

(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(43) 国際公開日
2009年9月11日(11.09.2009)

(10) 国際公開番号

WO 2009/110111 A1

(51) 国際特許分類:

G06F 11/30 (2006.01)

[JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 Tokyo (JP).

(21) 国際出願番号:

PCT/JP2008/060739

(72) 発明者; および

(22) 国際出願日:

2008年6月12日(12.06.2008)

(75) 発明者/出願人(米国についてのみ): 國分 俊介 (KOKUBU, Shunsuke) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP). 横口 肇(HIGUCHI, Tsuyoshi) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP).

(25) 国際出願の言語:

日本語

(26) 国際公開の言語:

日本語

(74) 代理人: 溝井 章司(MIZOI, Shoji); 〒2470056 神奈川県鎌倉市大船二丁目17番10号 N T A 大船ビル3階 溝井国際特許事務所 Kanagawa (JP).

(30) 優先権データ:

特願 2008-052815 2008年3月4日(04.03.2008) JP

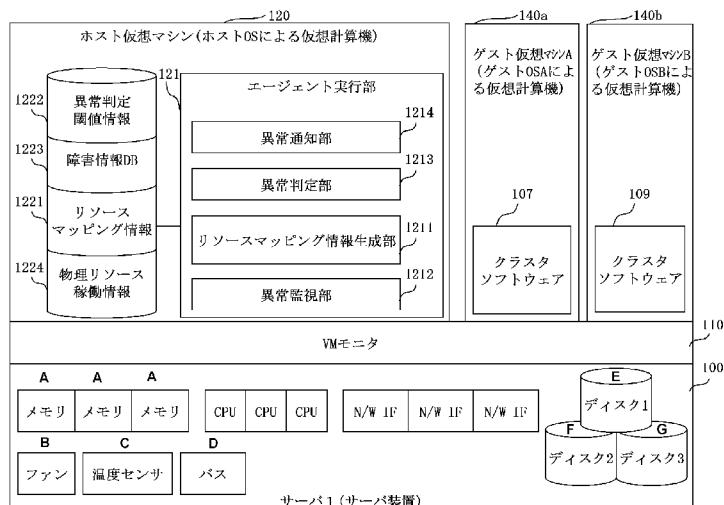
(71) 出願人(米国を除く全ての指定国について): 三菱電機株式会社 (Mitsubishi Electric Corporation)

[続葉有]

(54) Title: SERVER DEVICE, METHOD OF DETECTING FAILURE OF SERVER DEVICE, AND PROGRAM OF DETECTING FAILURE OF SERVER DEVICE

(54) 発明の名称: サーバ装置及びサーバ装置の異常検知方法及びサーバ装置の異常検知プログラム

[図1]



- 120 HOST VIRTUAL MACHINE (VIRTUAL COMPUTER BY HOST OS)
 1222 FAILURE DETERMINATION THRESHOLD INFORMATION
 1223 IMPEDIMENT INFORMATION DB
 1221 RESOURCE MAPPING INFORMATION
 1224 PHYSICAL RESOURCE OPERATING INFORMATION
 121 AGENT EXECUTION PART
 1214 FAILURE INFORMING PART
 1213 FAILURE DETERMINING PART
 1211 RESOURCE MAPPING INFORMATION GENERATING PART
 1212 FAILURE MONITOR PART
 140a GUEST VIRTUAL MACHINE A (VIRTUAL COMPUTER BY GUEST OS A)
 140b GUEST VIRTUAL MACHINE B (VIRTUAL COMPUTER BY GUEST OS B)
 107 CLUSTER SOFTWARE
 109 CLUSTER SOFTWARE
 110 VM MONITOR
 100 SERVER 1 (SERVER DEVICE)
 A MEMORY
 B FAN
 C TEMPERATURE SENSOR
 D BUS
 E DISK 1
 F DISK 2
 G DISK 3

(57) Abstract: Mapping of even a logical resource with a physical resource used by a host OS/guest OS respectively is enabled. An agent execution part (121) for detecting a failure of the physical resource includes a resource mapping information generating part (1211) for generating resource mapping information (1221) in which the physical resource of a server device (100) is associated with the logical resource, a failure monitor part (1212) for collecting physical resource operating information (1224) indicating operating conditions of the physical resource, a failure determining part (1213) for determining whether the physical resource operating information (1224) contains information of the physical resource under abnormal operating conditions and identifying,

[続葉有]



- (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, 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, JP, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (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, 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 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, 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)）

if the abnormal physical resource is found, a virtual computer where the failure occurs based on the information of the physical resource under the abnormal operating conditions and the resource mapping information (1221), and a failure informing part (1214) for informing the specified virtual computer depending on the information of the physical resource under the abnormal operating conditions.

(57) 要約: 論理的なリソースであってもホストOS／ゲストOS各々が使用する物理リソースとのマッピングを可能とすることを目的とする。物理リソースの異常を検知するエージェント実行部121が、論理リソースに対してサーバ装置100の物理リソースを対応付けたリソースマッピング情報1221を生成するリソースマッピング情報生成部1211と、物理リソースの稼働状態を示す物理リソース稼働情報1224を収集する異常監視部1212と、物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があるか否かを判定し、異常な物理リソースがある場合には、異常な稼働状態の物理リソースの情報とリソースマッピング情報1221とにに基づき異常が発生した仮想計算機を特定する異常判定部1213と、特定された仮想計算機に対して、異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知部1214とを備える。-

明細書

サーバ装置及びサーバ装置の異常検知方法及びサーバ装置の異常検知プログラム

技術分野

[0001] 本発明は、例えば、仮想環境においてハードウェアの障害(異常)を検知するエンジニア機能を備えたサーバ装置及びサーバ装置の異常検知方法に関する。

背景技術

[0002] 従来技術では、システムの可用性を高めるために、2台以上のサーバを冗長に構成し、一方の運用系サーバが故障や性能劣化等により動作不能になった場合に、もう片方の待機系サーバが処理を引き継ぐクラスタシステム(系切替システム)がある。一方、サーバリソースの有効活用や運用コストダウンを目的に、仮想化技術によってサーバ集約を行う事例が増加している。仮想化されたサーバによるクラスタシステムを構築する際、ハードウェアや仮想環境上の障害を監視し、特定のソフトウェアやOS(Operating・System)の制御を行う方法(特許文献1参照)や、ある閾値情報をベースに障害を事前予測することによって、予備システム用の仮想マシンを制御する方法(特許文献2参照)が開示されている。

特許文献1:特開2002-229806号公報

特許文献2:特開2004-030363号公報

発明の開示

発明が解決しようとする課題

[0003] 従来の物理サーバ間のクラスタシステムを、仮想化されたサーバ装置(仮想環境上)で利用する場合には以下の課題がある。

[0004] (ア)ゲストドメイン(ゲスト仮想マシン)からは管理ドメイン(ホスト仮想マシン)のリソースが把握できないため、ゲストドメインが動作するために必要な管理ドメインのリソースに障害が発生した場合に、それを検知することができない。

[0005] (イ)上記(ア)を解決するために、管理ドメイン上に障害を監視しゲストドメイン上のクラスタソフトウェアに通知する仕組みを導入したとしても、ドメイン(仮想マシン)から

は論理リソースしか認識できないために、障害内容によっては正しく障害内容の通知ができない。

- [0006] 上記の課題により、ハードウェアに障害や性能劣化が起きたときに、ゲストドメインのゲストOS(もしくはゲストOS上で動作するアプリケーション)が予期しない動作をする可能性があるにもかかわらず、それを適切に障害として検知することができず、データ破壊等の二次障害を引き起こした結果、初めて検知されることになる可能性がある。
- [0007] このとき、上記の課題を解決する方法として、特許文献1では、予めホストOS(ホストドメインのOS)が、物理リソースとホストOS／ゲストOSとのマッピング情報を保持することで、発生したハードウェア障害がどのゲストOSに影響があるかを判別することを可能としている。しかし、特許文献1で示しているマッピング情報は予め設計者により固定的に設定されており物理リソースが固定的なものを対象としているので、ホストOS／ゲストOSに割り当てられているリソースが論理的な表記(例えば、ブリッジに接続されている仮想ネットワークインターフェース名)の場合に対応できない。一方、先の特許文献2では、ホストOSやゲストOSの各々に配置されたエージェントが障害を検知しマネージャに通知し、マネージャが管理する閾値によって系切替制御を実施する構成であるが、上記課題を解決するには至らず、また、エージェント機能をホストOS／ゲストOSの各々すべてに配置しなければならず、処理の効率化の点で課題がある。
- [0008] 本発明は、例えば、上記のような課題を解決するためになされたものであり、論理的なリソースであってもホストOS／ゲストOSと各々が使用する物理リソースのマッピングを可能とする仕組みを提供する。また、物理リソースの障害や性能劣化を仮想環境の管理ドメイン上ののみで検知し、障害／性能劣化の内容に応じて、障害発生時に即時に該当のゲストOS、もしくは、ホストOSを停止させることで、他系のクラスタソフトウェアが系切り替えを実施可能とする仕組みを提供することを目的とする。

課題を解決するための手段

- [0009] 本発明に係るサーバ装置は、物理リソースを用いて複数の仮想計算機を稼働させるサーバ装置であって、上記

物理リソースのうち上記複数の仮想計算機各々が使用する物理リソースを論理リソースとして複数の仮想計算機を稼働させるサーバ装置において、
物理リソースの異常を検知するエージェント実行部を備え、
上記エージェント実行部は、
上記論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピング情報を生成するリソースマッピング情報生成部と、
上記リソースマッピング情報生成部が生成した上記リソースマッピング情報を記憶装置に記憶するリソースマッピング記憶部と、
物理リソースの稼働状態を示す物理リソース稼働情報を収集し記憶装置に記憶する異常監視部と、
上記異常監視部が収集した上記物理リソース稼働情報の中に異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定し、異常な稼働状態の物理リソースの情報がある場合には、上記異常な稼働状態の物理リソースの情報と上記リソースマッピング情報に基づいて上記異常な稼働状態の物理リソースに対応する論理リソースを使用している仮想計算機を処理装置により特定する異常判定部と、
上記異常判定部により特定された仮想計算機に対して、上記異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知部と
を備えることを特徴とする。

- [0010] 上記リソースマッピング情報生成部は、
上記リソースマッピング情報を定期的に生成することを特徴とする。
- [0011] 上記サーバ装置は、上記複数の仮想計算機各々ごとに、当該仮想計算機により使用される論理リソースに対して物理リソースを対応させるための情報である仮想計算機毎リソース管理情報を有する仮想計算機毎リソース管理ファイルを備え、
上記リソースマッピング情報生成部は、
上記仮想計算機毎リソース管理ファイルから物理リソースを含む上記仮想計算機毎リソース管理情報を取得して、取得した上記仮想計算機毎リソース管理情報に基づいて、上記複数の仮想計算機各々により使用される論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピングテーブルをリソースマッピング

情報として生成することを特徴とする。

- [0012] 上記サーバ装置は、リソースの種別ごとに、当該種別の論理リソースに対して当該種別の物理リソースを対応させるための情報であるリソース種別毎管理情報を有するリソース種別毎管理ファイルを備え、
上記リソースマッピング情報生成部は、
上記複数の仮想計算機各々により使用される論理リソースの種別に対応するリソース種別毎管理ファイルから上記リソース種別毎管理情報を取得して、取得した上記リソース種別毎管理情報に基づいて、上記複数の仮想計算機各々により使用される論理リソースに対して上記サーバ装置の物理リソースを対応付けてリソースマッピング情報を生成することを特徴とする。
- [0013] 上記エージェント実行部は、
仮想計算機のOS(Operating・System)のもとで動作するエージェントプログラムを実行し、
上記リソースマッピング情報生成部は、仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドを使用して、論理リソースが使用している物理リソースを調べることを特徴とする。
- [0014] 上記エージェント実行部は、さらに、
物理リソースの稼働状態が異常であるか否かを判定するために用いる閾値と、当該閾値により物理リソースの稼働状態が異常であると判定された場合に、稼働状態が異常であると判定された物理リソースに対応する論理リソースを使用している仮想計算機へ通知する異常通知情報を設定された異常判定閾値情報を予め記憶装置に記憶する異常判定閾値情報記憶部を備え、
上記異常通知部は、
上記異常判定閾値情報に設定された上記異常通知情報を基づいて通知を行うことを特徴とする。
- [0015] 上記複数の仮想計算機のなかで、1台の仮想計算機のみが上記エージェント実行部を備えたことを特徴とする。

[0016] 上記リソースマッピング情報生成部は、

上記論理リソースに対して上記サーバ装置の物理リソースを対応付けて予め作成され記憶装置に記憶されたリソースマッピングファイルを処理装置により取得し、取得したリソースマッピングファイルをリソースマッピング情報とすることを特徴とする。

[0017] 本発明に係るサーバ装置の異常検知方法は、

物理リソースを用いて複数の仮想計算機を稼働させるサーバ装置であつて、上記物理リソースのうち上記複数の仮想計算機各々が使用する物理リソースを論理リソースとして複数の仮想計算機を稼働させるサーバ装置の異常検知方法において、

エージェント実行部により物理リソースの異常を検知するエージェント実行ステップを備え、

上記エージェント実行ステップは、

リソースマッピング情報生成部が、上記論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピング情報を生成するリソースマッピング情報生成ステップと、

リソースマッピング記憶部が、上記リソースマッピング情報生成ステップにより生成された上記リソースマッピング情報を記憶装置に記憶するリソースマッピング記憶ステップと、

異常監視部が、物理リソースの稼働状態を示す物理リソース稼働情報を収集し記憶装置に記憶する異常監視ステップと、

異常判定部が、上記異常監視ステップにより収集された上記物理リソース稼働情報のなかに異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定し、異常な稼働状態の物理リソースの情報がある場合には、上記異常な稼働状態の物理リソースの情報と上記リソースマッピング情報とに基づいて上記異常な稼働状態の物理リソースに対応する論理リソースを使用している仮想計算機を処理装置により特定する異常判定ステップと、

異常通知部が、上記異常判定ステップにより特定された仮想計算機に対して、上記異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知ステップとを備えることを特徴とする。

- [0018] 本発明に係るサーバ装置の異常検知プログラムは、
上記サーバ装置の異常検知方法をコンピュータに実行させることを特徴とする。

発明の効果

- [0019] 本発明によれば、物理リソースの異常を検知するエージェント実行部が、論理リソースに対してサーバ装置の物理リソースを対応付けたリソースマッピング情報を生成するリソースマッピング情報生成部と、リソースマッピング情報を記憶装置に記憶するリソースマッピング記憶部と、物理リソースの稼働状態を示す物理リソース稼働情報を収集し記憶装置に記憶する異常監視部と、物理リソース稼働情報のなかに異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定し、異常な物理リソースがある場合には、異常な稼働状態の物理リソースの情報とリソースマッピング情報に基づいて、異常が発生した仮想計算機を処理装置により特定する異常判定部と、異常判定部により特定された仮想計算機に対して、異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知部とを備えているので、複数の仮想計算機の各々が使用する論理リソースとサーバ装置の物理リソースとのリソースマッピングが可能となり、適切な異常検知処理を行うことができるという効果を奏する。

発明を実施するための最良の形態

- [0020] 図1は、実施の形態1に係るサーバ装置100及びサーバ2装置200の外観の一例を示す図である。図1において、サーバ装置100及びサーバ2装置200は、システムユニット910、CRT(Cathode-Ray Tube)やLCD(液晶)の表示画面を有する表示装置901、キーボード902(Key·Board:K/B)、マウス903、FDD904(Flexible·Disk·Drive)、コンパクトディスク装置905(CDD)、プリンタ装置906、スキャナ装置907などのハードウェア資源を備え、これらはケーブルや信号線で接続されている。

- [0021] システムユニット910は、コンピュータであり、ファクシミリ機932、電話器931とケーブルで接続され、また、ローカルエリアネットワーク942(LAN)、ゲートウェイ941を介してインターネット940に接続されている。

- [0022] 図2は、以下に述べる実施の形態に係るサーバ装置100及びサーバ2装置200のハードウェア資源の一例を示す図である。

- [0023] 図2において、サーバ装置100及びサーバ2装置200は、プログラムを実行するCPU911(Central·Processing·Unit、中央処理装置、処理装置、演算装置、マイクロプロセッサ、マイクロコンピュータ、プロセッサともいう)を備えている。CPU911は、バス912を介してROM913、RAM914、通信ボード915(通信装置、送信装置、受信装置の一例)、表示装置901、キーボード902、マウス903、FDD904、CDD905、プリンタ装置906、スキャナ装置907、磁気ディスク装置920と接続され、これらのハードウェアデバイスを制御する。磁気ディスク装置920の代わりに、光ディスク装置、メモリカード読み書き装置などの記憶装置でもよい。
- [0024] RAM914は、揮発性メモリの一例である。ROM913、FDD904、CDD905、磁気ディスク装置920の記憶媒体は、不揮発性メモリの一例である。これらは、記憶装置あるいは記憶部の一例である。
通信ボード915、キーボード902、スキャナ装置907、FDD904などは、入力部、入力装置の一例である。
- [0025] また、通信ボード915、表示装置901、プリンタ装置906などは、出力部、出力装置の一例である。
- [0026] 通信ボード915は、図示はないが、ファクシミリ機、電話器、LAN等に接続されている。通信ボード915は、LANに限らず、インターネット、ISDN等のWAN(ワイドエリアネットワーク)などに接続されていても構わない。
- [0027] 磁気ディスク装置920には、オペレーティングシステム921(OS)、ウィンドウシステム922、VM(Virtual·Machine)モニタ9200等を含むプログラム群923、ファイル群924が記憶されている。プログラム群923のプログラムは、CPU911、オペレーティングシステム921、ウィンドウシステム922により実行される。
- [0028] 上記プログラム群923には、VMモニタ9200のほかにも以下に述べる実施の形態の説明において「～部」、「～手段」として説明する機能を実行するプログラムが記憶されている。プログラムは、CPU911により読み出され実行される。
- [0029] ファイル群924には、以下に述べる実施の形態の説明において、「～の判定結果」、「～の計算結果」、「～の処理結果」として説明する情報やデータや信号値や変数値やパラメータが、「～ファイル」、「～データベース」、「～データ」の各項目として記

憶されている。「～ファイル」、「～データベース」、「～データ」は、ディスクやメモリなどの記録媒体に記憶される。ディスクやメモリになどの記憶媒体に記憶された情報やデータや信号値や変数値やパラメータは、読み書き回路を介してCPU911によりメインメモリやキャッシングメモリに読み出され、抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示などのCPUの動作に用いられる。抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示のCPUの動作の間、情報やデータや信号値や変数値やパラメータは、メインメモリやキャッシングメモリやバッファメモリに一時的に記憶される。

- [0030] また、以下に述べる実施の形態の説明において説明するフローチャートの矢印の部分は主としてデータや信号の入出力を示し、データや信号値は、RAM914のメモリ、FDD904のフレキシブルディスク、CDD905のコンパクトディスク、磁気ディスク装置920の磁気ディスク、その他光ディスク、ミニディスク、DVD(Digital·Versatile·Disk)等の記録媒体に記録される。また、データや信号は、バス912や信号線やケーブルその他の伝送媒体によりオンライン伝送される。
- [0031] また、以下に述べる実施の形態の説明において「～部」として説明するものは、「～回路」、「～装置」、「～機器」、「手段」であってもよく、また、「～ステップ」、「～手順」、「～処理」であってもよい。すなわち、「～部」として説明するものは、ROM913に記憶されたファームウェアで実現されていても構わない。或いは、ソフトウェアのみ、或いは、素子・デバイス・基板・配線などのハードウェアのみ、或いは、ソフトウェアとハードウェアとの組み合わせ、さらには、ファームウェアとの組み合わせで実施されても構わない。ファームウェアとソフトウェアは、プログラムとして、磁気ディスク、フレキシブルディスク、光ディスク、コンパクトディスク、ミニディスク、DVD等の記録媒体に記憶される。プログラムはCPU911により読み出され、CPU911により実行される。すなわち、プログラムは、以下に述べる「～部」としてコンピュータを機能させるものである。あるいは、以下に述べる「～部」の手順や方法をコンピュータに実行させるものである。
- [0032] 実施の形態1.
本実施の形態では、ハードウェアの異常を検知するエージェント機能を備えるサーバ装置100について説明する。また、サーバ装置100と、サーバ装置100と同様の

構成を有するサーバ2装置200とを冗長に構成した冗長化システム800(系切替システム)について説明する。

[0033] 図3は、実施の形態1に係る冗長化システム800のシステム構成図である。図3を用いて、冗長化システム800のシステム構成について説明する。2台のマシンであるサーバ装置100とサーバ2装置200とは、LAN(Local・Area・Network)101に接続されている。

[0034] 実施の形態1に係るサーバ装置100は、ハードウェア資源(以下、物理リソースともいう)を用いて複数の仮想計算機(仮想マシンともいう)を稼働させるものである。サーバ装置100は、サーバ装置100の備える物理リソースのうち、複数の仮想計算機の各々が使用する物理リソースを論理リソースとして複数の仮想計算機を稼働せるものである。

[0035] サーバ装置100は、上述したようにハードウェア資源(例えば、CPU、ディスク(記憶装置)、ネットワークインターフェース(NW. I/F)、各種筐体ハードウェア等)を備えている。また、サーバ装置100の備えるOS上では、仮想化制御ソフトウェアであるVM(Virtual・Machine)モニタ110が動作している。

[0036] VMモニタ110とは、コンピュータのハードウェア資源(以下、物理リソースともいう)を一括で管理し、サーバ装置100のOSに対しては、物理リソースの一部を組み合わせたリソース(以下、論理リソースともいう)を用いて稼働する仮想マシン(仮想計算機あるいはドメインともいう)と呼ばれる仮想的なコンピュータとして振舞うソフトウェアである。仮想マシンは、仮想OSにより稼働するマシン(計算機)である。言い換えると、仮想マシンは、サーバ装置100の物理リソースを仮想的に定めた論理リソースを用いて仮想OSにより稼働する。このように、実施の形態1のサーバ装置100は、物理的には1台のサーバ装置100上でありながら、VMモニタ110を用いることにより複数の仮想OSを動作させ、あたかも複数の仮想マシン(仮想計算機)が稼働しているかのように振舞うことのできるサーバ装置である。

[0037] サーバ装置100のVMモニタ110上では、VMモニタ110を管理するホスト仮想マシン120(仮想計算機の一例)と2台のゲスト仮想マシンであるゲスト仮想マシンA140a(仮想計算機の一例)とゲスト仮想マシンB140b(仮想計算機の一例)とが仮想的

に稼働している。ここで、ホスト仮想マシン120はホストOSにより稼働する仮想マシンのことであり、以下、ホストOSにより稼働するホスト仮想マシン120をホストOSあるいはホストドメインと呼ぶ場合もある。同様に、ゲスト仮想マシンA140aはゲストOSAにより稼働する仮想マシンのことであり、以下、ゲストOSAにより稼働するゲスト仮想マシンA140aをゲストOSAあるいはゲストドメインAと呼ぶ場合もある。また、同様に、ゲスト仮想マシンB140bはゲストOSBにより稼働する仮想マシンのことであり、以下、ゲストOSBにより稼働するゲスト仮想マシンB140bをゲストOSBあるいはゲストドメインBと呼ぶ場合もある。また、ゲスト仮想マシンA140aとゲスト仮想マシンB140bとを合わせて、ゲスト仮想マシン140と呼ぶ場合もあり、ゲストOSAとゲストOSBとを合わせて、ゲストOSと呼ぶ場合もあるものとする。

- [0038] ホスト仮想マシン120(ホストOSにより稼働するホスト仮想マシン)は、サーバ装置100の物理リソース(ハードウェア資源)の異常、障害を検知するためのエージェント実行部121を備えている。また、ゲスト仮想マシンA140aは、既製のクラスタソフトウェア107を備えており、ゲスト仮想マシンB140bは、既製のクラスタソフトウェア109を備えている。クラスタソフトウェアとは、クラスタシステムにおいて系切替(多重化)制御を行うソフトウェアである。
- [0039] サーバ2装置200は、サーバ装置100と同様の構成をしている。すなわち、サーバ2装置200のOS上では、仮想化制御ソフトウェアであるVMモニタ210が動作しており、VMモニタ210上には、VMモニタ210を管理するホスト仮想マシン'220(ホストOS'により稼働する仮想マシン)と2台のゲスト仮想マシンであるゲスト仮想マシンA'240a(ゲストOSA'により稼働する仮想マシン)とゲスト仮想マシンB'240b(ゲストOSB'により稼働する仮想マシン)が動作している。ホスト仮想マシン'220は、サーバ2装置200の物理リソースの異常・障害を検知するエージェント実行部221を備えている。また、ゲスト仮想マシンA'240aは、既製のクラスタソフトウェア115を備えており、ゲスト仮想マシンB'240bは、既製のクラスタソフトウェア117を備えている。
- [0040] サーバ装置100と、サーバ装置100と同様の構成を有するサーバ2装置200とを冗長に構成した冗長化システム800は、一方の運用系サーバ(サーバ装置100)が故障や性能劣化等により動作不能になった場合に、もう片方の待機系サーバ(サーバ2

装置200)に系を切り替えて待機系サーバ(サーバ2装置200)が処理を引き継ぐクラスタシステム(多重化システムあるいは系切替システムともいう)である。

- [0041] 図4は、実施の形態1に係るサーバ装置100の備えるエージェント実行部121の機能ブロックの構成を示すブロック構成図である。特に記載がない場合は、サーバ2装置200の備えるエージェント実行部221も同様の構成であるものとする。
- [0042] エージェント実行部121は、サーバ装置100においてはホスト仮想マシン120のみに配置されている。サーバ2装置200においても同様に、エージェント実行部221は、ホスト仮想マシン'220のみに配置されている。
- [0043] エージェント実行部121は、リソースマッピング情報生成部1211、異常監視部1212、異常判定部1213、異常通知部1214を備えている。エージェント実行部121は、図示しないリソースマッピング情報記憶部にリソースマッピング情報1221を記憶装置に記憶させ、図示しない異常判定閾値情報記憶部に異常判定閾値情報1222を記憶装置に記憶させる。また、エージェント実行部121は、図示しない記憶部に、障害情報データベース1223、物理リソース稼働情報1224を記憶装置に記憶させる。
- [0044] リソースマッピング情報生成部1211は、サーバ装置100上で稼働する各仮想マシン(ホスト仮想マシン120、ゲスト仮想マシンA140a、ゲスト仮想マシンB140b)が使用する論理リソースに対して、サーバ装置100の物理リソースを対応付けたリソースマッピング情報1221を生成する。リソースマッピング情報生成部1211は、各仮想マシン(各ドメイン)の使用しているリソースと実際の物理リソースをマッピングし、リソースマッピング情報1221を生成する。リソースマッピング情報生成部1211により生成されたリソースマッピング情報1221は、リソースマッピング情報記憶部により記憶装置に記憶される。リソースマッピング情報生成部1211の、リソースマッピング情報生成処理については、後述する。
- [0045] 異常監視部1212は、物理リソースの稼働状態を示す物理リソース稼働情報1224を収集し記憶装置に記憶する。すなわち、異常監視部1212は、エージェント実行部121が動作しているサーバ装置100のCPUやディスク、ネットワークI/F(Interface)(ネットワークインターフェース(NW. I/F))等のハードウェア障害やディスク応答性能等の情報を収集して物理リソース稼働情報1224として記憶装置に記憶する。さら

に、異常監視部1212は、IPMI(Intelligent・Platform・Management・Interface)により、サーバ筐体の温度や電源、ファン、バス等の状態を監視してそれらの状態情報を収集し、物理リソース稼働情報1224として記憶装置に記憶する。IPMIとは、例えばサーバ装置100のサーバープラットフォームの状態(温度、電圧、ファン、バス等)の状態監視、復旧制御、リモート制御等を行うOSの標準インターフェース仕様である。

[0046] 異常判定閾値情報1222は、予め異常判定閾値情報記憶部により記憶装置に記憶されている情報である。異常判定閾値情報1222には、物理リソースの稼働状態が異常であるか否かを判定するために用いる閾値と、当該閾値により物理リソースの稼働状態が異常であると判定された場合に、稼働状態が異常であると判定された物理リソースに対応する論理リソースを使用している仮想マシン(仮想計算機)へ通知する異常通知情報とが設定されている。異常判定閾値情報1222の詳細については、後述する。

[0047] 異常判定部1213は、異常監視部1212が収集した物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定する。異常判定部1213は、物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があるか否かを異常判定閾値情報1222に基づいて判定する。すなわち、異常判定部1213は、異常監視部1212が収集した物理リソース稼働情報1224(監視情報)を、異常判定閾値情報1222に基づいて通知すべき異常かどうかを判定する。異常判定部1213は、物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があると判定した場合には、異常な稼働状態の物理リソースの情報とリソースマッピング情報1221に基づいて、異常な稼働状態の物理リソースに対応する論理リソースを使用している仮想マシン(仮想計算機)を処理装置により特定する。

[0048] 異常通知部1214は、異常判定部1213により異常な稼働状態の物理リソースに対応する論理リソースを利用している仮想マシンであるとして特定された仮想マシン(以下、障害発生仮想マシンという)に対して、異常な稼働状態の物理リソースの情報に応じた通知を行う。異常通知部1214は、後述する異常判定閾値情報1222に設定さ

れた異常通知情報1114に基づいて、異常な稼働状態の物理リソースの障害情報を応じた通知を行う。異常通知部1214は、物理リソース稼働情報1224(監視情報)が異常判定部1213により異常と判定された場合に、異常と判定された物理リソースの障害情報を障害情報データベース1223に記録して記憶装置に保存するとともに、異常判定部1213により障害のある仮想マシンと特定された障害発生仮想マシン(ホスト仮想マシン120あるいはゲスト仮想マシンA140aあるいはゲスト仮想マシンB140b)に対して、異常通知情報1114に基づいて障害情報に応じた通知を行う。

- [0049] 本実施の形態では、エージェント実行部121が、リソースマッピング情報1221を生成することが特徴のひとつである。また、エージェント実行部121が、ホスト仮想マシン120にのみ備えられていることが特徴のひとつである。エージェント実行部121は、ホスト仮想マシン120のみに備えられていても、リソースマッピング情報1221を備えることにより、サーバ装置100上で稼働する他の仮想マシンの論理リソースについても管理することができ、適切に障害発生仮想マシンを特定することができる。また、エージェント実行部121は、ホスト仮想マシン120のみに備えられていればよいので、サーバ装置100のエージェント機能の処理効率を向上させることができる。
- [0050] 図5は、実施の形態1におけるサーバ装置100の異常検知方法の処理動作を示すフロー図である。図5を用いて、実施の形態1におけるサーバ装置100の異常検知方法(異常検知プログラム)について説明する。以下で説明するOS(サーバ装置100のOS)及びホストOS及びゲストOS及びエージェント実行部121の各部は、以下で説明する各処理をCPU、記憶装置等のハードウェア資源を用いて実行する。
- [0051] まず、利用者により、あるいは、自動的にサーバ装置100が起動されると、ホストOSがCPUにより読み出され実行されホスト仮想マシン120が起動する(S101)。次に、サーバ装置100のCPUによりゲストOSAとゲストOSBとが読み出され実行され、ゲスト仮想マシンA140aとゲスト仮想マシンB140bとが起動する(S102)。各ゲスト仮想マシンの各ゲストOS上では、CPUにより既製のクラスタソフトウェアが読み出されて実行されることにより動作し、ゲストOSAとサーバ2装置200のゲストOSA'の組、及び、ゲストOSBとサーバ2装置200のゲストOSB'の組でそれぞれ冗長化構成がとられる。

[0052] 次に、エージェント実行部121が、CPUにより、ホスト仮想マシン120のホストOS上で起動される(S103)。エージェント実行部121は、ホスト仮想マシン120のホストOSのもとで動作するエージェントプログラムをCPUにより実行する。エージェントプログラムは、ホストOS上で常に動作するプログラム(常駐プログラム)としてCPUにより動作する。

[0053] <S104:リソースマッピング情報生成ステップ>

リソースマッピング情報生成部1211は、CPUを用いて、ホスト仮想マシン120及びゲスト仮想マシンA140a及びゲスト仮想マシンB140bが使用する論理リソースとサーバ装置100の物理リソースとのリソースマッピングの情報を取得してリソースマッピング情報1221を生成する(S104)。サーバ装置100は、例えば、各仮想マシン毎に、当該仮想マシンにより使用される論理リソースに対して物理リソースを対応させるための情報である仮想計算機毎リソース管理情報を有する仮想計算機毎リソース管理ファイルを記憶装置に備えており、リソースマッピング情報生成部1211は、仮想計算機毎リソース管理ファイルから物理リソースを含む仮想計算機毎リソース管理情報を取得して、取得した仮想計算機毎リソース管理情報に基づいて、各仮想マシンにより使用される論理リソースに対してサーバ装置100の物理リソースを対応付けたリソースマッピングテーブルをリソースマッピング情報1221として生成する。また、サーバ装置100は、例えば、リソースの種別ごとに、当該種別の論理リソースに対して当該種別の物理リソースを対応させるための情報であるリソース種別毎管理情報を有するリソース種別毎管理ファイルを記憶装置に備えており、リソースマッピング情報生成部1211は、各仮想マシンにより使用される論理リソースの種別に対応するリソース種別毎管理ファイルからリソース種別毎管理情報を取得して、取得したリソース種別毎管理情報に基づいて、各仮想マシンにより使用される論理リソースに対してサーバ装置100の物理リソースを対応付けてリソースマッピング情報1221を生成する。

[0054] このように、リソースマッピング情報生成部1211は、仮想マシンのOSに含まれたツール又はコマンド、あるいはエージェントプログラムに含まれたツール又はコマンドを使用して、各仮想マシンの論理リソースが使用している物理リソースをCPUを用いて調べてリソースマッピング情報1221を生成する。リソースマッピング情報記憶部は、

生成したリソースマッピング情報1221を記憶装置に記憶(保存)する。

[0055] リソースマッピング情報生成部1211は、CPUを用いて、定期的にリソースマッピング情報1221を収集・生成してリソースマッピング情報記憶部がリソースマッピング情報1221を記憶装置に記憶して更新する。すなわち、リソースマッピング情報1221は、定期的に更新される。このように、リソースマッピング情報生成・記憶処理は、CPUを用いて、定期的に実行される。あるいは、リソースマッピング情報生成・記憶処理は、エージェント実行部121が起動されて処理を開始した場合に、まず始めに実行される処理としてもよい。この場合は、エージェント実行部121が定期的に起動されれば、自動的にリソースマッピング情報生成・記憶処理も定期的に実行されることになる。あるいは、リソースマッピング情報生成部1211は、エージェント実行部121の処理とは独立して実行されても構わない。リソースマッピング情報生成部1211のリソースマッピング情報生成方法についての詳細な説明は後述する。

[0056] <S105:異常監視ステップ>

異常監視部1212は、CPUを用いて、定期的にハードウェア(物理リソース)を監視し、ハードウェア(物理リソース)の稼働状態を示す物理リソース稼働情報1224を収集する。異常監視部1212は、収集した物理リソース稼働情報1224を記憶装置に記憶する。(S105)。物理リソース稼働情報1224とは、例えば、上述したIPMIによる筐体情報(電源情報、CPU温度、バス情報、ファン稼働情報など)や、ハードディスクの書き込み/読み込みエラーや応答性能、ネットワークI/F(ネットワークインターフェース(NW. I/F))の応答性能などである。異常監視部1212は、CPUを用いて、物理リソース稼働情報1224を収集したことを異常判定部1213に通知する。

[0057] <S106～S108:異常判定ステップ>

<S106～S107:異常物理リソース検知ステップ>

異常判定部1213は、CPUを用いて、異常監視部1212が収集した物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があるか否かを判定する。異常判定部1213は、異常監視部1212により物理リソース稼働情報1224が収集されたことが通知されると、収集された物理リソース稼働情報1224のなかに異常(障害)があるかどうかを判断する(S106)。異常判定部1213は、予め異常判定閾値

情報記憶部により記憶装置に記憶されている異常判定閾値情報1222(異常判定閾値情報データベース)により定義された情報をもとにCPUを用いて障害・異常があるか否かを判定する(S107)。

[0058] 図6は、異常判定閾値情報1222のテーブル構成を示す図である。図6を用いて、異常判定部1213の異常判定処理の具体例について説明する。

[0059] 異常判定閾値情報1222は、異常な物理リソース稼働状態を識別する識別子を設定するID1111と、異常な稼働状態の対象となるハードウェア(物理リソース)を設定する対象ハードウェア1112と、異常な稼働状態と判定するための閾値を設定する異常判定閾値1113と、異常な稼働状態と判定した場合に、後述する異常が検知された仮想マシンを特定する処理(障害発生仮想マシン特定処理)において特定される障害発生仮想マシン(障害発生仮想OS)に対して行う異常検知時の通知内容を設定する異常知情報1114とから構成される。

[0060] 例えば、図6における異常な物理リソース稼働状態のID1111が「E00001」の情報は、CPU-1に関して温度が60度を超えた場合に異常であると判定するための情報であり、その際、CPU-1をリソース(論理リソース)として使用していると特定された仮想マシンに対しては、異常知情報1114「OSの停止」という内容の通知をすることを意味している。また、例えば、ID1111が「E00003」の情報は、ハードディスク「/dev/sda/」に関して読み込み応答時間(応答時間READ)が10秒を超えた場合に異常であると判定するための情報であり、その際、ハードディスク「/dev/sda/」を論理リソースとして使用していると特定された仮想マシンに対しては、「OSの停止」という内容の異常知情報1114を通知することを意味している。

[0061] 図5に戻り、異常判定部1213は、CPUを用いて、物理リソース稼働情報1224の各物理リソースの稼働状態を示す稼働情報のそれぞれを、異常判定閾値情報1222の各異常な物理リソース稼働状態(すなわち、各IDの情報)と比較することにより、物理リソース稼働情報1224のなかに異常な稼働状態の物理リソースの情報があるか否かを判定する。

[0062] 例えば、異常判定部1213は、記憶装置に記憶された物理リソース稼働情報1224のなかに「CPU-1の温度が63度」という情報があるとする。異常判定部1213は、C

PUを用いて、異常判定閾値情報1222を参照し、ID1111が「E00001」の「CPU-1」に関して温度が60度を超えた場合に異常であるとの情報に基づいて、物理リソース「CPU-1」は異常であると判定する。また、例えば、異常判定部1213は、収集された物理リソース稼働情報1224(ハードウェア稼働情報)なかに「ハードディスク「/dev/sda」の読み応答時間が20秒」という情報がある場合、CPUを用いて、異常判定閾値情報1222のID1111=「E00003」に基づくと「読み応答時間が10秒以上の場合に異常」と判定されるため、ここでハードディスク「/dev/sda」は異常(障害)であるとして認識する。

[0063] <S108:障害発生仮想マシン特定ステップ>

異常判定部1213により物理リソースの異常(障害)が検知されなかった場合(S107でNO)は、エージェント実行部121は、S104のリソースマッピング情報生成ステップへ処理を戻す。

[0064] 異常判定部1213により物理リソースの異常(障害)が検知された場合(S107でYES)は、異常判定部1213は、CPUを用いて、リソースマッピング情報1221に基づいて、異常(障害)が検知された物理リソースがどの仮想マシン(ホストOS/ゲストOS)に関連するかを抽出(特定)する(S108)。すなわち、異常判定部1213は、異常な物理リソース稼働状態の物理リソースを論理リソースとして使用している仮想マシン(以下、障害発生仮想マシン(障害発生ドメイン)と呼ぶ)を特定する。障害発生仮想マシンは、1つの場合もあるが、対象の物理リソースを複数の仮想マシンで共有している場合には複数の場合もある。異常判定部1213は、CPUを用いて、S106で検知した異常な物理リソース稼働状態のID1111とS108で特定した障害発生仮想マシンの情報を異常通知部1214へ出力する。また、あるいは、異常判定部1213は、CPUを用いて、S106で検知した異常な物理リソース稼働状態のID1111に対応する異常通知情報1114とS108で特定した障害発生仮想マシンの情報を異常通知部1214へ出力する。S108の障害発生仮想マシン特定処理についての詳細説明は後述する。

[0065] <S109～S110:異常通知ステップ>

異常通知部1214は、異常判定部1213により障害発生仮想マシン(障害発生ドメ

イン)が特定(抽出)されると、CPUを用いて、障害発生仮想マシンと異常(障害)のある物理リソースの異常状態とを対応付けて障害情報データベース1223として記憶装置に記録する(S109)。

- [0066] さらに、異常通知部1214は、CPUを用いて、障害発生仮想マシンに対して異常(障害)内容に応じた通知を実施する(S110)。異常通知部1214は、CPUを用いて、異常判定閾値情報1222から障害発生仮想マシンの異常な物理リソース稼働状態のID1111に対応する異常知情報1114の内容を取得する。あるいは、異常通知部1214は、異常判定部1213から異常な物理リソース稼働状態のID1111を入力して、入力したID1111を基に異常判定閾値情報1222から入力したID1111に対応する異常知情報1114を取得する。あるいは、異常通知部1214は、異常判定部1213から直接異常知情報1114を入力して取得する。異常通知部1214は、CPUを用いて、取得した異常知情報1114の内容を障害発生仮想マシンへ通知する。
- [0067] 例えば、障害発生仮想マシンに関連する異常な物理リソース稼働状態のID1111が「E00001」の場合は、異常通知部1214は、異常知情報1114「OSの停止」を障害発生仮想マシンに通知することを意味している。障害発生仮想マシンは、「OSの停止」が通知されると、通知内容に基づいてOSの停止を実施する。なお、障害によっては、障害発生仮想マシン自身が正常にOS停止できない場合がある。そのような場合は、例えば、障害発生仮想マシンがホストOSならば、OSパニックを起こさせて強制的に停止させる。また、例えば、障害発生マシンがゲストOSならば、エージェント実行部121は、VMモニタのコマンドを用いて該当の障害発生ゲストOSを強制的に停止させる。
- [0068] 図7は、実施の形態1に係る冗長化システム800の系切替時の動作の一例を示す図である。図7では、サーバ装置100のホスト仮想マシン120のホストOSが使用している「ハードディスク「/dev/sda」」が故障して、応答性能が10秒以上になったものとする。この場合、異常判定部1213は、上述した処理によりハードディスク「/dev/sda」は異常判定閾値情報1222のID1111「E00003」に基づいて異常であると判定する。また、異常判定部1213は、上述した処理により障害発生仮想マシンがホスト仮想マシン120であると特定する。異常通知部1214は、上述した処理により異常判

定閾値情報1222のID1111「E00003」の異常通知情報1114「OSの停止」を取得してホスト仮想マシン120に通知する。ホスト仮想マシン120は、受け取った通知内容に基づいてホストOSを停止する(S61)。ホストOSが停止したことによって、同一サーバ装置100上で稼働するゲスト仮想マシンA140aとゲスト仮想マシンB140bとが、それぞれゲストOSAとゲストOSBとを停止する(S62)。これにより、ゲストOSA上のクラスタソフトウェア107とゲストOSB上のクラスタソフトウェア109とが停止し、クラスタソフトウェア107, 109によりサーバ2装置200へ送信されていたハートビートが停止する(S63)。このハートビートの停止により、本実施の形態の冗長化システム800では、他系(待機系)(サーバ2装置200)のクラスタソフトウェア115, 117が適切に異常を検知することができるため、適切な系切り替え動作が可能となる(S64)。

- [0069] 図8は、エージェント実行部121が動作するホスト仮想マシンのホストOSにより認識できるディスク情報(ここでは、ゲスト仮想マシン140のディスク情報)とゲスト仮想マシンが実際に利用する物理的なディスク情報とのリソースマッピング情報生成処理を示すフロー図である。図9は、リソースマッピング情報の仮想マシン管理テーブル21のテーブル構成を示す図である。図10は、リソースマッピング情報のリソースマッピングテーブル13の構成を示す図である。図8～図10を用いて、エージェント実行部121のリソースマッピング情報生成部1211によるリソースマッピング情報生成処理の詳細動作について説明する。
- [0070] リソースマッピング情報1221は、以下に説明する仮想マシン管理テーブル21とリソースマッピングテーブル13とから構成されている。
- [0071] まず、図9のリソースマッピング情報の仮想マシン管理テーブル21について説明する。仮想マシン管理テーブル21には、新規に付与される管理ID211と、冗長化システム800における物理的なサーバを識別するハードウェア識別ID212と、仮想マシン(ドメイン)を識別するドメインID213と、ドメインIDに対応するドメインの名称を設定するドメイン名称214とがひとつの情報(1レコード)として設定される。すなわち、仮想マシン管理テーブル21は、仮想マシンと仮想マシンが稼働する物理的なサーバとを対応させるためのテーブルである。

- [0072] 次に、図10のリソースマッピング情報のリソースマッピングテーブル13について説明する。リソースマッピングテーブル13は、仮想マシン管理テーブル21で付与された管理ID211を設定する管理ID131と、その管理ID131により示される仮想マシンのリソース(論理リソース)に対してシーケンシャルにIDを付与するリソースID132と、リソースの種別を設定するリソース種別133と、対応するサーバ装置100物理リソースを設定する対応物理リソース名134、ホストOS上で認識されるリソースを設定するホストOS上の識別名135(論理リソース名)とから構成される。
- [0073] 図8を用いて、リソースマッピング情報生成部1211が、CPUを用いて、仮想マシン管理テーブル21及びリソースマッピングテーブル13に情報を設定することにより、リソースマッピング情報1221を生成する処理について説明する。リソースマッピング情報生成部1211は、CPUを用いてリソースマッピング情報生成プログラムを記憶装置より読み出して、CPUを用いてリソースマッピング情報生成プログラムを実行する。
- [0074] 図8では、ゲスト仮想マシン140のゲストOSのディスク情報(以下、論理ディスク情報という)とゲストOSが利用している物理的なディスク情報(物理ディスク情報)とをマッピングする方法について説明する。
- [0075] <ゲストOSのディスク情報と物理ディスク情報とのリソースマッピング情報生成処理>
- 前提処理として、リソースマッピング情報生成部1211は、サーバ(ハードウェア)を識別するハードウェア識別ID212として、サーバ名(ホスト名)やIPアドレス等を利用するものとする。まず、リソースマッピング情報生成部1211は、ハードウェア識別ID212として自己の動作するサーバのサーバ名「サーバ1(サーバ装置100)」を取得する(S201)。次に、リソースマッピング情報生成部1211は、サーバ装置100のVMモニタ110の管理ツールを用いて、サーバ装置100上で稼働する各仮想マシン(各ドメイン)を識別するドメインID213と各仮想マシンを識別するドメイン名称214とを取得する(S202、S203)。
- [0076] 例えば、リソースマッピング情報生成部1211は、ドメインID「0」=ドメイン名称「ホストOS」の情報を取得するものとする。リソースマッピング情報生成部1211は、新規に管理ID211を付加(取得)し、取得したハードウェア識別ID212とドメインID213とド

メイン名称214とを関連付けてリソースマッピング情報1221の仮想マシン管理テーブル21に登録する。リソースマッピング情報生成部1211は、新規に付与(取得)した管理ID「00001」に対してハードウェア識別ID「サーバ1(サーバ装置100)」とドメインID「0」とドメイン名称「ホストOS」とを関連付けて仮想マシン管理テーブル21に設定する(図9参照)。

- [0077] 次に、例えば、リソースマッピング情報生成部1211は、ドメインID「1」=ドメイン名称「ゲストOSA」の情報を取得するものとする。リソースマッピング情報生成部1211は、新規に管理ID211を付加(取得)し、取得したハードウェア識別ID212とドメインID213とドメイン名称214とを関連付けてリソースマッピング情報1221の仮想マシン管理テーブル21に登録する。リソースマッピング情報生成部1211は、新規に付与(取得)した管理ID「00002」に対してハードウェア識別ID「サーバ1(サーバ装置100)」とドメインID「1」とドメイン名称「ゲストOSA」とを関連付けて仮想マシン管理テーブル21に設定する(図9参照)。すなわち、リソースマッピング情報生成部1211は、管理ID211に「00002」、ハードウェア識別IDに「サーバ1(サーバ装置100)」、ドメインIDに「1」、ドメイン名称に「ゲストOSA」を設定する。
- [0078] このように、リソースマッピング情報生成部1211は、サーバ装置100上で稼働するすべての仮想マシンについて順に仮想マシンの各々と物理的なサーバと対応させる情報を仮想マシン管理テーブル21に設定する(S204)。リソースマッピング情報生成部1211は、既に仮想マシン管理テーブル21に同一の情報が設定されていれば、その情報をを利用して管理IDを取得する。
- [0079] リソースマッピング情報生成部1211は、S204で登録した仮想マシン管理テーブル21からひとつのゲストOSの管理ID211を取得する。リソースマッピング情報生成部1211は、取得した管理ID211の情報(ハードウェア識別ID212、ドメインID213、ドメイン名称214)を基に、該当のゲスト仮想マシンのゲストOS用VM設定ファイル(仮想計算機毎リソース管理情報を有する仮想計算機毎リソース管理ファイルの一例)を取得する(S205)。
- [0080] リソースマッピング情報生成部1211は、取得したゲストOS用VM設定ファイルから、対象ゲストOSが使用しているディスク情報(論理ディスク情報)(物理リソースを含む

上記仮想計算機毎リソース管理情報の一例)を取得して、取得したディスク情報が物理ディスク情報であるか否かをCPUを用いて判定する(S206)。例えば、リソースマッピング情報生成部1211は、対象ゲストOSが使用しているディスク情報が物理的な記載によるディスク情報である場合に物理ディスク情報であると判定する。

- [0081] リソースマッピング情報生成部1211は、取得したディスク情報が物理ディスク情報である場合は(S206でYES)、取得したディスク情報をそのままリソースマッピングテーブル13の対応物理リソース名134に設定する情報として取得する(S207)。リソースマッピング情報生成部1211は、取得したディスク情報が物理ディスク情報でない場合は(S206でNO)、S208に進む。S208では、リソースマッピング情報生成部1211は、取得した物理ディスク情報でないディスク情報がイメージファイル(イメージデータ)で指定されているか否かをCPUを用いて判定する(S208)。
- [0082] リソースマッピング情報生成部1211は、取得したディスク情報がイメージファイルで指定されている場合は(S208でYES)、dfコマンド等のOS管理ツールを用いて、イメージファイルが配置されている物理ディスク情報を取得する。リソースマッピング情報生成部1211は、取得した物理ディスク情報をゲストOSが使用している物理ディスク情報として取得する(S209)。リソースマッピング情報生成部1211は、取得したディスク情報が物理ディスク情報でもなくイメージファイルで指定されてもいい場合は(S208でNO)、エラー情報を出力してS205の処理に戻り、次のゲスト仮想マシン140のゲストOS用VM設定ファイルをチェックする処理を行う(S210)。
- [0083] また、例えば、リソースマッピング情報生成部1211は、S206にてVM設定ファイルにディスク情報が存在しない等の不正がある場合にも、エラー情報を出力してS205の処理に戻り、次のゲスト仮想マシン140のゲストOS用VM設定ファイルをチェックする処理を行う。
- [0084] S211では、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS205で取得した管理ID211を設定し、リソースID132に対象のゲスト仮想マシン140のリソースに対して例えばシーケンシャルに付与するIDを取得して設定し、リソース種別133にディスク情報のリソース種別を表す「HDD」を設定し、ホストOS上の識別名135にはS206で取得した対象のゲストOSが使用しているディ

スク情報(論理ディスク情報)を設定し、対応物理リソース名134にS207あるいはS209で取得した物理ディスク情報を設定する。リソースID132は、同一管理IDで管理されるリソースを一意に識別できる任意に付加されるIDであるものとする。このように、リソースマッピング情報生成部1211は、S205で取得した仮想マシン管理テーブル21の管理ID211と関連付けて、リソースマッピングテーブル13を登録する。

- [0085] S212にて、リソースマッピング情報生成部1211は、自己の動作するサーバ装置100上のすべてのゲスト仮想マシン140についてリソースマッピング情報生成処理が完了するまで、上記処理(S205～S212)を繰り返す。
- [0086] 具体例を用いて以下に説明する。例えば、S205において、リソースマッピング情報生成部1211は、管理ID211の「00002」を取得したものとする。S205において、リソースマッピング情報生成部1211は、管理ID211「00002」の示すゲストOSが「ゲストOSA」であるから、ゲストOSA用VM設定ファイルを取得する。リソースマッピング情報生成部1211は、取得したゲストOSA用VM設定ファイルから、ディスク情報を取得する。ここで、ゲストOSAのディスク情報は、イメージデータ「/dev/sdb/hdd.img」であるとする。そこで、リソースマッピング情報生成部1211は、S206からS208の処理を行い、ディスク情報がイメージデータであると判断して、dfコマンド等のOS管理ツールを用いて、イメージファイルが配置されている物理ディスク情報「/dev/sdb」を取得する(S209)。そして、S211において、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS204で取得した管理ID211「00002」を設定し、リソースID132にゲストOSAのリソースに対して付与するID「1」を設定し、リソース種別133にディスク情報のリソース種別を表す「HDD」を設定し、ホストOS上の識別名135にはS206で取得したゲストOSAのディスク情報「/dev/sdb/hdd.img」を設定し、対応物理リソース名134にS209で取得した物理ディスク情報「/dev/sdb」を設定する。
- [0087] 図11は、実施の形態1に係るホスト仮想マシン120のホストOSのディスク情報(論理ディスク情報)とホストOSが利用している物理的なディスク情報(物理ディスク情報)とのリソースマッピング情報生成処理を示すフロー図である。図11を用いて、ホスト

仮想マシン120のホストOSとホストOSが利用している物理的なディスク情報(物理ディスク情報)をマッピングする方法について説明する。

[0088] <ホストOSのディスク情報と物理ディスク情報とのリソースマッピング情報生成処理>

前提処理として、リソースマッピング情報生成部1211は、サーバ(ハードウェア)を識別するハードウェア識別ID212として、サーバ名(ホスト名)やIPアドレス等を利用するものとする。まず、リソースマッピング情報生成部1211は、ハードウェア識別ID212として自己の動作するサーバのサーバ名「サーバ1(サーバ装置100)」を取得する(S301)。次に、リソースマッピング情報生成部1211は、サーバ装置100のVMモニタ110の管理ツールを用いて、サーバ装置100上で稼働する各仮想マシン(各ドメイン)を識別するドメインID213と各仮想マシン(各ドメイン)を識別するドメイン名称214を取得する(S302)。例えば、リソースマッピング情報生成部1211は、サーバ装置100上で稼働するホスト仮想マシン120のドメインID「0」=ドメイン名称「ホストOS」の情報を取得するものとする。リソースマッピング情報生成部1211は、新規に管理ID211を取得して付加し、取得したハードウェア識別ID212とドメインID213とドメイン名称214とを関連付けてリソースマッピング情報1221の仮想マシン管理テーブル21に登録する。リソースマッピング情報生成部1211は、新規に付与した管理ID「00001」に対してハードウェア識別ID「サーバ1(サーバ装置100)」とドメインID「0」とドメイン名称「ホストOS」とを関連付けて仮想マシン管理テーブル21に設定する(図9参照)。このように、リソースマッピング情報生成部1211は、サーバ装置100上で稼働するすべての仮想マシンについて順に仮想マシンの各々と物理的なサーバと対応させる情報を仮想マシン管理テーブル21に設定する(S302)。リソースマッピング情報生成部1211は、既に仮想マシン管理テーブル21に同一の情報が設定されていれば、その情報をを利用して管理IDを取得する。

[0089] リソースマッピング情報生成部1211は、S304で登録した仮想マシン管理テーブル21からホストOSの管理ID211を取得する。S303において、リソースマッピング情報生成部1211は、ホストOSの管理ID211として「00001」を取得するものとする。リソースマッピング情報生成部1211は、OSの管理ツールを用いて、ホスト仮想マシン1

20のホストOSがマウントしている物理ディスク情報(例えば「/dev/sda」)を取得する(S303)。リソースマッピング情報生成部1211は、S303で取得した管理ID「00001」とS303で取得した物理ディスク情報(「/dev/sda」)とを関連付けて、リソースマッピングテーブル13に格納する(S304)。すなわち、S304において、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131に管理ID211「00001」を設定し、リソースID132にホストOSのリソースに対して付与するID「1」を設定し、リソース種別133にディスク情報のリソース種別を表す「HDD」を設定し、ホストOS上の識別名135にはホストOSがマウントしている物理ディスク情報「/dev/sda」を設定し、対応物理リソース名134にホストOSがマウントしている物理ディスク情報「/dev/sda」を設定する。すなわち、ホストOSがディスク情報として認識できる論理ディスク情報は、物理ディスク情報で表されている。

- [0090] 図12は、実施の形態1に係るゲスト仮想マシンのネットワークインターフェース情報に関するリソースマッピング情報生成処理を示すフロー図である。図12を用いて、ゲストOSとゲストOSが利用している物理ネットワークインターフェース情報のマッピング方法について説明する。
- [0091] <ゲストOSのネットワークインターフェース情報のリソースマッピング情報生成処理>
前提処理として、リソースマッピング情報生成部1211が、仮想マシン管理テーブル21に、管理ID211、ハードウェア識別ID212、ドメインID213、ドメイン名称214とを関連付けて登録する処理(S401～S404)は、図8のS201～S204の処理と同様であるので説明を省略する。
- [0092] リソースマッピング情報生成部1211は、S404で登録した仮想マシン管理テーブル21からひとつのゲストOSの管理ID211を取得する。リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、CPUによりOSのネットワーク管理ツール(ifconfigコマンドなど)(仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドの一例)を用いて、S404で取得した管理IDにより示される仮想マシン(ドメイン)を識別するドメインIDに関する仮想ネットワークインターフェース名のリストを取得する(S405)。ifconfigコマンドなどにより管理されるファイルは、リソース種別毎管理情報を有するリソース種別毎管理ファイルの一例である。例えば、

リソースマッピング情報生成部1211は、S404で取得した管理ID211「00002」から、ドメインID「1」の「ゲストOSA」に関する仮想ネットワークインターフェース名のリスト「vif1. 0」を取得する。これが、ゲストOSAで認識される仮想ネットワークインターフェース名(論理リソース)である。

- [0093] リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、OSのネットワーク管理ツール(brctlコマンドなど)(仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドの一例)を用いて、S405で取得した仮想ネットワークインターフェース名が接続されているブリッジインターフェースを取得する(S406)。例えば、リソースマッピング情報生成部1211は、OSのネットワーク管理ツール(brctlコマンドなど)を用いて、仮想ネットワークインターフェース名「vif1. 0」が接続されているブリッジインターフェースを取得する。
- [0094] リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、OSのネットワーク管理ツールを用いて、S406で取得したブリッジインターフェースに接続されている物理ネットワークインターフェース名を取得する(S407)。例えば、リソースマッピング情報生成部1211は、S406で取得した「vif1. 0」が接続されているブリッジインターフェースに接続されている物理ネットワークインターフェース名「peth0」を取得することができる。
- [0095] S408において、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS404で取得した管理ID211を設定し、リソースID132に対象のゲスト仮想マシン140のリソースに対して例えばシェンシャルに付与するIDを取得して設定し、リソース種別133にネットワークインターフェース情報のリソース種別を表す「N／W, I／F」を設定し、ホストOS上の識別名135にはS405で取得した対象のゲストOSが使用している仮想ネットワークインターフェース名(論理リソース)を設定し、対応物理リソース名134にS407で取得した物理ネットワークインターフェース名を設定する。リソースID132は、同一管理IDで管理されるリソースを一意に識別できる任意に付加されるIDであるものとする。このように、リソースマッピング情報生成部1211は、S404で取得した仮想マシン管理テーブル21の管理ID211と関連付けて、リソースマッピングテーブル13を登録する(S408)。例えば

、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS404で取得した管理ID211「00002」を設定し、リソースID132にゲストOSAのリソースに対して付与するID「2」(「1」はディスク情報のリソース)を設定し、リソース種別133にネットワークインターフェース情報のリソース種別を表す「N／W. I／F」を設定し、ホストOS上の識別名135にはS405で取得した対象のゲストOSが使用している仮想ネットワークインターフェース名「vif1. 0」を設定し、対応物理リソース名134にS407で取得した物理ネットワークインターフェース名「peth0」を設定する。

- [0096] S409にて、リソースマッピング情報生成部1211は、自己の動作するサーバ装置100上のすべてのゲスト仮想マシン140についてネットワークインターフェース情報のリソースマッピング情報生成処理が完了するまで、上記処理(S405～408)を繰り返す。
- [0097] 図13は、実施の形態1に係るホスト仮想マシン120のホストOSのネットワークインターフェース情報(論理ネットワークインターフェース情報)とホストOSが利用している物理的なネットワークインターフェース情報(物理ネットワークインターフェース情報)とのリソースマッピング情報生成処理を示すフロー図である。図13を用いて、ホスト仮想マシン120のホストOSとホストOSが利用している物理的なネットワークインターフェース情報(物理ネットワークインターフェース情報)をマッピングする方法について説明する。
- [0098] <ホストOSのネットワークインターフェース情報と物理ネットワークインターフェース情報とのリソースマッピング情報生成処理>
- 前提処理として、リソースマッピング情報生成部1211が、仮想マシン管理テーブル21に、管理ID211、ハードウェア識別ID212、ドメインID213、ドメイン名称214とを関連付けて登録する処理(S501～S502)は、図11のS301～S302の処理と同様であるので説明を省略する。
- [0099] リソースマッピング情報生成部1211は、S502で登録した仮想マシン管理テーブル21からホストOSの管理ID211を取得する。
- [0100] リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、CPUによりOSのネットワーク管理ツール(ifconfigコマンドなど)(仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドの一例)を用いて、取得した管理IDにより示されるホスト仮想マシン(ホストドメイン)を識別するドメインIDに関

する仮想ネットワークインターフェース名のリストを取得する(S503)。ifconfigコマンドなどにより管理されるファイルは、リソース種別毎管理情報を有するリソース種別毎管理ファイルの一例である。例えば、リソースマッピング情報生成部1211は、S502で取得した管理ID211「00001」から、ドメインID「0」の「ホストOS」に関する仮想ネットワークインターフェース名のリスト「vif0. 0」を取得する。これが、ホストOSで認識される仮想ネットワークインターフェース名(論理リソース)である。

- [0101] リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、OSのネットワーク管理ツール(brctlコマンドなど)(仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドの一例)を用いて、S503で取得した仮想ネットワークインターフェース名が接続されているブリッジインターフェースを取得する(S504)。例えば、リソースマッピング情報生成部1211は、OSのネットワーク管理ツール(brctlコマンドなど)を用いて、仮想ネットワークインターフェース名「vif0. 0」が接続されているブリッジインターフェースを取得する。
- [0102] リソースマッピング情報生成部1211は、ホスト仮想マシン120のホストOS上で、OSのネットワーク管理ツールを用いて、S504で取得したブリッジインターフェースに接続されている物理ネットワークインターフェース名を取得する(S505)。例えば、リソースマッピング情報生成部1211は、S504で取得した「vif0. 0」が接続されているブリッジインターフェースに接続されている物理ネットワークインターフェース名「peth0」を取得することができる。
- [0103] S506において、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS502で取得した管理ID211を設定し、リソースID132にホスト仮想マシン120のリソースに対して例えばシーケンシャルに付与するIDを取得して設定し、リソース種別133にネットワークインターフェース情報のリソース種別を表す「N／W. I／F」を設定し、ホストOS上の識別名135にはS503で取得したホストOSが使用している仮想ネットワークインターフェース名(論理リソース)を設定し、対応物理リソース名134にS505で取得した物理ネットワークインターフェース名を設定する。このように、リソースマッピング情報生成部1211は、S502で取得した仮想マシン管理テーブル21の管理ID211と関連付けて、リソースマッピングテーブル13を登録する(S506)。

例えば、リソースマッピング情報生成部1211は、リソースマッピングテーブル13の管理ID131にS502で取得した管理ID211「00001」を設定し、リソースID132にホストOSのリソースに対して付与するID「4」(図10において「1」～「3」はディスク情報のリソース)を設定し、リソース種別133にネットワークインターフェース情報のリソース種別を表す「N／W. I／F」を設定し、ホストOS上の識別名135にはS503で取得した対象のゲストOSが使用している仮想ネットワークインターフェース名「vif0. 0」を設定し、対応物理リソース名134にS505で取得した物理ネットワークインターフェース名「peth0」を設定する。

- [0104] 図14は、図12及び図13で説明したVMモニタ110上のそれぞれのホストOS／ゲストOSで認識されるネットワークインターフェースと仮想ネットワークインターフェース、ブリッジインターフェース、物理ネットワークインターフェースの接続関係図である。
- [0105] 本実施の形態に係るサーバ装置100において、上述したディスク情報及びネットワークインターフェース情報以外のリソース(例えば、CPUやメモリ、電源、ファン等)に関しては、すべてホストOSのリソース(論理リソース)としてリソースマッピングテーブル13にマッピングされているものとする。
- [0106] 次に、上述したリソースマッピング情報生成処理により生成されたリソースマッピング情報1221を用いて、図5のS108の異常判定ステップ(障害発生仮想マシン特定ステップ)について具体例を用いて説明する。
- [0107] 例えば、サーバ装置100のハードディスク「/dev/sda」に異常(障害)がある場合について説明する。異常判定部1213が、異常物理リソース検知ステップ(図5のS106～S107)において、異常判定閾値情報1222のID「E0003」の異常状態に基づいてサーバ装置100のハードディスク「/dev/sda」に異常(障害)があると判定する。異常判定部1213は、CPUを用いて、記憶装置に記憶されたリソースマッピング情報1221のリソースマッピングテーブル13の対応物理リソース名134を参照して、対応物理リソース名134が「/dev/sda」である管理ID131「00001」を取得する。異常判定部1213は、CPUを用いて、取得した管理ID131「00001」をもとに仮想マシン管理テーブル21を参照し、「00001」と等しい管理ID211「00001」を抽出する。このとき、仮想マシン管理テーブル21の管理ID211「00001」の情

報は、ハードウェア識別ID212「サーバ1(サーバ装置100)」、ドメインID「0」、ドメイン名称「ホストOS」である。したがって、異常判定部1213は、仮想マシン管理テーブル21より、サーバ装置100上の仮想マシン(ドメイン)(サーバ装置100上で動作するホストOS／ゲストOS)として「ホストOS」を抽出することができる。このように、異常判定部1213は、障害発生仮想マシンとしてホスト仮想マシン120を特定する。

[0108] 本実施の形態によれば、リソースマッピング情報生成部1211が、サーバ装置100条で稼働する各仮想マシン(各ドメイン)で使用する(認識される)リソースと物理リソースとを対応付けてリソースマッピング情報1221を生成することにより、エージェント実行部121がハードウェア障害を検知した場合において、検知した障害に関連するホスト仮想マシン120あるいはゲスト仮想マシン140(ホストOSまたはゲストOS)に対して適切な通知や停止動作を実行することができる。また、エージェント実行部121が、適切に検知した障害に関連するホスト仮想マシン120あるいはゲスト仮想マシン140(ホストOSまたはゲストOS)に対して適切な通知や停止動作を実行することにより、他系(待機系)側のサーバ2装置200上のクラスタソフトウェアがハートビート停止を検知し、適切な系切り替えが可能となる。

[0109] 実施の形態2.

実施の形態1では、エージェント実行部121の異常通知部1214は障害発生仮想マシンに対して「OSの停止」を通知する場合について説明した。実施の形態2では、エージェント実行部121の異常通知部1214が、ホスト仮想マシン120のホストOSやゲスト仮想マシン140a, 140bの各ゲストOS上のクラスタソフトウェア107, 109に、OSを停止させるための通知ではなく、例えば、異常を報告するための通知のみを実施する形態について説明する。

[0110] 仮想環境を有するサーバ装置等においては、処理負荷の集中によりハードディスクの読み込み時間が遅くなる等の、すぐにはシステム稼働上の障害は発生しないが警報として仮想マシンに通知したいという場合がある。すなわち、サーバ装置100の物理リソースの稼働状態が、OSを停止させるような「異常な物理リソース稼働状態」よりは「やや正常に近い物理リソース稼働状態」である場合である。このような場合には、エージェント実行部121は、直ちにOSを停止することはないが、「OSに対

する警報」を通知する。

- [0111] 本実施の形態における異常通知部1214の異常通知処理は、図6の異常判定閾値情報1222において、以下のように定義することで実施することができる。異常判定閾値情報1222の物理リソース稼働状態のID1111「E00007」において、異常判定閾値1113は、ディスクの読み込み応答時間が「10秒>応答時間READ>5秒」と設定されている。これは、物理リソース稼働状態のID1111「E00003」における異常判定閾値1113よりは、やや正常に近い値ではある。したがって、OSを停止させるほどではないが、OSに対して警報を通知するレベルであると設定する。すなわち、この場合（物理リソース稼働状態のID1111「E00007」）の、異常通知情報1114は、「OSへsyslogの通知」と設定する。例えば、障害発生仮想マシンがゲストOSの場合は、障害発生ゲストOSに対するsyslogの通知とともに、ホストOSに対してもsyslogの通知を行いたい場合がある。このような場合には、「ホストOSへsyslogの通知、障害発生仮想マシンのOSへsyslogの通知」というように具体的な通知先を指定して通知を行うように異常通知情報1114に設定することができる。

- [0112] これにより、物理リソース稼働状態がID1111「E00007」の場合は、異常通知部1214は、障害発生仮想マシンのOSもしくはクラスタソフトウェアに直接、もしくは、OSのログ管理システム(syslogやイベントログなど)の手段を用いて警報を通知することができる。警報の通知が行われた後のホストOSやゲストOSの動作に関しては、クラスタソフトウェアの定義された内容に基づいて行うことが可能となる。

- [0113] 本実施の形態によれば、障害の内容に応じてホストOSの停止や通知を実施するなどの処理を定義可能とすることで、既存のクラスタソフトウェアがエージェントから通知内容に基づき、クラスタソフトウェアの設定内容に基づいた系の制御動作を実施できる状況を作り出すことが可能となる。

- [0114] 実施の形態3.

実施の形態1では、リソースマッピング情報1221は、エージェント実行部121が自動で生成する手段について説明した。実施の形態3では、リソースマッピング情報を手動で定義する方法について説明する。

- [0115] 実施の形態1では、リソースマッピング情報生成部1211が自動的に、ホスト仮想マシン120及びゲスト仮想マシン140a, 140b(ホストOS／ゲストOS)が認識するディスク情報及びネットワークインターフェース情報と物理的なディスク情報及びネットワークインターフェース情報とのリソースマッピング情報生成方法について説明した。
- [0116] しかし、サーバ装置100のVMモニタ110(VM環境)では、ゲスト仮想マシン(ゲストOS)に対して、メモリやCPUの利用率でリソースを割り当てることがある。この場合は、ゲストOSの利用する論理リソースがサーバ筐体の中のどのスロット番号のメモリであるか、また、どのCPUのコアを割り当てられているか等を自動で判定することはできない。このため、自動的には明確にマッピングすることができない場合もある。
- [0117] 上記のような事態に対処するため、ユーザ(管理者あるいは設計者等)が、手動でリソースマッピング情報を定義する方法を提供する。ユーザによる手動でのリソースマッピング情報生成方法は、例えば、以下に示す方法で実施する。ユーザは、図9に示す仮想マシン管理テーブル21と図10に示すリソースマッピングテーブル13の内容を、予めCSV(Comma・Separated・Values)ファイル等で作成し、記憶装置に保存する。エージェント実行部121は、エージェント実行部121の起動時に、CPUを用いて、記憶装置に保存された仮想マシン管理テーブル21とリソースマッピングテーブル13との内容を作成したCSVファイル等を読み込み、仮想マシン管理テーブル21とリソースマッピングテーブル13にインポートして、リソースマッピング情報1221として記憶装置に記憶する。以上により、手動でリソースマッピング情報1221が生成され記憶装置に記憶される。以降の処理は、実施の形態1と同様である。
- [0118] 実施の形態1～3において、以下の特徴を有するサーバ装置100について説明した。
- [0119] 実施の形態1～3に係る仮想環境における冗長化手法、及び、それを用いたシステムは、仮想環境においてハードウェアの障害を検知するエージェントを備えたシステムにおいて、
- 前記エージェントは、
定期的に各ドメイン(ホストOSやゲストOS)の論理的なリソースと物理リソースをマッピングするリソースマッピング手段を備え、

ホストOS上でハードウェア稼働状態を監視し、筐体情報やCPU、メモリ、ハードディスク、ネットワークインターフェースカードのハードウェア情報を収集する異常監視手段を備え、

異常監視手段が収集したハードウェア稼働情報を、予め定義した異常判定閾値情報と前記リソースマッピング手段でマッピングしたリソースマッピング情報を基に、どのドメインに関連するハードウェア障害であるかを判定する異常判定手段を備え、

異常判定手段によって判定されたハードウェア障害情報に応じて、ホストOSやゲストOSへログ通知、もしくは、ホストOSの停止、または、ゲストOSの停止を行う異常通知手段を備えることにより、

検出したハードウェア障害に関連する該当ドメインへの障害通知、または、ドメインの停止を行うことを特徴とする。

- [0120] 前記エージェントが検出した障害内容に応じて、ホストOSやゲストOS停止させることで、他系の各ゲストOSに配置された既製クラスタソフトウェアが系切替を実施できる状況を作り出すことが可能であることを特徴とする。
- [0121] 前記エージェントにおける異常判定手段において、異常判定閾値情報に収集したハードウェア情報が障害であるか否かを識別するための閾値情報と、障害として判定された場合の通知内容やドメイン停止動作を定義可能であることを特徴とする。
- [0122] 前記エージェントのリソースマッピング手段において、定期的なリソースの自動マッピングだけでなく、手動でリソースマッピング情報を定義可能であることを特徴とする。
- [0123] 以上、実施の形態1～3について説明したが、これらのうち、2つ以上の実施の形態を組み合わせて実施しても構わない。あるいは、これらのうち、1つの実施の形態を部分的に実施しても構わない。あるいは、これらのうち、2つ以上の実施の形態を部分的に組み合わせて実施しても構わない。
- [0124] また、実施の形態1～3の説明において、エージェント実行部121におけるリソースマッピング情報生成部1211と、リソースマッピング情報記憶部と、異常監視部1212と、異常判定部1213と、異常通知部1214とはそれぞれ独立した機能ブロックとして構成されているが、ひとつの機能ブロックとしてもよい。また、リソースマッピング情報生成部1211とリソースマッピング情報記憶部とはひとつの機能ブロックとしても構わ

ない。あるいは、全ての機能ブロックを独立した機能ブロックとしても良い。あるいは、これらの機能ブロックを、他のどのような組み合わせで構成しても構わない。

[0125] 実施の形態1に係るサーバ装置及びサーバ装置の異常検知方法は、CPU(処理装置)、記憶装置等のハードウェアが用いられており、ソフトウェアによる情報処理がハードウェアを用いて具体的に実現されたものである。すなわち、上述した実施の形態1～3のサーバ装置及びサーバ装置の異常検知方法は、自然法則を利用したハードウェアの動作によりサーバ装置及びサーバ装置の異常検知方法の実現を図っているものであり、自然法則を利用した技術的創作に該当するものである。

図面の簡単な説明

[0126] [図1]実施の形態1に係るサーバ装置100及びサーバ2装置の外観の一例を示す図である。

[図2]サーバ装置100及びサーバ2装置のハードウェア資源の一例を示す図である。

[図3]実施の形態1に係る冗長化システム800のシステム構成図である。

[図4]実施の形態1に係るサーバ装置100の備えるエージェント実行部121の機能ブロックの構成を示すブロック構成図である。

[図5]実施の形態1におけるサーバ装置100の異常検知方法の処理動作を示すフロー図である。

[図6]異常判定閾値情報1222のテーブル構成を示す図である。

[図7]実施の形態1に係る冗長化システム800の系切替時の動作を示す図である。

[図8]エージェント実行部121が動作するホスト仮想マシンのホストOSにより認識できるディスク情報(ここでは、ゲスト仮想マシンのディスク情報)とゲスト仮想マシンが実際に利用する物理的なディスク情報とのリソースマッピング情報生成処理を示すフロー図である。

[図9]リソースマッピング情報の仮想マシン管理テーブルのテーブル構成を示す図である。

[図10]リソースマッピング情報のリソースマッピングテーブルの構成を示す図である。

[図11]実施の形態1に係るホスト仮想マシン120のホストOSのディスク情報(論理ディスク情報)とホストOSが利用している物理的なディスク情報(物理ディスク情報)とのリ

ソースマッピング情報生成処理を示すフロー図である。

[図12]実施の形態1に係るゲスト仮想マシンのネットワークインターフェース情報に関するリソースマッピング情報生成処理を示すフロー図である。

[図13]実施の形態1に係るホスト仮想マシン120のホストOSのネットワークインターフェース情報(論理ネットワークインターフェース情報)とホストOSが利用している物理的なネットワークインターフェース情報(物理ネットワークインターフェース情報)とのリソースマッピング情報生成処理を示すフロー図である。

[図14]図12及び図13で説明したVMモニタ110上のそれぞれのホストOS／ゲストOSで認識されるネットワークインターフェースと仮想ネットワークインターフェース、ブリッジインターフェース、物理ネットワークインターフェースの接続関係図である。

符号の説明

- [0127] 13 リソースマッピングテーブル、21 仮想マシン管理テーブル、100 サーバ装置、101 LAN、107, 109 クラスタソフトウェア、110 VMモニタ、115, 117 クラスタソフトウェア、120 ホスト仮想マシン、121 エージェント実行部、131 管理ID、132 リソースID、133 リソース種別、134 対応物理リソース名、135 ホストOS上の識別名、140 ゲスト仮想マシン、140a ゲスト仮想マシンA、140b ゲスト仮想マシンB、200 サーバ2装置、210 VMモニタ、211 管理ID、212 ハードウェア識別ID、213 ドメインID、214 ドメイン名称、220 ホスト仮想マシン'、240a ゲスト仮想マシンA'、240b ゲスト仮想マシンB'、221 エージェント実行部、800 元長化システム、901 表示装置、902 キーボード、903 マウス、904 FDD、905 CDD、906 プリンタ装置、907 スキナ装置、910 システムユニット、911 CPU、912 バス、913 ROM、914 RAM、915 通信ボード、920 磁気ディスク装置、921 OS、922 ウィンドウシステム、923 プログラム群、924 ファイル群、931 電話器、932 ファクシミリ機、942 LAN、940 インターネット、941 ゲートウェイ、111 ID、1112 対象ハードウェア、1113 異常判定閾値、1114 異常通知情報、1211 リソースマッピング情報生成部、1212 異常監視部、1213 異常判定部、1214 異常通知部、1221 リソースマッピング情報、1222 異常判定閾値情報、1223 障害情報データベース、1224 物理リソース稼働情報、9200 VMモニタ。

請求の範囲

- [1] 物理リソースを用いて複数の仮想計算機を稼働させるサーバ装置であつて、上記物理リソースのうち上記複数の仮想計算機各々が使用する物理リソースを論理リソースとして複数の仮想計算機を稼働させるサーバ装置において、
物理リソースの異常を検知するエージェント実行部を備え、
上記エージェント実行部は、
上記論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピング情報を生成するリソースマッピング情報生成部と、
上記リソースマッピング情報生成部が生成した上記リソースマッピング情報を記憶装置に記憶するリソースマッピング記憶部と、
物理リソースの稼働状態を示す物理リソース稼働情報を収集し記憶装置に記憶する異常監視部と、
上記異常監視部が収集した上記物理リソース稼働情報のなかに異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定し、異常な稼働状態の物理リソースの情報がある場合には、上記異常な稼働状態の物理リソースの情報と上記リソースマッピング情報に基づいて上記異常な稼働状態の物理リソースに対応する論理リソースを使用している仮想計算機を処理装置により特定する異常判定部と、
上記異常判定部により特定された仮想計算機に対して、上記異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知部と
を備えることを特徴とするサーバ装置。
- [2] 上記リソースマッピング情報生成部は、
上記リソースマッピング情報を定期的に生成することを特徴とする請求項1に記載のサーバ装置。
- [3] 上記サーバ装置は、上記複数の仮想計算機各々ごとに、当該仮想計算機により使用される論理リソースに対して物理リソースを対応させるための情報である仮想計算機毎リソース管理情報を有する仮想計算機毎リソース管理ファイルを備え、
上記リソースマッピング情報生成部は、
上記仮想計算機毎リソース管理ファイルから物理リソースを含む上記仮想計算機毎

リソース管理情報を取得して、取得した上記仮想計算機毎リソース管理情報に基づいて、上記複数の仮想計算機各々により使用される論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピングテーブルをリソースマッピング情報として生成することを特徴とする請求項2に記載のサーバ装置。

- [4] 上記サーバ装置は、リソースの種別ごとに、当該種別の論理リソースに対して当該種別の物理リソースを対応させるための情報であるリソース種別毎管理情報を有するリソース種別毎管理ファイルを備え、
上記リソースマッピング情報生成部は、
上記複数の仮想計算機各々により使用される論理リソースの種別に対応するリソース種別毎管理ファイルから上記リソース種別毎管理情報を取得して、取得した上記リソース種別毎管理情報に基づいて、上記複数の仮想計算機各々により使用される論理リソースに対して上記サーバ装置の物理リソースを対応付けてリソースマッピング情報を生成することを特徴とする請求項3に記載のサーバ装置。
- [5] 上記エージェント実行部は、
仮想計算機のOS(Operating・System)のもとで動作するエージェントプログラムを実行し、
上記リソースマッピング情報生成部は、仮想計算機のOSに含まれたツール又はエージェントプログラムに含まれたコマンドを使用して、論理リソースが使用している物理リソースを調べる
ことを特徴とする請求項1に記載のサーバ装置。
- [6] 上記エージェント実行部は、さらに、
物理リソースの稼働状態が異常であるか否かを判定するために用いる閾値と、当該閾値により物理リソースの稼働状態が異常であると判定された場合に、稼働状態が異常であると判定された物理リソースに対応する論理リソースを使用している仮想計算機へ通知する異常通知情報とが設定された異常判定閾値情報を予め記憶装置に記憶する異常判定閾値情報記憶部を備え、
上記異常通知部は、

上記異常判定閾値情報に設定された上記異常通知情報に基づいて通知を行うことを特徴とする請求項1に記載のサーバ装置。

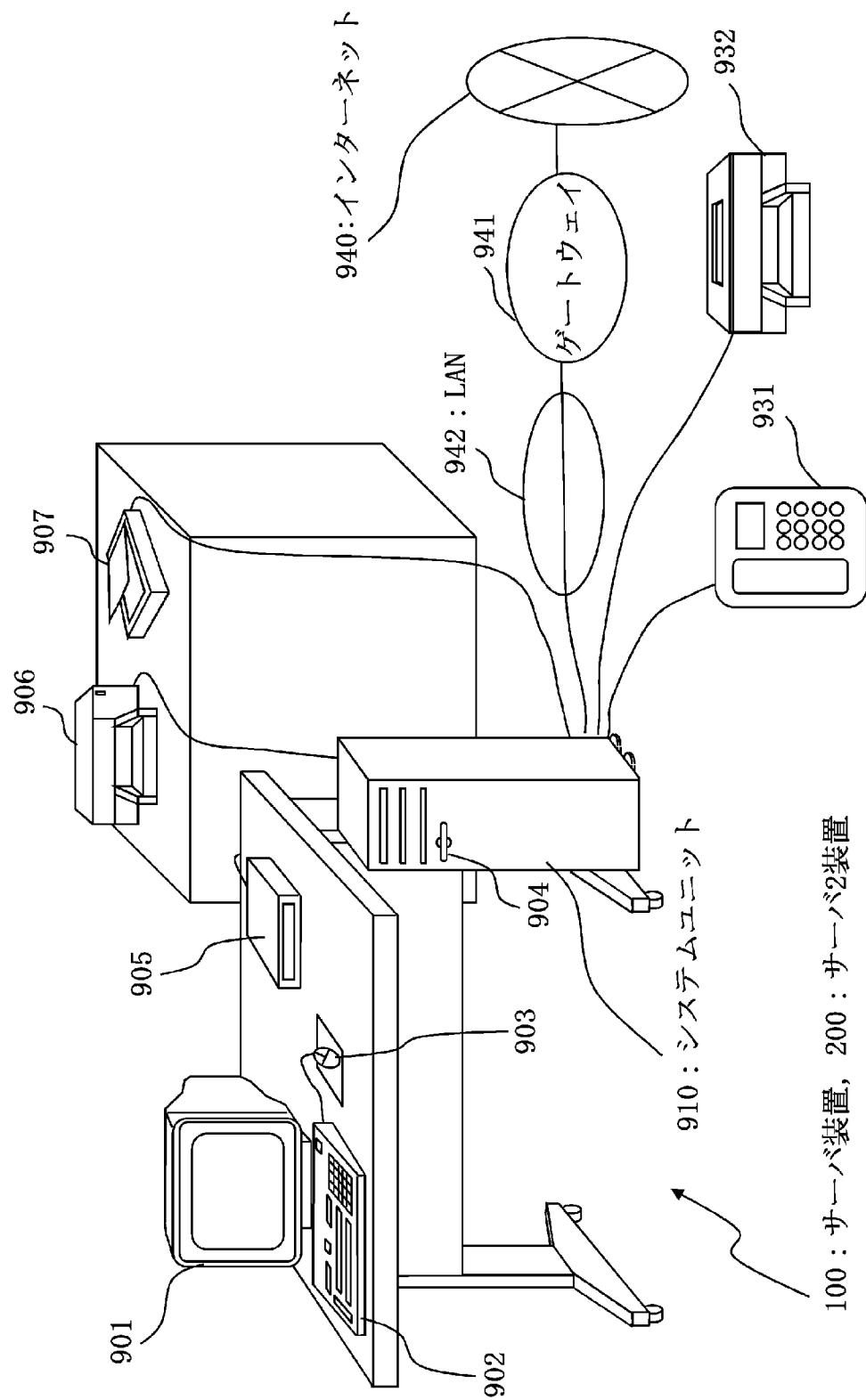
- [7] 上記複数の仮想計算機のなかで、1台の仮想計算機のみが上記エージェント実行部を備えたことを特徴とする請求項1に記載のサーバ装置。
- [8] 上記リソースマッピング情報生成部は、
上記論理リソースに対して上記サーバ装置の物理リソースを対応付けて予め作成され記憶装置に記憶されたリソースマッピングファイルを処理装置により取得し、取得したリソースマッピングファイルをリソースマッピング情報とすることを特徴とする請求項1に記載のサーバ装置。
- [9] 物理リソースを用いて複数の仮想計算機を稼働させるサーバ装置であって、上記物理リソースのうち上記複数の仮想計算機各々が使用する物理リソースを論理リソースとして複数の仮想計算機を稼働させるサーバ装置の異常検知方法において、
エージェント実行部により物理リソースの異常を検知するエージェント実行ステップを備え、
上記エージェント実行ステップは、
リソースマッピング情報生成部が、上記論理リソースに対して上記サーバ装置の物理リソースを対応付けたリソースマッピング情報を生成するリソースマッピング情報生成ステップと、
リソースマッピング記憶部が、上記リソースマッピング情報生成ステップにより生成された上記リソースマッピング情報を記憶装置に記憶するリソースマッピング記憶ステップと、
異常監視部が、物理リソースの稼働状態を示す物理リソース稼働情報を収集し記憶装置に記憶する異常監視ステップと、
異常判定部が、上記異常監視ステップにより収集された上記物理リソース稼働情報のなかに異常な稼働状態の物理リソースの情報があるか否かを処理装置により判定し、異常な稼働状態の物理リソースの情報がある場合には、上記異常な稼働状態の物理リソースの情報と上記リソースマッピング情報とにに基づいて上記異常な稼働状態の物理リソースに対応する論理リソースを使用している仮想計算機を処理装置により

特定する異常判定ステップと、

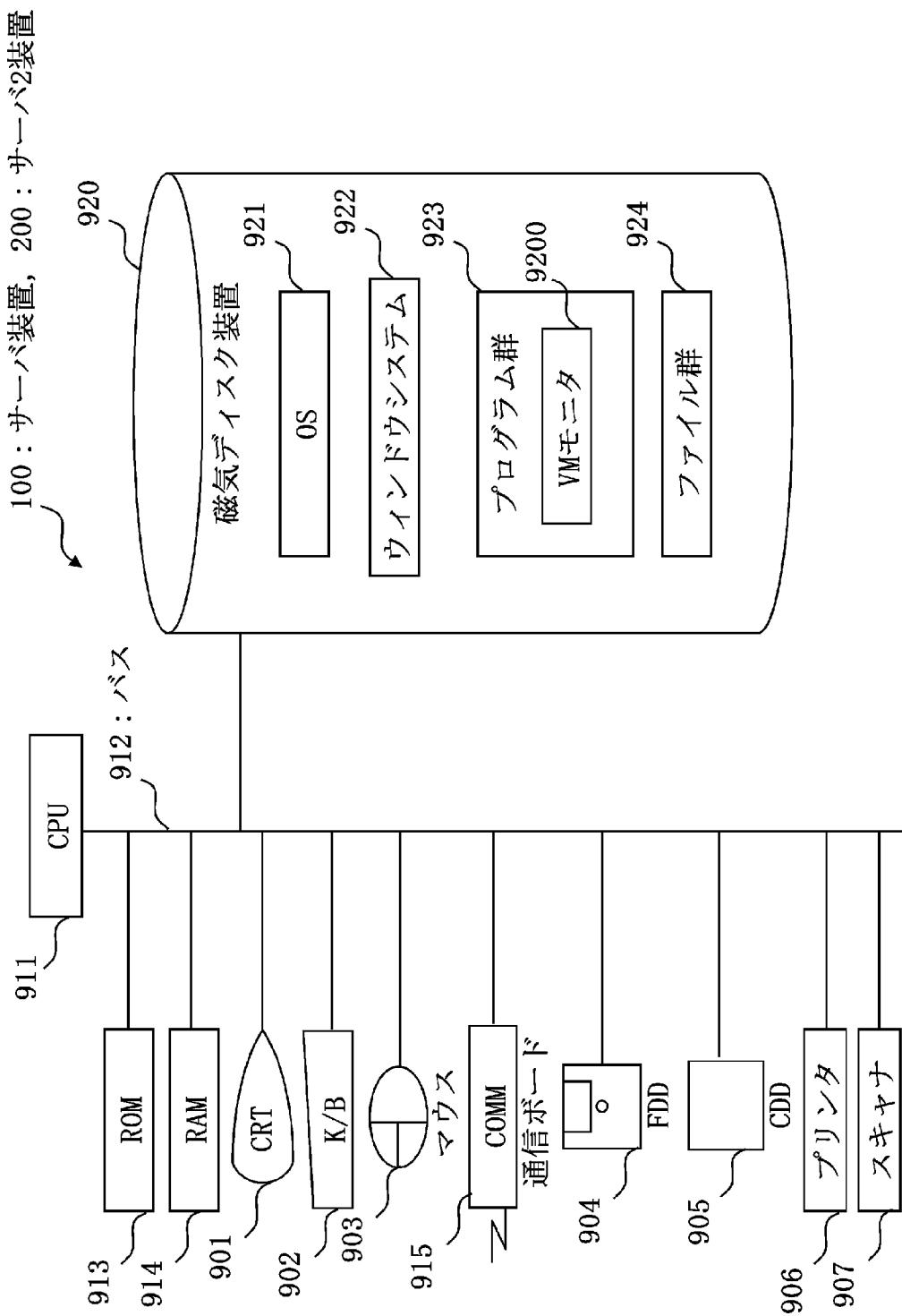
異常通知部が、上記異常判定ステップにより特定された仮想計算機に対して、上記異常な稼働状態の物理リソースの情報に応じた通知を行う異常通知ステップとを備えることを特徴とするサーバ装置の異常検知方法。

- [10] 請求項9記載のサーバ装置の異常検知方法をコンピュータに実行させることを特徴とするサーバ装置の異常検知プログラム。

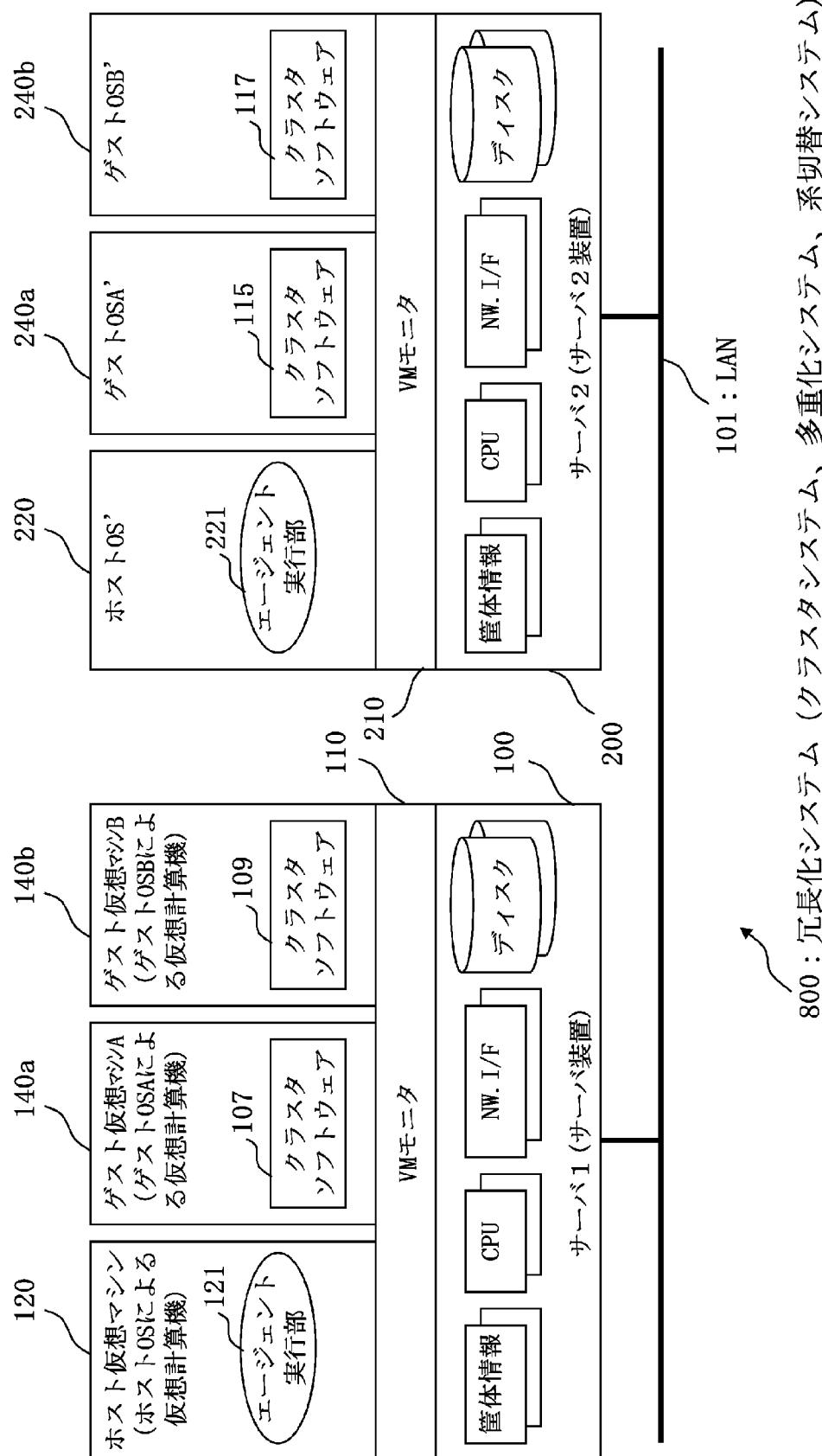
[図1]



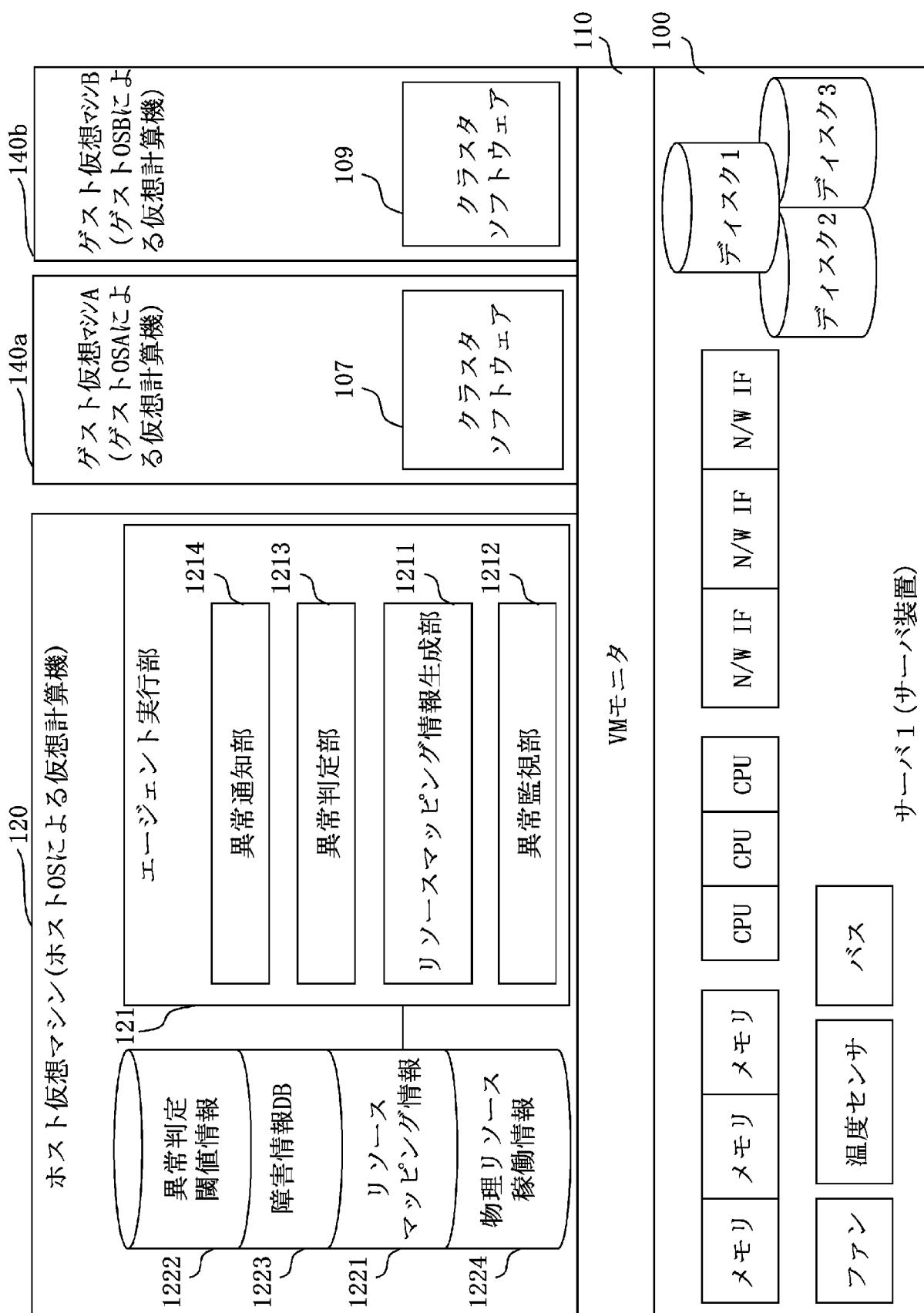
[図2]



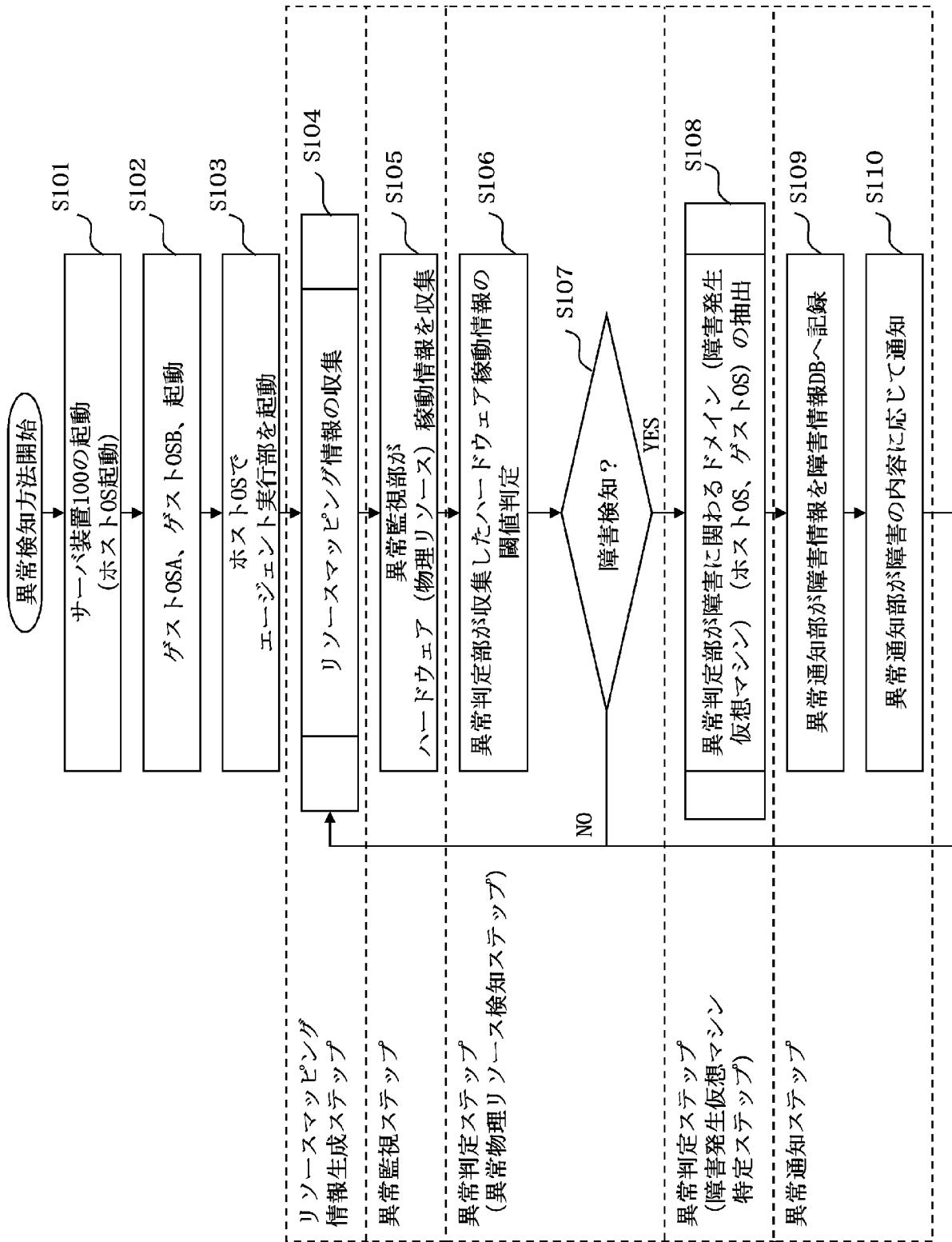
[図3]



[図4]



[図5]

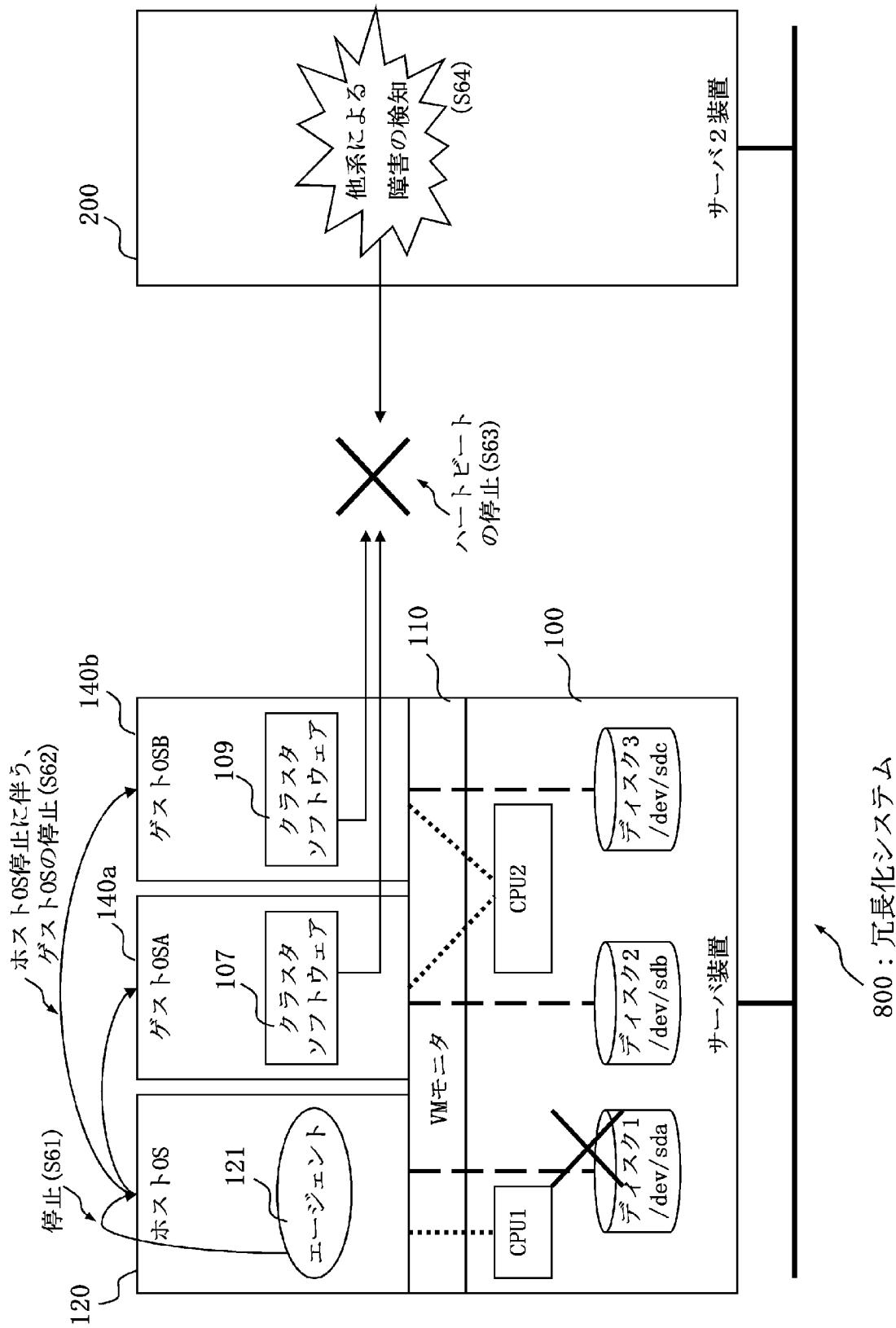


[図6]

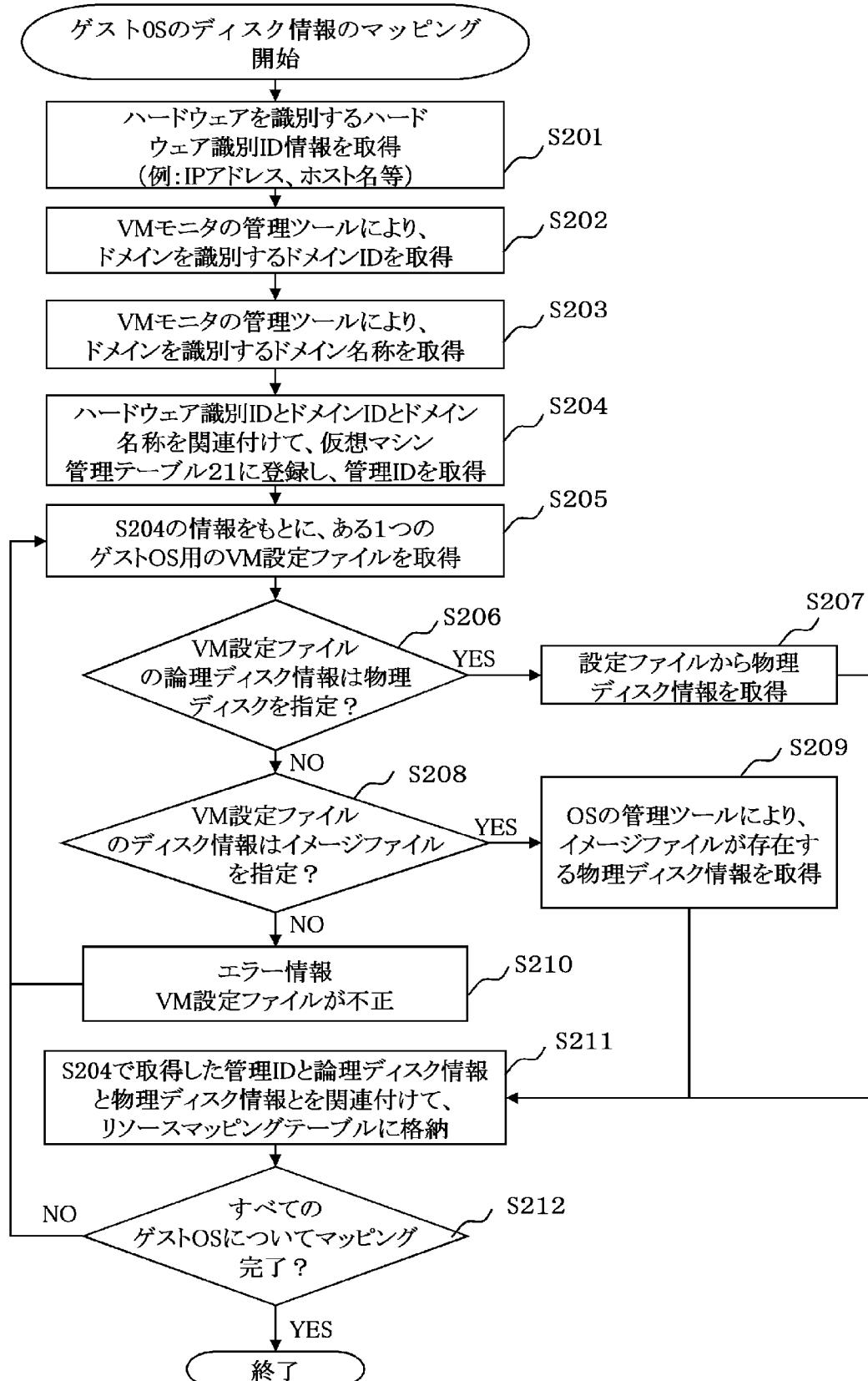
1222：異常判定閾値情報

ID	対象	異常判定閾値	異常通知情報 (異常検知時の通知内容)
E00001	CPU-1	温度>60度	障害発生仮想マシンのOSの停止
E00002	CPU-2	温度>60度	障害発生仮想マシンのOSの停止
E00003	/dev/sda	応答時間READ>10秒	障害発生仮想マシンのOSの停止
E00004	/dev/sda	応答時間WRITE>10秒	障害発生仮想マシンのOSの停止
E00005	/dev/sdb	停止	障害発生仮想マシンのOSの停止
E00006	/dev/sdc	停止	障害発生仮想マシンのOSの停止
E00007	/dev/sdb	10秒>応答時間READ>5秒	ホストOSへsyslog通知 障害発生仮想マシンのOSへ syslog通知
E00008	/dev/sdb	応答時間READ>10秒	障害発生仮想マシンのOSの停止
• • •	• • •	• • •	• • •

[义7]



[図8]



[図9]

21：仮想マシン管理テーブル（リソースマッチピング情報）

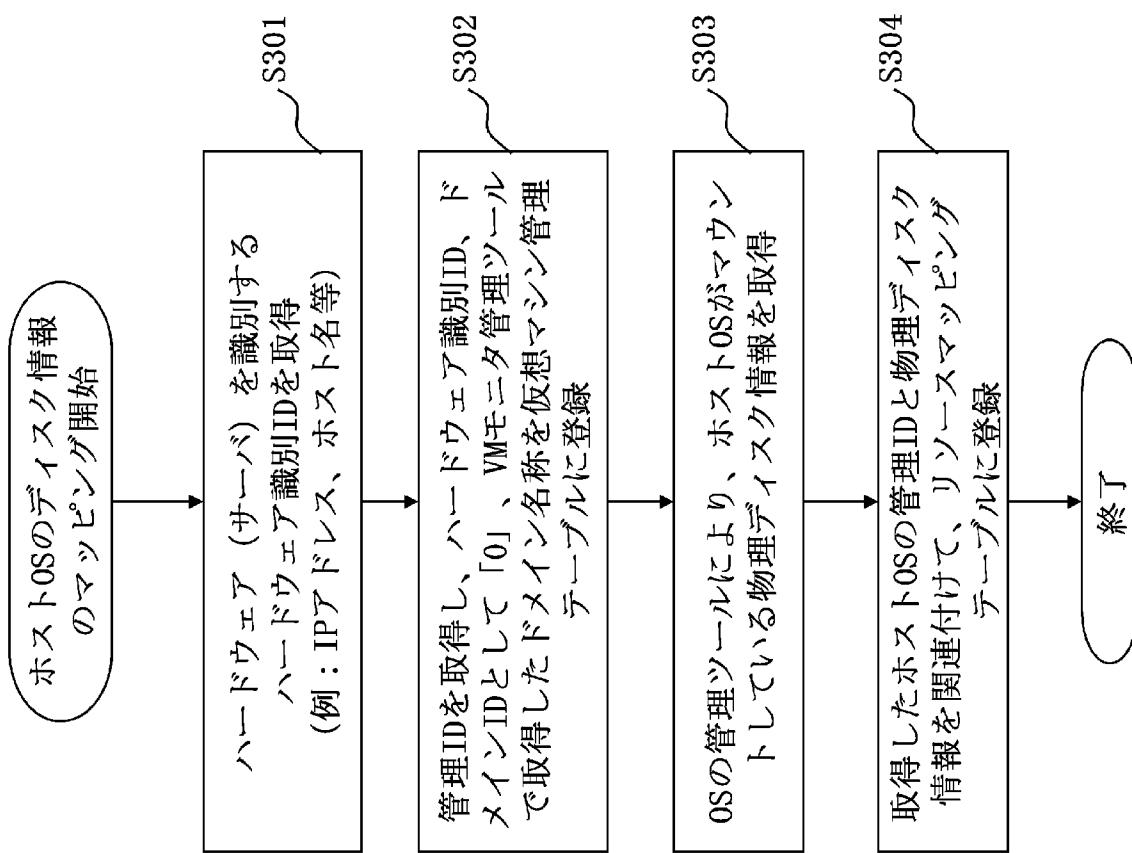
211 管理ID	212 ハードウェア識別ID	213 ドメインID	214 ドメイン名称
00001 サーバ1	0	ホストOS	
00002 サーバ1	1	ゲストOSA	
00003 サーバ1	2	ゲストOSB	
00004 サーバ2	0	ホストOS'	
00005 サーバ2	1	ゲストOSA'	
00006 サーバ2	2	ゲストOSB'	
...

[図10]

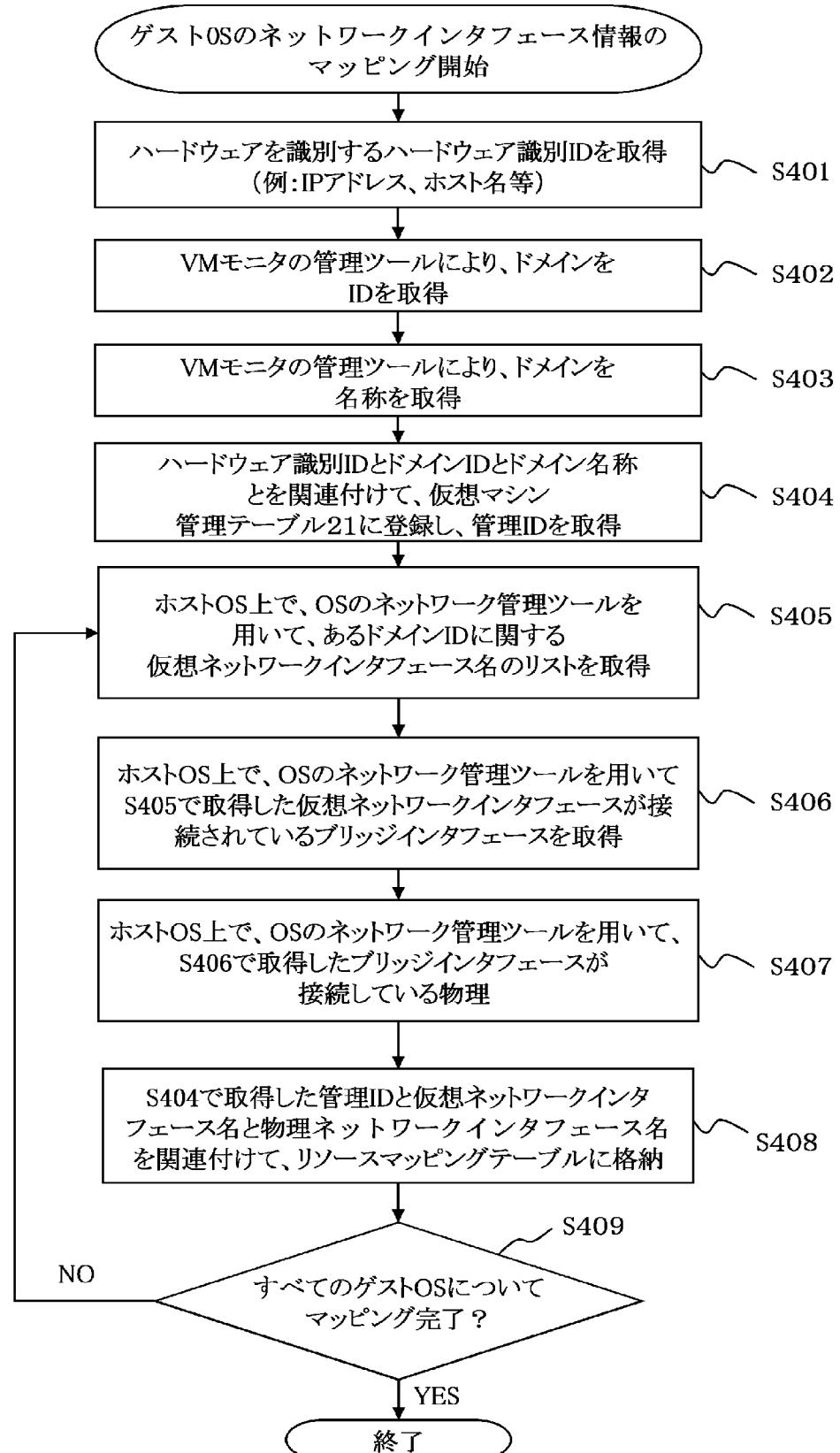
13：リソースマッチングテーブル（リソースマッピング情報）

管理ID	リソースID	リソース種別	対応物理リソース名 (論理リソース)	ホストOS上の識別名 (論理リソース)
00001	1	HDD	/dev/sda	/dev/sda
00001	2	HDD	/dev/sdb	/dev/sdb
00001	3	HDD	/dev/sdc	/dev/sdc
00001	4	N/W. I/F	peth0	vif0.0
00002	1	HDD	/dev/sdb	/dev/sdb/hdd.img
00002	2	N/W. I/F	peth0	vif1.0
00003	1	HDD	/dev/sdc	/dev/sdc
00003	2	N/W. I/F	peth0	vif2.0
⋮	⋮	⋮	⋮	⋮

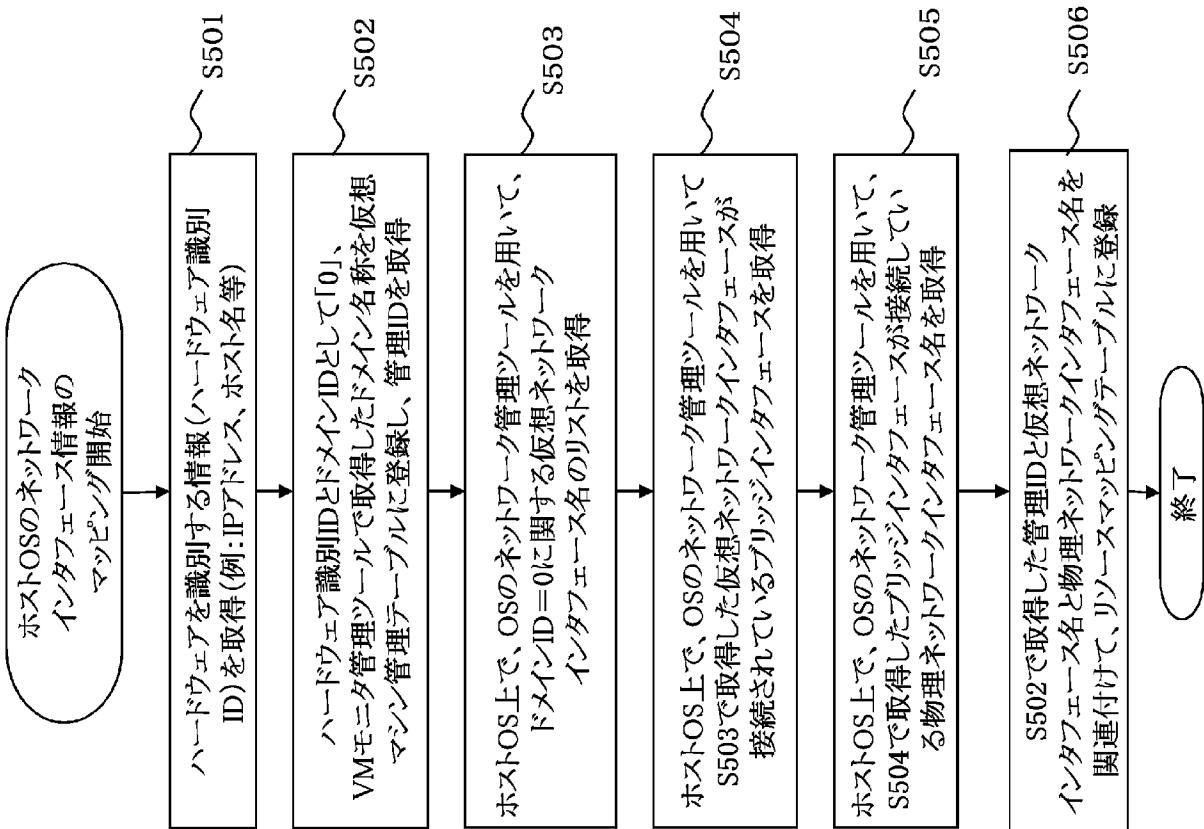
[図11]



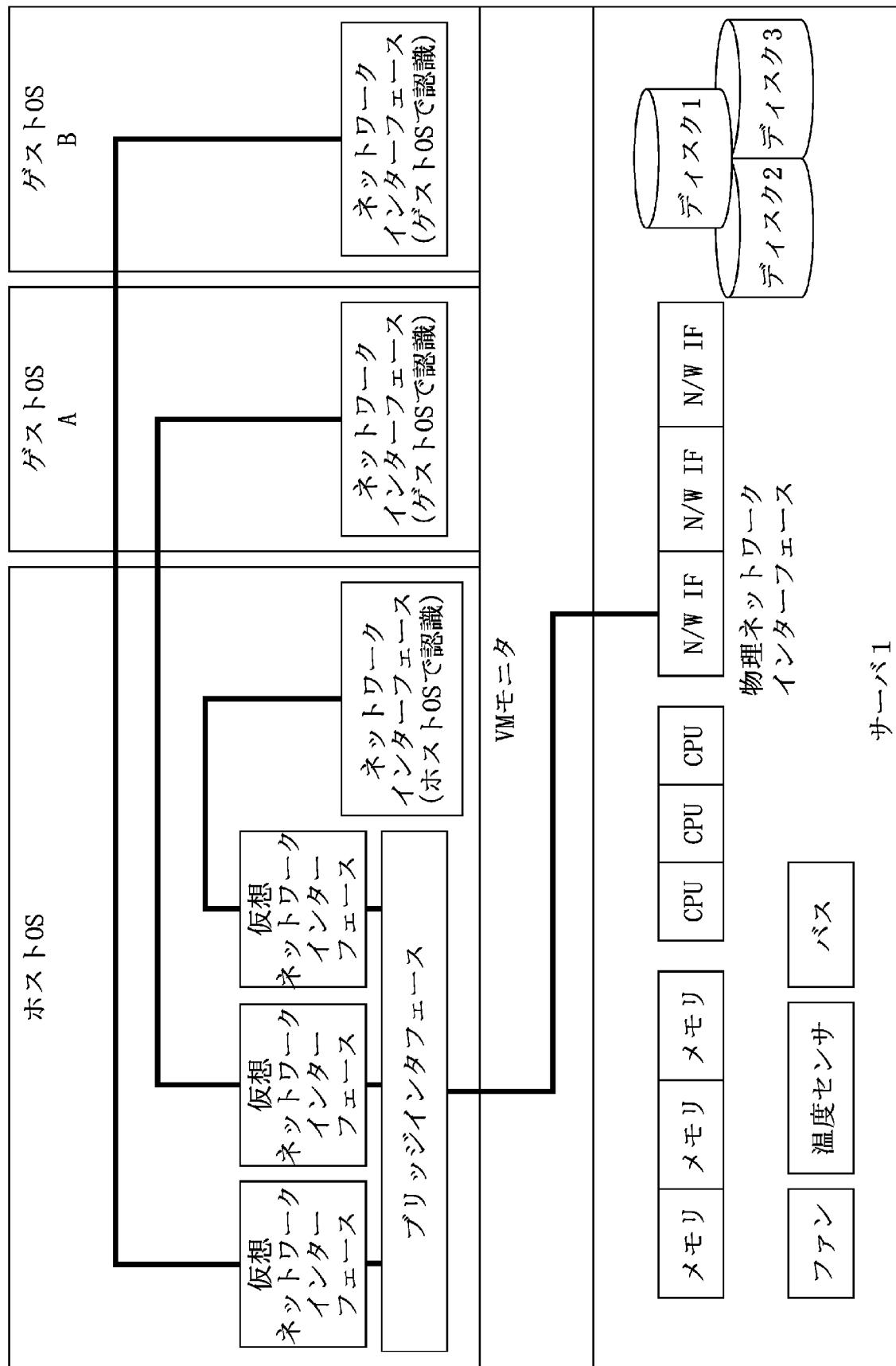
[図12]



[図13]



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/060739

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/30 (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)
G06F11/30

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 1996-2008
 Kokai Jitsuyo Shinan Koho 1971-2008 Toroku Jitsuyo Shinan Koho 1994-2008

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.
Y	JP 2007-233687 A (NEC Corp.), 13 September, 2007 (13.09.07), Par. Nos. [0042] to [0082] (Family: none)	1-10
Y	JP 2002-229806 A (Hitachi, Ltd.), 16 August, 2002 (16.08.02), Par. Nos. [0073] to [0081] & US 2002/0108074 A1	1-10
Y	JP 2003-108420 A (Hitachi, Ltd.), 11 April, 2003 (11.04.03), Par. Nos. [0037] to [0055] & US 2003/0061331 A1	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"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	"&" document member of the same patent family

Date of the actual completion of the international search
11 July, 2008 (11.07.08)

Date of mailing of the international search report
22 July, 2008 (22.07.08)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/060739

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2005-234861 A (Mitsubishi Electric Corp.), 02 September, 2005 (02.09.05), Full text; all drawings (Family: none)	1-10
A	JP 2004-295811 A (Hitachi, Ltd.), 21 October, 2004 (21.10.04), Full text; all drawings & US 2003/0061331 A1	1-10

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F11/30(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. G06F11/30

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	1971-2008年
日本国実用新案登録公報	1996-2008年
日本国登録実用新案公報	1994-2008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 請求の範囲の番号
Y	JP 2007-233687 A (日本電気株式会社) 2007.09.13、段落0042-0082 (ファミリーなし)	1-10
Y	JP 2002-229806 A (株式会社日立製作所) 2002.08.16、段落0073-0081 & US 2002/0108074 A1	1-10

 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。

* 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

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

国際調査を完了した日

11.07.2008

国際調査報告の発送日

22.07.2008

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）

多胡 滋

5B

3562

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 請求の範囲の番号
Y	JP 2003-108420 A (株式会社日立製作所) 2003. 04. 11、段落0037-0055 & US 2003/0061331 A1	1-10
A	JP 2005-234861 A (三菱電機株式会社) 2005. 09. 02、全文、全図 (ファミリーなし)	1-10
A	JP 2004-295811 A (株式会社日立製作所) 2004. 10. 21、全文、全図 & US 2003/0061331 A1	1-10