(19)日本国特許庁(JP)

再 公 表 特 許(A1)

(11)国際公開番号
WO2012/063294

発行日　平成26年5月12日(2014.5.12)

(43)国際公開日　平成24年5月18日(2012.5.18)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 11/20	(2006.01)	G06F	11/20	310A	5B034
G06F 9/46	(2006.01)	G06F	9/46	350	

審査請求 有　予備審査請求 未請求　（全 34 頁）

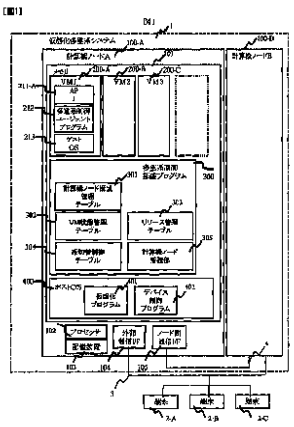
出願番号　特願2012-542705 (P2012-542705)
(21)国際出願番号　PCT/JP2010/006654
(22)国際出願日　平成22年11月12日(2010.11.12)
(81)指定国　AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,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,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,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

(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　100100310
弁理士　井上　学
(74)代理人　100098660
弁理士　戸田　裕二
(74)代理人　100091720
弁理士　岩崎　重美
(72)発明者　金　成昊
神奈川県横浜市戸塚区吉田町２９２番地
株式会社日立製作所システム開発研究所内
(72)発明者　西島　英児
神奈川県横浜市戸塚区吉田町２９２番地
株式会社日立製作所システム開発研究所内

Fターム(参考)　5B034 BB01　CC01　DD02

(54)【発明の名称】計算機システム

(57)【要約】
物理計算機上で搭載されている複数の仮想マシンが同一の動作モードとして動作するよう、各仮想マシンの主系/従系を切り替える。
計算機システムは複数の計算機ノードを有し、各計算機ノードは複数の仮想計算機と仮想計算機を制御する制御基盤部とを有する。各仮想計算機は、自計算機ノードとは別の他の計算機ノード上で動作している他の仮想計算機と多重化グループを構成し、いずれか一方が主系として動作し他方が従系として動作する。制御基盤部は、各仮想計算機について、主系/従系のいずれとして動作しているのかを管理し、各仮想計算機各々の状態を監視している。制御基盤部は、自計算機ノード上で主系仮想計算機として動作する仮想計算機の障害を検知した場合に、予め定められた規則に従って、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判定する。



1 Virtualization multiplexing system
2-A, 2-B, 2-C Terminal
105-A Computer node A
105-B Computer node B
101 Memory
102 Processor
103 Storage Device
104 External communication interface
105 Inter-node communication interface
212 Multiplexing control agent program
213 Host OS
300 Multiplexing control base program
301 Computer node structure management table
302 VM state management table
303 Resource management table
304 System switch control table
305 Computer node monitoring unit
400 Host OS
401 Virtualization program
402 Device control program

【特許請求の範囲】
【請求項１】
複数の計算機ノードを有する計算機システムであって、
前記複数の計算機ノードは各々、複数の仮想計算機と、当該複数の仮想計算機を制御する制御基盤部とを有しており、
各仮想計算機は、当該仮想計算機が動作している自計算機ノードとは別の他の計算機ノード上で動作している他の仮想計算機と多重化グループを構成し、当該仮想計算機若しくは当該他の仮想計算機のいずれか一方が主系仮想計算機として動作し他方が従系仮想計算機として動作しており、
前記制御基盤部は、当該制御基盤部が備えられている自計算機ノードの複数の仮想計算機各々について、主系仮想計算機若しくは従系仮想計算機のいずれとして動作しているのかを管理し、当該複数の仮想計算機各々の状態を監視しており、
前記制御基盤部は、自計算機ノード上で主系仮想計算機として動作する仮想計算機の障害を検知した場合に、予め定められた規則に従って、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判定し、判定結果に従って切替の制御を行うことを特徴とする計算機システム。
【請求項２】
請求項１記載の計算機システムであって、
前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数と従系仮想計算機として動作している仮想計算機の数との比率に応じて、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判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２記載の計算機システムであって、
前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数が、当該自計算機ノード上で動作している仮想計算機の総数の半数未満である場合に、障害が生じた仮想計算機とともに当該自計算機ノード上で主系仮想計算機として動作している他の仮想計算機も、主系仮想計算機から従系仮想計算機へ切り替えるよう制御を行うことを特徴とする計算機システム。
【請求項４】
請求項２記載の計算機システムであって、
前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数が、当該自計算機ノード上で動作している仮想計算機の総数の半数以上である場合には、障害が生じた仮想計算機を主系仮想計算機から従系仮想計算機へ切り替えるよう制御し、
前記制御基盤部は、前記障害が回復した場合に、前記仮想計算機を従系仮想計算機から主系仮想計算機へ戻すよう制御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記載の計算機システムであって、
前記制御基盤部は、自計算機ノード上で動作する仮想計算機を主系仮想計算機から従系仮想計算機へ切り替える場合に、当該仮想計算機へ従系仮想計算機への切替要求を通知すると共に、当該仮想計算機と多重化グループを構成している他の仮想計算機を従系仮想計算機から主系仮想計算機へ切り替えるために、当該他の仮想計算機が動作する他の計算機ノードへ通知を送信することを特徴とする計算機システム。
【請求項６】
請求項１記載の計算機システムであって、
前記制御基盤部は、前記自計算機ノード上で動作する各仮想計算機が使用するハードウェアリソースを管理しており、主系仮想計算機として動作する複数の仮想計算機に共有されるハードウェアリソースの障害を検知した場合には、当該複数の仮想計算機を共に主系

仮想計算機から従系仮想計算機へ切り替えるよう制御することを特徴とする計算機システム。

【請求項７】

請求項１記載の計算機システムであって、

前記制御基盤部は、前記自計算機ノード上で動作する複数の仮想計算機が論理的なグループを構成している場合に、当該複数の仮想計算機の一に障害が生じた場合は、当該論理的なグループを構成している他の仮想計算機も共に主系仮想計算機から従系仮想計算機へ切り替えるよう制御することを特徴とする計算機システム。

【請求項８】

複数の計算機ノードを有する計算機システムにおいて実行される方法であって、

前記複数の計算機ノードは各々、複数の仮想計算機と、当該複数の仮想計算機を制御する制御基盤部とを有しており、

各仮想計算機は、当該仮想計算機が動作している自計算機ノードとは別の他の計算機ノード上で動作している他の仮想計算機と多重化グループを構成し、当該仮想計算機若しくは当該他の仮想計算機のいずれか一方が主系仮想計算機として動作し他方が従系仮想計算機として動作しており、

前記制御基盤部は、当該制御基盤部が備えられている自計算機ノードの複数の仮想計算機各々について、主系仮想計算機若しくは従系仮想計算機のいずれとして動作しているのかを管理し、当該複数の仮想計算機各々の状態を監視しており、

前記制御基盤部は、自計算機ノード上で主系仮想計算機として動作する仮想計算機の障害を検知した場合に、予め定められた規則に従って、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判定し、判定結果に従って切替の制御を行うことを特徴とする方法。

【請求項９】

請求項８記載の方法であって、

前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数と従系仮想計算機として動作している仮想計算機の数との比率に応じて、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判定することを特徴とする方法。

【請求項１０】

請求項９記載の方法であって、

前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数が、当該自計算機ノード上で動作している仮想計算機の総数の半数未満である場合に、障害が生じた仮想計算機とともに当該自計算機ノード上で主系仮想計算機として動作している他の仮想計算機も、主系仮想計算機から従系仮想計算機へ切り替えるよう制御を行うことを特徴とする計算機システム。

【請求項１１】

請求項９記載の方法であって、

前記制御基盤部は、前記自計算機ノード上で主系仮想計算機として動作している仮想計算機の数が、当該自計算機ノード上で動作している仮想計算機の総数の半数以上である場合には、障害が生じた仮想計算機を主系仮想計算機から従系仮想計算機へ切り替えるよう制御し、

前記制御基盤部は、前記障害が回復した場合に、前記仮想計算機を従系仮想計算機から主系仮想計算機へ戻すよう制御することを特徴とする方法。

【請求項１２】

請求項８記載の方法であって、

前記制御基盤部は、自計算機ノード上で動作する仮想計算機を主系仮想計算機から従系仮想計算機へ切り替える場合に、当該仮想計算機へ従系仮想計算機への切替要求を通知すると共に、当該仮想計算機と多重化グループを構成している他の仮想計算機を従系仮想計算機から主系仮想計算機へ切り替えるために、当該他の仮想計算機が動作する他の計算機

ノードへ通知を送信することを特徴とする方法。
【請求項１３】
請求項８記載の方法であって、
前記制御基盤部は、前記自計算機ノード上で動作する各仮想計算機が使用するハードウェアリソースを管理しており、主系仮想計算機として動作する複数の仮想計算機に共有されるハードウェアリソースの障害を検知した場合には、当該複数の仮想計算機を共に主系仮想計算機から従系仮想計算機へ切り替えるよう制御することを特徴とする方法。
【請求項１４】
請求項８記載の計算機システムであって、
前記制御基盤部は、前記自計算機ノード上で動作する複数の仮想計算機が論理的なグループを構成している場合に、当該複数の仮想計算機の一に障害が生じた場合は、当該論理的なグループを構成している他の仮想計算機も共に主系仮想計算機から従系仮想計算機へ切り替えるよう制御することを特徴とする方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、サーバ仮想化環境において可用性を上げるために複数の計算機ノードを多重化して構成された計算機システムに関し、特に計算機ノード上のホストマシンおよび各仮想マシンの障害監視と計算機ノードの系切り替えを行う方法に関する。
【背景技術】
【０００２】
サーバ仮想化技術は、１台の物理計算機上で複数の仮想マシンを実行し、物理計算機の台数を削減することで、運用コスト、計算機の設置スペース及び消費電力等を削減できるメリットがあり、近年、注目を浴びている。ただし、物理計算機で運用していたシステムと同様な信頼性の確保のため、仮想化環境を提供する物理計算機の多重化とともに、各物理計算機で制御・運用される仮想マシンの多重系構成も要求されている。
【０００３】
特許文献１には、物理計算機上の各仮想マシン単位で、障害を検知し系を切替える方法が開示されている。特許文献１では、図２０に示すように、物理計算機２００１と２０３１が、ＬＡＮを介して多重化されており、物理計算機Ａ２００１はホストマシン２００２および仮想マシン２００５と２０１５を備えたサーバ仮想化環境を備えており、物理計算機Ｂ２０３１も同様な構成を備えている。また、障害検知方法および系切替え方法として、物理計算機Ａ２００１上の仮想マシンＡ２００５上にゲストクラスタプログラム２００７、仮想マシンＢ２０１５上にゲストクラスタプログラム２０１７、および、ホストマシン２００２上にホストクラスタプログラム２００３を設けている。物理計算機Ｂ２０３１も同様な構成である。
【０００４】
ゲストクラスタプログラム２００７と２０１７または２０３７と２０４７は、アプリケーションプログラム２００６、２０１６、２０３６、２０４６の稼働状況を監視しつつ、ホストクラスタプログラム２００３または２０３３に監視結果を通知する。また、ゲストクラスタプログラム２００７、２０１７は、ホストクラスタプログラム２００３から系切替え命令を受けた場合には、各々対応する仮想マシン２００５、２０１５に対する主系／従系モードの系切替えを実施する。仮想マシンが主系モードの時には、 アプリケーションが入力されたデータに対して処理を行い処理結果を出力する。一方、仮想マシンが従系モードの時には、 アプリケーションが入力されたデータに対して処理を行うが処理結果の出力は行わない。また、ゲストクラスタプログラム２０３７、２０４７も同様に、ホストクラスタプログラム２０３３から系切替え命令を受けた場合には、各々対応する仮想マシン２０３５、２０４５に対する主系／従系モードの系切替えを実施する。一方、ホストクラスタプログラム２００３または２０３３は、同じ物理計算機上で稼動している各仮想マシン（仮想マシン２００５と２０１５または仮想マシン２０３５と２０４５）の稼働状

況を監視し、例えば、主系モードの仮想マシンＡ２００５が停止していることを検知した場合には、仮想マシンＡ２０３５を従系モードから主系モードへ切り替える。

【先行技術文献】

【特許文献】

【０００５】

【特許文献１】特開２００８－２６９３３２号公報

【発明の概要】

【発明が解決しようとする課題】

【０００６】

特許文献１では、物理計算機上の仮想マシン単位での障害を検知し障害発生時には主系と従系との間の系切り替えを実施する。例えば、図２０において、物理計算機Ａ２００１上の各仮想マシン２００５と２０１５が全て主系モードで、物理計算機Ｂ２０３１上の各仮想マシン２０３５と２０４５が全て従系モードの場合に、物理計算機Ａ２００１上の仮想マシンＡ２００５が障害とする。この場合には、物理計算機Ｂ上の仮想マシンＡ２０３５が従系モードから主系モードに切り替わって稼働する。この場合物理計算機Ａには主系モードの仮想マシンと従系モードの仮想マシンが混在した状態となる。物理計算機Ｂも同様に主系モードおよび従系モードの仮想マシンが混在した状態となる。

【０００７】

しかしながら、物理計算機に対してハードウェア保守を行いたい場合には、物理計算機上の全ての仮想マシンが従系モード（または停止モード）でなければ、物理計算機を一旦停止させることができない。なぜなら、主系モードの仮想マシンは実際の処理を行った結果を出力しているから、処理中の仮想マシンを停止することができない。つまり、物理計算機上に主系モードの仮想マシンおよび従系モードの仮想マシンが混在している状態では、ハードウェア保守を行うことができない問題がある。

【０００８】

そこで、仮想化計算機の多重系構成制御において、物理計算機上に搭載されている複数の仮想マシンがなるべく同じの動作モードとなるように、各仮想マシンの系切替えを実施する技術を提供する。

【課題を解決するための手段】

【０００９】

計算機システムは複数の計算機ノードを有し、各計算機ノードは複数の仮想計算機と仮想計算機を制御する制御基盤部とを有する。各仮想計算機は、自計算機ノードとは別の他の計算機ノード上で動作している他の仮想計算機と多重化グループを構成し、いずれか一方が主系仮想計算機として動作し他方が従系仮想計算機として動作している。制御基盤部は、各仮想計算機について、主系/従系仮想計算機のいずれとして動作しているのかを管理し、各仮想計算機各々の状態を監視している。制御基盤部は、自計算機ノード上で主系仮想計算機として動作する仮想計算機の障害を検知した場合に、予め定められた規則に従って、障害が生じた仮想計算機とともに自計算機ノード上で動作する他の仮想計算機も主系仮想計算機から従系仮想計算機へ切り替えるかどうか判定する。

【発明の効果】

【００１０】

仮想化計算機の多重系構成制御において、物理計算機上に搭載されている複数の仮想マシンがなるべく同じの動作モードとなるように、各仮想マシンの系切替えを実施することができる。

【図面の簡単な説明】

【００１１】

【図１】仮想化多重システムの構成例を示す図である。

【図２】計算機ノードの構成例を示す図である。

【図３】計算機ノード構成管理テーブルの構成例を示す図である。

【図４】ＶＭ状態管理テーブルの構成例を示す図である。

【図５】リソース管理テーブルの構成例を示す図である。
【図６】系切替制御テーブルの構成例を示す図である。
【図７】ＶＭ１を対象とした系切替制御テーブルの例を示す図である。
【図８】計算機ノード間通信フォーマットの一例を示す図である。
【図９】多重系制御エージェントプログラムと多重系制御基盤プログラム間の通信フォーマットの一例を示す図である。
【図１０】仮想化多重系システムで実行される処理の一例を示すフローチャートである。
【図１１】リソース監視処理の一例を示す図である。
【図１２】リソース監視処理の一例を示すフローチャートである。
【図１３】系切替通知処理の一例を示すフローチャートである。
【図１４】メッセージ送信処理の一例を示すフローチャートである。
【図１５】ノード監視処理の一例を示すフローチャートである。
【図１６】生死監視処理の一例を示すフローチャートである。
【図１７】全ＶＭ切替処理の一例を示すフローチャートである。
【図１８】主系処理の一例を示すフローチャートである。
【図１９】従系処理の一例を示すフローチャートである。
【図２０】計算機システムの一例を示す図である。
【発明を実施するための形態】
【００１２】
以下、実施の形態を図面を用いて説明する。なお、本発明が実施の形態に制限されることは無く、本発明の思想に合致するあらゆる応用例が本発明に該当する。
【実施例１】
【００１３】
＜図１：全体構成例＞
図１は、仮想化多重システムの全体構成例である。
【００１４】
仮想化多重系システム１は一つまたは複数の端末２と外部ネットワーク３を介して接続される。仮想化多重系システム１は二台の計算機ノード１００（図１では計算機ノードＡ　１００－Ａと計算機ノードＢ　１００－Ｂ）と内部ネットワーク４より構成されている。
【００１５】
計算機ノード１００はメモリ１０１、プロセッサ１０２、記憶装置１０３、外部通信インタフェース１０４（図中、インタフェースをＩ／Ｆと略記）、ノード間通信Ｉ／Ｆ１０５を備える。計算機ノードＡ　１００－Ａと計算機ノードＢ　１００－Ｂは共に同様の構成を有する。
【００１６】
メモリ１０１内にはホストオペレーティングシステム４００（図中、ホストＯＳと略記）、多重系制御基盤プログラム３００と一つまたは複数の仮想マシン環境２００（図中、ＶＭと略記）が保持される。ホストＯＳ４００内には仮想化プログラム４０１とデバイス制御プログラム４０２が保持される。多重系制御基盤プログラム３００は計算機ノード構成管理テーブル３０１、ＶＭ状態管理テーブル３０２、リソース管理テーブル３０３、系切替制御テーブル３０４と計算機ノード監視部３０５から構成される。ＶＭ２００はアプリケーション２１１（図中、ＡＰと略記）、多重系制御エージェントプログラム２１２、ゲストオペレーティングシステム２１３（図中、ゲストＯＳと略記）から構成される。
【００１７】
尚、図１では端末２として三台の端末２－Ａ、２－Ｂ、２－Ｃが仮想化多重系システム１に接続されており、また計算機ノードＡ　１００－Ａ内にはＶＭ２００としてＶＭ１　２００－Ａ、ＶＭ２　２００－Ｂ、ＶＭ３　２００－Ｃが備えられている例を示したが、特に台数は限定されるものではない。
【００１８】

仮想化多重系システム１内の計算機ノードＡ　１００－Ａ、計算機ノードＢ　１００－Ｂは内部ネットワーク４を介して接続される。計算機ノード１００では、外部ネットワーク３は外部通信Ｉ／Ｆ１０４に接続し、内部ネットワーク４はノード間通信Ｉ／Ｆ１０５に接続する。

【００１９】

端末２は外部ネットワーク３を介して仮想化多重系システム１に要求を送信し、仮想化多重系システム１は要求を処理した結果を端末２に返信する。仮想化多重系システム１内の計算機ノードＡ　１００－Ａ、計算機ノードＢ　１００－Ｂは、一方が「主系ノード」、他方が「従系ノード」として動作する。端末２からの要求は計算機ノードＡ　１００－Ａ、計算機ノードＢ　１００－Ｂ両方で受取られ、処理されるが、正常な状態で動作する場合には、処理結果は「主系ノード」である計算機ノード１００だけが端末２に返信する。外部通信Ｉ／Ｆ１０４などの故障で主系ノードの処理継続に支障が生じた時に、計算機ノード１００間の主系ノード／従系ノードの関係は切替る。

【００２０】

計算機ノード１００の記憶装置１０３はホストＯＳ４００、多重系制御基盤プログラム３００（プログラムと各種テーブル情報）、ＶＭ２００を構成するＡＰ２１１、多重系制御エージェントプログラム２１２、ゲストＯＳ２１３を格納する。プロセッサ１０２は記憶装置１０３からホストＯＳ４００、多重系制御基盤プログラム３００、ＶＭ２００構成物をメモリ１０１に展開、実行し、また外部通信Ｉ／Ｆ１０４やノード間通信Ｉ／Ｆ１０５からの割込みを処理する。外部通信Ｉ／Ｆ１０４は外部ネットワーク３を通じて端末２とのデータの送受信を行う。ノード間通信Ｉ／Ｆ１０５は内部ネットワーク４を通じて他の計算機ノード１００との間でデータの送受信を行う。

【００２１】

仮想化プログラム４０１はＶＭ２００を作成、管理するためのプログラムであり、デバイス制御プログラム４０２は外部通信Ｉ／Ｆ１０４、ノード間通信Ｉ／Ｆ１０５や記憶装置１０３にアクセスするためのプログラムである。ホストＯＳ４００は仮想化プログラム４０１やデバイス制御プログラム４０２を用いて、多重系制御基盤プログラム３００やＶＭ２００の実行を制御し、外部通信Ｉ／Ｆ１０４、ノード間通信Ｉ／Ｆ１０５や記憶装置１０３へのアクセス制御を行う。

【００２２】

ＶＭ２００は固有のプログラムやオペレーティングシステムが動作することが可能な、仮想化プログラム４０１により作成された仮想的な実行環境である。複数のＶＭ２００が存在する場合に、異なるＶＭ２００内で動作するプログラムやオペレーティングシステムは同じでも良いし、異なっていても良い。また、異なるＶＭ２００内で動作するプログラムやオペレーティングシステムが互いに直接影響を与え合うことはない。ＶＭ２００は仮想化多重系システム１内で一意の識別子を持ち、例えば、図１中のＶＭ２００は「ＶＭ１」、「ＶＭ２」、「ＶＭ３」の識別子を保有する。また異なる計算機ノード１００内で動作する任意のＶＭ２００の対（ペア）は、「主系ＶＭ」／「従系ＶＭ」の関係を持つ。通常、主系ノードの計算機ノード１００内のＶＭ２００は「主系ＶＭ」となり、従系ノードの計算機ノード１００内のＶＭ２００は「従系ＶＭ」となる。「主系ＶＭ」／「従系ＶＭ」の関係を持つＶＭ２００の対は、二つの単体ＶＭ２００間で構成される場合もあるし、一つの計算機ノード１００内の複数のＶＭグループが協同してアプリケーションを実行してサービスを提供する場合には複数のＶＭ２００から構成されるＶＭグループ間で構成される場合もある。主系ＶＭの動作に異常が生じた場合には、「主系ＶＭ」／「従系ＶＭ」の関係を入れ替える、系切替処理が行われる。

【００２３】

ＡＰ２１１は端末２からの要求を処理する。ＡＰ２１１は自身が動作するＶＭ２００が主系ＶＭ、従系ＶＭのいずれであっても端末２からの要求を読取り、演算処理を行うが、端末２への応答動作は異なる。主系ＶＭで動作するＡＰ２１１は演算結果を端末２へ返答するが、従系ＶＭで動作するＡＰ２１１は演算結果を返答しない。ＡＰ２１１は「系切替

」を行うことで端末２への返答動作を切替える。
【００２４】
多重系制御エージェントプログラム２１２は多重系制御基盤プログラム３００との間で情報通信を行うプログラムであり、ＡＰ２１１の動作状態を監視して多重系制御基盤プログラム３００に通知したり、多重系制御基盤プログラム３００からの指示に従ってＡＰ２１１の系切替を行ったりする。ＡＰ２１１の動作状態を監視する方法の詳細は図１１、図１２と共に後述する。
【００２５】
ゲストＯＳ２１３はＶＭ２００の環境内でＡＰ２１１や多重系制御エージェントプログラム２１２の動作を制御する。
【００２６】
多重系制御基盤プログラム３００は、他の計算機ノード１００内で動作する多重系制御基盤プログラム３００との間で情報を交換することで仮想化多重系システム１内のＶＭ２００の状態を判断し、ＶＭ２００の「主系ＶＭ」／「従系ＶＭ」の系切替を制御する。主系ノードである計算機ノード１００内で動作する多重系制御基盤プログラム３００はＶＭ２００の状態判断と系切替の実行を指示し、従系ノードの計算機ノード１００内で動作する多重系制御基盤プログラム３００は主系ノードの生死状態を監視する。計算機ノード１００の監視は、主系ノード、従系ノードの双方から行う方式も有り得るが、本実施形態では従系ノードからのみ実行する。しかし、本発明の思想がこれに限定されるものでは無い。多重系制御基盤プログラム３００のＶＭ２００の状態判断の詳細は図１３と共に後述する。系切替や生死状態監視処理についての詳細は、計算機ノード監視部３０５の処理説明として、図１４、図１５、図１６、図１７、図１８、図１９と共に後述する。
【００２７】
計算機ノード構成管理テーブル３０１は多重系制御基盤プログラム３００が計算機ノード１００の主系ノード／従系ノードなどの状態を管理するためのテーブルである。詳細は図３と共に後述する。
【００２８】
ＶＭ状態管理テーブル３０２は多重系制御基盤プログラム３００が計算機ノード１００内のＶＭ２００の主系ＶＭ／従系ＶＭなどの状態を管理するためのテーブルである。詳細は図４と共に後述する。
【００２９】
リソース管理テーブル３０３は多重系制御基盤プログラム３００が、ＶＭ２００が使用しているメモリ１０１や実行しているＡＰ２１１、他のＶＭ２００との関係などの「リソース」を管理するためのテーブルである。詳細は図５と共に後述する。
【００３０】
系切替制御テーブル３０４は多重系制御基盤プログラム３００が、ＶＭ２００が使用するリソースに異常が発生した時などにおける、障害内容とこれに対する対応動作を記載したテーブルである。詳細は図６、図７と共に後述する。
【００３１】
計算機ノード監視部３０５は多重系制御基盤プログラム３００が他の計算機ノード１００内で動作する多重系制御基盤プログラム３００との間で通信を行い、生死状態監視処理を行う処理である。計算機ノード監視部３０５は多重系制御基盤プログラム３００を構成するプログラムではあるが、多重系制御基盤プログラム３００の処理とは独立した契機で活動する。詳細は図１４、図１５、図１６、図１７、図１８、図１９と共に後述する。
＜図２：仮想化環境構成例＞
図２は計算機ノードの構成の例を示す図である。アプリケーションやオペレーティングシステムの動作環境であるＶＭ２００は仮想的な計算機の実行環境であり、実際には、計算機ノード１００内のメモリ１０１やプロセッサ１０２などの現存するリソースを割当てることで構築される。
【００３２】

図２は計算機ノード１００のメモリ１０１、プロセッサ１０２、外部通信Ｉ／Ｆ１０４等のリソースを複数のＶＭ２００やホストＯＳ４００への割当てた例を示した図である。

【００３３】

図２中において、計算機ノードＡ　１００－Ａ内には複数のＶＭ２００（ＶＭ１　２００－Ａ、ＶＭ２　２００－Ｂ、ＶＭ３　２００－Ｃ）が作成されている。メモリ１０１は複数のメモリ領域に分割することが可能で、それぞれのメモリ領域が個別に各ＶＭ２００やホストＯＳ４００に割り当てられたり、共有されたりする。プロセッサ１０２は「マルチコア」のプロセッサで、コア単位でＶＭ２００やホストＯＳ４００に割り当てられたり、共有されたりする。外部通信Ｉ／Ｆ１０４は計算機ノード１００内には複数存在し、それぞれの外部通信Ｉ／Ｆ１０４がＶＭ２００やホストＯＳ４００に割り当てられたり、共有されたりする。

【００３４】

仮想環境ＶＭ１　２００－Ａは、プロセッサ１０２のＶＭ１占有コア＃１　１０２－Ａ、ＶＭ１占有メモリ１０１－ＡとＶＭ１とＶＭ２間共有メモリ１０１－Ｂのメモリ領域、ＶＭ１とＶＭ２共有外部通信Ｉ／Ｆ１０４－Ａのリソースを使用して構築される。仮想環境ＶＭ１　２００－Ａは、プロセッサ１０２のＶＭ２占有コア＃２　１０２－Ｂ、ＶＭ２占有メモリ１０１－ＣとＶＭ１とＶＭ２共有メモリ１０１－Ｂのメモリ領域、ＶＭ１とＶＭ２共有外部通信Ｉ／Ｆ１０４－Ａのリソースを使用して構築される。

ＶＭ１　２００－Ａ、ＶＭ２　２００－Ｂに割当てられていないリソースであるプロセッサ１０２のホストＯＳ占有コア＃０　１０２－Ｃは、ホストＯＳ４００のみが使用するプロセッサ・コアで、ＶＭ２００から使用されることは無い。また外部通信Ｉ／Ｆ１０４－ＢはホストＯＳ４００と、ＶＭ１　２００－ＡとＶＭ２　２００－Ｂ以外のＶＭ２００が使用する。

【００３５】

計算機ノードＢ　１００－Ｂ内にも複数のＶＭ２０１（ＶＭ４　２０１－Ａ、ＶＭ５　２０１－Ｂ、ＶＭ６　２０１－Ｃ）が作成されており、各ＶＭ２０１やホストＯＳ４００には計算機ノードＢ　１００－Ｂ内のリソースが割り当てられる。

＜図３：計算機ノード構成管理テーブル＞

図３は計算機ノード構成管理テーブル３０１の構成例を示す図である。計算機ノード構成管理テーブル３０１は多重系制御基盤プログラム３００内に保持され、仮想化多重系システム１内の計算機ノード１００の状態情報を格納する。仮想化多重系システム１内の各計算機ノードＡ　１００－Ａ、計算機ノードＢ　１００－Ｂは、同じ内容の計算機ノード構成管理テーブル３０１を保持する。

【００３６】

計算機ノード識別子３１１には仮想化多重系システム１内の計算機ノード１００を識別する一意の情報を格納する。

【００３７】

主／従系状態フラグ３１２には各計算機ノード１００が「主系ノード」、「従系ノード」のいずれの状態にあるかを識別する情報を格納する。対象の計算機ノード１００が主系ノードの場合は「Ｍ」を格納し、従系ノードの場合は「Ｓ」を格納する。

【００３８】

マルチキャストアドレス３１３には計算機ノード１００にアクセスするための、内部ネットワーク４のマルチキャストアドレスを格納する。

【００３９】

生死状態フラグ３１４には計算機ノード１００の正常／異常状態を識別する情報を格納する。計算機ノード１００が「正常状態」の場合には「０」値を格納し、「異常状態」の場合には「１」を格納する。本発明の実施の形態では正常／異常状態を「０／１」値で表すが、正常／異常状態の識別が可能な情報であれば格納情報は「０／１」値に限らない。

【００４０】

監視周期３１５は計算機ノード１００の状態を監視する周期時間を格納する。従系ノー

ドの多重系制御基盤プログラム３００の計算機ノード監視部３０５は、監視周期３１５の時間周期で、主系ノードの計算機ノード１００の状態を取得する。主系ノードは従系ノードの計算機ノード１００の状態取得要求に応じて、状態情報を返信する。

【００４１】

生死確定カウンタ３１６は計算機ノード１００の状態取得要求に対して応答が無かった場合に、相手計算機ノード１００が異常状態にあることを判断するためのしきい値を格納する。生死確定カウンタ３１６の回数分、状態取得要求に対する相手計算機ノード１００からの返信が連続して無かった場合は、相手計算機ノード１００が異常状態にあると判断する。

＜図４：ＶＭ状態管理テーブル＞

図４は本発明の実施の形態のＶＭ状態管理テーブル３０２の構成を示す図である。ＶＭ状態管理テーブル３０２は多重系制御基盤プログラム３００内に保持され、計算機ノード１００内に作成されているＶＭ２００の状態情報を保持する。

【００４２】

ＶＭ状態管理テーブル３０２に格納される状態情報は計算機ノード１００毎に異なる。計算機ノードＡ　１００－ＡのＶＭ状態管理テーブル３０２はＶＭ１　２００－Ａ、ＶＭ２　２００－Ｂ、ＶＭ３　２００－Ｃの状態情報を保持する。計算機ノードＢ　１００－ＢのＶＭ状態管理テーブル３０２はＶＭ４　２０１－Ａ、ＶＭ５　２０１－Ｂ、ＶＭ６　２０１－Ｃの状態情報を保持する。

管理ＶＭ総数３２１には一台の計算機ノード１００内に作成されているＶＭ２００の総数を格納する。

【００４３】

障害ＶＭ総数３２２には一台の計算機ノード１００内に作成されているＶＭ２００内、異常状態にあるＶＭ２００の個数を格納する。障害ＶＭ総数３２２は「０」以上、管理ＶＭ総数３２１の値以下の値となる。

【００４４】

ＶＭ識別子３２３には仮想化多重系システム１内でＶＭ２００を一意に識別する情報を格納する。行３２８はＶＭ１　２００－Ａの状態情報を格納した行であることを示す。

【００４５】

ホスト計算機ノード識別子３２４はＶＭ２００が作成されている計算機ノード１００の識別情報を格納する。

【００４６】

主／従系状態フラグ３２５にはＶＭ２００の「主系ＶＭ」、「従系ＶＭ」を識別する情報を格納する。対象のＶＭ２００が主系ＶＭの場合は「Ｍ」を格納し、従系ＶＭの場合は「Ｓ」を格納する。

【００４７】

生死状態フラグ３２６にはＶＭ２００の正常／異常状態を識別する情報を格納する。ＶＭ２００が「正常状態」の場合には「０」値を格納し、「異常状態」の場合には「１」を格納する。本発明の実施の形態では正常／異常状態を「０／１」値で表すが、正常／異常状態の識別が可能な情報であれば格納情報は「０／１」値に限らない。

【００４８】

主／従系構成ＶＭ情報３２７には、ＶＭ識別子３２３のＶＭ２００と主系ＶＭ／従系ＶＭの関係に対応するＶＭ２００の情報を格納する。ホスト計算機ノード識別子３２７－Ａには対応するＶＭ２００が作成されている計算機ノード１００の識別情報を格納し、ＶＭ識別子３２７－Ｂには対応するＶＭ２００を一意に特定する識別情報を格納する。

＜図５：リソース管理テーブル＞

図５はリソース管理テーブル３０３の構成例を示す図である。リソース管理テーブル３０３は多重系制御基盤プログラム３００内に保持され、ＶＭ２００が使用する個別リソースの状態情報を格納する。リソース管理テーブル３０３に格納される状態情報は計算機ノード１００毎に異なる。計算機ノードＡ　１００－ＡのリソースＶ

ＭＭ１　２００－Ａ、ＶＭ２　２００－Ｂ、ＶＭ３　２００－Ｃが使用するリソースの状態情報を保持する。計算機ノードＢ　１００－Ｂのリソース管理テーブル３０３はＶＭ４　２０１－Ａ、ＶＭ５　２０１－Ｂ、ＶＭ６　２０１－Ｃが使用するリソースの状態情報を保持する。

【００４９】

ＶＭ識別子３３１にはＶＭ２００を一意に識別する情報を格納し、リソース識別子３３２にはＶＭ２００が使用するリソースを一意に特定する情報を格納する。行３３６はＶＭ１　２００－ＡがＡＰ１　２１１－Ａをリソースとして使用していることを意味する。

【００５０】

監視周期３３３には使用リソースの正常／異常状態を監視するための周期時間を格納する。リソース状態の取得をイベントなどの契機で行い、周期的に監視しない場合は無効値である「０」を格納する。

【００５１】

リソース生死状態フラグ３３４には使用リソースの正常／異常状態を識別する情報を格納する。リソースの状態が「正常状態」の場合には「０」値を格納し、「異常状態」の場合には「１」を格納する。本発明の実施の形態では正常／異常状態を「０／１」値で表すが、正常／異常状態の識別が可能な情報であれば格納情報は「０／１」値に限らない。

【００５２】

リソース詳細情報３３５には所有状態３３５－Ａと個別情報３３５－Ｂの項目が含まれる。所有状態３３５－ＡにはＶＭ２００のリソースを使用する形態を格納し、個別情報３３５－Ｂにはリソース状態の取得手段を示す情報を格納する。

【００５３】

例えば図５における所有状態３３５－Ａの値、「占有」は使用リソースをＶＭ１　２００－Ａのみで使用することを意味し（行３３６）、「共有」は使用リソースをＶＭ１　２００－Ａと他ＶＭ２００とで使用することを意味し（行３３８）、「論理／SERVER」はＶＭ１　２００－ＡがＶＭ２　２００－Ｂと論理的な繋がりのあるグループを形成して「サーバ」の関係になっていることを意味し（行３３９）、「論理／CLIENT」はＶＭ２　２００－ＢがはＶＭ１　２００－Ａと論理的な繋がりのあるグループを形成して「クライアント」の関係になっていることを意味する（行３４０）。

【００５４】

例えば図５における個別情報３３５－Ｂの値、「２１２」は使用リソースＡＰ１　２１１－Ａの状態を多重系制御エージェントプログラム２１２が取得することを意味し（行３３６）、「ＥＶＴ」は使用リソースＶＭ１占有メモリ１０１－Ａ状態の取得はハードウェア状態を検知する計算機ノード１００のハードウェア機能などが送出するシグナル信号を検出することで行う事を意味する（行３３７）。図５の例では、多重系制御エージェントプログラム２１２は監視周期３３３に格納された値の周期で使用リソースＡＰ１　２１１－Ａの状態を取得するが（行３３６）、「ＥＶＴ」の場合はハードウェア機能は使用リソースＶＭ１占有メモリ１０１－Ａの状態に変化が生じた時にだけシグナル信号を送出するので監視周期３３３の格納値は使用しない（行３３７）。

＜図６：系切替制御テーブル＞

図６は系切替制御テーブル３０４の構成例を示す図である。系切替制御テーブル３０４は多重系制御基盤プログラム３００内に保持され、障害発生時に多重系制御基盤プログラム３００がＶＭ２００の主系ＶＭ／従系ＶＭの系切替を行うための規則を格納する。仮想化多重系システム１内の各計算機ノード１００は同じ内容の系切替制御テーブル３０４を保持している。

【００５５】

障害内容３４１には計算機ノード１００やＶＭ２００の障害発生部位を指定する情報を格納する。図６における「計算機ノード障害」は計算機ノード１００自体の動作継続に支障が生じる部位の指定情報であることを示す。ＶＭ２００の障害発生部位は、全てのＶＭ２００を対象に、使用リソースの所有状態種別毎にリソース識別子を格納する。

【００５６】
系切替対象ＶＭ３４２には、障害内容３４１に障害発生時に系切替対象となるＶＭ２００のＶＭ識別子を格納する。図６における「全ＶＭ」は障害が生じた計算機ノード１００内に存在する全てのＶＭ２００を系切替対象とすることを意味する。「該当ＶＭのみ」は障害が発生したリソースを使用するＶＭ２００だけが系切替対象となることを意味し、実際には当該ＶＭ２００の識別子が格納される。「関連全ＶＭ」は障害が発生したリソースを共用する複数のＶＭ２００、または障害が発生したリソースを使用しているＶＭ２００及び当該ＶＭ２００と論理的なグループ関係にある複数のＶＭ２００が一緒に系切替対象となることを意味し、実際には系切替対象となる複数のＶＭ２００の識別子が格納される。

【００５７】
過半数判断ロジック適用要否３４３にはＶＭ２００の系切替実行時に「過半数判断ロジック」の適用の有無を指定する情報を格納する。「要」は「過半数判断ロジック」を適用することを意味し、「否」は「過半数判断ロジック」を適用しないことを意味する。「過半数判断ロジック」は主系ＶＭ／従系ＶＭの系切替対象となるＶＭ２００を再評価する処理であり、ＶＭ状態管理テーブル３０２の障害ＶＭ総数３２２と系切替対象ＶＭ３４２のＶＭ２００の合計数が、ＶＭ状態管理テーブル３０２の管理ＶＭ総数３２１の半数以上であるかを判断し、系切替対象のＶＭ２００を決定する。過半数であれば「過半数判断ロジック」は全ＶＭを系切替対象と決定し、半数未満であれば系切替対象ＶＭ３４２に格納されたＶＭ２００のみを系切替対象とする。

【００５８】
Ｆａｉｌｂａｃｋ処理適用可否３４４には系切替を行ったＶＭ２００に「フェイルバック」処理の適用の可否を指定する情報を格納する。「可」は系切替を行ったＶＭ２００をフェイルバックしても良いことを意味し、「否」は系切替を行ったＶＭ２００をフェイルバックしてはいけないことを意味する。「フェイルバック」とはＶＭ２００の主系ＶＭ／従系ＶＭの状態を系切替実行前に戻す処理のことである。主系ＶＭのＶＭ２００に障害が発生すると多重系制御基盤プログラム３００により系切替が実行され、ＶＭ２００は従系ＶＭに切替る。「フェイルバック」処理は、系切替後の従系ＶＭのＶＭ２００を、後に障害回復した際に再度系切替を実行して、主系ＶＭに戻す処理を行う。

＜図７：ＶＭ１障害内容と系切替制御内容例＞

図７は計算機ノードＡ　１００－Ａに作成されたＶＭ１　２００－Ａを対象とした系切替制御テーブル３０４の例を示した図である。ＶＭ１　２００－Ａは、占有リソースとしてＡＰ１　２１１－ＡとＶＭ１占有メモリ１０１－Ａを、共有リソースとしてＶＭ１とＶＭ２共有メモリ１０１－ＢとＶＭ１とＶＭ２共有外部通信Ｉ／Ｆ１０４－Ａを使用し、論理リソースとしてＶＭ２　２００－Ｂとグループ関係にある。障害内容３４１には「ＶＭ１占有リソース障害」としてＡＰ１　２１１－ＡとＶＭ１占有メモリ１０１－Ａを格納し（行３５２）、「ＶＭ１共有リソース障害」としてＶＭ１とＶＭ２共有メモリ１０１－ＢとＶＭ１とＶＭ２共有外部通信Ｉ／Ｆ１０４－Ａを格納し（行３５３）、「ＶＭ１論理リソース障害」としてＶＭ２　２００－Ｂを格納する。

【００５９】
行３５１はＶＭ１　２００－Ａが存在する計算機ノードＡ　１００－Ａが障害により動作継続不可となった場合の系切替規則で、計算機ノードＡ　１００－Ａ内の全ＶＭ２００が系切替対象となり、「過半数判断ロジック」と「フェイルバック」処理は適用しない。

【００６０】
行３５２は占有リソースのＡＰ１　２１１－ＡとＶＭ１占有メモリ１０１－Ａが障害で使用不能となった場合の系切替規則で、ＶＭ１　２００－Ａのみが系切替対象となり、系切替実行時に「過半数判断ロジック」を適用し、占有リソースの障害回復時には「フェイルバック」処理を行う。

【００６１】
行３５３は共有リソースのＶＭ１とＶＭ２共有メモリ１０１－ＢとＶＭ１とＶＭ２共有

外部通信Ｉ／Ｆ１０４－Ａが障害で使用不能となった場合の系切替規則で、共に同じリソースを使用するＶＭ１　２００－ＡとＶＭ２　２００－Ｂが系切替対象となり、系切替実行時に「過半数判断ロジック」を適用し、共有リソースの障害回復時には「フェイルバック」処理を行う。

【００６２】

行３５４はグループ関係にあるＶＭ２　２００－Ｂが障害で動作不能となった場合の系切替規則で、同一グループ内のＶＭ１　２００－ＡとＶＭ２　２００－Ｂが系切替対象となり、系切替実行時に「過半数判断ロジック」を適用し、ＶＭ２　２００－Ｂの障害回復時には「フェイルバック」処理を行う。

＜図８：計算機ノード間通信フォーマット＞

図８は計算機ノード間通信フォーマット５００を示す図である。計算機ノード間通信フォーマット５００は異なる計算機ノード１００の多重系制御基盤プログラム３００の間で送受信される通信データのフォーマットで、計算機ノード１００の動作状況を取得する「生死監視」、及び主系ＶＭ／従系ＶＭの系切替要求である「フェイルバック」、「異常ＶＭ系切替」、「全ＶＭ系切替」の用途で用いられる。

【００６３】

計算機ノード識別子５０１には通信データを生成した計算機ノード１００を特定する識別子を格納する。通信データ生成時には、送受信される当該通信データを識別する情報をシーケンス番号５０２に格納し、また当該通信データ生成時の時刻情報を送信タイムスタンプ５０３に格納する。

【００６４】

通信種類５０４は通信データの用途を識別する情報を格納する。通信用途を識別する情報として、生死５０４－Ａ、ＦａｉｌＢａｃｋ５０４－Ｂ、異常ＶＭ５０４－Ｃ、全ＶＭ５０４－Ｄのうちの一つを使用する。従系ノードから主系ノードへ「生死監視」を行う場合には生死５０４－Ａを使用し、主系ノードから従系ノードへ異常状態になった特定の主系ＶＭの系切替を指示する「異常ＶＭ系切替」の場合には異常ＶＭ５０４－Ｃを使用し、全ての主系ＶＭの系切替を指示する「全ＶＭ系切替」の場合には全ＶＭ５０４－Ｄを使用し、ＶＭ２００の状態が回復した際に切替っていた主系ＶＭ／従系ＶＭの状態を元に戻す「フェイルバック」を指示する場合にはＦａｉｌＢａｃｋ５０４－Ｂを使用する。

【００６５】

計算機ノード１００の間でデータ通信を行う際には、相手計算機ノード１００に自計算機ノード１００内のＶＭ２００の状態を通知するために、ＶＭ状態管理テーブル内容５０５にはＶＭ状態管理テーブル３０２の内容を格納し、リソース管理テーブル内容５０６にはリソース管理テーブル３０３の内容を格納する。

「生死監視」などの用途での計算機ノード１００間でデータ通信では、相手計算機ノード１００状態の異常発生などにより、発信したデータ通信に対する応答が返ってこないなどの応答異常が発生する場合がある。応答異常カウント５０７には応答異常の発生回数を格納する。

＜図９：多重系制御エージェントプログラムと多重系制御基盤プログラム間の通信フォーマット＞

図９は多重系制御エージェントプログラムと多重系制御基盤プログラム間の通信フォーマット６００を示す図である（以降、プログラム間通信フォーマット６００と称する）。プログラム間通信フォーマット６００は計算機ノード１００内の多重系制御基盤プログラム３００と多重系制御エージェントプログラム２１２の間で送受信されるデータ通信のフォーマットである。

【００６６】

ＶＭ識別子６０１には通信元／先のＶＭ２００を特定する情報を格納し、シーケンス番号６０２には送受信される通信データを識別する情報を格納し、送信タイムスタンプ６０３には通信データ生成時の時刻情報を格納する。

【００６７】

多重系制御基盤プログラム３００と多重系制御エージェントプログラム２１２の間でＶＭ２００が使用するリソース情報の通知を行う場合には、リソース管理テーブル内容６０４にリソース管理テーブル３０３の内容を格納する。

【００６８】

多重系制御基盤プログラム３００と多重系制御エージェントプログラム２１２の間で主系ＶＭ／従系ＶＭの系切替要求を通知したい場合には、系切替要求６０５に値を設定する。本発明の形態では、系切替要求がある場合には「１」を設定し、系切替要求が無い場合には「０」を設定するが、系切替要求の有無が識別可能であれば「０／１」値に限られるものではない。

＜図１０：仮想化多重系システムの全体フローチャート＞

図１０は仮想化多重系システム１で実行される処理の全体フローチャートである。仮想化多重系システム１の全体処理は、多重系制御基盤プログラム３００の起動処理及び系切替通知処理Ｓ１１６、多重系制御エージェントプログラム２１２の起動処理及びリソース監視処理Ｓ１０６と系切替処理、計算機ノード監視部３０５のメッセージ送信処理Ｓ１１２及びノード監視処理Ｓ１２３より成り立つ。

【００６９】

ホストＯＳ４００及びプロセッサ１０２は多重系制御基盤プログラム３００及びＶＭ２００を計算機ノード１００のメモリ１０１内に展開すると、多重系制御基盤プログラム３００の処理を開始する。計算機ノード監視部３０５及びＶＭ２００は多重系制御基盤プログラム３００の処理の中から処理を開始される。

【００７０】

多重系制御基盤プログラム３００は処理が開始されると、初めに計算機ノード監視部３０５の起動を行う（Ｓ１１１）。

【００７１】

次に多重系制御基盤プログラム３００は記憶装置１０３からＶＭ状態管理テーブル３０２とリソース管理テーブル３０３をメモリ１０１内に読込み（Ｓ１１２）、系切替制御テーブル３０４もメモリ１０１内に読込む（Ｓ１１３）。

【００７２】

多重系制御基盤プログラム３００は制御下にある各ＶＭ２００のメモリ領域にリソース管理テーブル３０３をコピーし（Ｓ１１４）、ＶＭ２００の処理を起動する（Ｓ１１５）。

【００７３】

多重系制御基盤プログラム３００はＳ１１５の後、系切替通知処理Ｓ１１６を実行する。

【００７４】

計算機ノード監視部３０５は多重系制御基盤プログラム３００によって立ち上げられ処理を開始すると、多重系制御基盤プログラム３００からの「ＶＭ切替」または「フェイルバック」実行の送信要求の有無を判定する（Ｓ１２１）。

【００７５】

多重系制御基盤プログラム３００からの送信要求が存在する場合（Ｓ１２１の結果が「Ｙ」）はメッセージ送信処理Ｓ１１２を実行し、その後ノード監視処理Ｓ１２３を実行する。多重系制御基盤プログラム３００からの送信要求が存在しない場合（Ｓ１２１の結果が「Ｎ」）はノード監視処理Ｓ１２３を実行する。

【００７６】

計算機ノード監視部３０５はノード監視処理Ｓ１２３の実行終了後は、再び多重系制御基盤プログラム３００からの送信要求の有無判定Ｓ１２１から処理を実行する。

【００７７】

多重系制御エージェントプログラム２１２は処理を開始すると、自ＶＭ２００のメモリ領域内のリソース管理テーブル３０３を読込む（Ｓ１０１）。

【００７８】

多重系制御エージェントプログラム２１２は読込んだリソース管理テーブル３０３から、ＶＭ識別子３３１の格納値が自ＶＭ２００の識別子と一致し、かつ個別情報３３５－Ｂの格納値が「多重系制御エージェントプログラム２１２」である個別リソース管理情報を選別する。そして選別した個別リソース管理情報のリソース識別子３３２と監視周期３３３から状態監視対象のリソースと監視周期を取得する。（Ｓ１０２）

次に多重系制御エージェントプログラム２１２は自ＶＭ２００の状態を多重系制御基盤プログラム３００に通知する。多重系制御エージェントプログラム２１２は状態監視対象リソースの状態を取得、自ＶＭ２００のメモリ領域内のリソース管理テーブル３０３を更新し、プログラム間通信フォーマット６００の通信データを作成して多重系制御基盤プログラム３００に送信する。（Ｓ１０３）

次に多重系制御エージェントプログラム２１２は多重系制御基盤プログラム３００からの主系ＶＭ／従系ＶＭの系切替要求の有無を確認する（Ｓ１０４）。

【００７９】

系切替要求が有る場合（Ｓ１０４の結果が「Ｙ」）、多重系制御エージェントプログラム２１２は系切替処理Ｓ１０５を実行の後、再びＳ１０４を実行する。系切替処理Ｓ１０５は自ＶＭ２００内で動作するＡＰ２１１に対して系切替を要求する処理である。

【００８０】

系切替要求が無い場合（Ｓ１０４の結果が「Ｎ」）、多重系制御エージェントプログラム２１２はリソース監視処理Ｓ１０６を実行の後、再びＳ１０４を実行する。

＜図１１：リソース監視処理概要＞

図１１はリソース監視処理Ｓ１０６の概要を示す図である。本実施形態では、多重系制御エージェントプログラム２１２のリソース状態は、ＡＰ２１１の実行状態を取得することで判断する。

【００８１】

ＡＰ２１１が動作する時には、ＶＭ２００内の任意の場所（例えば／ｖａｒ／ｒｕｎ）にＡＰ２１１の実行状態やエラー情報を出力するＡＰｄｏｇ７００ファイルが確保される。ＡＰｄｏｇ７００ファイルはＡＰ２１１以外のプログラムからも参照が可能なファイルである。ＡＰ２１１は自身の実行状態を定期的にＡＰｄｏｇ７００ファイルに出力し、多重系制御エージェントプログラム２１２はＡＰｄｏｇ７００ファイルを参照することでＡＰ２１１の実行状態の監視を行う。

【００８２】

多重系制御エージェントプログラム２１２のリソース状態の監視は、例えばＡＰ２１１リソースの稼動状態はＡＰｄｏｇ７００が定期的に更新されていることを確認することで認識し、外部通信Ｉ／Ｆ１０４リソースの異常はＡＰ２１１がＡＰｄｏｇ７００に出力した端末２との外部通信異常情報を取得することで行い、論理グループ関係にある他のＶＭ２００リソースの異常はＡＰ２１１がＡＰｄｏｇ７００に出力した他のＡＰ２１１との内部通信異常情報を取得することで行う。

リソース状態の変化を検知した多重系制御エージェントプログラム２１２は、リソース状態を多重系制御基盤プログラム３００に通知する。

＜図１２：リソース監視処理のフローチャート＞

図１２はリソース監視処理Ｓ１０６の一例を示すフローチャートである。多重系制御エージェントプログラム２１２は監視対象リソースに異常を検出すると、異常状態を多重系制御基盤プログラム３００に通知する。

【００８３】

多重系制御エージェントプログラム２１２は起動すると、初めにＡＰｄｏｇ７００を参照してリソース状態の監視を開始し（Ｓ２０１）、監視対象リソースの異常の有無を判断する（Ｓ２０２）。

【００８４】

監視対象リソースの内のいずれかに異常が認められた場合（Ｓ２０２の結果が「Ｙ」）、ＶＭ２００内に保持されているリソース管理テーブル３０３の該当リソースのリソース

生死状態フラグ３３４に死状態である「１」を設定し（Ｓ２０３）、リソース管理テーブル３０３をプログラム間通信フォーマット６００のリソース管理テーブル内容６０４に設定して通信データを作成して多重系制御基盤プログラム３００に送信して（Ｓ２０４）、リソース監視処理Ｓ１０６を終了する（Ｓ２０５）。

【００８５】

監視対象リソースの内のいずれにも異常が認められない場合（Ｓ２０２の結果が「Ｎ」）は、リソース監視処理Ｓ１０６を終了する（Ｓ２０５）。

【００８６】

リソース監視処理Ｓ１０６を終了した後、多重系制御エージェントプログラム２１２は、多重系制御基盤プログラム３００からの主系ＶＭ／従系ＶＭの系切替要求の有無を確認するために図１０に示したＳ１０４を起動する。

＜図１３：系切替通知処理のフローチャート＞

図１３は本発明の実施の形態の系切替通知処理Ｓ１１６のフローチャートである。多重系制御基盤プログラム３００は系切替通知処理Ｓ１１６で多重系制御エージェントプログラム２１２からのＶＭ２００状態の通知を取得し、計算機ノード監視部３０５に主系ＶＭ／従系ＶＭの系切替要求を行う。

【００８７】

多重系制御基盤プログラム３００は系切替通知処理Ｓ１１６を開始すると、初めに多重系制御エージェントプログラム２１２からのＶＭ２００状態の通知、及び計算機ノード１００やホストＯＳ４００からのハードウェア状態検出の通知を待つ（Ｓ３０１）。

【００８８】

ＶＭ２００状態やハードウェア状態の通知内容を解釈し、計算機ノード１００内のあるＶＭ２００に新たに異常が認められた場合（Ｓ３０２の結果が「Ｙ」）、ＶＭ状態管理テーブル３０２の該当ＶＭ２００の主／従系状態フラグ３２５には従系ＶＭ状態を意味する「Ｓ」を設定して、生死状態フラグ３２６には死状態を意味する「１」を設定する（Ｓ３１１）。

【００８９】

そしてＶＭ状態管理テーブル３０２の障害ＶＭ総数３２２の値を「１」増加させる（Ｓ３１２）。

【００９０】

次に多重系制御基盤プログラム３００は計算機ノード構成管理テーブル３０１の自計算機ノード１００に該当する主／従系状態フラグ３１２を参照し、自計算機ノード１００の主系ノード／従系ノード状態を判別する（Ｓ３１３）。

【００９１】

主／従系状態フラグ３１２の値が「Ｓ」で自計算機ノード１００が従系ノード状態である場合（Ｓ３１３の結果が「Ｎ」）、多重系制御基盤プログラム３００は何も処理することなく、再びＶＭ２００状態やハードウェア状態の通知を待つためにＳ３０１を起動する。

【００９２】

主／従系状態フラグ３１２の値が「Ｍ」で自計算機ノード１００が主系ノード状態である場合（Ｓ３１３の結果が「Ｙ」）、多重系制御基盤プログラム３００は異常ＶＭ２００の主系ＶＭ／従系ＶＭの系切替要求を行うために計算機ノード間通信フォーマット５００の通信パケットを作成する（Ｓ３３１）。

【００９３】

多重系制御基盤プログラム３００は、異常ＶＭ２００のリソース障害内容が「過半数判断ロジック」の適用を要するものである場合には、ＶＭ状態管理テーブル３０２の障害ＶＭ総数３２２と異常ＶＭ２００の数量の和が管理ＶＭ総数３２１の半数以上であるかの判断を行う（Ｓ３３２）。

【００９４】

障害ＶＭ２００の総数が過半数となる場合には（Ｓ３３２の結果が「Ｙ」）、正常なＶ

Ｍ２００も含めて全てのＶＭ２００を系切替対象とするために、計算機ノード間通信フォーマット５００の通信パケットの通信種類５０４には全ＶＭ５０４－Ｄの識別値を設定する（Ｓ３４１）。

【００９５】

障害ＶＭ２００の総数が過半数とならない場合には（Ｓ３３２の結果が「Ｎ」）、異常ＶＭ２００のみを系切替対象とするために、計算機ノード間通信フォーマット５００の通信パケットの通信種類５０４には異常ＶＭ５０４－Ｃの識別値を設定する（Ｓ３５１）。

【００９６】

計算機ノード間通信フォーマット５００の通信パケットの作成が完了すると、多重系制御基盤プログラム３００は計算機ノード監視部３０５に送信要求を行う（Ｓ３４２）。

【００９７】

次に異常ＶＭ２００を停止させるためにリセットを行い（Ｓ３４３）、リソース管理テーブル３０３を参照して異常ＶＭ２００が使用しているリソースに対してもリセットを行い（Ｓ３４４）、停止させた異常ＶＭ２００を再起動させるためにＳ１１５を起動する。

【００９８】

ＶＭ２００状態やハードウェア状態の通知内容から計算機ノード１００内のあるＶＭ２００に異常が認められない場合（Ｓ３０２の結果が「Ｎ」）、当該ＶＭ２００は「死状態」から回復したものであるかを確認するために、ＶＭ状態管理テーブル３０２に記録された当該ＶＭ２００の生死状態フラグ３２６の値を参照する（Ｓ３２１）。

【００９９】

生死状態フラグ３２６の値が「０」で当該ＶＭ２００が「正状態」であると判断されていた場合には（Ｓ３２１の結果が「Ｎ」）、当該ＶＭ２００の状態に変化は起きていないため、多重系制御基盤プログラム３００は再び多重系制御エージェントプログラム２１２からの通知や計算機ノード１００／ホストＯＳ４００からのハードウェア状態検出の通知を待つために、Ｓ３０１を起動する。

【０１００】

生死状態フラグ３２６の値が「１」で当該ＶＭ２００が「死状態」であると判断されていた場合には（Ｓ３２１の結果が「Ｙ」）、当該ＶＭ２００は状態が正常に回復したものと判断されるため、多重系制御エージェントプログラム２１２から受信したプログラム間通信フォーマット６００の通信データのリソース管理テーブル内容６０４の該当ＶＭ２００の状態を、多重系制御基盤プログラム３００内のＶＭ状態管理テーブル３０２に反映する（Ｓ３２２）。

【０１０１】

次に当該ＶＭ２００の主系ＶＭ／従系ＶＭの状態のフェイルバック要求を行うために、計算機ノード構成管理テーブル３０１の主／従系状態フラグ３１２を参照して、自計算機ノード１００の主系ノード／従系ノード状態を判別する（Ｓ３２３）。

【０１０２】

主／従系状態フラグ３１２の値が「Ｓ」で自計算機ノード１００が従系ノードであった場合には（Ｓ３２３の結果が「Ｎ」）、従系ノードが自らフェイルバック要求は行わないため、再び多重系制御エージェントプログラム２１２からの通知や計算機ノード１００／ホストＯＳ４００からのハードウェア状態検出の通知を待つためにＳ３０１を起動する。

【０１０３】

主／従系状態フラグ３１２の値が「Ｍ」で自計算機ノード１００が主系ノードであった場合には（Ｓ３２３の結果が「Ｙ」）、計算機ノード間通信フォーマット５００の通信パケットを作成し（Ｓ３２４）、フェイルバック要求であるため通信種類５０４にはＦａｉｌＢａｃｋ５０４－Ｂの識別値を設定し（Ｓ３２５）、計算機ノード監視部３０５に作成した通信パケットの送信要求を行い（Ｓ３２６）、再び多重系制御エージェントプログラム２１２からの通知や計算機ノード１００／ホストＯＳ４００からのハードウェア状態検出の通知を待つためにＳ３０１を起動する。

＜図１４：メッセージ送信処理のフローチャート＞

図１４はメッセージ送信処理Ｓ１１２のフローチャートである。計算機ノード監視部３０５は多重系制御基盤プログラム３００から受けた他計算機ノード１００への「フェイルバック」要求、「異常ＶＭ切替」要求、「全ＶＭ切替」要求を、メッセージ送信処理Ｓ１１２で発信する。

【０１０４】

主系ＶＭ／従系ＶＭの「全ＶＭ切替」が実行される場合、切替対象となるＶＭ２００の中には本来切替を必要としない正常状態のＶＭ２００が含まれることもある。「全ＶＭ切替」を実行するために、正常ＶＭ２００の系切替を行うための事前処理はメッセージ送信処理Ｓ１１２内で実行する。

【０１０５】

計算機ノード監視部３０５はメッセージ送信処理Ｓ１１２を開始すると、初めに通信要求の種類が「全ＶＭ切替」とその他を特定するために、計算機ノード間通信フォーマット５００の通信要求パケットの通信種類５０４を確認する（Ｓ４０１）。

【０１０６】

通信種類５０４の値が全ＶＭ５０４－Ｄの識別値ではなく通信要求が「全ＶＭ切替」以外の場合（Ｓ４０１の結果が「Ｎ」）、系切替対象は状態が異常に転じた、または状態が回復したＶＭ２００に限られるため、Ｓ４１１を起動して通信要求パケットを仮想化多重系システム１内の他の計算機ノード１００にマルチキャスト送信する。

【０１０７】

通信種類５０４の値が全ＶＭ５０４－Ｄの識別値であり通信要求が「全ＶＭ切替」の場合は（Ｓ４０１の結果が「Ｙ」）、正常状態のＶＭ２００を系切替するための処理を起動する。

【０１０８】

計算機ノード監視部３０５はＶＭ状態管理テーブル３０２の主／従系状態フラグ３２５の値が「Ｍ」の主系ＶＭのＶＭ２００を検索する（Ｓ４０２）。

【０１０９】

検索された該当ＶＭ２００に系切替を要求するために、プログラム間通信フォーマット６００の系切替要求６０５に「１」を設定した通信データを作成し（Ｓ４０３）、該当ＶＭ２００の多重系制御エージェントプログラム２１２に対して送信する（Ｓ４０４）。

【０１１０】

次にＶＭ状態管理テーブル３０２の全てのＶＭ２００の主／従系状態フラグ３２５に従系ＶＭを意味する「Ｓ」を設定し、また生死状態フラグ３２６を死状態である「１」に変更する（Ｓ４０５）。

【０１１１】

そしてＶＭ状態管理テーブル３０２の障害ＶＭ総数３２２の値を「０」にリセットして（Ｓ４０６）、「全ＶＭ切替」の通信要求パケットを仮想化多重系システム１内の他の計算機ノード１００にマルチキャスト送信する（Ｓ４１１）。

【０１１２】

Ｓ４１１で通信要求パケットを送信し終えた後は、メッセージ送信処理Ｓ１１２を終了し（Ｓ４２１）、次のノード監視処理Ｓ１２３を起動する。

＜図１５：ノード監視処理のフローチャート＞

図１５はノード監視処理Ｓ１２３のフローチャートである。主系ノード及び従系ノードの計算機ノード１００内で動作する多重系制御基盤プログラム３００の計算機ノード監視部３０５の間では、「生死監視」及び「フェイルバック」、「異常ＶＭ切替」、「全ＶＭ切替」要求の種別の計算機ノード間通信が行われる。計算機ノード監視部３０５は受信した計算機ノード間通信データの種別を判別し、種別に応じた処理の起動を行う。

【０１１３】

計算機ノード監視部３０５はノード監視処理Ｓ１２３の処理を開始すると、初めに他の計算機ノード１００からの計算機ノード間通信パケットを受信しているかを判別する。従系ノードの計算機ノード監視部３０５は、「生死監視」に対する応答、および「異常ＶＭ

切替」、「全ＶＭ切替」、「フェイルバック」要求の計算機ノード間通信データを受信する。主系ノードの計算機ノード監視部３０５は、「生死監視」の計算機ノード１００状態取得要求、および「異常ＶＭ切替」、「全ＶＭ切替」、「フェイルバック」に対する応答の計算機ノード間通信データを受信する。（Ｓ５０１）。

【０１１４】

計算機ノード間通信パケットを受信している場合（Ｓ５０１の結果が「Ｙ」）、計算機ノード監視部３０５は計算機ノード構成管理テーブル３０１の主／従系状態フラグ３１２を参照し、自計算機ノード１００の主系ノード／従系ノード状態を判別する。主／従系状態フラグ３１２が「Ｍ」で自計算機ノード１００が主系ノードの場合（Ｓ５０２の結果が「Ｙ」）は主系処理Ｓ５０３を起動し、主／従系状態フラグ３１２が「Ｓ」で自計算機ノード１００が従系ノードの場合（Ｓ５０２の結果が「Ｎ」）は従系処理Ｓ５０４を起動する。（Ｓ５０２）

主系処理Ｓ５０３または従系処理Ｓ５０４が終了すると、計算機ノード監視部３０５はノード監視処理Ｓ１２３を終了する（Ｓ５３１）。

【０１１５】

計算機ノード間通信データを受信していない場合（Ｓ５０１の結果が「Ｎ」）、計算機ノード監視部３０５は計算機ノード構成管理テーブル３０１の主／従系状態フラグ３１２を参照し、自計算機ノード１００の主系ノード／従系ノード状態を判別する。

【０１１６】

主／従系状態フラグ３１２が「Ｍ」で自計算機ノード１００が主系ノードの場合（Ｓ５１１の結果が「Ｎ」）はノード監視処理Ｓ１２３を終了し、主／従系状態フラグ３１２が「Ｓ」で自計算機ノード１００が従系ノードの場合（Ｓ５１１の結果が「Ｙ」）は「生死監視」要求に対する応答が返ってこないことであるため次のＳ５１２を起動する。（Ｓ５１１）

Ｓ５１２では、「生死監視」要求に対する応答の無い計算機ノード１００の計算機ノード構成管理テーブル３０１の生死確定カウンタ３１６の値と、計算機ノード間通信フォーマット５００の応答異常カウント５０７の値を比較し、当該計算機ノード１００の生死状態の判別を行う。

【０１１７】

応答異常カウント５０７の値が生死確定カウンタ３１６の値未満の場合（Ｓ５１２の結果が「Ｎ」）は当該計算機ノード１００がまだ正常状態であると判断して生死監視処理Ｓ５１３を起動し、応答異常カウント５０７の値が生死確定カウンタ３１６の値以上の場合（Ｓ５１２の結果が「Ｙ」）は当該計算機ノード１００が異常状態にあると判断して全ＶＭ切替処理Ｓ５２１を起動する。

【０１１８】

生死監視処理Ｓ５１３または全ＶＭ切替処理Ｓ５２１が終了すると、計算機ノード監視部３０５はノード監視処理Ｓ１２３を終了する（Ｓ５３１）。

【０１１９】

ノード監視処理Ｓ１２３の終了後、計算機ノード監視部３０５は多重系制御基盤プログラム３００からの送信要求の有無判定Ｓ１２１を起動する。

＜図１６：生死監視処理のフローチャート＞

図１６は本発明の実施の形態の生死監視処理Ｓ５１３のフローチャートである。従系ノードである計算機ノード１００の多重系制御基盤プログラム３００の計算機ノード監視部３０５は生死監視処理Ｓ５１３で「生死監視」要求を実行する。

【０１２０】

従系ノードの計算機ノード監視部３０５は生死監視処理Ｓ５１３を起動すると、計算機ノード構成管理テーブル３０１の主系ノードである計算機ノード１００の監視周期３１５を参照する（Ｓ６０１）。

【０１２１】

前回の「生死監視」要求からの経過時間が監視周期３１５を超えていない場合（Ｓ６０

１の結果が「Ｎ」）は生死監視処理Ｓ５１３を終了する。

【０１２２】

経過時間が監視周期３１５を超えている場合（Ｓ６０１の結果が「Ｙ」）は「生死監視」要求を行うために次のＳ６０２を起動する。

【０１２３】

主系ノードである計算機ノード１００に送る計算機ノード間通信フォーマット５００のパケットを作成するために、自計算機ノード１００のＶＭ状態管理テーブル３０２をＶＭ状態管理テーブル内容５０５に、リソース管理テーブル３０３の内容をリソース管理テーブル内容５０６に設定する（Ｓ６０２）。そして通信種類５０４には生死５０４－Ａの識別値を設定する（Ｓ６０３）。

【０１２４】

次に作成したパケットを、仮想化多重系システム１内の主系ノードである計算機ノード１００に対してマルチキャストで送信する（Ｓ６０４）。

【０１２５】

「生死監視」要求のパケット送信が完了すると生死監視処理Ｓ５１３を終了する。

＜図１７：全ＶＭ切替処理のフローチャート＞

図１７は全ＶＭ切替処理Ｓ５２１のフローチャートである。従系ノードである計算機ノード１００で動作する多重系制御基盤プログラム３００は、全ＶＭ切替処理Ｓ５２１を実行して、当該計算機ノード１００内の全てのＶＭ２００を主系ＶＭ状態に、当該計算機ノード１００を主系ノード状態に変更する。

【０１２６】

多重系制御基盤プログラム３００は全ＶＭ切替処理Ｓ５２１を起動すると、初めに計算機ノード構成管理テーブル３０１の自計算機ノード１００に該当する主／従系状態フラグ３１２に「Ｍ」を設定し、主系ノード状態に変更する（Ｓ７０１）。

【０１２７】

次に自計算機ノード１００内の全てのＶＭ２００の系切替を行うために、プログラム間通信フォーマット６００の系切替要求６０５に系切替の実行を指定する「１」を設定した、通信データを作成する（Ｓ７０２）。

【０１２８】

次にＳ７０３で自計算機ノード１００内の全てのＶＭ２００の多重系制御エージェントプログラム２１２に対して作成したプログラム間通信フォーマット６００の通信データを送信した後、Ｓ７０４で全ＶＭ切替処理Ｓ５２１を終了する。

＜図１８：主系処理のフローチャート＞

図１８は主系処理Ｓ５０３のフローチャートである。ノード監視処理Ｓ１２３のＳ５０２から起動された主系処理Ｓ５０３は、従系ノードである計算機ノード１００からの「生死監視」要求、および「フェイルバック」、「異常ＶＭ切替」、「全ＶＭ切替」の応答に対する処理を行う。

【０１２９】

処理を開始した主系処理Ｓ５０３は、初めに計算機ノード間通信フォーマット５００の受信パケットの通信種類５０４を参照し（Ｓ８０１）、「生死監視」要求の受信パケットであるかを判別する。

【０１３０】

受信パケットの通信種類５０４が生死５０４－Ａの識別値である場合（Ｓ８０２の結果が「Ｙ」）、Ｓ８１１を起動して自計算機ノード１００の状態の返信処理を開始する。受信パケットの通信種類５０４が生死５０４－Ａの識別値でなかった場合（Ｓ８０２の結果が「Ｎ」）、受信パケットの種別を更に特定するためにＳ８２１を起動する。（Ｓ８０２）

Ｓ８１１では計算機ノード間通信フォーマット５００の返信パケットに自計算機ノード１００のＶＭ状態管理テーブル３０２の内容をＶＭ状態管理テーブル内容５０５に設定し、リソース管理テーブル３０３の内容をリソース管理テーブル内容５０６に設定する。

【０１３１】
次に作成した「生死監視」要求に対する返信パケットを、Ｓ８１２で従系ノードである計算機ノード１００にマルチキャストで送信した後、主系処理Ｓ５０３を終了する。
【０１３２】
Ｓ８２１では受信パケットが「フェイルバック」の応答であるかを判別する。受信パケットの通信種類５０４がＦａｉｌＢａｃｋ５０４－Ｂの識別値である場合（Ｓ８２１の結果が「Ｙ」）、主系ノードである計算機ノード１００での「フェイルバック」処理を行うためにＳ８３１を起動する。受信パケットの通信種類５０４がＦａｉｌＢａｃｋ５０４－Ｂの識別値でない場合（Ｓ８２１の結果が「Ｎ」）、受信パケットの種別を更に特定するためにＳ８２２を起動する。
【０１３３】
Ｓ８３１ではＶＭ状態管理テーブル３０２の主／従系状態フラグ３２５を参照して、値が「Ｓ」である従系ＶＭのＶＭ２００を探索し、これを主系ＶＭへの系切替対象とする。
【０１３４】
次にプログラム間通信フォーマット６００の系切替要求６０５に系切替要求である「１」を設定した通信データを作成する（Ｓ８３２）。
【０１３５】
次のＳ８３３において、Ｓ８３１で探索した系切替対象のＶＭ２００の多重系制御エージェントプログラム２１２にＳ８３２で作成した通信データを送信した後、主系処理Ｓ５０３を終了する。
【０１３６】
Ｓ８２２では受信パケットが「全ＶＭ切替」の応答であるかを判別する。受信パケットの通信種類５０４が全ＶＭ５０４－Ｄの識別値である場合（Ｓ８２２の結果が「Ｙ」）、Ｓ８２３を起動して計算機ノード構成管理テーブル３０１の自計算機ノード１００に該当する主／従系状態フラグ３１２に「Ｓ」を設定することで自計算機ノード１００を従系ノード状態に変更した後、主系処理Ｓ５０３を終了する。
【０１３７】
受信パケットの通信種類５０４が全ＶＭ５０４－Ｄの識別値でない場合（Ｓ８２２の結果が「Ｎ」）は主系処理Ｓ５０３を終了する。
＜図１９：従系処理のフローチャート＞
図１９は従系処理Ｓ５０４のフローチャートである。ノード監視処理Ｓ１２３のＳ５０２から起動された従系処理Ｓ５０４は、主系ノードである計算機ノード１００からの「生死監視」への応答、及び「フェイルバック」、「異常ＶＭ切替」、「全ＶＭ切替」要求に対する処理を行う。
【０１３８】
処理を開始した従系処理Ｓ５０４は、初めに計算機ノード間通信フォーマット５００の受信パケットの通信種類５０４を参照し（Ｓ９０１）、「生死監視」への応答の受信パケットであるかを判別する。受信パケットの通信種類５０４が生死５０４－Ａの識別値である場合（Ｓ９０２の結果が「Ｙ」）はＳ９１１を起動して「生死監視」の処理を行い、通信種類５０４が生死５０４－Ａの識別値で無い場合（Ｓ９０２の結果が「Ｎ」）は受信パケットの種別を更に特定するためにＳ９２１を起動する。（Ｓ９０２）
Ｓ９１１において、「生死監視」要求に対する返答が合ったことから主系ノードの計算機ノード１００は正常状態にあると判断できるため、計算機ノード間通信フォーマット５００の通信パケットの応答異常カウント５０７の値を「０」にリセットした後、従系処理Ｓ５０４を終了する。
【０１３９】
Ｓ９２１では受信パケットが「フェイルバック」要求であるかを判別する。受信パケットの通信種類５０４がＦａｉｌＢａｃｋ５０４－Ｂの識別値である場合（Ｓ９２１の結果が「Ｙ」）、従系ノードである計算機ノード１００での「フェイルバック」処理を行うためにＳ９４１を起動する。受信パケットの通信種類５０４がＦａｉｌＢａｃｋ５０４－Ｂ

の識別値で無い場合（Ｓ９２１の結果が「Ｎ」）、受信パケットの種別を更に特定するためにＳ９２２を起動する。

【０１４０】

Ｓ９４１では計算機ノード間通信フォーマット５００の受信パケットのＶＭ状態管理テーブル内容５０５に格納されているＶＭ状態管理テーブル３０２情報を参照して、主／従系状態フラグ３２５の値が「Ｓ」である従系ＶＭ状態にある主系ノードの計算機ノード１００内のＶＭ２００を検索する。

【０１４１】

次に、自計算機ノード１００の識別子及び自計算機ノード１００内のＶＭ２００の識別子を、Ｓ９４１で検索したＶＭ２００の主／従系構成ＶＭ情報３２７のホスト計算機ノード識別子３２７－Ａ及びＶＭ識別子３２７－Ｂと比較し、系切替に該当する自計算機ノード１００内のＶＭ２００を検索する（Ｓ９４２）。

【０１４２】

次に該当ＶＭ２００に系切替要求を行うために、Ｓ９４３でプログラム間通信フォーマット６００の系切替要求６０５に系切替要求である「１」を設定した通信データを作成し、Ｓ９４４で該当ＶＭ２００の多重系制御エージェントプログラム２１２に作成した通信データを送信する。

【０１４３】

次に主系ノードの計算機ノード１００に「フェイルバック」の応答を行うために、Ｓ９４５で自計算機ノード１００のＶＭ状態管理テーブル３０２とリソース管理テーブル３０３の内容を計算機ノード間通信フォーマット５００のＶＭ状態管理テーブル内容５０５とリソース管理テーブル内容５０６に格納した通信パケットを作成し、次のＳ９４６では通信パケットの通信種類５０４にＦａｉｌＢａｃｋ５０４－Ｂの識別値を設定する。

【０１４４】

次のＳ９４７で、作成した通信パケットを仮想化多重系システム１内の主系ノードである計算機ノード１００にマルチキャストで送信した後、従系処理Ｓ５０４を終了する。

【０１４５】

Ｓ９２２では受信パケットが「全ＶＭ切替」要求であるかを判別する。受信パケットの通信種類５０４が全ＶＭ５０４－Ｄの識別値である場合（Ｓ９２２の結果が「Ｙ」）、従系ノードである計算機ノード１００での「全ＶＭ切替」処理を行うために全ＶＭ切替処理Ｓ５２１を起動する。受信パケットの通信種類５０４が全ＶＭ５０４－Ｄの識別値で無い場合（Ｓ９２２の結果が「Ｎ」）、従系ノードである計算機ノード１００での「異常ＶＭ切替」処理を行うためにＳ９２３を起動する。

【０１４６】

全ＶＭ切替処理Ｓ５２１の終了後、主系ノードの計算機ノード１００に「全ＶＭ切替」の応答を行うために、Ｓ９３２で自計算機ノード１００のＶＭ状態管理テーブル３０２とリソース管理テーブル３０３の内容を計算機ノード間通信フォーマット５００のＶＭ状態管理テーブル内容５０５とリソース管理テーブル内容５０６に格納した通信パケットを作成し、次のＳ９３３では通信パケットの通信種類５０４に全ＶＭ５０４－Ｄの識別値を設定する。

【０１４７】

次のＳ９３４で、作成した通信パケットを仮想化多重系システム１内の主系ノードである計算機ノード１００にマルチキャストで送信した後、従系処理Ｓ５０４を終了する。

【０１４８】

Ｓ９２３では、「異常ＶＭ切替」処理を行うために、計算機ノード間通信フォーマット５００の受信パケットのＶＭ状態管理テーブル内容５０５に格納されているＶＭ状態管理テーブル３０２情報を参照して、主／従系状態フラグ３２５の値が「Ｓ」である従系ＶＭ状態にある主系ノードの計算機ノード１００内のＶＭ２００を検索する。

【０１４９】

次に、自計算機ノード１００の識別子及び自計算機ノード１００内のＶＭ２００の識別

子を、Ｓ９２３で検索したＶＭ２００の主／従系構成ＶＭ情報３２７のホスト計算機ノード識別子３２７－Ａ及びＶＭ識別子３２７－Ｂと比較し、系切替に該当する自計算機ノード１００内のＶＭ２００を検索する（Ｓ９２４）。

【０１５０】

次に該当ＶＭ２００に系切替要求を行うために、Ｓ９２５でプログラム間通信フォーマット６００の系切替要求６０５に系切替要求である「１」を設定した通信データを作成し、次のＳ９２６で該当ＶＭ２００の多重系制御エージェントプログラム２１２に作成した通信データを送信し、従系処理Ｓ５０４を終了する。

【０１５１】

以上の実施形態によれば、物理計算機上で障害検知した主系モードの仮想マシンが搭載済みの仮想マシンよりも半数未満の場合には、仮想マシン毎に対して系切替え（フェイルオーバ）を行った後にフェイルバックする。一方、物理計算機上で障害検知した主系モードの仮想マシンが搭載済みの仮想マシンよりも半数以上の場合には、全ての仮想マシンに対して系切替えを行う。以上によって、単一の主系モードの仮想マシンが障害を起こしたとしても、複数の主系モードの仮想マシンが障害を起こしたとしても、つまり、いずれの場合にも、最終的には物理計算機の片方に全ての仮想マシンが主系モードで稼働するようになり、利用者にとって物理計算機の全体の動作モードが管理しやすくなる。その結果、物理計算機上の全ての仮想マシンを主系モードで稼働させながら、もう一方の物理計算機に対してハードウェアやソフトウェアの保守を行うことが可能となる。

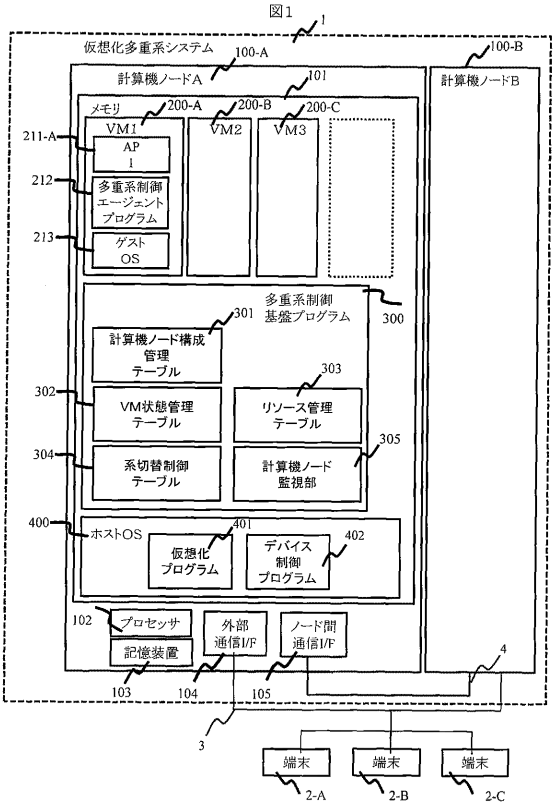
【符号の説明】

【０１５２】

１…仮想化多重系システム
２…端末
３…外部ネットワーク
４…内部ネットワーク
１００…計算機ノード
１０１…メモリ
１０２…プロセッサ
１０３…記憶装置、
１０４…外部通信Ｉ／Ｆ
１０５…ノード間通信Ｉ／Ｆ
２００…ＶＭ
２０１…ＶＭ
２１１…ＡＰ
２１２…多重系制御エージェントプログラム
２１３…ゲストＯＳ
３００…多重系制御基盤プログラム
３０１…計算機ノード構成管理テーブル
３０２…ＶＭ状態管理テーブル
３０３…リソース管理テーブル
３０４…系切替制御テーブル
３０５…計算機ノード監視部
４００…ホストＯＳ
４０１…仮想化プログラム
４０２…デバイス制御プログラム
７００…ＡＰｄｏｇ

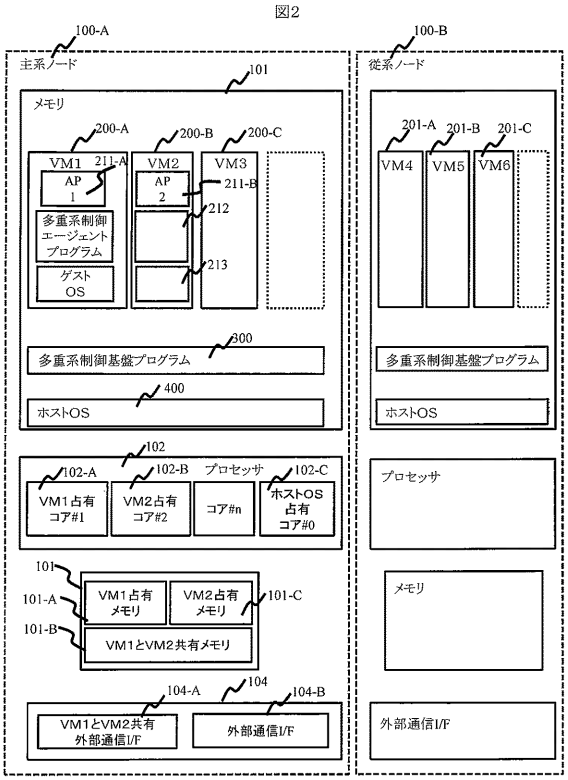
【図１】

図1



【図２】

図2



【図３】

図3

計算機ノード構成管理テーブル 301

311 計算機ノード識別子	312 主／従系状態フラグ	313 マルチキャストアドレス	314 生死状態フラグ	315 監視周期[ms]	316 生死確定カウンタ
100-A	M	226.94.1.1	0	1000	3
100-B	S	226.94.1.1	0	1000	3

【図４】

図4

VM状態管理テーブル 302

321 管理VM総数	3		322 障害VM総数	0	
323 VM識別子	324 ホスト計算機ノード識別子	325 主／従系状態フラグ	326 生死状態フラグ	327 主／従系構成VM情報 327-A ホスト計算機ノード識別子	327-B VM識別子
328 VM1	100-A	M	0	100-B	VM4
VM2	100-A	M	0	100-B	VM5
VM3	100-A	M	0	100-B	VM6

【図５】

図5

リソース管理テーブル 303

331 VM識別子	332 リソース識別子 211	333 監視周期[ms]	334 リソース生死状態フラグ	335 リソース詳細情報 335-A 所有状態	335-B 個別情報
336 VM1	AP1	100	0	占有	212
337 VM1	101-A	0	0	占有	EVT
338 VM1	101-B	0	0	共有	EVT
VM1	104-A	50	0	共有	212
VM1	VM2	100	0	論理/SERVER	212
339 VM2	AP2	100	0	占有	212
VM2	101-C	0	0	占有	EVT
VM2	101-B	0	0	共有	EVT
VM2	104-A	50	0	共有	212
340 VM2	VM1	100	0	論理/CLIENT	212

【図６】

図6

系切替制御テーブル 304

341 障害内容	342 系切替対象VM	343 過半数判断ロジック適用要否	344 Failback処理適用可否
計算機ノード障害	全VM	否	否
VM占有リソース障害	該当VMのみ	要	可
VM共有リソース障害	関連全VM	要	可
VM論理リソース障害	関連全VM	要	可

【図７】

図7

304

VM1障害内容と系切替制御内容例

	障害内容		系切替対象VM	過半数判断ロジック適用要否	Failback処理適用可否
351	計算機ノード障害		全VM切替	否	否
352	VM1占有リソース障害	AP1	VM1切替	要	可
		101-A			
353	VM1共有リソース障害	101-B	VM1、VM2切替	要	可
		104-A			
354	VM1論理リソース障害	VM2	VM1、VM2切替	要	可

【図８】

図8

500

計算機ノード間通信フォーマット

501 計算機ノード識別子	502 シーケンス番号	503 送信タイムスタンプ	504 通信種類				505 VM状態管理テーブル内容	506 リソース管理テーブル内容	507 応答異常カウント
			504-A 生死	系切替					
				Fail Back 504-B	異常VM 504-C	全VM 504-D			
200-A	001	HHMMSS					302	303	0

【図９】

図9

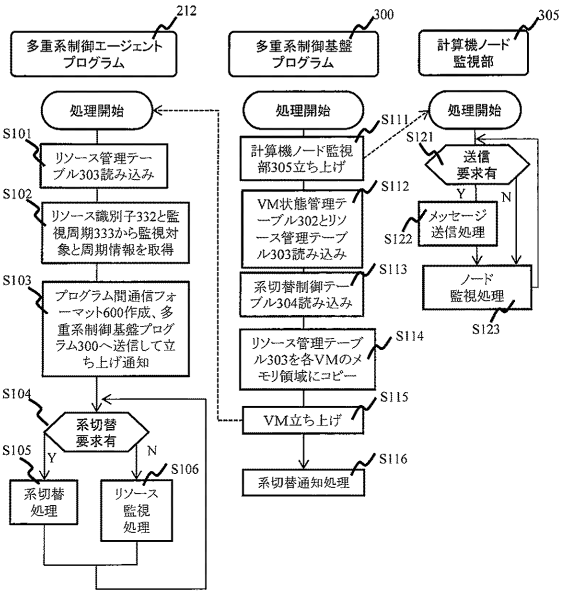
600

プログラム間通信フォーマット

601 VM識別子	602 シーケンス番号	603 送信タイムスタンプ	604 リソース管理テーブル内容	605 系切替要求
323	001	HHMMSS (例:081203)	303	0

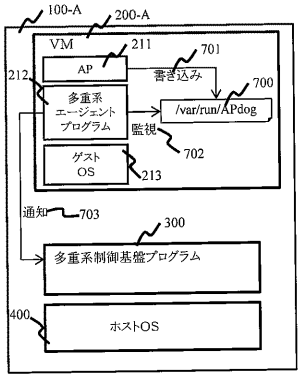
【図１０】

図10



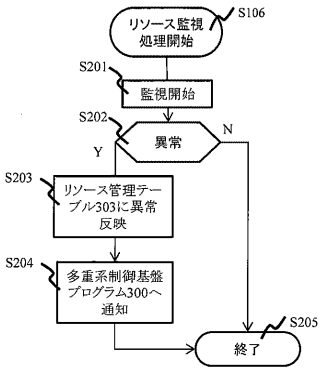
【図１１】

図11



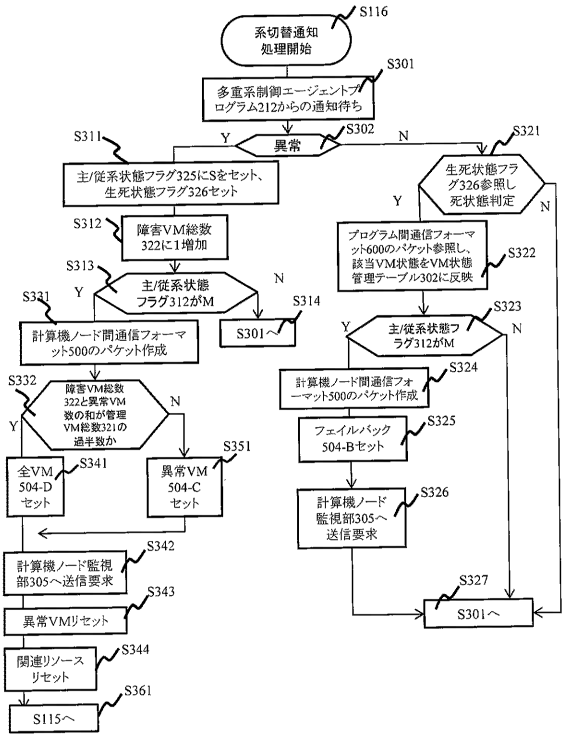
【図１２】

図12



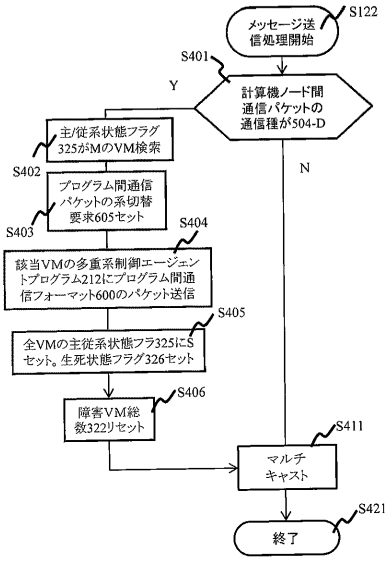
【図１３】

図13



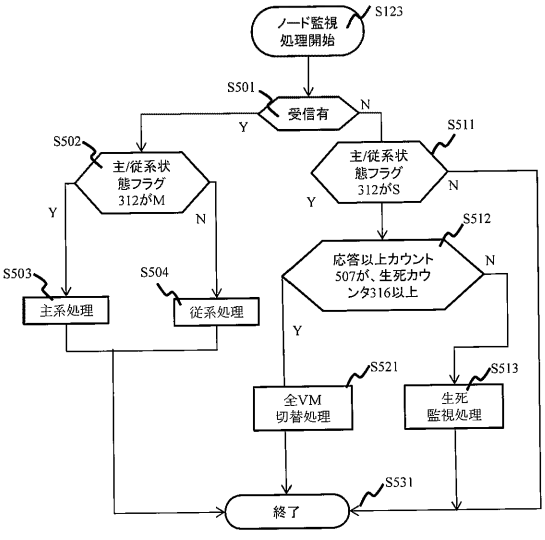
【図１４】

図14



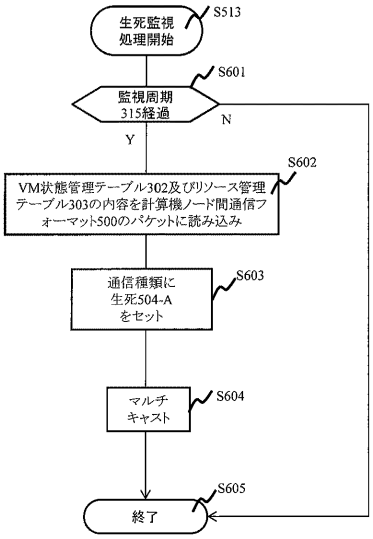
【図１５】

図15



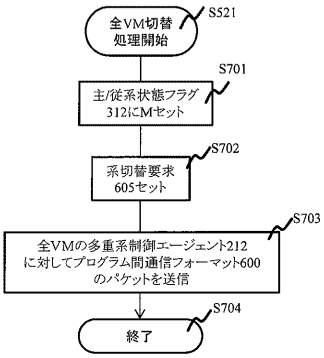
【図１６】

図16



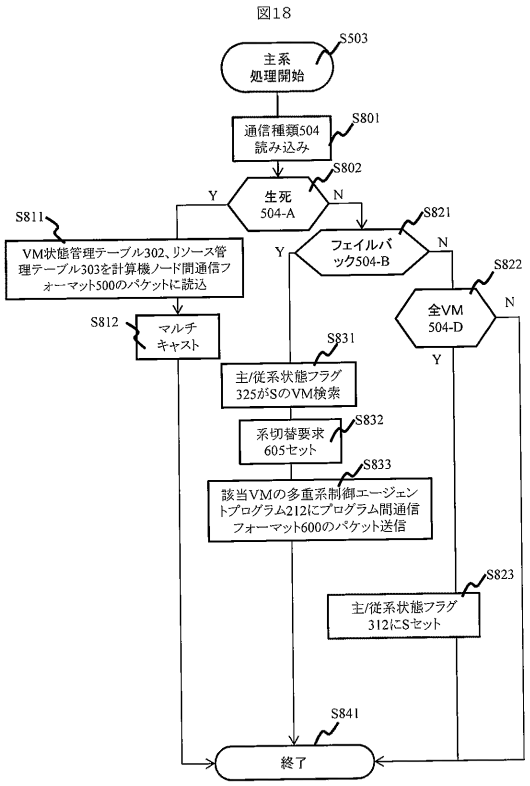
【図１７】

図17



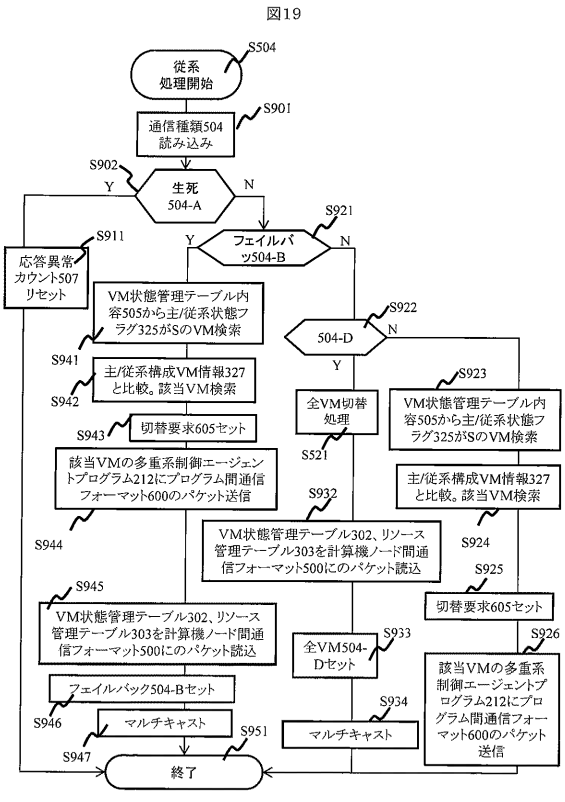
【図１８】

図18



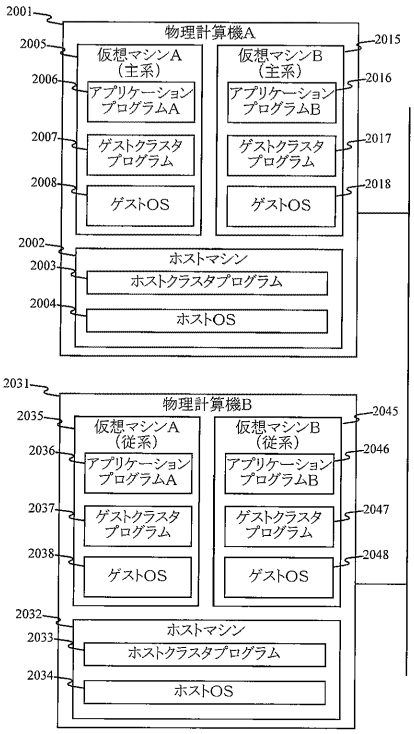
【図１９】

図19



【図２０】

図20


2001
物理計算機A
2005
仮想マシンA
（主系）
2015
仮想マシンB
（主系）
2006
アプリケーション
プログラムA
2016
アプリケーション
プログラムB
2007
ゲストクラスタ
プログラム
2017
ゲストクラスタ
プログラム
2008
ゲストOS
2018
ゲストOS
2002
ホストマシン
2003
ホストクラスタプログラム
2004
ホストOS
2031
物理計算機B
2035
仮想マシンA
（従系）
2045
仮想マシンB
（従系）
2036
アプリケーション
プログラムA
2046
アプリケーション
プログラムB
2037
ゲストクラスタ
プログラム
2047
ゲストクラスタ
プログラム
2038
ゲストOS
2048
ゲストOS
2032
ホストマシン
2033
ホストクラスタプログラム
2034
ホストOS

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2010/006654

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/20(2006.01)i, *G06F9/46*(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)
G06F11/16-11/20, G06F9/46-9/54

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 1996-2011
Kokai Jitsuyo Shinan Koho 1971-2011 Toroku Jitsuyo Shinan Koho 1994-2011

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 A	JP 2008-269332 A (Hitachi, Ltd.), 06 November 2008 (06.11.2008), paragraphs [0023] to [0036]; fig. 1, 2 & US 2008/0263390 A1	1,5-8,12-14 2-4,9-11
Y A	JP 2009-211517 A (NEC Corp.), 17 September 2009 (17.09.2009), paragraphs [0016] to [0021], [0026] to [0031]; fig. 1, 2 (Family: none)	1,5,6,8, 12,13 2-4,9-11
Y A	JP 2005-327279 A (International Business Machines Corp.), 24 November 2005 (24.11.2005), paragraph [0036] & US 2005/0268298 A1 & CN 1696902 A	1,5,7,8, 12,14 2-4,9-11

☒ 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
"E" earlier application or patent but published on or after the international filing date
"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)
"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
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 February, 2011 (03.02.11)	15 February, 2011 (15.02.11)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2010/006654

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 2009-187090 A (NEC Corp.), 20 August 2009 (20.08.2009), entire text; all drawings (Family: none)	1-14

特 許 協 力 条 約

ＰＣＴ

国際調査報告

(法８条、法施行規則第40、41条)
〔ＰＣＴ18条、ＰＣＴ規則43、44〕

出願人又は代理人 の書類記号　341050897PCT	今後の手続きについては、様式ＰＣＴ／ＩＳＡ／２２０ 及び下記５を参照すること。	
国際出願番号 ＰＣＴ／ＪＰ２０１０／００６６５４	国際出願日 (日.月.年)　１２．１１．２０１０	優先日 (日.月.年)
出願人（氏名又は名称） 株式会社日立製作所		

国際調査機関が作成したこの国際調査報告を法施行規則第41条（ＰＣＴ18条）の規定に従い出願人に送付する。
この写しは国際事務局にも送付される。

この国際調査報告は、全部で 3 ページである。

☐ この調査報告に引用された先行技術文献の写しも添付されている。

１．国際調査報告の基礎
ａ．言語に関し、この国際調査は以下のものに基づき行った。
☑ 出願時の言語による国際出願
☐ 出願時の言語から国際調査のための言語である＿＿＿＿＿＿＿＿ 語に翻訳された、
この国際出願の翻訳文（ＰＣＴ規則12.3(a)及び23.1(b)）

ｂ．☐ この国際調査報告は、ＰＣＴ規則91の規定により国際調査機関が認めた又は国際調査機関に通知された明らかな誤りの訂正を考慮して作成した（ＰＣＴ規則43.6の2(a)）。

ｃ．☐ この国際出願は、ヌクレオチド又はアミノ酸配列を含んでいる（第Ｉ欄参照）。

２．☐ 請求の範囲の一部の調査ができない（第Ⅱ欄参照）。

３．☐ 発明の単一性が欠如している（第Ⅲ欄参照）。

４．発明の名称は
☑ 出願人が提出したものを承認する。
☐ 次に示すように国際調査機関が作成した。

＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

５．要約は
☑ 出願人が提出したものを承認する。
☐ 第Ⅳ欄に示されているように、法施行規則第47条第１項（ＰＣＴ規則38.2）の規定により国際調査機関が作成した。出願人は、この国際調査報告の発送の日から１月以内にこの国際調査機関に意見を提出することができる。

６．図面に関して
ａ．要約書とともに公表される図は、
第 1 図とする。 ☑ 出願人が示したとおりである。
☐ 出願人は図を示さなかったので、国際調査機関が選択した。
☐ 本図は発明の特徴を一層よく表しているので、国際調査機関が選択した。

ｂ．☐ 要約とともに公表される図はない。

国際調査報告

国際出願番号 ＰＣＴ／ＪＰ２０１０／００６６５４

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, G06F9/46(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. G06F11/16－11/20, G06F9/46－9/54

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	１９７１－２０１１年
日本国実用新案登録公報	１９９６－２０１１年
日本国登録実用新案公報	１９９４－２０１１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
Y A	JP 2008-269332 A（株式会社日立製作所）2008.11.06, 段落[0023]－[0036]，第1,2図 & US 2008/0263390 A1	1，5－8， 12－14 2－4，9－11
Y A	JP 2009-211517 A（日本電気株式会社）2009.09.17, 段落[0016]－[0021]，[0026]－[0031]，第1,2図 （ファミリーなし）	1，5，6，8， 12，13 2－4，9－11

☑ 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 ０３．０２．２０１１	国際調査報告の発送日 １５．０２．２０１１
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） ５Ｂ ３８５６ 北元 健太 電話番号 ０３－３５８１－１１０１ 内線 ３５４５

様式ＰＣＴ／ＩＳＡ／２１０（第２ページ）（２００９年７月）

国際調査報告 国際出願番号 ＰＣＴ／ＪＰ２０１０／００６６５４

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
Y	JP 2005-327279 A（インターナショナル・ビジネス・マシーンズ・コーポレーション）2005.11.24，段落[0036]	1，5，7，8，12，14
A	& US 2005/0268298 A1 & CN 1696902 A	2－4，9－11
A	JP 2009-187090 A（日本電気株式会社）2009.08.20，全文，全図（ファミリーなし）	1－14