

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(10) 国際公開番号

WO 2011/11245 A1

PCT

(43) 国際公開日
2011年9月15日(15.09.2011)

- (51) 国際特許分類:
G06F 11/20 (2006.01) G06F 13/14 (2006.01)
- (21) 国際出願番号: PCT/JP2010/063276
- (22) 国際出願日: 2010年8月5日(05.08.2010)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2010-055544 2010年3月12日(12.03.2010) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).
- (72) 発明者: および
- (75) 発明者/出願人 (米国についてのみ): 若松 峻彦 (WAKAMATSU, Takahiko) [JP/JP]; 〒2448555 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 Kana-

gawa (JP). 大西 洋司 (ONISHI, Yoji) [JP/JP]; 〒2448555 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 Kanagawa (JP).

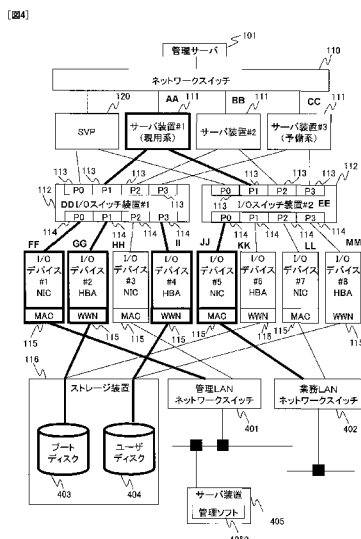
(74) 代理人: 後藤 政喜 (GOTO, Masaki); 〒1000013 東京都千代田区霞が関三丁目3番1号尚友会館 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[続葉有]

(54) Title: COMPUTER SYSTEM, CONTROL METHOD OF COMPUTER SYSTEM, AND STORAGE MEDIUM ON WHICH PROGRAM IS STORED

(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を格納した記憶媒体



(57) Abstract: Disclosed is a control method of a computer system wherein a management server, comprising configuration management information for managing I/O switches for connecting a plurality of computers with a plurality of I/O devices, controls assignment of the I/O devices for the computers; wherein the management server acquires an identifier of an I/O device that has been assigned to a first computer and stores thereof in the configuration management information, receives a switch from the first computer to a second computer, stops the first computer, assigns the I/O device that had been assigned to the first computer to the second computer, activates the second computer, and rewrites the identifier of a specific I/O device among the I/O devices that have been switched to the second computer to a pre-set virtual identifier.

(57) 要約: 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。

- 101 MANAGEMENT SERVER
- 110 NETWORK SWITCH
- AA SERVER DEVICE #1 (CURRENTLY USED SYSTEM)
- BB SERVER DEVICE #2
- CC SERVER DEVICE #3 (AUXILIARY SYSTEM)
- DD I/O SWITCH DEVICE #1
- EE I/O SWITCH DEVICE #2
- FF I/O DEVICE #1 NIC
- GG I/O DEVICE #2 HBA
- HH I/O DEVICE #3 NIC
- II I/O DEVICE #4 HBA
- JJ I/O DEVICE #5 NIC
- KK I/O DEVICE #6 HBA
- LL I/O DEVICE #7 NIC
- MM I/O DEVICE #8 HBA
- 116 STORAGE DEVICE
- 401 MANAGEMENT LAN NETWORK SWITCH
- 402 BUSINESS LAN NETWORK SWITCH
- 403 BOOT DISK
- 404 USER DISK
- 405 SERVER DEVICE
- 4050 MANAGEMENT SOFTWARE

WO 2011/11245 A1



(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF,

BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

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を格納した記憶媒体

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デバイスが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、例えば、特許文献1に記載されるように、PCIバスの切り替えが容易となり、柔軟にI/O構成を変更することが可能となる。

[0003] 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によって、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

[0004] 特許文献1：特開2005-301488号公報

発明の概要

発明が解決しようとする課題

[0005] 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。

[0006] 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

[0007] 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。

発明の効果

[0008] したがって、本発明により、管理者はI/Oスイッチ（PCI-Express Switc

h) 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

[0009] [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 構成情報テーブルを示す説明図である。

[図8] 本発明の実施形態を示し、仮想識別子テーブルを示す説明図である。

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

[図10] 本発明の実施形態を示し、サーバ障害回復部で行われる処理の一例を示すフローチャートである。

[図11] 本発明の実施形態を示し、I/O スイッチ切り替え部で行われる処理の一例を示すフローチャートである。

[図12] 本発明の実施形態を示し、デバイス識別子取得選択部で行われる処理の一例を示すフローチャートである。

[図13] 本発明の実施形態を示し、デバイス識別子書き換え部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

[0010] 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。

- [0011] 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111で現用系のサーバ装置111と予備系（または待機系）のサーバ装置111を構成し、現用系と予備系でI/Oデバイス115を切り換え可能なI/Oスイッチ装置112を共用し、管理サーバ101からの指示に応じて現用系と予備系を切り替える。
- [0012] 管理サーバ101は本実施形態の計算機システムにおける制御の中心である。管理サーバ101は、I/O構成管理部102と各種テーブル（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を実行する。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、およびデバイス識別子書き換え部107を有する。
- [0013] 管理サーバ101は、ネットワークスイッチ110を介して、複数のサーバ装置111、複数のI/Oスイッチ装置112、ファームウェア層の Service Processor（以後、SVPと表記する）120に接続される。I/Oスイッチ装置112は、サーバ装置111及びSVP120と接続する複数の上流ポート113と、複数のI/Oデバイス115と接続する複数の下流ポート114を備えて、サーバ装置111及びSVP120とI/Oデバイス115とを接続する。複数のI/Oデバイス115のうちいくつかは、ストレージ装置116と接続されるHBA（Host Bus Adaptor）で構成され、サーバ装置111からストレージ装置116をアクセスすることができる。
- [0014] また、複数のI/Oデバイス115のうちいくつかは、管理LANスイッチ401、業務LANスイッチ402に接続されるNIC（Network interface Card）で構成され、サーバ装置111から管理LANスイッチ401と業務LANスイッチ402にアクセスすることができる。
- [0015] なお、複数のサーバ装置111は添え字#1～#3で個々のサーバ装置111を識別し、また、複数のI/Oスイッチ装置112は、添え字#1、#

2で識別し、上流ポート113と下流ポート114はそれぞれ0~3の添え字で識別し、I/Oデバイス115は#1~#8で識別する。

- [0016] 管理LANスイッチ401は管理ソフトウェア4050（図4参照）が稼動するサーバ装置405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405の管理ソフトウェア4050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。
- [0017] 業務LANスイッチ402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。
- [0018] 管理サーバ101は、サーバ装置111やI/Oスイッチ装置112やI/Oデバイス115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103は、サーバ装置111に接続されたI/Oデバイス115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115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HBAのWWN（World Wide Name）等である。
- [0019] サーバ障害回復部104は、サーバ装置111やI/Oスイッチ装置112やI/Oデバイス115の障害を検知し、検知された障害を回復する機能を有する。I/Oスイッチ切り替え部105は、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の情報を取得し、I/Oスイッチ装置112の切り替えを行う機能を有する。
- [0020] デバイス識別子取得選択部106は、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の情報を取得し、取得した情報に基づいて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107は、デバイス識別子取得選択部106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。
- [0021] サーバ管理テーブル108は、サーバ装置111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。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。

[0022] 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 を切り替え、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。

[0023] 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、メモリ 2 0 1 とプロセッサ 2 0 2 とディスクインターフェース 2 0 3 とネットワークインターフェース 2 0 4 から構成される。メモリ 2 0 1 内には、サーバ管理テーブル 1 0 8、サーバ I/O 構成情報テーブル 1 0 9、デバイス識別子取得プログラム 1 2 1、デバイス識別子書き換えプログラム 1 2 2 が格納される。

[0024] 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、およびデバイス識別子書き換え部 1 0 7 を含む。メモリ内の I/O 構成管理部 1 0 2、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 は、プロセッサ 2 0 2 に読み込まれて実行される。ディスクインターフェース 2 0 3 は、管理サーバ 1 0 1 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 2 0 4 は、ネットワークス

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よい。

[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、および BMC (Base board Management Controller) 304 を有する。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。BMC 304 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 SVP 120 に障害を通知する機能を有する。BMC 304 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。

[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、SVP 120 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。SVP 120 は、各サーバ装置 111 の BMC 304 を監視し、BMC 304 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。SVP 120 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 BMC 304 に対して起動やリセット等を指令する。

[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。

[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、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。

[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る。

[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。

[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。

[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。

[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witch 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、予備系のサーバ装置 # 3 へ接続する。

[0039] つまり、サーバ装置 111 と I/O スwitch 装置 112 間の経路は、図 5 に示す経路 501 から経路 503 へ、経路 502 から経路 504 へと変更される。このとき、I/O スwitch 装置 112 と I/O デバイス 115 間の経路は変更されないことが重要である。

[0040] 次に管理サーバ 101 は、予備系のサーバ装置 # 3 を起動し、管理 LAN スwitch 401 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。

[0041] このとき、管理サーバ 101 は、管理 LAN スwitch 401 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witch 402 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 115 が HBA 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。

[0042] 図 6 は、サーバ管理テーブル 108 を示す。カラム 1101 は、サーバ装置識別子を示す。カラム 1102 は、サーバ装置 111 のプロセッサ構成、カラム 1103 はメモリ容量が格納される。カラム 1104 は、当該サーバ装置 111 が接続されている I/O スwitch 装置 112 の識別子が格納される。

[0043] カラム 1105 は当該サーバ装置 111 が接続されている I/O スwitch 装置 112 の上流ポート 113 のポート番号が格納される。カラム 1106 は当該サーバ装置 111 に割り当てられている I/O デバイス 115 が接続された下流ポート 114 のポート番号が格納される。

- [0044] サーバ管理テーブル108によって、サーバ装置#1~#3（図中HOST1~3）に割り当てられたI/Oデバイス115のI/Oスイッチ装置112の識別子と、下流ポート114のポート番号と、上流ポート113のポート番号の対応関係が保持される。
- [0045] 図7は、サーバI/O構成情報テーブル109を示す。カラム1202はI/Oスイッチ装置112の識別子が格納される。カラム1202は、I/Oスイッチ装置112の下流ポート114のポート番号が格納される。カラム1203は下流ポート114に接続されたI/Oデバイス115の種類が格納される。カラム1204は、I/Oデバイス115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1205は、サーバ装置111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1203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。
- [0046] デバイス識別子は、管理対象のI/Oデバイス115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111に接続されているI/Oデバイス115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115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。
- [0047] サーバI/O構成情報テーブル109を管理することによって、一つのサーバ装置111に対して複数のI/O構成を管理することができる。
- [0048] 図8は、仮想識別子テーブル123を示す説明図である。仮想識別子テーブル123は、I/Oスイッチ装置112に接続されているI/Oデバイス115の固有の識別子を格納するデバイス識別子として格納するカラム1231と、管理サーバ101が設定した仮想デバイス識別子を格納するカラム1232から構成される。
- [0049] 仮想デバイス識別子は、サーバ装置111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405に通知するため、I/Oデバイス115に固

有のデバイス識別子に代わって I/O デバイス 115 に付与する識別子である。

[0050] 図 9 は、管理サーバ 101 のデバイス識別子検出部 103 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 101 がサーバ装置 111 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 111 の起動、停止や I/O デバイス 115 の変更などである。

[0051] ステップ 1301 では、管理サーバ 101 のデバイス識別子検出部 103 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、サーバ装置 111 の指定デバイス識別子を取得する。ステップ 1302 では、デバイス識別子検出部 103 は、サーバ装置 111 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 1303 へ進み、指定デバイス識別子がなければ処理を終了する。

[0052] ステップ 1303 では、デバイス識別子検出部 103 がサーバ装置 111 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 111 に I/O デバイス (NIC) 115 が接続されている場合には、MAC 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 111 に接続されている複数の I/O デバイス 115 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。

[0053] ステップ 1304 では、デバイス識別子検出部 103 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 109 に格納する。

[0054] 上記処理により、デバイス識別子検出部 103 は、各サーバ装置 111 から管理ネットワークに接続されている I/O デバイス 115 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 109 の指定デバイス識別子 1205 に格納する。なお、サーバ装置 111 は、デバイス識別子検出部 103 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 115 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111は、管理LANスイッチ401に接続されているI/Oデバイス#1のMACを管理サーバ101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101に通知しない。また、サーバ装置111は、管理ネットワークの所定の装置（例えば、サーバ装置405）と通信可能なI/Oデバイス115を、管理ネットワークに接続されているI/Oデバイス115として判定することができる。

[0055] 上記処理は、管理サーバ101が管理対象とするサーバ装置111の全てについて繰り返して行うことができる。

[0056] なお、管理サーバ101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101が管理ネットワークからI/Oデバイス115のデバイス識別子を取得するようにしてもよい。

[0057] 図10は、サーバ障害回復部104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104は、SVP120からサーバ装置111の障害の通知を受信すると図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よく、公知または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を更新する。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。

[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。

[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。

[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 1

11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。

[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] 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。

[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。
- [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。
- [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定する。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。

[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のポート番号を、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112に送信する。

[0074] 図12は、デバイス識別子取得選択部106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1406で行われる処理の詳細である。

[0075] ステップ1601では、デバイス識別子取得選択部106はデバイス識別子取得プログラム121により全てのサーバ装置111に接続されているI/Oデバイス115のデバイス識別子を全て取得する。

[0076] ステップ1602では、デバイス識別子取得選択部106は上記ステップ1601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109に格納する。ステップ1603では、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の指定デバイス識別子を取得する。

[0077] ステップ1604では、デバイス識別子取得選択部106はステップ1602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123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1605では、ステップ1604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1232を書き換え対象として選択する。

[0078] 図13は、デバイス識別子書き換え部107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1407で行われる処理の詳細である。

[0079] ステップ1701では、デバイス識別子書き換え部107が、デバイス識別子取得選択部106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定する。デバイス識別子取得選択部106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1702で、デバイス識別子取得選択部106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106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115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111をサーバ装置405の管理ソフトウェア4050に認識させる。一方、その他のI/Oデバイス115については、現用系のサーバ装置111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111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116へのアクセスを行うことができる。

[0080] 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111へ切り替える例を示したが、現用系のサーバ装置111の保守などで管理サーバ101が予備系のサーバ装置111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4050からアクセスされるI/Oデバイス115のデバイス識別子を、管理サーバ101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よい。この場合、サーバ障害回復部104は、サーバ切り替え部として機能し、管理サーバ101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111から予備系のサーバ装置111への切り替えを実行する。

[0081] また、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処理については、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、SVP 120がBMC 304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よい。

[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を実行してもよい。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。

[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保持する構成管理情報であればよい。

[0084] 以上、本発明を添付の図面を参照して詳細に説明したが、本発明はこのような具体的構成に限定されるものではなく、添付した請求の範囲の趣旨内における様々な変更及び同等の構成を含むものである。

産業上の利用可能性

[0085] 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。

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 の計算機の識別子と、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、を有し、
- 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は、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は、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され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 スイッチに接続された前記計算機と I / O デバイスとの接続関係、および、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保持し、

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 デバイスである場合に、前記特定の I / O デバイスを示す情報を前記構成管理情報に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場合には、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、前記各計算機の稼働状態を管理する第 3 の計算機をさらに備え、

前記デバイス識別子書き換え部は、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、前記特定の 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前記第 3 の計算機に送信することを特徴とする計算機システム。

[請求項5]

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、サービスを提供する前記複数の計算機を接続する第2のネットワークと、をさらに備え、

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デバイスと判定し、

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] 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それぞれ予め設定する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と、前記計算機を前記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とを有する計算機システムにおいて、前記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第1の計算機の識別子と、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格納ステップと、

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うちの第2の計算機への切り換えを受け付ける受付ステップと、

前記管理サーバが、前記第1の計算機を停止させる停止ステップと

、
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送信ステップと、

前記管理サーバが、前記第 2 の計算機を起動させる起動ステップと

、
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された仮想識別子に書き換える書き換えステップと、

を含むことを特徴とする計算機システムの制御方法。

[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 スイッチに接続された前記計算機と I / O デバイスとの接続関係、および、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保持し

、
前記格納ステップでは、

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、

当該 I / O デバイスが特定の I / O デバイスである場合に、前記管理サーバが、前記特定の I / O デバイスを示す情報を前記構成管理情報に設定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 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、

前記受付ステップでは、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場合に、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]

請求項 7 に記載の計算機システムの制御方法であって、

前記計算機システムは、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の計算機をさらに備え、

前記書き換えステップでは、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、前記管理サーバは、前記特定の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前記第3の計算機に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、

前記計算機システムは、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、サービスを提供する前記複数の計算機のを接続する第2のネットワークと、をさらに備え、

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、

前記書き換えステップでは、前記管理サーバは、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デバイスと判定し、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、

前記書き換えステップでは、前記管理サーバは、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それぞれ予め設定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と、前記計算機を前記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とを有する計算機システムにおいて、前記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が格納された記憶媒体であって、

前記複数の計算機のうちの第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うちの第2の計算機への切り換えを受け付ける手順と、

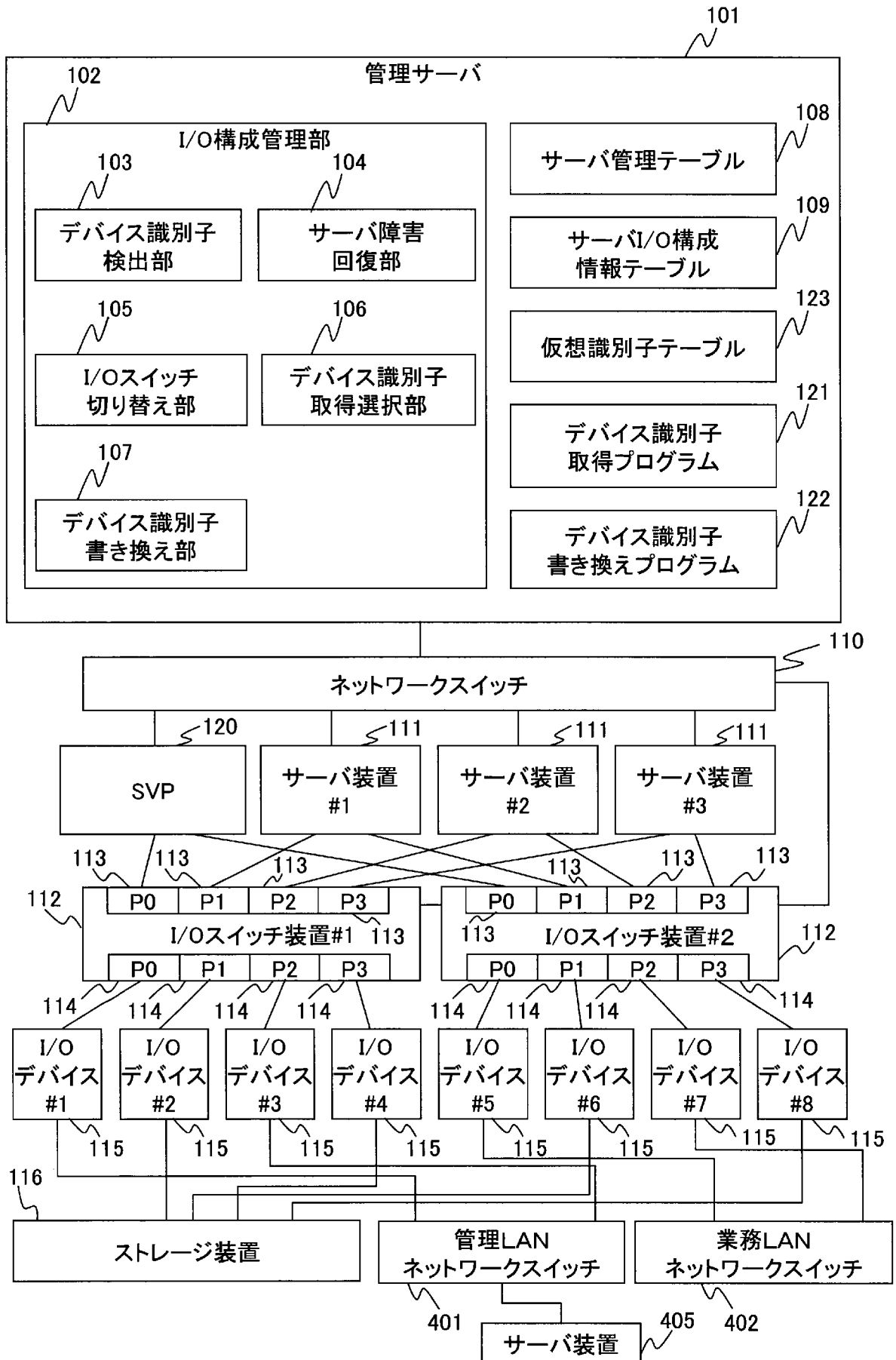
前記第1の計算機を停止させる手順と、

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、

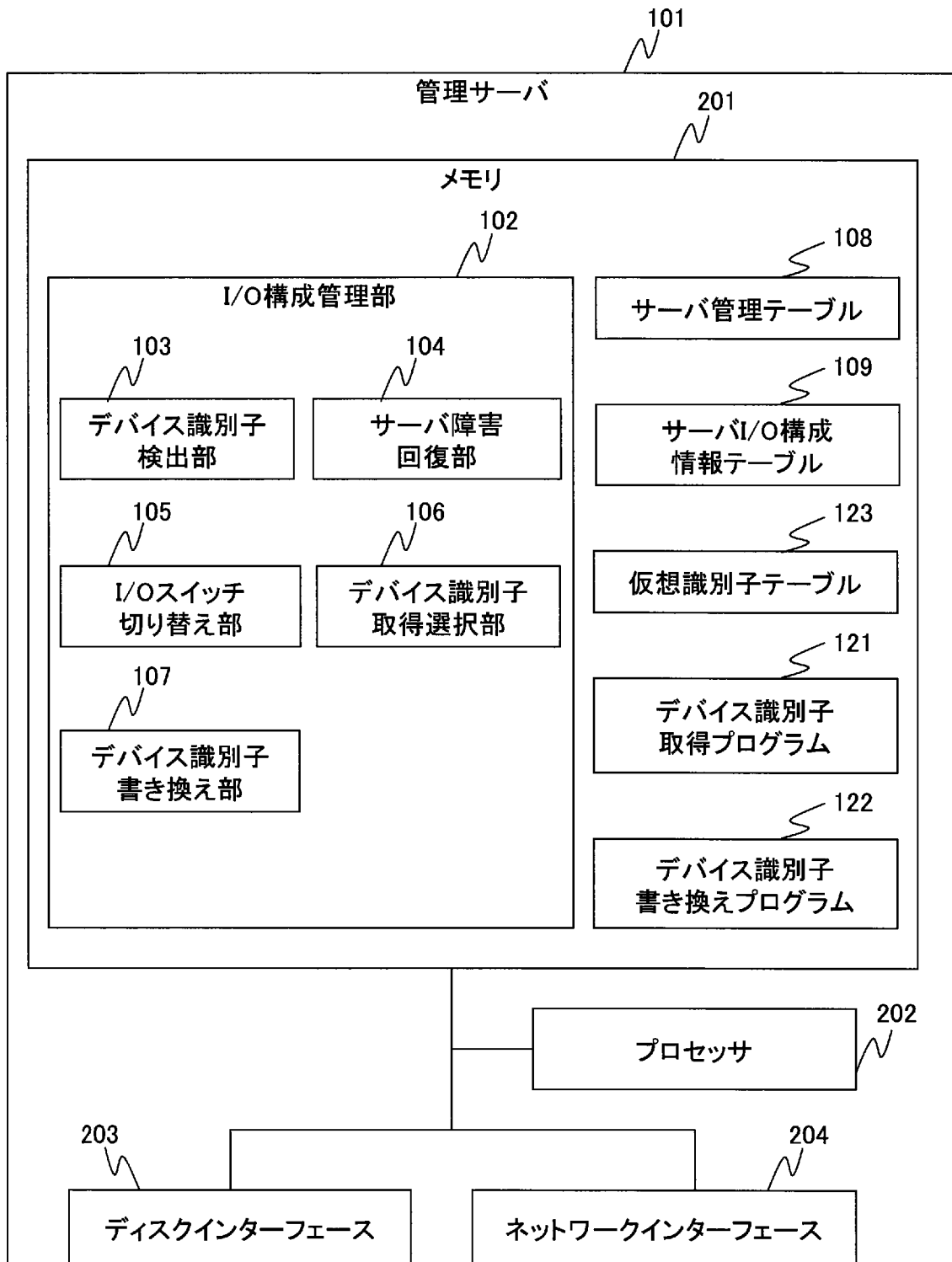
前記第2の計算機を起動させる手順と、

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され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を格納した記憶媒体。

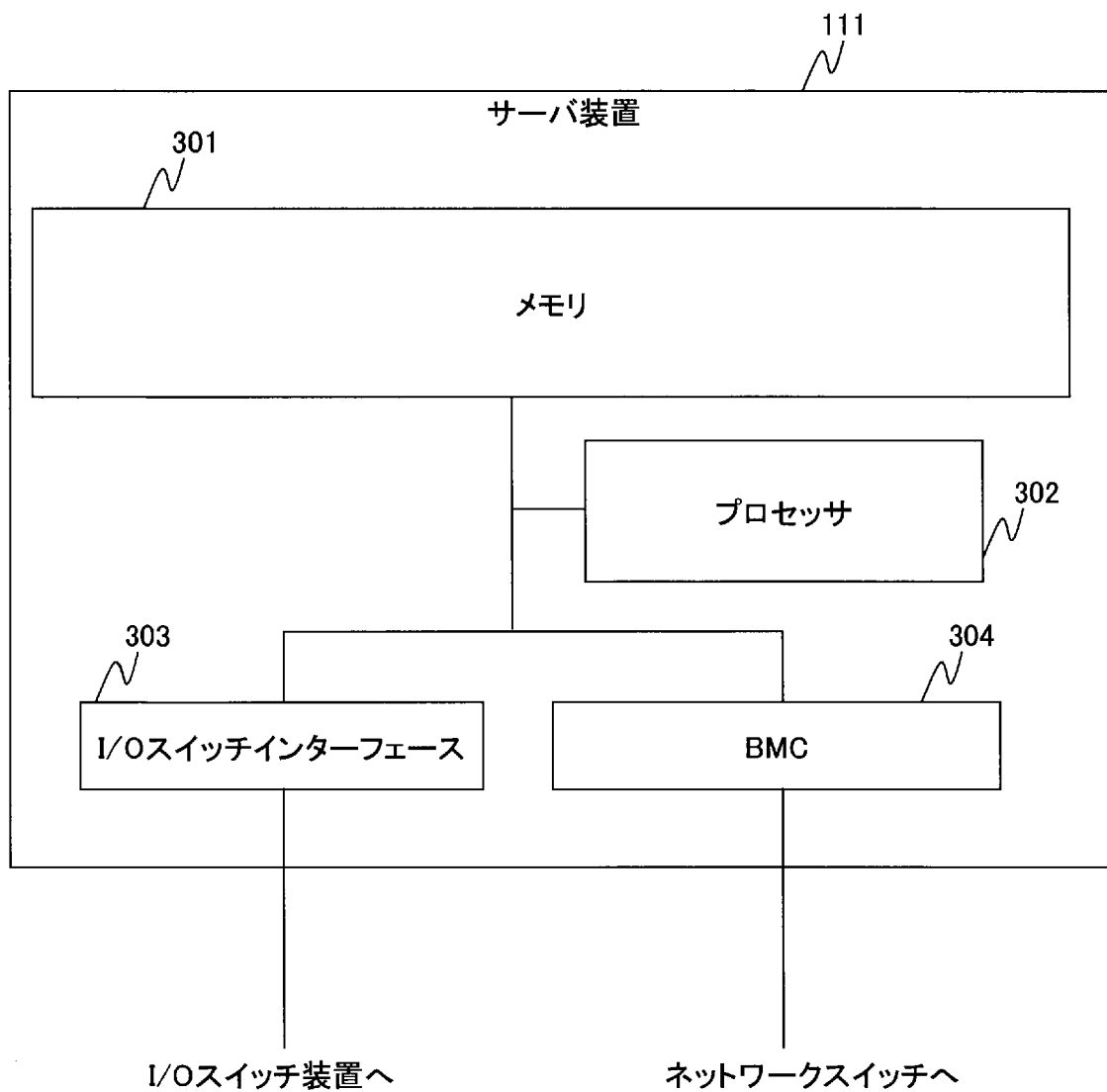
[図1]



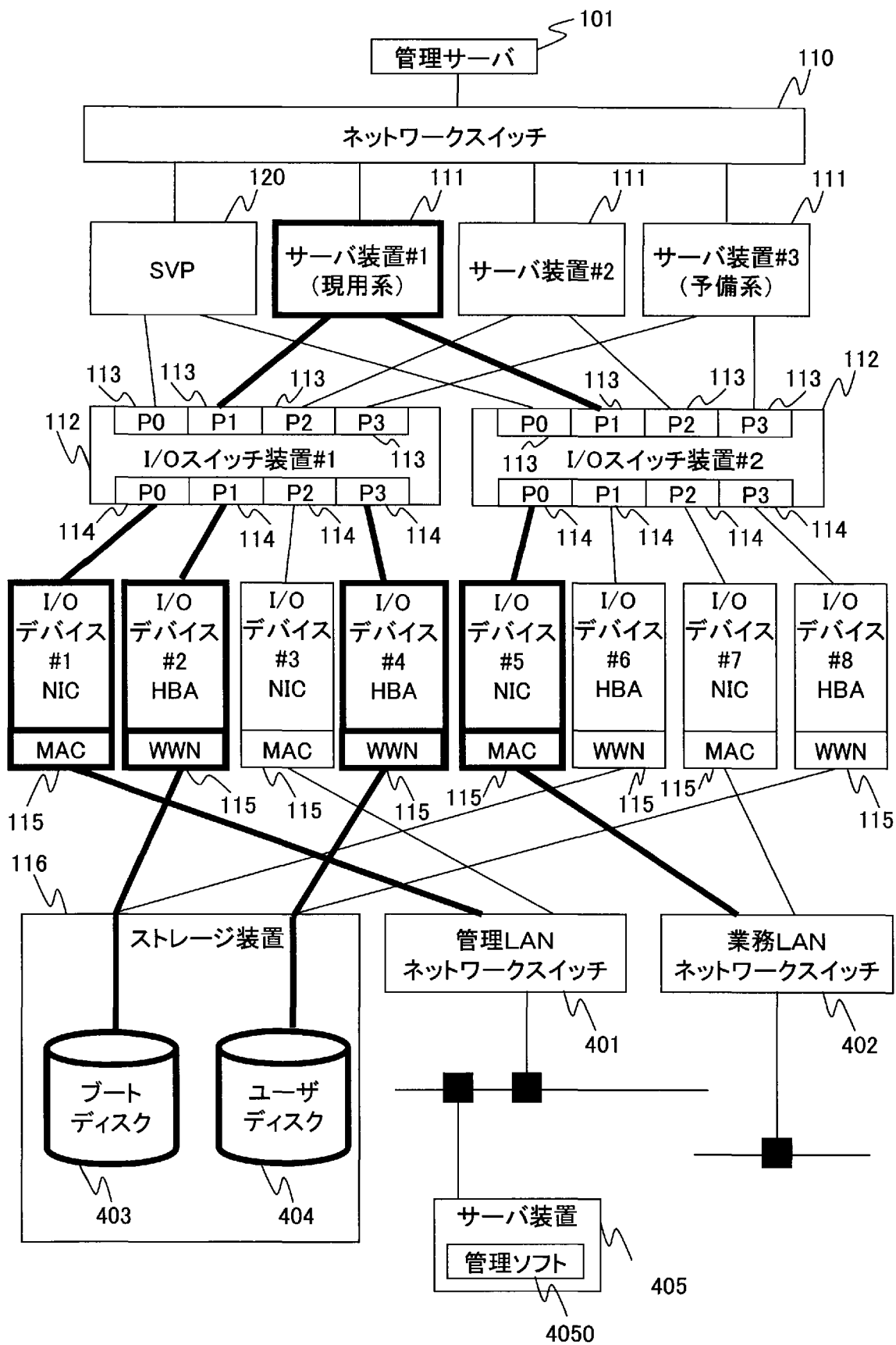
[図2]



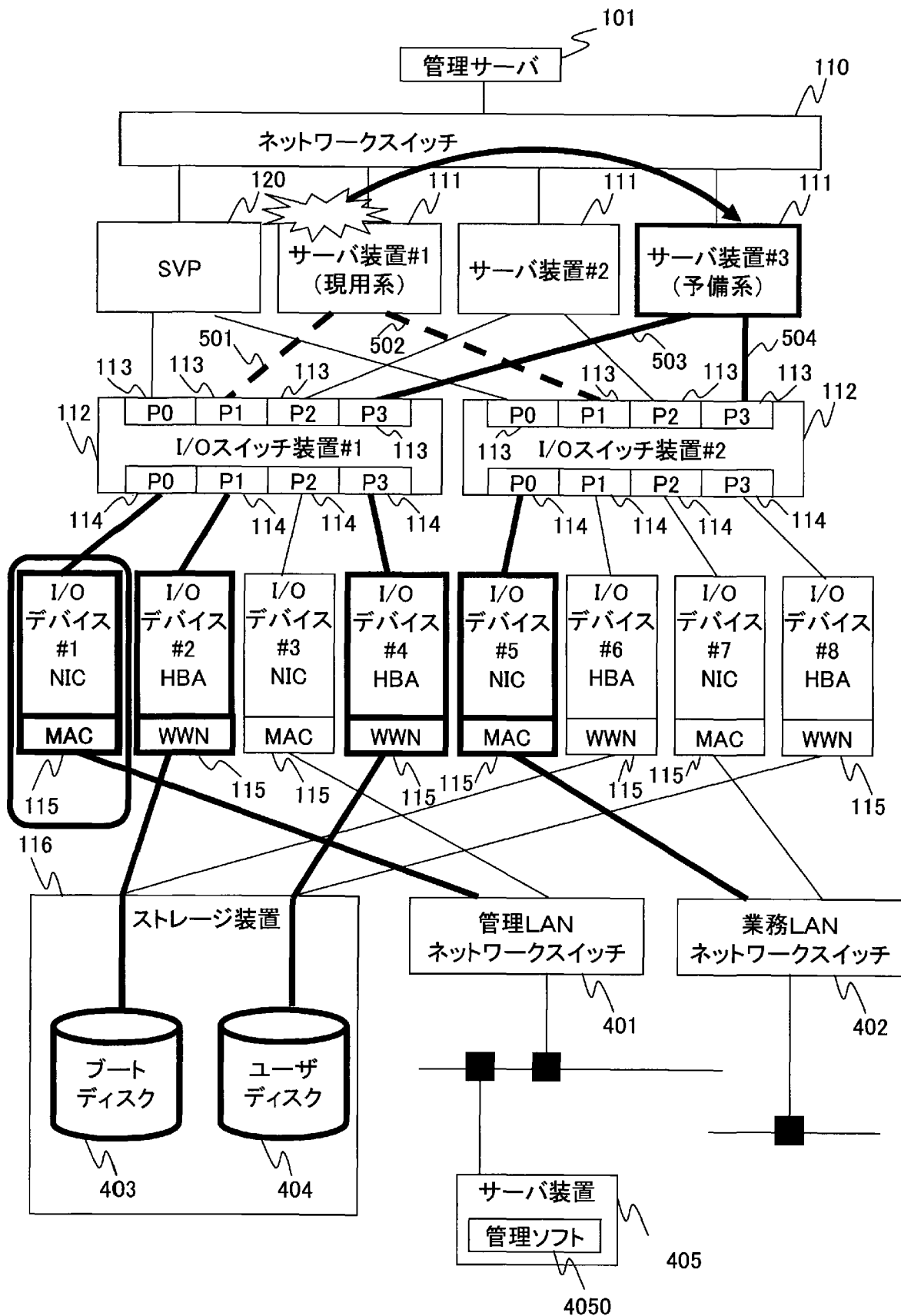
[図3]



[図4]



[図5]



[図6]

サーバ管理テーブル ¹⁰⁸

¹¹⁰¹ サーバ装置 識別子	¹¹⁰² プロセッサ 構成	¹¹⁰³ メモリ 容量	¹¹⁰⁴ サーバ接続 I/Oスイッチ 識別子	¹¹⁰⁵ 割り当て I/Oスイッチ 上流ポート番号	¹¹⁰⁶ 割り当て I/Oスイッチ 下流ポート番号
HOST1	プロセッサ1 2GHz×2	4GB	SW1	P1	P0,P1,P3
			SW2	P1	P0
HOST2	プロセッサ1 1GHz	2GB	SW1	P2	P2
			SW2	P2	P2
HOST3	プロセッサ1 1GHz	2GB	SW1	P3	-
			SW2	P3	-

[図7]

サーバI/O構成情報テーブル ¹⁰⁹

¹²⁰¹ I/Oスイッチ 識別子	¹²⁰² ポート番号	¹²⁰³ 接続デバイス	¹²⁰⁴ デバイス識別子	¹²⁰⁵ 指定デバイス識別子
SW1	P0	NIC	MAC1	MAC1
SW1	P1	HBA	WWN1	WWN1
SW1	P2	NIC	MAC2	-
SW1	P3	HBA	WWN2	WWN2
SW2	P0	NIC	MAC4	MAC3
SW2	P1	HBA	WWN3	-
SW2	P2	NIC	MAC4	-
SW2	P3	HBA	WWN4	-

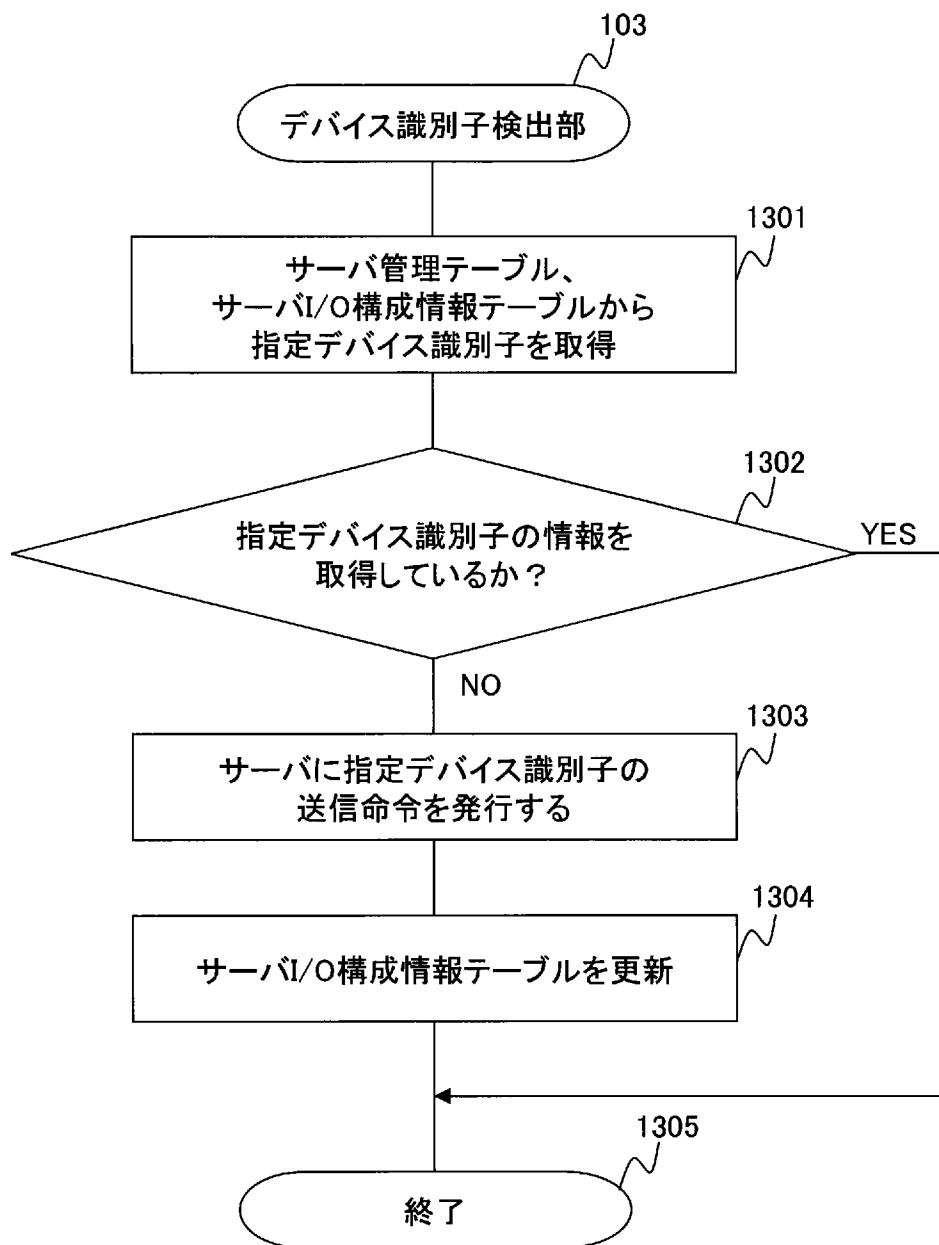
[図8]

123
仮想識別子テーブル

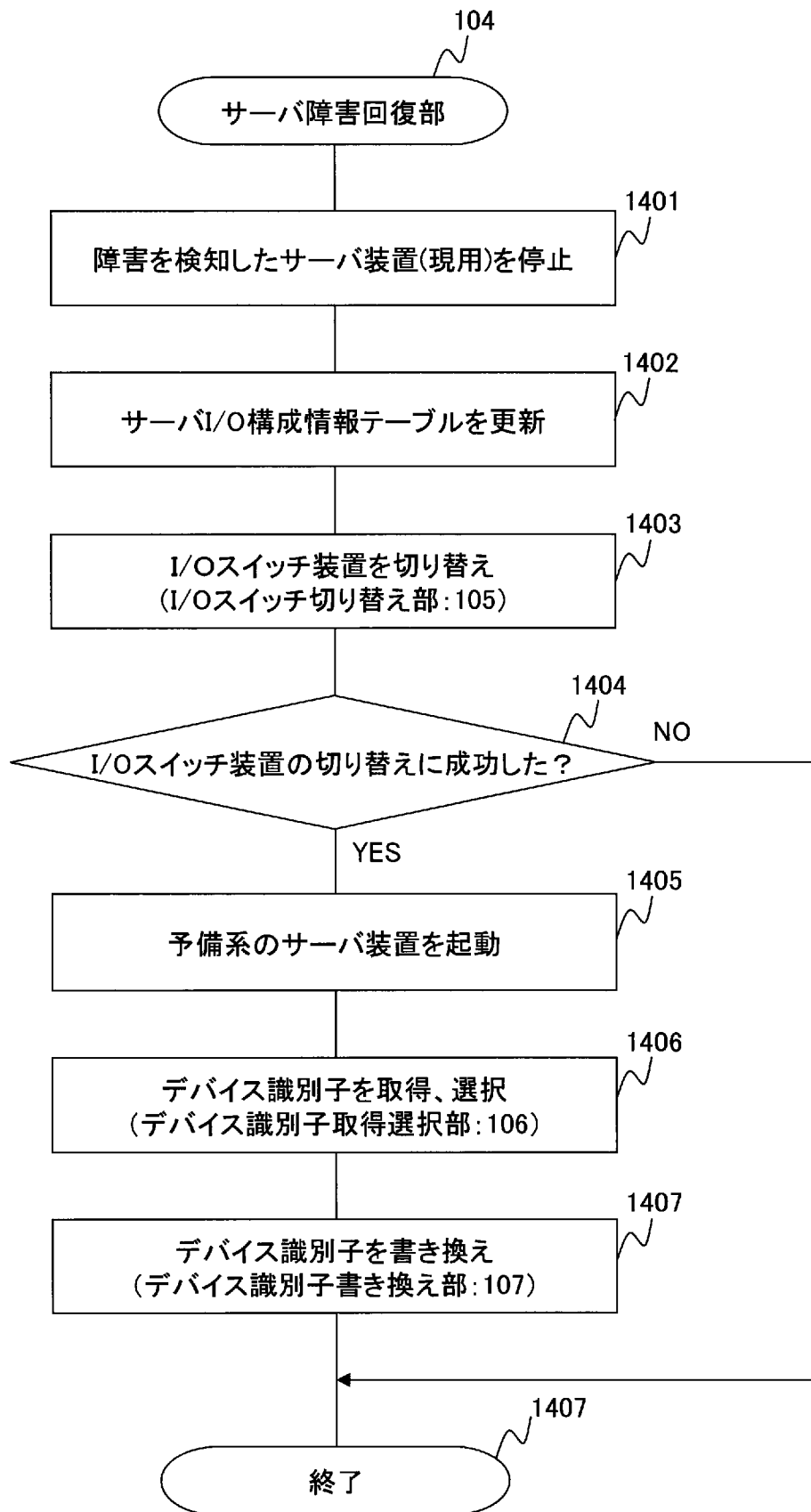
1231 1232

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

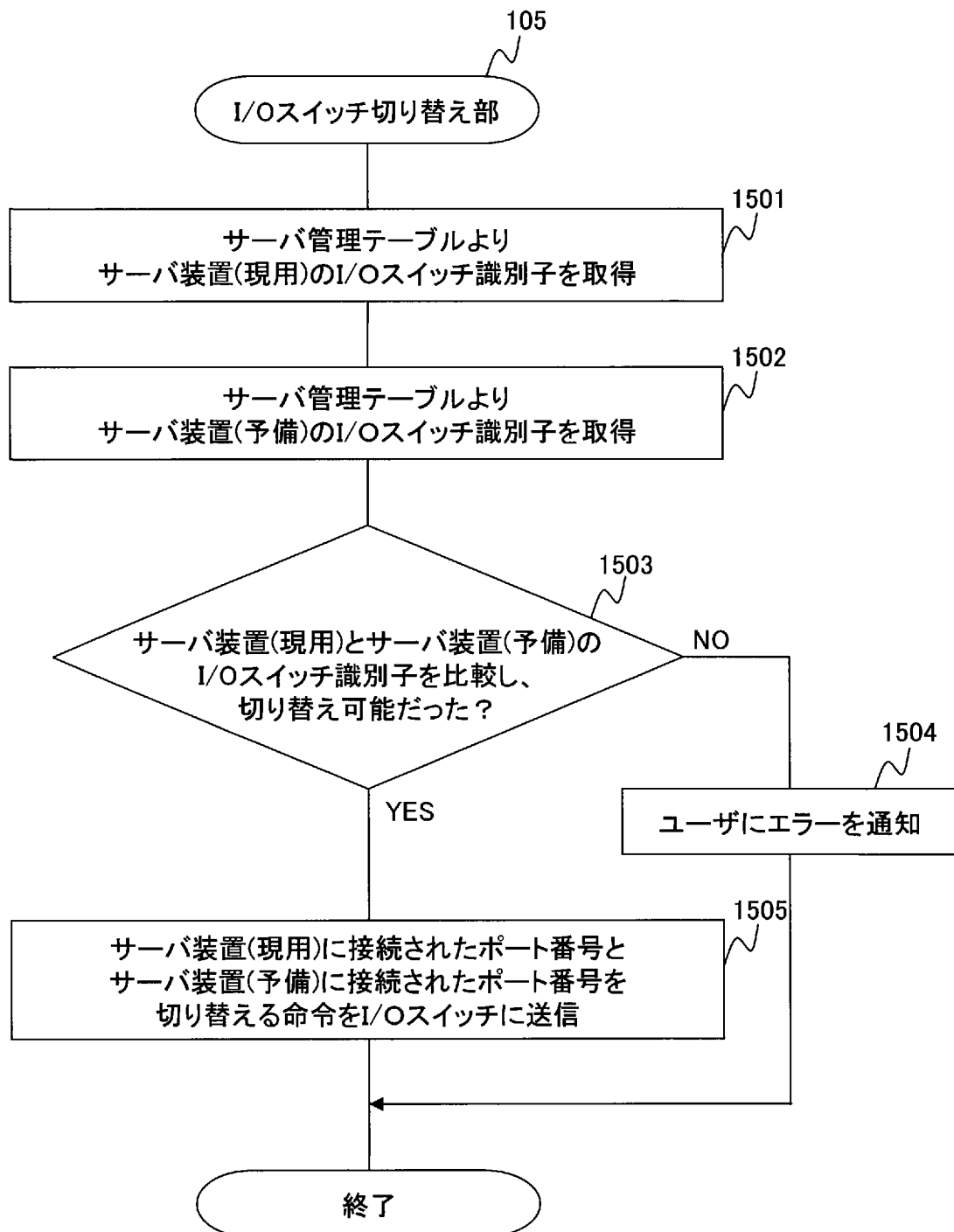
[図9]



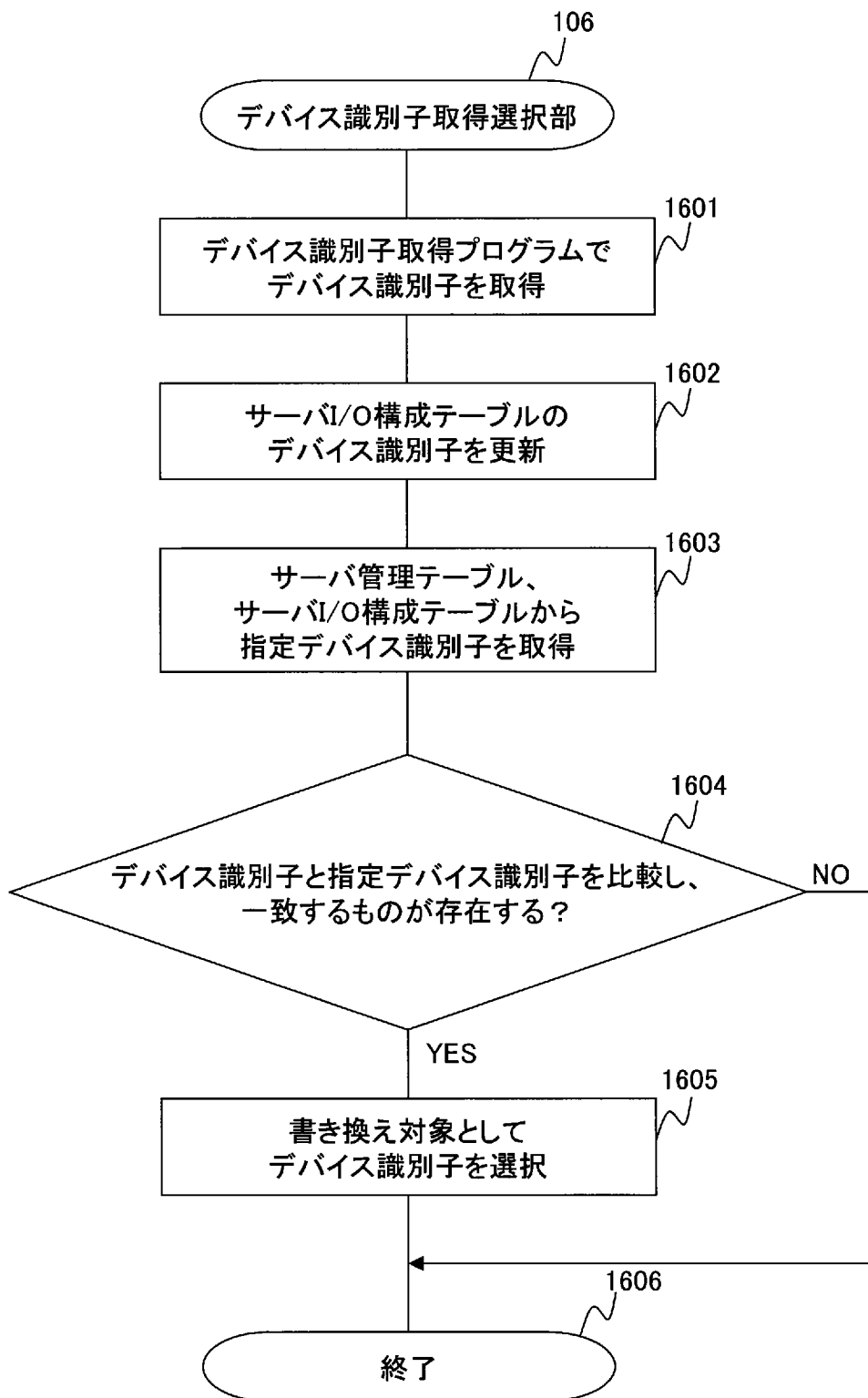
[図10]



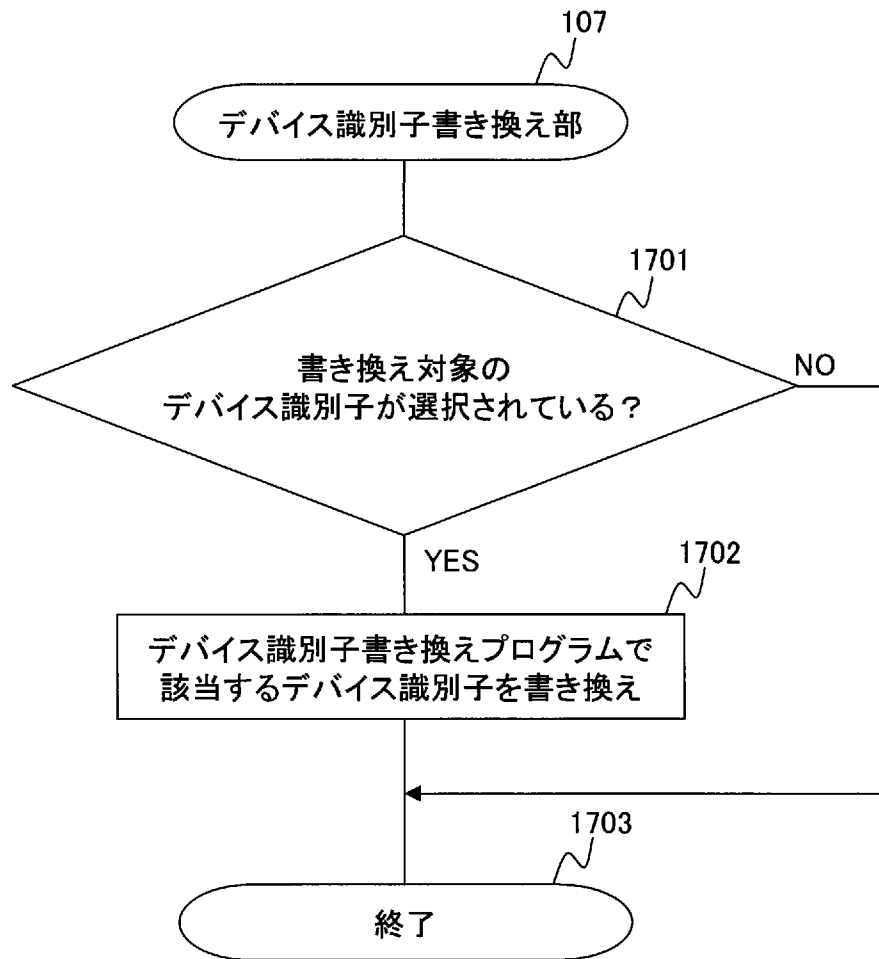
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/063276

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/20 (2006.01) i, G06F13/14 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 G06F11/20, G06F13/14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2010
 Kokai Jitsuyo Shinan Koho 1971-2010 Toroku Jitsuyo Shinan Koho 1994-2010

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2008-310489 A (Hitachi, Ltd.), 25 December 2008 (25.12.2008), paragraphs [0020] to [0022], [0028]; fig. 1, 6, 10 & US 2008/0313362 A1	1-13
A	JP 2007-164394 A (Hitachi, Ltd.), 28 June 2007 (28.06.2007), paragraphs [0036] to [0037]; fig. 1, 11 to 12 & US 2007/0174659 A1	1-13
A	JP 2003-234752 A (Nippon Telegraph And Telephone Corp.), 22 August 2003 (22.08.2003), paragraphs [0024] to [0028]; fig. 4 (Family: none)	1-13

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
 04 October, 2010 (04.10.10)

Date of mailing of the international search report
 12 October, 2010 (12.10.10)

Name and mailing address of the ISA/
 Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))
 Int.Cl. G06F11/20(2006.01)i, G06F13/14(2006.01)i

B. 調査を行った分野
 調査を行った最小限資料 (国際特許分類 (IPC))
 Int.Cl. G06F11/20, G06F13/14

最小限資料以外の資料で調査を行った分野に含まれるもの
 日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2010年
 日本国実用新案登録公報 1996-2010年
 日本国登録実用新案公報 1994-2010年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2008-310489 A (株式会社日立製作所) 2008.12.25, 段落【0020】-【0022】、【0028】、第1, 6, 10図 & US 2008/0313362 A1	1-13
A	JP 2007-164394 A (株式会社日立製作所) 2007.06.28, 段落【0036】-【0037】、第1, 11-12図 & US 2007/0174659 A1	1-13

C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

<p>* 引用文献のカテゴリー 「A」特に関連のある文献ではなく、一般的技術水準を示すもの 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」口頭による開示、使用、展示等に言及する文献 「P」国際出願日前で、かつ優先権の主張の基礎となる出願</p>	<p>の日の後に公表された文献 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」同一パテントファミリー文献</p>
---	---

国際調査を完了した日 04.10.2010	国際調査報告の発送日 12.10.2010
--------------------------	--------------------------

国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 塚田 肇	5B	3652
	電話番号 03-3581-1101 内線 3545		

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2003-234752 A (日本電信電話株式会社) 2003.08.22, 段落【024】 - 【0028】、第4図 (ファミリーなし)	1 - 13