

(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2010-146087
(P2010-146087A)

(43) 公開日 平成22年7月1日(2010.7.1)

(51) Int.Cl.
G06F 11/20 (2006.01)

F I
G06F 11/20 310B

テーマコード(参考)
5B034

審査請求 未請求 請求項の数 6 O L (全 17 頁)

(21) 出願番号 特願2008-319872 (P2008-319872)
(22) 出願日 平成20年12月16日 (2008.12.16)

(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 110000350
ポレール特許業務法人
(72) 発明者 中山 大輔
神奈川県秦野市堀山下1番地 株式会社日立製作所エンタープライズサーバ事業部内
(72) 発明者 岡澤 宏一
神奈川県秦野市堀山下1番地 株式会社日立製作所エンタープライズサーバ事業部内
(72) 発明者 山内 雅彦
神奈川県秦野市堀山下1番地 株式会社日立製作所エンタープライズサーバ事業部内

最終頁に続く

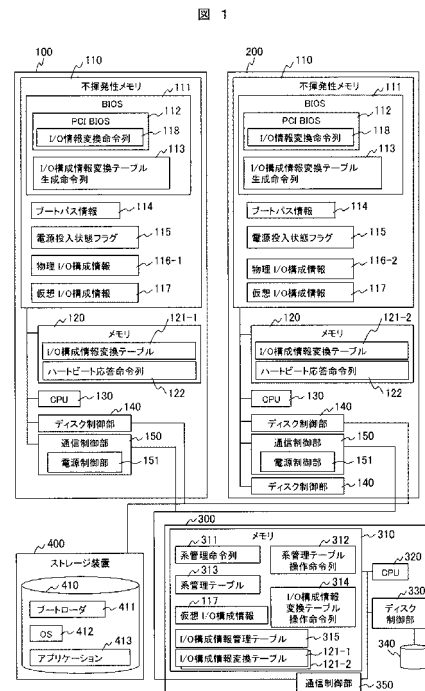
(54) 【発明の名称】 系切替計算機システムの管理方法

(57) 【要約】

【課題】系切替機能を持つ計算機システムは、現用系計算機と待機系計算機に接続されているI/Oの位置情報が異なる場合、現用系から待機系に切り替えたとき、動作不能になるという問題を解決する。

【解決手段】各計算機に、OSレベルで認識できるI/Oの仮想的な位置情報を物理的な位置情報に変換するための手段を設ける。管理計算機は、現用系計算機のもつI/Oの仮想位置情報又は変換テーブルを待機系計算機に引き継ぐ。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

現用系計算機と、待機系計算機と、前記現用系計算機の監視を行い、現用系計算機の障害を検出した場合には計算機稼働を前記待機系計算機に切り替える管理計算機とを備える系切替計算機システムの管理方法において、

前記現用系計算機は、I/Oの仮想位置情報を含むOSが認識するためのI/Oの構成情報を記憶手段に格納する仮想I/O構成情報を有し、前記仮想I/O構成情報に含まれる前記仮想位置情報に対応する物理位置情報を取得し、前記仮想位置情報をI/Oの物理位置情報に変換するためのI/O構成情報変換テーブルを作成し、I/Oにアクセスする際に、前記I/O構成情報変換テーブルを参照して前記仮想位置情報に対応する前記物理位置情報に変換し、

10

前記管理計算機は、前記現用系計算機から前記仮想I/O構成情報を受け取り前記待機系計算機へ送るとともに、前記現用系計算機の障害を検出したとき、前記待機系計算機の電源を投入し、

前記待機系計算機は、前記管理計算機からの起動要求を受けて当該計算機を起動し、受け取った仮想I/O構成情報から前記の手順により前記I/O構成情報変換テーブルを作成し、作成されたI/O構成情報変換テーブルの参照を含む前記の変換処理を引き継ぐことを特徴とする系切替計算機システムの管理方法。

【請求項 2】

現用系計算機と、待機系計算機と、前記現用系計算機の監視を行い、現用系計算機の障害を検出した場合には計算機稼働を前記待機系計算機に切り替える管理計算機とを備える系切替計算機システムの管理方法において、

20

前記現用系計算機は、I/Oの仮想位置情報を含むOSが認識するためのI/Oの構成情報を記憶手段に格納する仮想I/O構成情報を有し、前記仮想I/O構成情報に含まれる前記仮想位置情報に対応する物理位置情報を取得し、前記仮想位置情報をI/Oの物理位置情報に変換するためのI/O構成情報変換テーブルを作成し、I/Oにアクセスする際に、前記I/O構成情報変換テーブルを参照して前記仮想位置情報に対応する前記物理位置情報に変換し、

前記管理計算機は、前記現用系計算機から前記I/O構成情報変換テーブルを受け取り前記待機系計算機へ送るとともに、前記現用系計算機の障害を検出したとき、前記待機系計算機を起動し、

前記待機系計算機は、前記管理計算機からの起動要求を受けて当該計算機を起動し、前記管理計算機から受け取った前記I/O構成情報変換テーブルの参照を含む前記の変換処理を引き継ぐことを特徴とする系切替計算機システムの管理方法。

30

【請求項 3】

前記仮想位置情報及び前記物理位置情報は、各々各I/Oのバス番号、デバイス番号およびファンクション番号を含むことを特徴とする請求項 1 または請求項 2 記載の系切替計算機システムの管理方法。

【請求項 4】

前記現用系計算機と前記待機系計算機の各々は、通常の計算機起動か系切替時の計算機起動かを区別するためのフラグを記憶手段上に有し、

前記現用系計算機と前記待機系計算機の各々は、前記フラグが通常の計算機起動を示しているときには、I/Oの物理的な構成情報を記憶手段に格納する物理I/O構成情報から同一内容の前記仮想I/O構成情報を作成し、作成された仮想I/O構成情報から対応する前記I/O構成情報変換テーブルを作成し、

40

前記待機系計算機は、前記フラグが系切替時の計算機起動を示しているときには、前記管理計算機から受け取った仮想I/O構成情報から対応する前記I/O構成情報変換テーブルを作成することを特徴とする請求項 1 記載の系切替計算機システムの管理方法。

【請求項 5】

前記現用系計算機と前記待機系計算機の各々は、通常の計算機起動か系切替時の計算機起動かを区別するためのフラグを記憶手段上に有し、

前記現用系計算機と前記待機系計算機の各々は、前記フラグが通常の計算機起動を示し

50

ているときには、I/Oの物理的な構成情報を記憶手段に格納する物理I/O構成情報から同一内容の前記仮想I/O構成情報を作成し、作成された仮想I/O構成情報から対応する前記I/O構成情報変換テーブルを作成し、

前記待機系計算機は、前記フラグが系切替時の計算機起動を示しているときには、前記管理計算機から受け取った前記I/O構成情報変換テーブルの参照を含む前記の変換処理を引き継ぐことを特徴とする請求項2記載の系切替計算機システムの管理方法。

【請求項6】

前記管理計算機は、前記現用系計算機と前記待機系計算機の各々について、障害の有無を記憶する系管理テーブルを有し、前記現用系計算機と前記待機系計算機の各々が障害なし、障害あり、復旧により障害なしに応じて前記系管理テーブルの前記障害の有無の状態を更新することを特徴とする請求項1または請求項2記載の系切替計算機システムの管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、系切替機能を有する計算機システムに関する。

【背景技術】

【0002】

現代社会において、計算機システムは欠かせないものとなっている。銀行のATMシステムなどの24時間連続稼働を想定している計算機システムでは、障害が発生した場合には、速やかに業務を再開する必要がある。

【0003】

従来技術として、業務を処理する現用系計算機に障害が発生したら、待機系計算機に業務を引き継がせることにより、業務を再開するまでの時間を短くする系切替技術がある。系切替技術は特許文献1に開示されている。

【0004】

また、非特許文献1によると、パーソナルコンピュータやサーバ分野の計算機のI/Oを接続するインターフェースとしては、PCIバスが一般に普及している。I/Oは計算機の持つPCIスロットに接続して使用される。計算機で動作するソフトウェアは、各I/Oに対して、I/Oが接続されたPCIスロットと1対1に対応する位置情報を割り当て、I/Oを管理している。

【0005】

【特許文献1】特開2007-94611号公報

【非特許文献1】「改訂新版PCIバス&PCI-Xバスの徹底研究」、CQ出版株式会社、2004年発行

【発明の開示】

【発明が解決しようとする課題】

【0006】

従来のOSが格納されたストレージ装置を共有する系切替システムは、現用系計算機と待機系計算機に接続されているI/Oの位置情報を同じにする必要がある。

【0007】

I/Oの位置情報が異なる場合、待機系計算機でOSが起動する際に、OSが格納されたストレージ装置に接続されているI/Oの位置情報が現用系計算機と待機系計算機とで一致しないため、OSが起動しない等の問題が発生するからである。

【0008】

そのため、異なるI/Oの位置情報を持つ現用計算機が複数存在する場合には、異なるI/Oの位置情報を持つ現用系計算機ごとに待機系計算機を用意することになり、待機系計算機の導入コストが高くなるという問題がある。

【0009】

本発明の目的は、上記課題を解決し、I/Oの位置情報やI/Oの数が異なる現用系計算機が

10

20

30

40

50

複数存在する場合でも、異なるI/Oの位置情報や数を持つ現用系計算機ごとに待機系計算機を用意せずに済む系切替制御方式を提供することを目的としている。

【課題を解決するための手段】

【0010】

本発明は、現用系計算機と、待機系計算機と、現用系計算機の監視を行い、現用系計算機の障害を検出した場合には計算機稼働を待機系計算機に切り替える管理計算機とを備える系切替計算機システムの管理技術において、現用系計算機は、I/Oの仮想位置情報を含むOSが認識するためのI/Oの構成情報を記憶手段に格納する仮想I/O構成情報を有し、仮想I/O構成情報に含まれる仮想位置情報に対応する物理位置情報を取得し、仮想位置情報をI/Oの物理位置情報に変換するためのI/O構成情報変換テーブルを作成し、I/Oにアクセスする際に、I/O構成情報変換テーブルを参照して仮想位置情報に対応する物理位置情報に変換し、管理計算機は、現用系計算機から仮想I/O構成情報を受け取り待機系計算機へ送るとともに、現用系計算機の障害を検出したとき、待機系計算機の電源を投入し、待機系計算機は、管理計算機からの起動要求を受けて当該計算機を起動し、受け取った仮想I/O構成情報から前記の手順によりI/O構成情報変換テーブルを作成し、作成されたI/O構成情報変換テーブルの参照を含む前記の変換処理を引き継ぐ系切替計算機システムの管理技術の特徴とする。

10

【0011】

なお管理計算機が現用系計算機からI/O構成情報変換テーブルを受け取り、I/O構成情報変換テーブルの段階で待機系計算機に引き継いでもよい。

20

【発明の効果】

【0012】

本発明は、異なるI/Oの位置情報、種類情報、数を持つ現用系計算機ごとに待機系計算機を用意することなく系切替システムを成立させることで、待機系計算機の導入コストを減らすことができる。

【発明を実施するための最良の形態】

【0013】

以下、本発明を実施するための実施例について図面に基づいて詳細に説明する。

【0014】

図1は、本実施例の計算機システムのブロック図である。図1において、管理計算機300は、メモリ310とCPU320とディスク制御部330と記憶装置340を備え、他の計算機と通信を行う通信制御部350を備える。メモリ310は、系管理命令列311、系管理テーブル操作命令列312、系管理テーブル313、I/O構成情報変換テーブル操作命令列314、I/O構成情報管理テーブル315、および仮想I/O構成情報117を格納している。I/O構成情報管理テーブル315は、I/O構成情報変換テーブル121を含んでいる。なお管理計算機300は、図示しない表示装置を備えている。

30

【0015】

1つまたは複数で構成される現用系計算機100及び待機系計算機200は、不揮発性メモリ110、メモリ120、CPU130、ストレージ装置400にアクセスするためのディスク制御部140、および他の計算機と通信を行う通信制御部150を備える。

40

【0016】

不揮発性メモリ110は、PCI BIOS112とI/O構成情報変換テーブル生成命令列113を備えたBIOS111、ブートパス情報114、電源投入状態フラグ115、物理I/O構成情報116、および仮想I/O構成情報117を格納している。PCI BIOS112は、I/O情報変換命令列118を含む。

【0017】

メモリ120は、I/O構成情報変換テーブル121とハートビート応答命令列122を格納している。

【0018】

通信制御部150は、電源制御部151を備えている。

50

【 0 0 1 9 】

電源投入状態フラグ 1 1 5 は、どのような状態で電源を投入したかを示す。通常の電源投入ならば「 0 」を格納し、系切替による電源投入ならば、「 0 」以外のコードを格納する。

【 0 0 2 0 】

ストレージ装置 4 0 0 は、1つまたは複数のディスクドライブを備え、現用系計算機 1 0 0 と待機系計算機 2 0 0 からアクセス可能な記憶領域としてボリューム 4 1 0 が設定される。

【 0 0 2 1 】

ボリューム 4 1 0 は、ブートローダ 4 1 1、OS (オペレーティングシステム) 4 1 2、およびアプリケーション 4 1 3 を格納している。

10

【 0 0 2 2 】

図 2 は、系管理テーブル 3 1 3 の一実施例を示している。系管理テーブル 3 1 3 は、テーブル番号 5 0 0、コンピュータ識別子 5 0 1、系の種類 5 0 2、障害状態 5 0 3、および稼動状態 5 0 4 の列で構成されている。コンピュータ識別子 5 0 1 は、現用系計算機 1 0 0、待機系計算機 2 0 0 と管理計算機 3 0 0 とを接続する通信回線上でその計算機を識別する IP アドレスによって指定される。

【 0 0 2 3 】

図 3 と図 4 は、物理 I/O 構成情報 1 1 6 の一実施例を示している。物理 I/O 構成情報 1 1 6 は、計算機に物理的に接続された I/O の構成情報であり、テーブル番号 6 0 0、1 6 0 0 と、バス番号 6 0 1、1 6 0 1 と、デバイス番号 6 0 2、1 6 0 2 と、ファンクション番号 6 0 3、1 6 0 3 と、ベンダ ID 6 0 4、1 6 0 4 と、デバイス ID 6 0 5、1 6 0 5 と、クラスコード 6 0 6、1 6 0 6 の列で構成されている。ここで各テーブルとは、物理 I/O 構成情報 1 1 6 の 1 行を指し、各 I/O デバイスに関する構成情報を格納する。本実施例では、第 1 行目の I/O デバイスはディスク制御部 1 4 0 を指し、第 2 行目の I/O デバイスは通信制御部 1 5 0 を指し、第 3 行目の I/O デバイスは 2 番目のディスク制御部 1 4 0 を指す。ファンクション番号 6 0 3 は、1 つのデバイスが複数のポートをもつ場合にそれらを識別するための番号である。待機系計算機 2 0 0 は、現用系計算機 1 0 0 のもつ I/O をすべて包含するものとする。

20

【 0 0 2 4 】

図 5 は、仮想 I/O 構成情報 1 1 7 の一実施例を示している。仮想 I/O 構成情報 1 1 7 は、各 OS によって認識される I/O の構成情報であり、テーブル番号 7 0 0、バス番号 7 0 1、デバイス番号 7 0 2、ファンクション番号 7 0 3、ベンダ ID 7 0 4、デバイス ID 7 0 5、およびクラスコード 7 0 6 の列で構成されている。

30

【 0 0 2 5 】

図 6 と図 7 は、現用系計算機 1 0 0 と待機系計算機 2 0 0 の持つ物理 I/O 構成情報 1 1 6 と仮想 I/O 構成情報 1 1 7 を元に作成される I/O 構成情報変換テーブル 1 2 1 の一実施例である。I/O 構成情報変換テーブル 1 2 1 は、仮想 I/O 構成情報 1 1 7 の内容を包含している。I/O 構成情報変換テーブル 1 2 1 は、PCI BIOS 1 1 2 によって参照されるテーブルであり、テーブル番号 8 0 0、1 7 0 0 と、バス番号 8 0 1、1 7 0 1 と、デバイス番号 8 0 2、1 7 0 2 と、ファンクション番号 8 0 3、1 7 0 3 と、ベンダ ID 8 0 4、1 7 0 4 と、デバイス ID 8 0 5、1 7 0 5 と、クラスコード 8 0 6、1 7 0 6 と、変換先番号 8 0 7、1 7 0 7 の列で構成されている。

40

【 0 0 2 6 】

以上の情報の中で、バス番号 6 0 1、7 0 1、8 0 1、1 6 0 1、1 7 0 1 とデバイス番号 6 0 2、7 0 2、8 0 2、1 6 0 2、1 7 0 2 とファンクション番号 6 0 3、7 0 3、8 0 3、1 6 0 3、1 7 0 3 の列は、I/O の位置情報を示している。

【 0 0 2 7 】

ベンダ ID 6 0 4、7 0 4、8 0 4、1 6 0 4、1 7 0 4 の列は、I/O の製造会社情報を格納する。デバイス ID 6 0 5、7 0 5、8 0 5、1 6 0 5、1 7 0 5 の列は、I/O の識別

50

情報を格納する。クラスコード 606、706、806、1606、1706の列は、I/Oの種類を格納する。変換先番号 807、1707は、各行のI/Oの位置情報に対応する物理I/O構成情報 116のテーブル番号を格納する。

【0028】

なお通常の電源投入によって作成されるI/O構成情報変換テーブル 121 - 2は、物理I/O構成情報 116 - 2の通り3行で構成されるが、系切替後のI/O構成情報変換テーブル 121 - 2は、I/O構成情報変換テーブル 121 - 1を引き継ぐため、図7に示すようにI/O構成情報変換テーブル 121 - 1と同じテーブルに変更される。

【0029】

図8は、管理計算機 300の持つI/O構成情報管理テーブル 315の一実施例である。I/O構成情報管理テーブル 315は、引継元と引継先の計算機の組合せによって使用するI/O構成情報変換テーブル 121を参照するためのテーブルであり、テーブル番号 1800、引継元のコンピュータ識別子 1801、引継先のコンピュータ識別子 1802、I/O構成情報変換テーブル 1803の列で構成されている。引継元のコンピュータ識別子 1801は現用系計算機のいずれか、引継先のコンピュータ識別子 1802は待機系計算機を示す。I/O構成情報変換テーブル 1803は、引継元計算機のI/O構成情報変換テーブル 121を格納する。

【0030】

図9は、I/O構成情報管理テーブル 315を作成する管理計算機 300の処理手順を示すフローチャートである。

【0031】

ステップ 1501では、管理者により、現用系計算機 100と待機系計算機 200の電源が投入された後、管理計算機 300は、各計算機から物理I/O構成情報 116を読み出し、表示装置上に表示する。

【0032】

ステップ 1502では、I/O構成情報変換テーブル操作命令列 314は、I/O構成情報管理テーブル 315を作成する。I/O構成情報変換テーブル操作命令列 314は、管理者の指示により、引継元計算機と引継先計算機との組合せを各々IPアドレスにより作成し、I/O構成情報管理テーブル 315の引継元のコンピュータ識別子 1801と引継先のコンピュータ識別子 1802に書き込む。次にI/O構成情報変換テーブル操作命令列 314は、各引継元計算機からI/O構成情報変換テーブル 121を読み出し、I/O構成情報管理テーブル 315のI/O構成情報変換テーブル 1803の対応する欄に書き込む。

【0033】

管理者は、各計算機のI/O構成情報変換テーブル 121とその物理I/O構成情報 116とを比較し、必要に応じて引継元のI/O構成情報変換テーブル 121に修正を加えるようにI/O構成情報変換テーブル操作命令列 314に指示する。I/O構成情報変換テーブル操作命令列 314は、修正後のI/O構成情報変換テーブル 121をI/O構成情報変換テーブル 1803に格納する。

【0034】

例えば待機系計算機 200が現用系計算機 100の稼動を引き継いだとすると、引継後は待機系計算機 200は、引継元となり、別の引継先に引き継ぐことが起こり得る。元のI/O構成情報変換テーブル 121 - 2は、3行で構成されているが、引継後のI/O構成情報変換テーブル 121 - 2は、I/O構成情報変換テーブル 121 - 1の通り2行で構成されているため、I/O構成情報変換テーブル 121 - 2をI/O構成情報変換テーブル 121 - 1に合わせて2行で構成する必要がある。このように、待機系計算機が「現用系」かつ「引継元」となった後も、この計算機に障害が発生した場合の引継先計算機についての対策をとる必要がある。

【0035】

図10は、現用系計算機 100および待機系計算機 200で実行されるハートビート応答命令列 122の処理手順を示すフローチャートである。

10

20

30

40

50

【0036】

ステップ900では、ハートビート応答命令列122は、管理計算機300からのハートビート信号を受信したかを判定する。受信していないならステップ900へ戻り、受信したならステップ901へ進む。

【0037】

ステップ901では、ハートビート応答命令列122は、管理計算機300の通信制御部350にハートビート応答信号を送信し、ステップ900へ戻る。

【0038】

図11～図13は、系管理命令列311の処理手順を示すフローチャートである。ステップ1000では、系管理命令列311は、系管理テーブル313を参照し、系の種類502が「現用系」となっている行のコンピュータ識別子501が示すIPアドレス、すなわち現用系計算機100、にハートビートを送信する。

10

【0039】

ステップ1001では、タイマーをスタートする。

【0040】

ステップ1002では、系管理命令列311は、現用系計算機100からのハートビート応答を受信できたかを判定する。受信できたなら、ステップ1000へ戻り、受信できなかったなら、ステップ1003へ進む。

【0041】

ステップ1003では、系管理命令列311は、タイマーが制限時間を超えたかどうかを判定する。制限時間を超えていないなら、ステップ1002へ戻り、制限時間を超えたなら、ステップ1004へ進む。

20

【0042】

ステップ1004では、系管理命令列311は、系管理テーブル313を参照し、系の種類502が「現用系」となっている行のコンピュータ識別子501が示すIPアドレスに現用系計算機の仮想I/O構成情報117を要求するためのパケットを送信する。

【0043】

ステップ1005では、系管理命令列311は、現用系計算機100の仮想I/O構成情報117を受信できたかどうかを判定する。受信できていないなら、ステップ1005へ戻る。受信できたなら、ステップ1006へ進む。

30

【0044】

なお現用系計算機の電源オン後のシーケンスで仮想I/O構成情報117が作成されるから、系管理命令列311がこの時点を検知し、現用系計算機が障害となる前にステップ1004, 1005を実行して仮想I/O構成情報117を取得しておいてもよい。

【0045】

図12に移り、ステップ1006では、系管理命令列311は、受信した仮想I/O構成情報117を管理計算機300のメモリ310に格納する。

【0046】

ステップ1007では、系管理命令列311は、系管理テーブル313を参照し、系の種類502が「現用系」となっている行のコンピュータ識別子501が示すIPアドレスに電源切断を要求するためのパケットを送信する。電源切断要求の代わりに、現用系計算機の停止要求でもよい。

40

【0047】

ステップ1008では、系管理命令列311は、系管理テーブル313を参照し、電源切断の要求パケットの送信先IPアドレスがある行の障害状態503を「障害あり」、稼働状態504を「停止」という情報に変更する。

【0048】

ステップ1009では、系管理命令列311は、系管理テーブル313を参照し、障害状態503が「障害なし」で、系の種類502が「待機系」となっている行のコンピュータ識別子が示すIPアドレスを取得する。

50

【0049】

ステップ1010では、系管理命令列311は、I/O構成情報管理テーブル315を参照して、引継元のコンピュータ識別子が、ステップ1008で参照したIPアドレスであり、引継先のコンピュータ識別子がステップ1009で参照したIPアドレスである行で指定されているI/O構成情報変換テーブル1803に設定されるI/O構成情報変換テーブル121を取得する。

【0050】

ステップ1011では、系管理命令列311は、待機系へ送信する現用系のI/O構成情報変換テーブル121が格納されていたかどうかを判定し、格納されていなければ、ステップ1012へ進み、格納されていれば、ステップ1013へ進む。ステップ1502で、現用系のI/O構成情報変換テーブル121をI/O構成情報管理テーブル315に格納する処理が行なわれるが、I/O構成情報変換テーブル121の格納処理を実行できず、ステップ1011でI/O構成情報変換テーブル121が格納されていないことが生じ得る。

10

【0051】

ステップ1012では、系管理命令列311は、ステップ1009で取得した待機系のIPアドレスに管理計算機300が保持する仮想I/O構成情報117を送信し、ステップ1014へ進む。待機系計算機200は、この仮想I/O構成情報117を不揮発性メモリ110に格納する。

【0052】

ステップ1013では、系管理命令列311は、ステップ1009で取得した待機系のIPアドレスにステップ1010で取得したI/O構成情報変換テーブル121を送信し、ステップ1014へ進む。

20

【0053】

図13に移り、ステップ1014では、系管理命令列311は、系管理テーブル313を参照し、系の種類502が「待機系」となっている行のコンピュータ識別子501の示すIPアドレスに系切替による電源投入を要求するためのパケットを送信する。

【0054】

待機系計算機は、図14の初期手順を実行し、待機系計算機を起動する。このとき、系管理命令列311は、待機系計算機の電源投入状態フラグを0でない値に設定するよう要求する。

30

【0055】

ステップ1015では、系管理命令列311は、系管理テーブル313の系の種類502が「待機系」となっている行の稼動状態504を「稼動」に変更する。

【0056】

ステップ1016では、系管理命令列311は、ステップ1008とステップ1009で対象とした計算機それぞれについて、系管理テーブル313の系の種類502の「現用系」を「待機系」という情報に変更し、「待機系」を「現用系」という情報に変更し、ステップ1000へ戻る。

【0057】

系管理命令列311は、待機系計算機200に対しても上記ステップ1000～1003を実行する。待機系計算機200の障害を検出したとき、系管理命令列311は、待機系計算機200に対して上記ステップ1007、1008を実行する。

40

【0058】

図14は、現用系計算機と待機系計算機の電源投入又は計算機起動からサービス提供までの処理手順を示すフローチャートである。

【0059】

ステップ1101では、計算機は、BIOS111のハードウェア初期化とハードウェア診断テスト処理を行う。

【0060】

ステップ1102では、計算機は、I/O構成情報変換テーブル生成命令列113を呼び

50

出して実行する。このステップは、本発明の特徴とするものである。

【0061】

ステップ1103では、計算機は、ブートローダ411のカーネルロード処理を行う。
ステップ1104では、計算機は、OS412の起動処理を行う。

【0062】

ステップ1105では、計算機は、アプリケーション413の起動処理を行う。

【0063】

図15～図17は、現用系計算機100および待機系計算機200で実行されるI/O構成情報変換テーブル生成命令列113の処理手順を示すフローチャートである。

【0064】

ステップ1200では、I/O構成情報変換テーブル生成命令列113は、PCIバス空間のバス番号0～255、デバイス番号0～31、ファンクション番号0～7から1つずつ順に番号を選択する。

【0065】

ステップ1201では、I/O構成情報変換テーブル生成命令列113は、全てのI/Oデバイス、すなわちバス番号、デバイス番号、ファンクション番号が選択されたかどうかを判定する。最後のI/Oデバイスの番号の組(255, 31, 7)が選択されたとき、ステップ1201の選択完了となる。選択されていないなら、ステップ1202へ進み、選択されたなら、図16のステップ1205へ進む。

【0066】

ステップ1202では、I/O構成情報変換テーブル生成命令列113は、選択されたバス番号、デバイス番号、ファンクション番号、ベンダIDのオフセット番号を使って、選択されたI/Oデバイスのコンフィグレーションレジスタの内容を読み込む。

【0067】

ステップ1203では、I/O構成情報変換テーブル生成命令列113は、ステップ1202でアクセスしたコンフィグレーションレジスタ値のベンダIDが「0xFFFF」かどうかを判定する。値「0xFFFF」は、I/Oデバイスが存在しないことを示す。「0xFFFF」であるなら、ステップ1200へ戻り、「0xFFFF」でないなら、ステップ1204へ進む。

【0068】

ステップ1204では、I/O構成情報変換テーブル生成命令列113は、選択されたI/Oデバイスの位置情報、すなわちバス番号、デバイス番号、ファンクション番号とコンフィグレーションレジスタ値のベンダID、デバイスID、クラスコードを物理I/O構成情報116に格納し、ステップ1200へ戻る。

【0069】

図16に移り、ステップ1205では、I/O構成情報変換テーブル生成命令列113は、電源投入状態フラグ115が0かどうかを判定する。0であるなら、ステップ1206へ進み、0でないなら、ステップ1217へ進む。

【0070】

ステップ1217では、I/O構成情報変換テーブル生成命令列113は、管理計算機300からI/O構成情報変換テーブル121が送信されており、存在するかどうかを判定し、送信されていないければ、ステップ1207へ進み、送信されていれば、1213へ進む。

【0071】

ステップ1206では、I/O構成情報変換テーブル生成命令列113は、物理I/O構成情報116と同じ内容の仮想I/O構成情報117を作成する。

【0072】

ステップ1207では、I/O構成情報変換テーブル生成命令列113は、仮想I/O構成情報117から1行を選択する。

【0073】

ステップ1208では、I/O構成情報変換テーブル生成命令列113は、仮想I/O構成情

10

20

30

40

50

報 1 1 7 の行が全て選択されたかどうかを判定する。全ての行が選択されていないなら図 1 7 のステップ 1 2 0 9 へ進み、全ての行が選択されたのならステップ 1 2 1 3 へ進む。

【 0 0 7 4 】

ステップ 1 2 1 3 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、ブートパス情報 1 1 4 からディスク制御部 1 4 0 のバス番号、デバイス番号、ファンクション番号を取得する。

【 0 0 7 5 】

ステップ 1 2 1 4 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、取得した仮想的な位置情報を物理的な位置情報に変換する。すなわち、取得したバス番号、デバイス番号、ファンクション番号を、I/O構成情報変換テーブル 1 2 1 のバス番号、デバイス番号、ファンクション番号と比較して一致した行の変換先番号が指す物理I/O構成情報 1 1 6 のバス番号、デバイス番号、ファンクション番号に変換する。

10

【 0 0 7 6 】

ステップ 1 2 1 5 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、変換後のバス番号、デバイス番号、ファンクション番号で特定されるディスク制御部 1 4 0 を介してストレージ装置内のブートルoader 4 1 1 をメモリ上に読み込む。

【 0 0 7 7 】

図 1 7 に移り、ステップ 1 2 0 9 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、選択した行のベンダID、デバイスID、クラスコードと物理I/O構成情報 1 1 6 の全ての行のベンダID、デバイスID、クラスコードを比較する。

20

【 0 0 7 8 】

ステップ 1 2 1 0 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、ステップ 1 2 0 9 で一致する行があったかどうかを判定する。一致する行がないなら、ステップ 1 2 1 1 へ進み、一致する行があればステップ 1 2 1 2 へ進む。

【 0 0 7 9 】

ステップ 1 2 1 1 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、エラーメッセージを発行して処理を終了する。物理的なI/Oデバイスが仮想的なI/Oデバイスを包含しているとすれば、ステップ 1 2 1 1 は生じない。

【 0 0 8 0 】

ステップ 1 2 1 2 では、I/O構成情報変換テーブル生成命令列 1 1 3 は、ステップ 1 2 0 7 で選択した仮想I/O構成情報 1 1 7 の行をI/O構成情報テーブル 1 2 1 に格納し、ステップ 1 2 1 0 で一致した物理I/O構成情報 1 1 6 のテーブル番号をI/O構成情報変換テーブル 1 2 1 の変換先番号 8 0 7 , 1 7 0 7 に格納し、ステップ 1 2 0 7 へ戻る。

30

【 0 0 8 1 】

図 1 8 は、PCI BIOS 1 1 2 が行うコンフィグレーションレジスタの読み込み処理の処理手順を示すフローチャートである。

【 0 0 8 2 】

ステップ 1 3 0 1 では、PCI BIOS 1 1 2 中のI/O情報変換命令列 1 1 8 は、呼び出し元から渡されたI/Oの仮想的な位置情報（バス番号、デバイス番号、ファンクション番号）をI/O構成情報変換テーブル 1 2 1 と比較して、一致した行の変換先番号が指している物理I/O構成情報 1 1 6 のバス番号、デバイス番号、ファンクション番号に変換する。ここで呼び出し元とは、I/O構成情報変換テーブル生成命令列 1 1 3 のステップ 1 2 0 2、または通常のI/OアクセスをするときのBIOS 1 1 1 である。

40

【 0 0 8 3 】

ステップ 1 3 0 2 では、PCI BIOS 1 1 2 は、変換されたI/Oのバス番号、デバイス番号、ファンクション番号と呼び出し元から渡されたオフセット番号を使って作成したコンフィグレーションアドレスをI/Oポートのアドレス「0xCF8」に書き込む。

【 0 0 8 4 】

ステップ 1 3 0 3 では、PCI BIOS 1 1 2 は、I/Oポートのアドレス「0xCFC」に格納されているコンフィグレーションレジスタの値を読み出す。

50

【 0 0 8 5 】

ステップ 1 3 0 4 では、PCI BIOS 1 1 2 は、読み出したコンフィグレーションレジスタの値を呼び出し元へ返す。

【 0 0 8 6 】

図 1 9 は、系切替後の復旧作業およびこれに伴う系管理テーブル操作命令列 3 1 2 の処理手順を示すフローチャートである。

【 0 0 8 7 】

ステップ 1 4 0 1 では、保守員が障害の発生した計算機の部品交換等の復旧作業を行う。

【 0 0 8 8 】

ステップ 1 4 0 2 では、保守員からの指示により、系管理テーブル操作命令列 3 1 2 は、系管理テーブル 3 1 3 を参照し、復旧した計算機について、障害状態 5 0 3 の「障害あり」を「障害なし」に変更する。

【 0 0 8 9 】

このように、障害が起きた現用系計算機 1 0 0 を系切替対象の待機系計算機 2 0 0 へと切り替える場合に、現用系計算機 1 0 0 と待機系計算機 2 0 0 の I/O の位置情報や数が異なっても、問題なく、現用系計算機 1 0 0 の業務を待機系計算機 2 0 0 に引き継ぐことが可能になる。

【 0 0 9 0 】

なお現用系計算機が複数台ある場合についても、各々の現用系計算機について上記の処理手順が適用される。

【 産業上の利用可能性 】

【 0 0 9 1 】

以上のように、本発明によれば現用系計算機と待機系計算機を備え、現用系計算機の障害時に待機系計算機へ切り替える系切替計算機システムに適用することができる。

【 図面の簡単な説明 】

【 0 0 9 2 】

【 図 1 】 実施例の系切替システムのブロック図である。

【 図 2 】 系管理テーブルの一実施例である。

【 図 3 】 物理 I/O 構成情報の一実施例である。

【 図 4 】 物理 I/O 構成情報の一実施例である。

【 図 5 】 仮想 I/O 構成情報の一実施例である。

【 図 6 】 I/O 構成情報変換テーブルの一実施例である。

【 図 7 】 I/O 構成情報変換テーブルの一実施例である。

【 図 8 】 I/O 構成情報管理テーブルの一実施例である。

【 図 9 】 I/O 構成情報管理テーブル作成の一実施例を示すフローチャートである。

【 図 1 0 】 ハートビート応答命令列の一実施例を示すフローチャートである。

【 図 1 1 】 系管理命令列の一実施例を示すフローチャートである。

【 図 1 2 】 系管理命令列の一実施例を示すフローチャート（続き）である。

【 図 1 3 】 系管理命令列の一実施例を示すフローチャート（続き）である。

【 図 1 4 】 現用系計算機と待機系計算機の起動からサービス提供までの一実施例を示すフローチャートである。

【 図 1 5 】 I/O 構成情報変換テーブル生成命令列の一実施例を示すフローチャートである。

。

【 図 1 6 】 I/O 構成情報テーブル変換命令列の一実施例を示すフローチャート（続き）である。

【 図 1 7 】 I/O 構成情報テーブル変換命令列の一実施例を示すフローチャート（続き）である。

【 図 1 8 】 コンフィグレーションレジスタの読み込み処理の一実施例を示すフローチャートである。

10

20

30

40

50

【図19】系切替後の復旧処理の一実施例を示すフローチャートである。

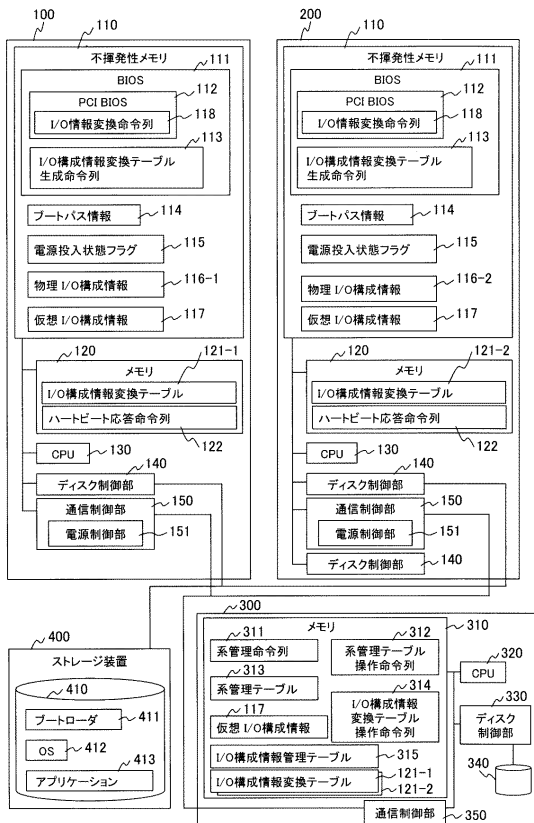
【符号の説明】

【0093】

100：現用系計算機、111：BIOS、113：I/O構成情報変換テーブル生成命令列、
 115：電源投入状態フラグ、116：物理I/O構成情報、117：仮想I/O構成情報、
 118：I/O情報変換命令列、121：I/O構成情報変換テーブル、200：待機系計算機、
 300：管理計算機、311：系管理命令列、312：系管理テーブル操作命令列、313：系管理テーブル、
 314：I/O構成情報変換テーブル操作命令列、315：I/O構成情報管理テーブル。

【図1】

図1



【図2】

図2

313: 系管理テーブル

テーブル番号	コンピュータ識別子	系の種類	障害状態	稼動状態
1	IPアドレス1	現用系	障害あり	停止
2	IPアドレス2	待機系	障害なし	稼動

【図3】

図3

116-1: 物理 I/O構成情報

テーブル番号	バス番号	デバイス番号	ファンクション番号	ベンダID	デバイスID	クラスコード
1	0	10	0	1000	AAAA	FC
2	0	20	0	1000	BBBB	NIC

【 図 4 】

図 4

116-2: 物理 I/O構成情報

1600	1601	1602	1603	1604	1605	1606
テーブル 番号	バス 番号	デバイス 番号	ファンクション 番号	ベンダ ID	デバイス ID	クラス コード
1	1	10	0	1000	AAAA	FC
2	1	20	0	1000	BBBB	NIC
3	1	30	0	1000	AAAA	FC

【 図 5 】

図 5

117: 仮想 I/O構成情報

700	701	702	703	704	705	706
テーブル 番号	バス 番号	デバイス 番号	ファンクション 番号	ベンダ ID	デバイス ID	クラス コード
1	0	10	0	1000	AAAA	FC
2	0	20	0	1000	BBBB	NIC

【 図 6 】

図 6

121-1: I/O構成情報変換テーブル

800	801	802	803	804	805	806	807
テーブル 番号	バス 番号	デバイス 番号	ファンクション 番号	ベンダ ID	デバイス ID	クラス コード	変換先 番号
1	0	10	0	1000	AAAA	FC	1
2	0	20	0	1000	BBBB	NIC	2

【 図 7 】

図 7

121-2: I/O構成情報変換テーブル

1700	1701	1702	1703	1704	1705	1706	1707
テーブル 番号	バス 番号	デバイス 番号	ファンクション 番号	ベンダ ID	デバイス ID	クラス コード	変換先 番号
1	0	10	0	1000	AAAA	FC	1
2	0	20	0	1000	BBBB	NIC	2

【 図 8 】

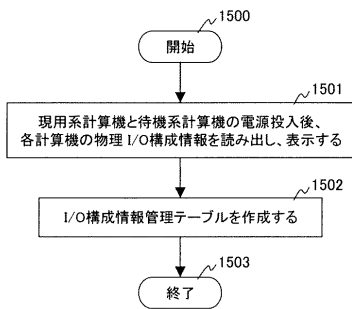
図 8

315: I/O構成情報管理テーブル

1800	1801	1802	1803
テーブル番号	引継元の コンピュータ識別子	引継先の コンピュータ識別子	I/O構成情報変換テーブル
1	IPアドレス1	IPアドレス2	121-1
2	IPアドレス2	IPアドレス1	121-2

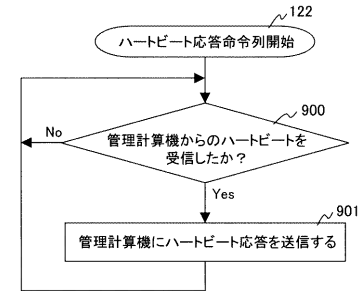
【 図 9 】

図 9

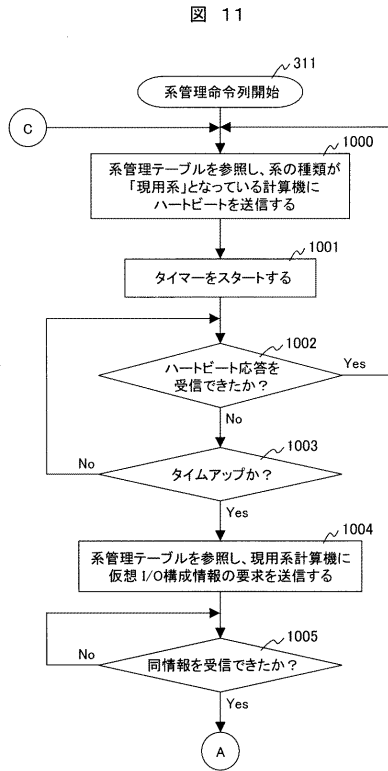


【 図 10 】

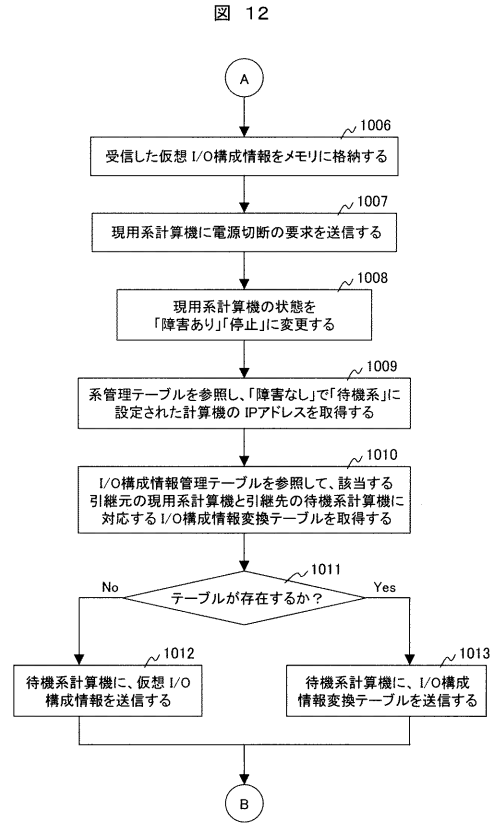
図 10



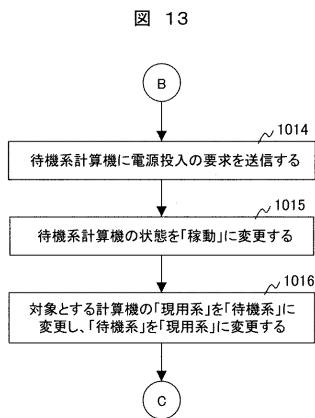
【 図 1 1 】



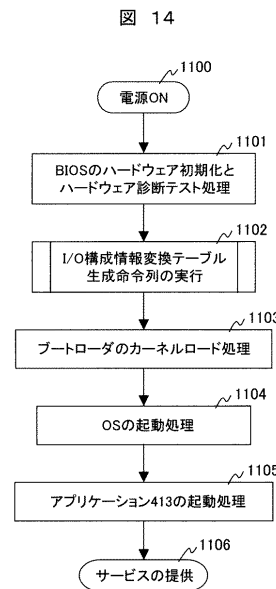
【 図 1 2 】



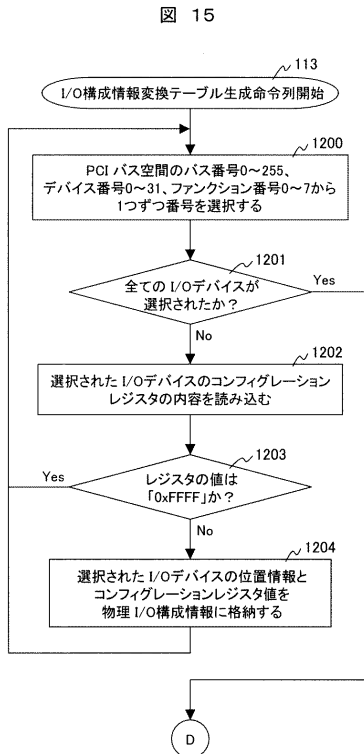
【 図 1 3 】



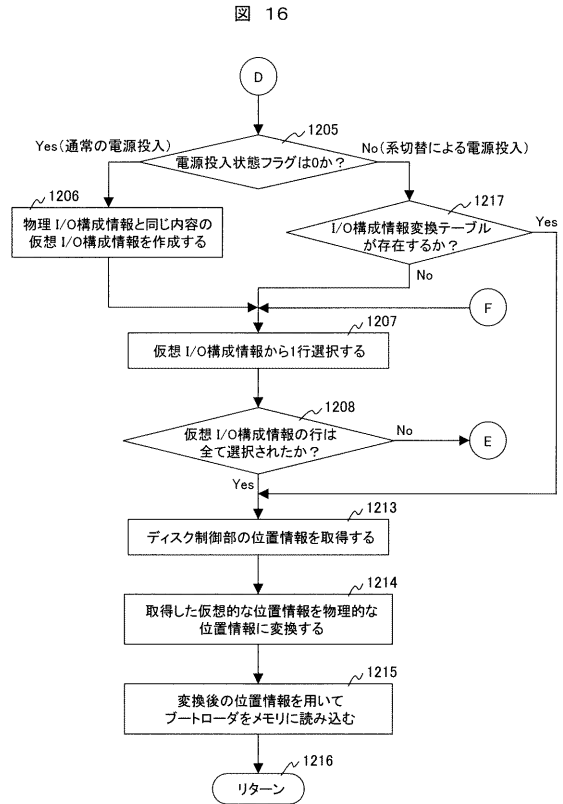
【 図 1 4 】



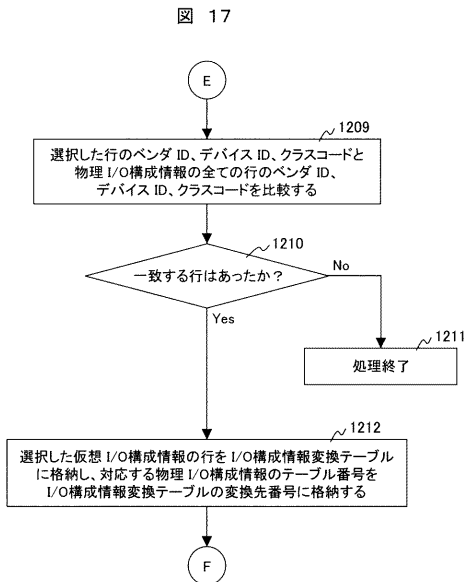
【 図 1 5 】



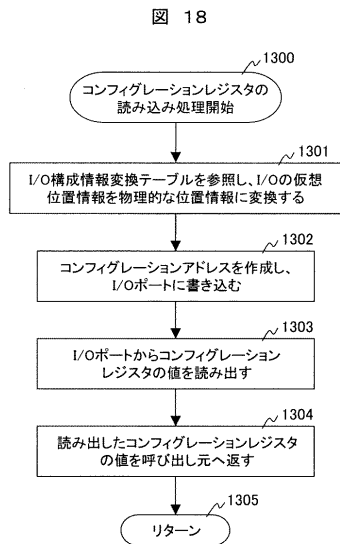
【 図 1 6 】



【 図 1 7 】

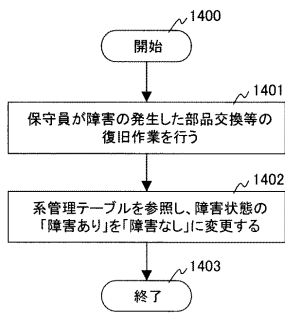


【 図 1 8 】



【 図 19 】

図 19



フロントページの続き

(72)発明者 岩永 美穂

神奈川県秦野市堀山下1番地 株式会社日立製作所エンタ - プライズサ - バ事業部内

Fターム(参考) 5B034 BB03 BB17 CC01 DD05