発生前と同様に予備系のサーバ装置111にアクセスすることができる。

【0069】

なお、管理ネットワークに接続されているI/Oデバイス115をVLANにより隔離した場合には、デバイス識別子を仮想デバイス識別子に書き換えた後に、VLANの設定を変更して管理ネットワークに接続させればよい。

30

【0070】

図11は、I/Oスイッチ切り替え部105で行われる処理の一例を示すフローチャートである。この処理は、上記図10のステップ1403で行われる処理の詳細である。

【0071】

ステップ1501では、I/Oスイッチ切り替え部105が、サーバ管理テーブル108、およびサーバI/O構成情報テーブル109から、障害が発生したサーバ装置111に接続されるI/Oスイッチ装置112のI/O識別子を取得する。

【0072】

ステップ1502では、I/Oスイッチ切り替え部105が、サーバ管理テーブル108、およびサーバI/O構成情報テーブル109から、予備系のサーバ装置111に接続されているI/Oスイッチ装置112のI/O識別子を取得する。ステップ1503では、現用系のサーバ装置111に接続されるI/Oスイッチ装置112のI/Oスイッチ識別子の全てが、予備系のサーバ装置111に接続されるI/Oスイッチ装置112のI/Oスイッチ識別子に含まれるか比較し、I/Oスイッチ装置112の切り替え可能か判別する。この比較は、スイッチ切り替えの判定条件となるため、非常に重要である。I/Oスイッチ装置112の切り替えが不可能な場合のステップ1504では、ユーザ（または管理サーバ101の管理者）にエラーを通知する。

40

【0073】

一方、I/Oスイッチ装置112の切り替えが可能な場合のステップ1505では、現

50

用系のサーバ装置 1 1 1 に接続される I / O スイッチ装置 1 1 2 のポート番号を、予備系のサーバ装置 1 1 1 に接続される I / O スイッチ装置 1 1 2 のポート番号に書き換える命令を全ての I / O スイッチ装置 1 1 2 に送信する。

【 0 0 7 4 】

図 1 2 は、デバイス識別子取得選択部 1 0 6 で行われる処理の一例を示すフローチャートである。この処理は、上記図 1 0 のステップ 1 4 0 6 で行われる処理の詳細である。

【 0 0 7 5 】

ステップ 1 6 0 1 では、デバイス識別子取得選択部 1 0 6 はデバイス識別子取得プログラム 1 2 1 により全てのサーバ装置 1 1 1 に接続されている I / O デバイス 1 1 5 のデバイス識別子を全て取得する。

10

【 0 0 7 6 】

ステップ 1 6 0 2 では、デバイス識別子取得選択部 1 0 6 は上記ステップ 1 6 0 1 で取得したデバイス識別子をサーバ I / O 構成情報テーブル 1 0 9 に格納する。ステップ 1 6 0 3 では、サーバ管理テーブル 1 0 8、およびサーバ I / O 構成情報テーブル 1 0 9 から、障害が発生した現用系のサーバ装置 1 1 1 に接続される I / O スイッチ装置 1 1 2 の指定デバイス識別子を取得する。

【 0 0 7 7 】

ステップ 1 6 0 4 では、デバイス識別子取得選択部 1 0 6 はステップ 1 6 0 2 で取得した指定デバイス識別子を検索キーとして、仮想識別子テーブル 1 2 3 を検索し、一致するデバイス識別子が存在するか判定する。この検索は、書き換え対象のデバイス識別子の有無を判定するため、非常に重要な意味を持つ。ステップ 1 6 0 5 では、ステップ 1 6 0 4 で一致したデバイス識別子に対応する仮想デバイス識別子 1 2 3 2 を書き換え対象として選択する。

20

【 0 0 7 8 】

図 1 3 は、デバイス識別子書き換え部 1 0 7 で行われる処理の一例を示すフローチャートである。この処理は、上記図 1 0 のステップ 1 4 0 7 で行われる処理の詳細である。

【 0 0 7 9 】

ステップ 1 7 0 1 では、デバイス識別子書き換え部 1 0 7 が、デバイス識別子取得選択部 1 0 6 で書き換え対象のデバイス識別子が選択されているか否かを判別する。デバイス識別子取得選択部 1 0 6 で書き換え対象のデバイス識別子が選択されている場合は、ステップ 1 7 0 2 で、デバイス識別子取得選択部 1 0 6 が書き換え対象のデバイス識別子を仮想デバイス識別子に書き換える。このとき、デバイス識別子取得選択部 1 0 6 で書き換え対象のデバイス識別子のみを書き換え、他のデバイス識別子は全て書き換えないところが重要である。つまり、管理ネットワークに接続されている I / O デバイス 1 1 5 のデバイス識別子のみを仮想デバイス識別子に書き換えることで、起動した予備系のサーバ装置 1 1 1 をサーバ装置 4 0 5 の管理ソフトウェア 4 0 5 0 に認識させる。一方、その他の I / O デバイス 1 1 5 については、現用系のサーバ装置 1 1 1 で用いていたデバイス識別子をそのまま使用することで、予備系のサーバ装置 1 1 1 は切り替え以前と同一の環境でサービスの提供やストレージ装置 1 1 6 へのアクセスを行うことができる。

30

【 0 0 8 0 】

なお、上記では、障害が発生したときに予備系のサーバ装置 1 1 1 へ切り替える例を示したが、現用系のサーバ装置 1 1 1 の保守などで管理サーバ 1 0 1 が予備系のサーバ装置 1 1 1 への切り換えを指示した場合にも、上述したように管理ソフトウェア 4 0 5 0 からアクセスされる I / O デバイス 1 1 5 のデバイス識別子を、管理サーバ 1 0 1 で予め設定された仮想デバイス識別子に書き換えるようにしても良い。この場合、サーバ障害回復部 1 0 4 は、サーバ切り替え部として機能し、管理サーバ 1 0 1 の図示しないコンソールなどからの指令によって、現用系のサーバ装置 1 1 1 から予備系のサーバ装置 1 1 1 を指定して切り替えを実行する。

40

【 0 0 8 1 】

また、I / O デバイス 1 1 5 のデバイス識別子を仮想デバイス識別子に書き換える処理

50

については、上述したように管理サーバ101が予備系のサーバ装置111へ指令する他、管理サーバ101がSVP120にデバイス識別子と仮想デバイス識別子を指示し、SVP120がBMC304を経由して対象のI/Oデバイス115のデバイス識別子を仮想デバイス識別子に書き換えるようにしても良い。

【0082】

また、上記では管理サーバ101と、MACアドレスでサーバ装置111の物理位置を管理する管理ソフトウェア4050を実行するサーバ装置405を異なる計算機で構成した例を示したが、管理サーバ101で管理ソフトウェア4050を実行しても良い。この場合、管理サーバ101に複数のネットワークインターフェースを設けて、ネットワークスイッチ110と管理LANスイッチ401にそれぞれ接続すればよい。

10

【0083】

また、上記ではサーバ装置111とI/Oスイッチ装置112とポートの関係を保持するサーバ管理テーブル108と、I/Oスイッチ装置112のポートとI/Oデバイスの情報(種類及びデバイス識別子)とサーバ装置111の関係を保持するサーバI/O構成情報テーブル109と、デバイス識別子と仮想デバイス識別子を保持する仮想識別子テーブル123を分離した例を示したが、I/Oスイッチ装置112のポート毎に接続されたサーバ装置111とI/Oデバイスの情報と仮想識別子の関係を保持する構成管理情報であればよい。

【産業上の利用可能性】

【0084】

以上のように、本発明はPCI-Express Switchを備えて複数の計算機でI/Oデバイスを共用する計算機システムに適用することができる。

20

【符号の説明】

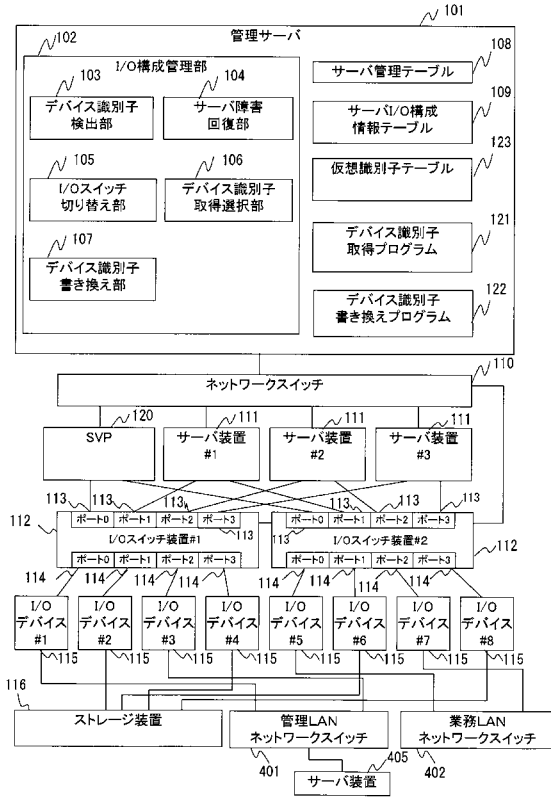
【0085】

- 101 管理サーバ
- 102 I/O構成管理部
- 103 デバイス識別子検出部
- 104 サーバ障害回復部
- 105 I/Oスイッチ切り替え部
- 106 デバイス識別子取得選択部
- 107 デバイス識別子書き換え部
- 108 サーバ管理テーブル
- 109 サーバI/O構成情報テーブル
- 110 ネットワークスイッチ
- 111 サーバ装置
- 112 I/Oスイッチ装置
- 113 上流ポート
- 114 下流ポート
- 115 I/Oデバイス
- 116 ストレージ装置
- 120 SVP
- 304 BMC

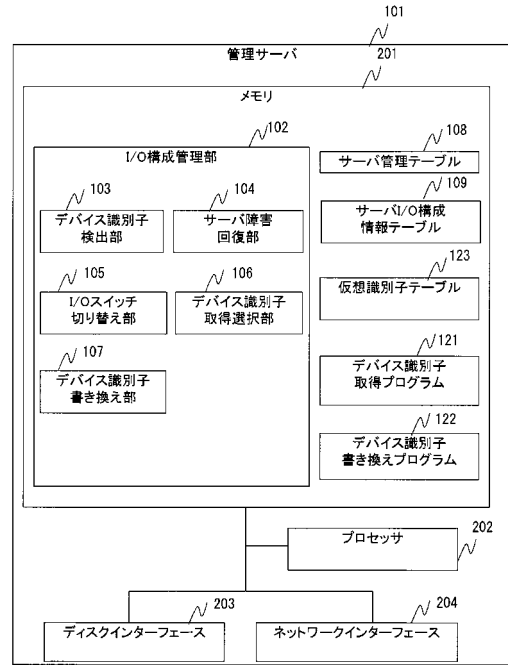
30

40

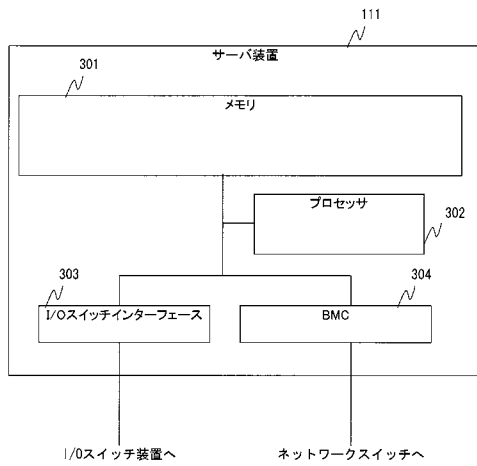
【図1】



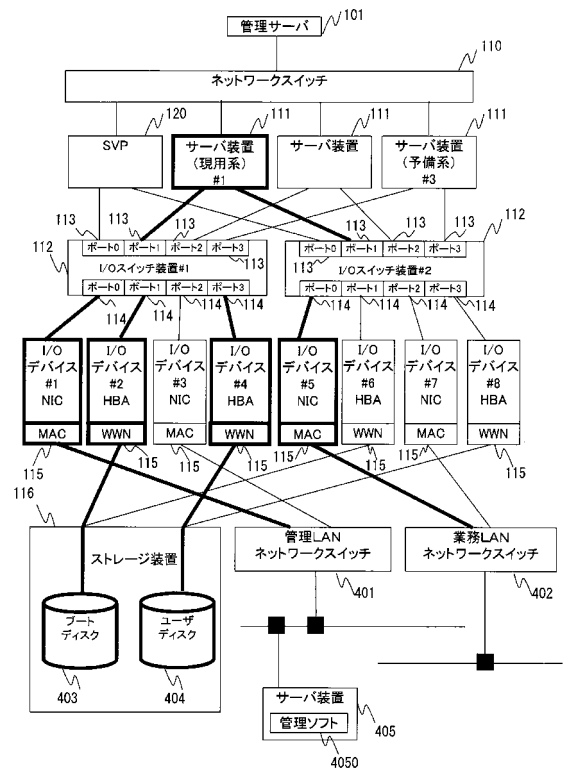
【図2】



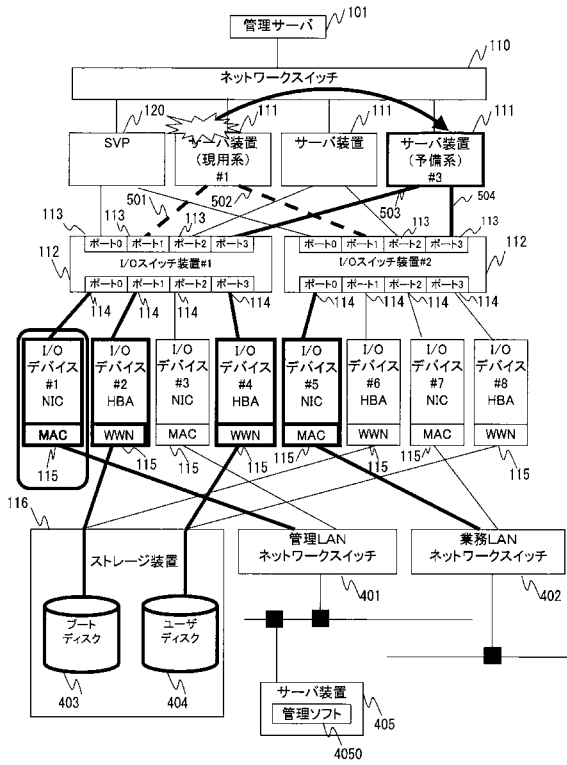
【図3】



【図4】



【図5】



【図6】

サーバ管理テーブル 1108

サーバ装置識別子 1101	プロセッサ構成 1102	メモリ容量 1103	サーバ接続 I/Oスイッチ識別子 1104	割り当て I/Oスイッチ上流ポート番号 1105	割り当て I/Oスイッチ下流ポート番号 1106
HOST1	プロセッサ1 2GHz×2	4GB	SW1	ポート1	ポート0.1.3
			SW2	ポート1	ポート0
HOST2	プロセッサ1 1GHz	2GB	SW1	ポート2	ポート2
			SW2	ポート2	ポート2
HOST3	プロセッサ1 1GHz	2GB	SW1	ポート3	-
			SW2	ポート3	-

【図7】

サーバI/O構成情報テーブル 1109

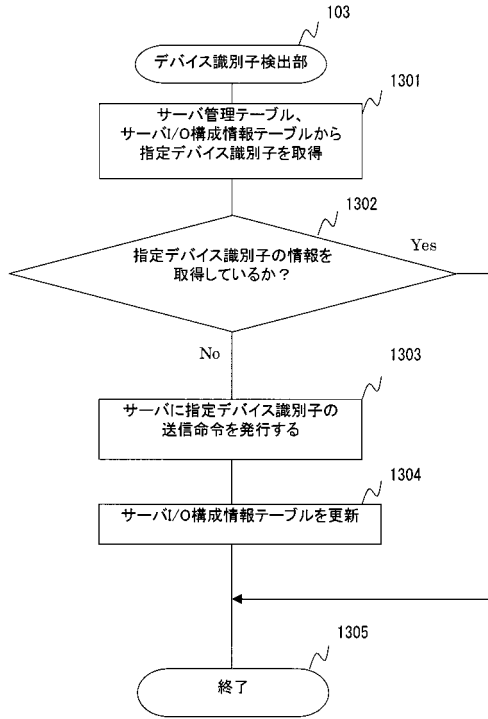
I/Oスイッチ識別子 1201	ポート番号 1202	接続デバイス 1203	デバイス識別子 1204	指定デバイス識別子 1205
SW1	0	NIC	MAC1	MAC1
SW1	1	HBA	WWN1	WWN1
SW1	2	NIC	MAC2	-
SW1	3	HBA	WWN2	WWN2
SW2	0	NIC	MAC4	MAC3
SW2	1	HBA	WWN3	-
SW2	2	NIC	MAC4	-
SW2	3	HBA	WWN4	-

【図8】

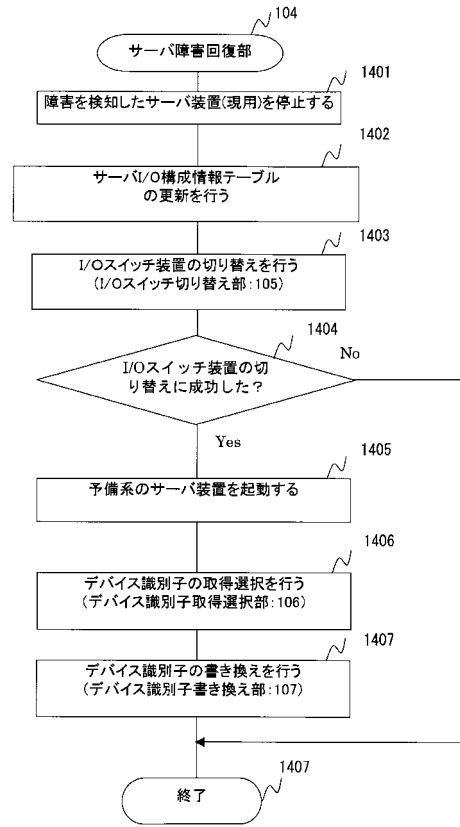
仮想識別子テーブル 1230

デバイス識別子 1231	仮想デバイス識別子 1232
WWN1	WWN11
WWN2	WWN12
MAC1	MAC11
WWN3	WWN13
MAC2	MAC12
MAC3	MAC13
WWN4	WWN14
MAC4	MAC14
MAC5	MAC15

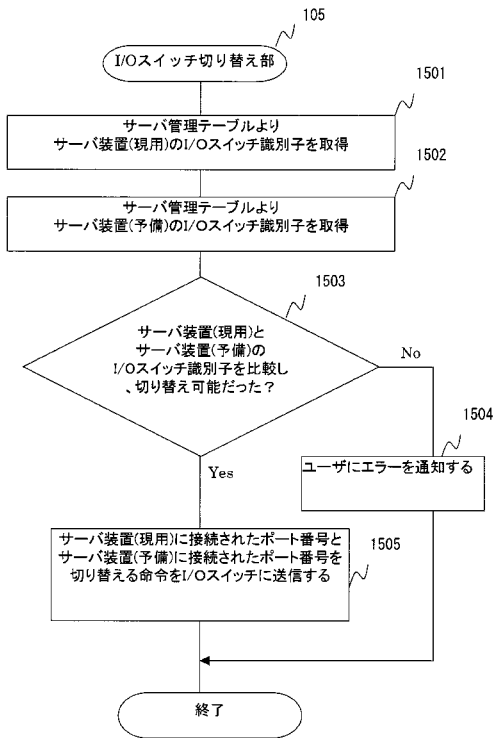
【 図 9 】



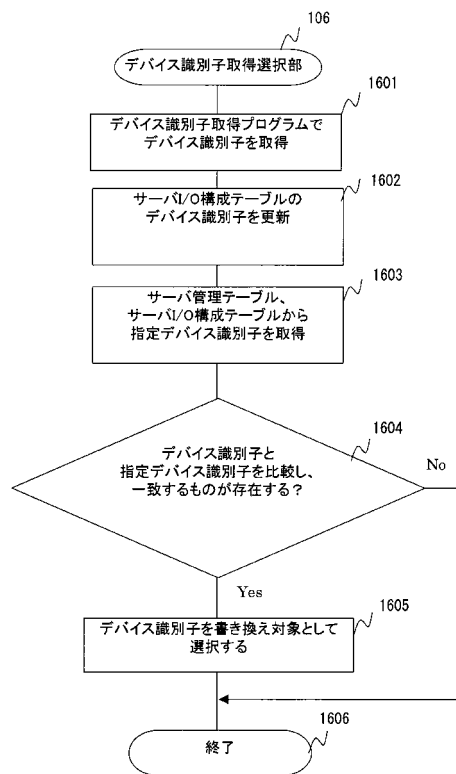
【 図 1 0 】



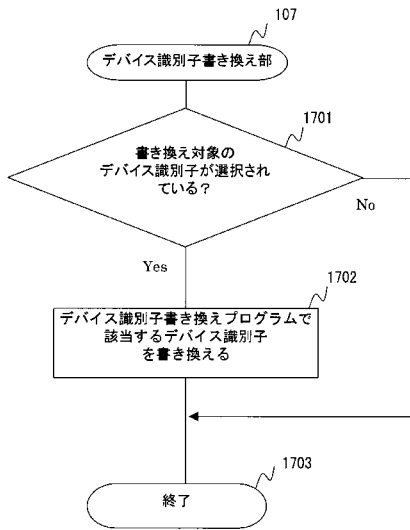
【 図 1 1 】



【 図 1 2 】



【 図 1 3 】



フロントページの続き

(72)発明者 大西 洋司

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア事業部内

Fターム(参考) 5B014 EA05 HC01

5B034 BB02 CC01 DD01 DD02