

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4575740号
(P4575740)

(45) 発行日 平成22年11月4日(2010.11.4)

(24) 登録日 平成22年8月27日(2010.8.27)

(51) Int.Cl.	F I
G06F 12/00 (2006.01)	G06F 12/00 545A
G06F 3/06 (2006.01)	G06F 12/00 514E
	G06F 12/00 531M
	G06F 12/00 531R
	G06F 12/00 535Z
請求項の数 9 (全 22 頁) 最終頁に続く	

(21) 出願番号	特願2004-289361 (P2004-289361)	(73) 特許権者	591008605 株式会社日本デジタル研究所 東京都江東区新砂1丁目2番3号
(22) 出願日	平成16年9月30日(2004.9.30)	(74) 代理人	100089118 弁理士 酒井 宏明
(65) 公開番号	特開2006-106901 (P2006-106901A)	(72) 発明者	土本 一生 東京都江東区新砂1-2-3 株式会社日本デジタル研究所内
(43) 公開日	平成18年4月20日(2006.4.20)	(72) 発明者	倉部 正明 東京都江東区新砂1-2-3 株式会社日本デジタル研究所内
審査請求日	平成19年9月28日(2007.9.28)	(72) 発明者	金子 正雄 東京都江東区新砂1-2-3 株式会社日本デジタル研究所内
		最終頁に続く	

(54) 【発明の名称】 仮想サーバ

(57) 【特許請求の範囲】

【請求項1】

ネットワークを介して相互に接続された複数のコンピュータと、前記各コンピュータにそれぞれ接続された記憶手段とで構築される仮想サーバにおいて、

前記各コンピュータにそれぞれ接続された記憶手段を利用して、単一の仮想記憶領域を構成し、

前記ネットワークに接続されたコンピュータの表示手段が、前記単一の仮想記憶領域を表示可能とし、

前記複数のコンピュータはそれぞれ、

コンピュータ内の各部に電源が供給される第1のモードと、バックアップ動作を実行するための部位にのみ電源が供給される第2のモードと、前記ネットワークを介して通信を行うための部位にのみ電源が供給される第3のモードである、3つのモードの間でのモード遷移を管理する電源管理手段と、

すべてのモードにおいて電源が供給され、前記ネットワークを介して他のコンピュータと通信を行うLANコントローラ手段と、

すべてのモードにおいて電源が供給され、前記電源管理手段または他のコンピュータからの指示に基づいて、自コンピュータの各部に対する電源供給を制御する電源制御手段と

前記電源制御手段の制御に基づいて、自コンピュータの各部へ電源を供給する電源供給手段と、

を備え、

前記複数のコンピュータのうちの第3のモードで動作中の特定のコンピュータの電源SWがONされた場合に、

電源SWがONされたコンピュータの電源制御手段が、自コンピュータの電源管理手段に対して電源を供給するよう自コンピュータの電源供給手段を制御することにより、当該電源管理手段を起動し、

前記電源を供給された電源管理手段が、自コンピュータ内の各部へ電源を供給するための指示を、自コンピュータの電源制御手段に対して行うことにより、自コンピュータの動作モードを第3のモードから第1のモードに遷移させ、また、第3のモードで動作中の他のコンピュータを第3のモードから第2のモードに遷移させるための信号である特殊電源ON信号を、前記LANコントローラ手段を経由してネットワーク上に送信し、

一方、ネットワークを介して特殊電源ON信号を受信した前記第3のモードで動作中の他のコンピュータの電源制御手段が、自コンピュータの電源管理手段に電源を供給するよう自コンピュータの電源供給手段を制御することにより、当該電源管理手段を起動し、

前記電源管理手段を起動した他のコンピュータの電源管理手段が、バックアップ動作を実行するための部位へ電源を供給するための指示を、自コンピュータの電源制御手段に対して行うことにより、自コンピュータの動作モードを第3のモードから第2のモードに遷移させる、

ことを特徴とする仮想サーバ。

【請求項2】

前記複数のコンピュータのうちの第1のモードで動作中のコンピュータの電源SWがOFFされた場合、または当該コンピュータより電源OFFコマンドが発行された場合に、

前記第1のモードで動作中のコンピュータにおいて、電源SWのOFFまたは電源OFFコマンドの発行を検出した電源制御手段が、その旨を自コンピュータの電源管理手段へ通知し、

前記通知を受けた電源管理手段が、自コンピュータのLANコントローラ手段を経由して、他のコンピュータに対してモード状態の問い合わせを行い、

前記他のコンピュータの電源管理手段が、自コンピュータのLANコントローラ手段を経由して、自コンピュータのモード状態を前記第1のモードで動作中のコンピュータに対して返信し、

前記第1のモードで動作中のコンピュータの電源管理手段が、前記他のコンピュータから返信されたモード状態を確認した結果、自身以外の全てのコンピュータが第2のモードのときは、前記他のコンピュータを第2のモードから第3のモードに遷移させるための信号である特殊電源OFF信号を、前記LANコントローラ手段を経由してネットワーク上に送信し、また、前記ネットワークを介して通信を行うための部位以外への電源の供給を停止するための指示を、自コンピュータの電源制御手段に対して行うことにより、自コンピュータの動作モードを第1のモードから第3のモードに遷移させ、

一方、ネットワークを介して特殊電源OFF信号を受信した前記他のコンピュータの電源管理手段が、それぞれ、前記ネットワークを介して通信を行うための部位以外への電源の供給を停止するための指示を、自コンピュータの電源制御手段に対して行うことにより、自コンピュータの動作モードを第2のモードから第3のモードに遷移させる、

ことを特徴とする請求項1に記載の仮想サーバ。

【請求項3】

前記ネットワークに接続されたコンピュータの表示手段が、前記単一の仮想記憶領域とともに、当該単一の仮想記憶領域を管理するサーバを表示することを特徴とする請求項1または2に記載の仮想サーバ。

【請求項4】

前記各記憶手段は、それぞれ特定領域を有しており、前記各特定領域で前記単一の仮想記憶領域を形成し、

前記ネットワークを介して接続される外部装置からデータの書き込みの要求があった場

10

20

30

40

50

合、または、1つのコンピュータでデータの書き込み要求が発生した場合は、前記各コンピュータは、同一のデータを前記各特定領域にそれぞれ略リアルタイムに書き込み、前記各特定領域で同一のデータを保持することを特徴とする請求項1、2または3に記載の仮想サーバ。

【請求項5】

前記ネットワークを介して接続される外部装置からデータの読み込み要求、または、自機の読み込み要求が発生した場合は、前記各コンピュータは、前記各特定領域からそれぞれ独自のタイミングでデータの読み込みを行うことを特徴とする請求項4に記載の仮想サーバ。

【請求項6】

前記記憶手段はディスク装置からなり、当該ディスク装置は複数の領域を有し、前記特定領域は、前記複数の領域のうちの1つの領域に割り当てられることを特徴とする請求項4または5に記載の仮想サーバ。

【請求項7】

前記記憶手段は、複数のディスク装置からなり、前記特定領域はディスク装置の全領域に割り当てられることを特徴とする請求項4または5に記載の仮想サーバ。

【請求項8】

前記コンピュータの前記特定領域にエラーが発生した場合には、当該エラーが発生したコンピュータからの要求に応じて、他のコンピュータが、当該特定領域のデータのリカバリーを行うことを特徴とする請求項4～7のいずれか1つに記載の仮想サーバ。

【請求項9】

前記各コンピュータのうち障害が発生したコンピュータがある場合には、当該障害が発生した以外のコンピュータで、前記仮想サーバを構成することを特徴とする請求項1～8のいずれか1つに記載の仮想サーバ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、仮想サーバ、そのコンピュータ、およびコンピュータが実行するためのプログラムに関し、詳細には、ネットワーク内で仮想的な単一の記憶領域を形成する仮想サーバ、そのコンピュータ、およびコンピュータが実行するためのプログラムに関する。

【背景技術】

【0002】

複数のクライアントとサーバを使用して、仮想的にファイルを共有する技術としては、例えば、特許文献1～4が公知である。また、サーバ/クライアントシステムにおいて、各クライアントで生成されるデータをサーバに保存し、その共有データを各クライアントで利用するシステムがあるが、サーバに障害が発生した場合には、ネットワークを利用した業務を継続することが困難になるという問題がある。

【0003】

サーバの障害の耐性を高めるために、ハードウェアを二重化する技術があるが、サーバのコストが高くなってしまいう問題がある。また、ストレージを集中化し、ネットワークからのファイル共有アクセスを提供するNAS(Network Attached Storage)が公知であるが、ストレージ機能を実装した専用装置が必要であるため、コストが高くなるという問題がある。この場合、信頼性・可用性を確保するためにNAS装置内においてハードウェアの二重化が必要になるため、この部分に信頼性の高いハードウェアを装備する場合には、さらに、コストが高くなってしまいう問題がある。

【0004】

さらに、ハードウェアRAIDのように、複数のハードディスクにデータを書き込む冗長化により、DISKの1台が故障してもデータの紛失を防ぐ装置が公知であるが、冗長化を制御するための専用ハードウェアおよび複数のディスクを装備する必要があるため、同様にコストが高くなってしまいう問題がある。

10

20

30

40

50

【 0 0 0 5 】

【特許文献1】特開2003-186727号公報

【特許文献2】特開平11-45203号公報

【特許文献3】特開平10-3421号公報

【特許文献4】特開2003-6032号公報

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 6 】

本発明は、上記課題に鑑みてなされたものであり、安価でかつ高い信頼性を有するストレージシステムを実現することが可能な仮想サーバ、そのコンピュータ、およびコンピュータが実行するためのプログラムを提供することを目的とする。

10

【課題を解決するための手段】

【 0 0 0 7 】

上述した課題を解決し、目的を達成するために、本発明は、ネットワークを介して相互に接続された複数のコンピュータと、前記各コンピュータにそれぞれ接続された記憶手段とで構築される仮想サーバにおいて、前記各コンピュータにそれぞれ接続された記憶手段を利用して、単一の仮想記憶領域を構成し、前記ネットワーク上に、前記単一の仮想記憶領域を表示することを特徴とする。これにより、複数台の低コストなコンピュータを使用して、仮想サーバを実現することにより、負荷を複数台のコンピュータに分散でき、安価でかつ信頼性の高いストレージシステムを提供することができる。

20

【 0 0 0 8 】

また、本発明の好ましい態様によれば、前記ネットワーク上に、前記単一の仮想記憶領域とともに、当該単一の仮想記憶領域を管理するサーバを表示することが望ましい。これにより、外部装置では、仮想サーバの実体がどこにあるかを特に意識することなく、通常のサーバに対するのと同様の操作感で仮想サーバにアクセスすることができる。

【 0 0 0 9 】

また、本発明の好ましい態様によれば、前記各記憶手段は、それぞれ特定領域を有しており、前記各特定領域で前記単一の仮想記憶領域を形成し、前記ネットワークを介して接続される外部装置からデータの書き込みの要求があった場合、または、1つのコンピュータでデータの書き込み要求が発生した場合は、前記各コンピュータは、同一のデータを前記各特定領域にそれぞれ略リアルタイムに書き込み、前記各特定領域で同一のデータを保持することが望ましい。これにより、自機の記憶手段の特定領域にデータを書き込むことにより、他のコンピュータの特定領域にもデータが書き込む冗長化によりデータの紛失を防ぐことができる。

30

【 0 0 1 0 】

また、本発明の好ましい態様によれば、前記ネットワークを介して接続される外部装置からデータの読み出し要求、または、自機の読み込み要求が発生した場合は、前記各コンピュータは、前記各特定領域からそれぞれ独自のタイミングでデータの読み込みを行うことが望ましい。これにより、各コンピュータが自機の記憶手段内の特定領域を独自のタイミングにてデータの読み込みを行うことができ、データの読み込みを高速に行うことができる。

40

【 0 0 1 1 】

また、本発明の好ましい態様によれば、前記記憶手段はディスク装置からなり、当該ディスク装置は複数の領域を有し、前記特定領域は、前記複数の領域のうちの1つの領域に割り当てられることが望ましい。これにより、ディスク装置の資源を有効に活用することが可能となる。

【 0 0 1 2 】

また、本発明の好ましい態様によれば、前記記憶手段は、複数のディスク装置からなり、前記特定領域はディスク装置の全領域に割り当てられることが望ましい。これにより、システムの運用の安全性を保つことができ、また、ホットスワップを構成することも可能

50

となる。

【0013】

また、本発明の好ましい態様によれば、前記各コンピュータのうち障害が発生したコンピュータがある場合には、当該障害が発生した以外のコンピュータで、前記仮想サーバを構成することが望ましい。これにより、故障コンピュータを切り離して修理作業を行っている期間においても、他のコンピュータにて業務を継続をすることができ、ユーザの負担をできるだけ軽減し、可用性を向上させた仮想サーバを提供することができる。なお、ここで障害とは、コンピュータに障害が発生して現に故障した場合はもちろん、当該コンピュータが保守作業中である場合において当該コンピュータのデータへアクセスできない場合を含むものである（以下、同様である）。

10

【0014】

また、本発明の好ましい態様によれば、前記コンピュータの前記特定領域にエラーが発生した場合には、当該エラーが発生したコンピュータからの要求に応じて、他のコンピュータが、当該特定領域のデータのリカバリーを行うことが望ましい。これにより、データのリカバリーが容易になり仮想サーバに直ぐに復帰することができる。

【0015】

また、本発明の好ましい態様によれば、前記各コンピュータは、仮想サーバを構成するコンピュータの電源を一元的に管理可能に構成されていることが望ましい。これにより、ネットワークを介して他のコンピュータの電源管理を行うことにより、仮想サーバを構成する複数のコンピュータの稼働状態に応じて、電源ON/OFFを一元的に管理することができ、低消費電力で仮想サーバの効率性・信頼性を向上させることができる。

20

【0016】

また、本発明の好ましい態様によれば、前記各コンピュータは、使用者により電源SWがONされたことを検出してコンピュータ内の各部に電源を供給する第1モードと、RAID機能を実行するための部位にのみ電源が供給される第2モードと、前記ネットワークを介して通信を行うための部位にのみ電源が供給される第3モードと、を有することが望ましい。これにより、システムを運用するために必要な部位にのみ通電でき、より低消費電力化を図ることができる。

【0017】

上述した課題を解決し、目的を達成するために、本発明は、各々が記憶手段を有する他の1または複数のコンピュータとネットワークを介して接続され、前記記憶手段と自機の記憶手段を利用して前記ネットワーク内で仮想サーバを構成するコンピュータにおいて、前記各コンピュータにそれぞれ接続された記憶手段を利用して構成される単一の仮想記憶領域を前記ネットワーク上に表示する表示手段を備えたことを特徴とする。

30

【0018】

また、本発明の好ましい態様によれば、前記表示手段は、前記ネットワーク上に、前記単一の仮想記憶領域とともに、当該単一の仮想記憶領域を管理するサーバを表示することが望ましい。

【0019】

また、本発明の好ましい態様によれば、前記単一の仮想記憶領域を構成する特定領域を有する前記記憶手段と、前記ネットワークと接続するためのインターフェイスと、前記ネットワークを介して接続される外部装置からのデータの書き込み要求が発生した場合、または、自機でデータの書き込み要求が発生した場合に、前記他のコンピュータに当該データの書き込みを依頼し、また、前記特定領域に当該データを書き込む第1の書込手段と、前記他のコンピュータからデータの書き込み要求を受け取った場合に、前記特定領域に当該データを書き込む第2の書込手段と、を備えたことを特徴とする。

40

【0020】

また、本発明の好ましい態様によれば、前記第1および第2の書込手段は、前記特定領域にデータを書き込み中に、他の書き込み要求があった場合には、排他処理を行うことが望ましい。

50

【 0 0 2 1 】

また、本発明の好ましい態様によれば、前記記憶手段の特定領域から独自のタイミングでデータの読み込みを行う読出手段を備えたことが望ましい。

【 0 0 2 2 】

また、本発明の好ましい態様によれば、前記仮想サーバを構成するコンピュータの電源を一元的に管理する電源管理手段を備えたことが望ましい。

【 0 0 2 3 】

また、本発明の好ましい態様によれば、前記電源管理手段は、使用者により電源 S W が O N されたことを検出してコンピュータ内の各部に電源を供給する第 1 モードと、バックアップ動作を実行するための部位にのみ電源が供給される第 2 モードと、前記ネットワークを介して通信を行うための部位にのみ電源が供給される第 3 モードと、を遷移させることが望ましい。

10

【 0 0 2 4 】

上述した課題を解決し、目的を達成するために、本発明のコンピュータが実行するためのプログラムは、各々が記憶手段を有する他の 1 または複数のコンピュータとネットワークを介して接続され、前記記憶手段と自機の記憶手段を利用して前記ネットワーク内で仮想サーバを構成するコンピュータに、前記各コンピュータにそれぞれ接続された記憶手段を利用して構成される単一の仮想記憶領域を前記ネットワーク上に表示する表示工程を実行させることを特徴とする。

【 0 0 2 5 】

また、本発明の好ましい態様によれば、前記表示工程では、前記ネットワーク上に、前記単一の仮想記憶領域とともに、当該単一の仮想記憶領域を管理するサーバを表示することが望ましい。

20

【 0 0 2 6 】

また、本発明の好ましい態様によれば、さらに、コンピュータに、前記ネットワークを介して接続される外部装置からのデータの書き込み要求が発生した場合、または、自機でデータの書き込み要求が発生した場合に、前記他のコンピュータに当該データの書き込みを依頼し、また、前記単一の仮想記憶領域を構成する自機の記憶手段の特定領域に当該データを書き込む第 1 の書込工程と、前記他のコンピュータからデータの書き込み要求を受け取った場合に、前記特定領域に当該データを書き込む第 2 の書込工程と、をコンピュータに実行させることが望ましい。

30

【 0 0 2 7 】

また、本発明の好ましい態様によれば、前記第 1 および第 2 の書込工程では、前記特定領域にデータを書き込み中に、他の書き込み要求があった場合には、排他処理を行うことが望ましい。

【 0 0 2 8 】

また、本発明の好ましい態様によれば、さらに、コンピュータに、前記記憶手段の特定領域から独自のタイミングでデータの読み込みを行う読出工程を実行させることが望ましい。

【 0 0 2 9 】

また、本発明の好ましい態様によれば、さらに、コンピュータに、前記仮想サーバを構成するコンピュータの電源を一元的に管理する電源管理工程を実行させることが望ましい。

40

【 0 0 3 0 】

また、本発明の好ましい態様によれば、前記電源管理工程では、使用者により電源 S W が O N されたことを検出してコンピュータ内の各部に電源を供給する第 1 モードと、バックアップ動作を実行するための部位にのみ電源が供給される第 2 モードと、前記ネットワークを介して通信を行うための部位にのみ電源が供給される第 3 モードと、を遷移させることが望ましい。

【 発明の効果 】

50

【0031】

本発明によれば、各コンピュータにそれぞれ接続された記憶手段を利用して、単一の仮想記憶領域を構成し、ネットワーク上に単一の仮想記憶領域を表示することとしているので、複数台の低コストなコンピュータを使用して、仮想サーバを実現することにより、負荷を複数台のコンピュータに分散でき、安価でかつ信頼性の高いストレージシステムを提供することが可能な仮想サーバ、そのコンピュータ、およびコンピュータが実行するためのプログラムを提供することが可能とな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】

【0032】

以下に、この発明にかかる仮想サーバ、そのコンピュータ、およびコンピュータが実行するためのプログラムを適用した実施例を、図面を参照しつつ詳細に説明する。なお、この実施例によりこの発明が限定されるものではない。また、下記実施例における構成要素には、当業者が容易に想定できるものまたは実質的に同一のものが含まれる。

【実施例】

【0033】

[仮想サーバの全体構成]

図1は、本発明の一実施例に係る仮想サーバ100の構成を示す図である。図1において、ネットワーク10には、仮想サーバ100と、外部装置20・・・が相互に接続されている。仮想サーバ100は、複数、例えば3台のコンピュータ200A、B、Cがネットワーク10を介して相互に接続されて構築されている。

【0034】

各コンピュータ200A、B、Cには、その起動状態(モード)を識別するための表示を行う通常起動用LED206およびスタンバイ用LED207が設けられている。以下、コンピュータ200A、B、Cを特に区別する必要がない場合は、コンピュータ200と表記する。ネットワーク10は、例えばLAN(ローカルエリアネットワーク)で構成することができる。このネットワーク10は、有線に限られるものではなく無線でもよい。

【0035】

コンピュータ200A、B、Cには、それぞれ記憶手段であるディスク装置300A、B、Cが接続されている。ディスク装置は、内蔵タイプでも外付けタイプのいずれでも良い。以下、ディスク装置300A、B、Cを特に区別する必要がない場合は、ディスク装置300と表記する。ディスク装置300A、B、Cは、それぞれ、OS(例えば、Windows(登録商標)OS、Unix(登録商標)、Linux(登録商標)等)やアプリケーション等が格納される第1領域301A、B、Cと、RAID専用の領域である第2領域(以後「特定領域」という)302A、B、Cと、コンピュータ200A、B、Cのマシン固有の情報が格納される第3領域303A、B、Cとを有している。

【0036】

この特定領域302A、B、Cは、コンピュータ200A、B、Cにより共有される仮想的な単一の記憶領域(仮想記憶領域)400を構成する。以下の説明では、RAID1について説明するが、本発明はこれに限られるものではなく、RAID0、2、3、4、5についても適用可能である。なお、この領域の割り当ては一例であって、ディスク装置300の領域の割り当ては特に限定されるものではない。

【0037】

本実施例に係る仮想サーバ100の第1の特徴は、ネットワーク10を介して接続されたコンピュータ200A、B、Cで仮想的にサーバが構成される点にある。

【0038】

本実施例に係る仮想サーバ100の第2の特徴は、仮想サーバ100内の各コンピュータ200A、B、Cが有するディスク装置300A、B、Cを利用して仮想記憶領域400を構成している点にある。仮想記憶領域400を構成する複数の特定領域302A、302B、302Cのデータを、各コンピュータ200A、B、Cで共有することができる

10

20

30

40

50

。すなわち、各コンピュータ200A、B、Cは、それぞれ特定領域302A、302B、302Cに同一のデータを書き込み、各特定領域302A、302B、302Cで同一のデータを保持することができる(RAID1の機能)。

【0039】

本実施例に係る仮想サーバ100の第3の特徴は、仮想記憶領域400を構成する複数の特定領域302A、302B、302のいずれか1つに障害が発生しても、残りの特定領域のデータから障害が発生した特定領域のデータが復元でき、データの可用性を向上させることができる点にある。

【0040】

本実施例に係る仮想サーバ100の第4の特徴は、各コンピュータ200A、B、Cでは、特定領域302A、302B、302Cを単一な仮想記憶領域400として認識させる点にある。具体的には、各コンピュータ200A、B、Cでは、特定領域302A、302B、302Cを単一な仮想記憶領域400としてユーザーインターフェイス(表示装置205)に表示する。使用者は、このユーザーインターフェイスの表示を見て、このディスク領域に対してファイルの書き込みや読み込みの操作を行うことができる。

【0041】

本実施例に係る仮想サーバ100の第5の特徴は、外部装置20に対して、仮想サーバ100を構成する各コンピュータ200A、B、Cの特定領域302A、302B、302Cを単一な仮想記憶領域(ディスク領域)400として表示させる点にある。具体的には、外部装置20では、仮想サーバ100を構成する各コンピュータ200A、B、Cの特定領域302A、302B、302Cを単一な仮想記憶領域400としてユーザーインターフェイスに表示する。使用者は、このユーザーインターフェイスの表示を見て、このディスク領域に対してファイルの書き込みや読み込みの操作を行うことができる。

【0042】

本実施例に係る仮想サーバ100の第6の特徴は、低消費電力化および高効率性・高信頼性のために、各コンピュータ200A、B、Cが、仮想サーバ100内のコンピュータ200A、B、Cの電源を一元的に管理している点である。

【0043】

なお、図1の構成においては、第1領域301、特定領域302、および第3領域303を1つのディスク装置300に割り当てているが、記憶手段を複数のディスク装置で構成し、特定領域302を第1領域301、第2領域303と物理的に別のディスク装置に割り当てることにしても良い。

【0044】

図2は、第1領域301と第3領域303を第1ディスク装置310に割り当て、特定領域302を第2ディスク装置311に割り当てた場合の仮想サーバ100の構成例を示す図である。すなわち、本実施例の記憶手段は、単一のディスク装置で構成しても、複数のディスク装置で構成することに良い。ディスク装置310、311は、ホットスワップ可能に構成されるのが望ましい。

【0045】

[コンピュータの構成]

図3は、コンピュータ200の構成を示すブロック図である。コンピュータ200は、例えば、PC(パーソナルコンピュータ)やWS(ワークステーション)で構成することができる。コンピュータ200は、図3に示すように、ネットワーク10に接続するためのLANコントローラ201と、コンピュータの各部の動作を制御する制御部202と、電源装置208を制御する電源制御装置203と、キーボード、マウス、タブレット等からなる入力装置204と、LCDやCRTからなる表示装置205と、通常起動状態にある場合に点灯する通常起動用LED206と、スタンバイ状態にある場合に点灯するスタンバイ用LED207と、コンピュータの各部に電源を供給する電源装置208とを備えている。

【0046】

LANコントローラ201は、ネットワーク10に接続するLANカード等から構成されており、Wake up LAN機能を有している。制御部202は、メインボードに搭載されており、プログラムやデータを記憶するROM(Read Only Memory)、一時的にデータを保管するRAM(Random Access Memory)、及びプログラムやデータを当該ROMやディスク装置300から読み込んで実行するCPU等で構成される。

【0047】

制御部202は、例えば、プログラムを実行することにより、ディスク装置300に対するデータの書き込み/読み込みの制御等を行うデータ管理部221として機能する。データ管理部221は、RAID機能を有している。データ管理部221は、具体的には、
(1)ネットワーク10上に、単一の仮想記憶領域400およびそれを管理するサーバを表示する機能、
(2)外部装置20からデータの書き込み要求があった場合、または、自機でデータの書き込み要求が発生した場合に、前記他のコンピュータに当該データの書き込みを依頼し、また、ディスク装置300の特定領域302に当該データを書き込む機能、
(3)他のコンピュータ200からデータの書き込み要求を受け取った場合に、ディスク装置300の特定領域302に当該データを書き込む機能、
(4)外部装置20からのデータの読み込み要求があった場合、または、自機で発生したデータの読み込み要求が発生した場合に、ディスク装置300の特定領域302から独自のタイミングでデータの読み込みを行う機能を有している。

【0048】

また、制御部202は、例えば、プログラムを実行することにより、コンピュータ内部の電源管理を行うとともに、仮想サーバ100を構成するコンピュータ200A、B、Cの電源をネットワーク10を介して一元管理する電源管理部222として機能する。

【0049】

電源管理部222は、電源制御装置203を制御するとともに、電源制御装置203から入力される起動情報に基づいてコンピュータ200の起動状態(通常起動状態、スタンバイ状態、電源OFF状態)を遷移させる。電源管理部222は、仮想サーバ100を構成する他のコンピュータ200に特殊電源ON/OFF信号を送出する。

【0050】

また、電源管理部222は、使用者がコンピュータのJOB操作画面に設けられた電源OFFを実行するJOBを選択指示した場合において、ソフトウェア処理により電源OFFコマンドを受け付けて、電源制御装置203に通知する。これに応じて、電源制御装置203は、電源SWがOFFされた場合と同様の処理を行う。

【0051】

電源制御装置203は、電源用のボード等に搭載されており、例えば、CPU、ファームウェアが記憶されたROM、LANコントローラ201等と通信を行うためのポート制御部等で構成することができる。電源制御装置203は、電源管理部222の指示に従って、電源装置208の制御、第1および第2スイッチ209、210のON/OFF、および電源状態表示用LED206、207の点灯を制御する。また、電源制御装置203は、使用者により押下される機械スイッチである電源SWのON/OFF、ネットワーク10を介して入力される特殊電源ON/OFF信号を検出する。

【0052】

電源装置208は、商用電源に接続されており、AC-DC変換を行って、コンピュータ用のDC電源を生成する。電源装置208は、電源制御装置203の制御に従って、電源ライン231、232を介して、コンピュータ200の各部に電源を供給する。電源ライン231は、常時電源が供給されるラインであり、電源SWがOFFの場合においても電源装置208が商用電源に接続されている限り給電が行われる。この電源ライン231には、LANコントローラ201および電源制御装置203が接続されている。これにより、電源SWがOFFの場合でも、ネットワーク10を介して他のコンピュータと通信することができる。電源ライン232は、電源制御装置203の制御に従って、給電が行わ

10

20

30

40

50

れる。この電源ライン 232 には、制御部 202、入力装置 204、表示装置 205、およびディスク装置 300 が接続されている。入力装置 204 および表示装置 205 に接続される電源ライン 232 には、第 1 および第 2 スイッチ 209、210 が設けられている。

【0053】

[仮想記憶領域 400 および仮想サーバ 100 の表示方法]

図 4 は、仮想サーバ 100 を構成するコンピュータ 200 A、B、C において、表示装置 205 A、B、C に表示される管理装置情報の一例を示す図である。同図に示す例は、コンピュータ 200 A の場合を示している。同図においては、名称「Aグループ」、管理装置名称「共通データ領域（仮想記憶領域 400）」、空き容量「900MB」と、名称「コンピュータ 1（コンピュータ 200 A）」、管理装置名称「CD-R（未装着）」と、名称「コンピュータ 1（コンピュータ 200 A）」、管理装置名称「DISK」、空き容量「1500MB」とが表示されている。ここで、Aグループは、仮想サーバ 100 を構成するグループを示しており、特定領域 302 A、B、C は、各コンピュータ 200 A、B、C では単一のディスク領域（共有データ領域）として表示される。同図に示す例では、各特定領域 302 A、B、C は 900MB であり、各コンピュータ 200 A、B、C では、900MB の単一のディスク領域として表示される。これにより、アプリケーションから特定領域 302 A、B、C を表示する場合には、使用者は、データの実体がどこにあるかを特に意識することなく、自機 DISK と同様の操作感でアクセスすることができる。

【0054】

図 5 は、外部装置 20 において、その表示装置に表示されるネットワーク上のコンピュータリスト情報の一例を示す図である。同図においては、コンピュータ名称「コンピュータ 1（コンピュータ 200 A）」、グループ名称「Aグループ」、管理装置名称「DISK 1（ディスク装置 300 A）」と、コンピュータ名称「コンピュータ 2（コンピュータ 200 B）」、グループ名称「Aグループ」、管理装置名称「DISK 2（ディスク装置 300 B）」と、コンピュータ名称「コンピュータ 3（コンピュータ 200 C）」、グループ名称「Aグループ」、管理装置名称「DISK 3（ディスク装置 300 C）」と、コンピュータ名称「サーバ A（仮想サーバ 100）」、グループ名称「Aグループ」、管理装置名称「共通データ領域（仮想記憶領域 400）」と、コンピュータ名称「外部 1（外部装置 20）」と、コンピュータ名称「外部 2（外部装置 20）」とが表示されている。

【0055】

このように、共通データ領域（仮想記憶領域 400）を管理する仮想サーバ 100 は、ネットワーク 10 上に追加されているように表示される。これにより、外部装置 20 の使用者は、仮想サーバ 100 の実体がどこにあるかを特に意識することなく、通常のサーバに対するのと同様の操作感で仮想サーバ 100 にアクセスすることができる。なお、仮想サーバ 100 を構成する各コンピュータ 200 A、B、C では、サーバとしての機能も汎用サーバと同様の機能を実装することにも良い。

【0056】

[データの書き込み・読み込み動作]

各コンピュータ 200 A、B、C が、ディスク装置 300 A、B、C の特定領域 302 A、B、C に対して、データの書き込み・読み込む動作を説明する。図 6 は、コンピュータ 200 A、B、C が、ディスク装置 300 A、B、C の特定領域 302 A、B、C に対して、データの書き込み・読み込む動作の概略を説明するためのフローチャートである。

【0057】

図 6 において、コンピュータ 200 A、B、C のいずれか 1 つでデータの更新が発生した場合には（A1 の「Yes」）、当該コンピュータ 200 は、特定領域 302 を占有して、データを書き込むことを他のコンピュータ 200 に通知する（A2）。通知を受信した全てのコンピュータ 200 の用意ができた場合には（A3 の「Yes」）、各コンピュータ 200 でそれぞれの特定領域 302 でデータの更新を開始する（A4）。全てのコン

10

20

30

40

50

コンピュータ200でデータ更新が完了した場合には(A5の「Yes」)、データの更新を終了する。他方、コンピュータ200A、B、Cのいずれにもデータの更新が発生しない場合には(A1の「No」)、各コンピュータ200A、B、Cは、独自のタイミングで各特定領域302A、B、Cで読み取る(A6の「No」)。

【0058】

図7は、各コンピュータ200A、B、Cが、ディスク装置300A、B、Cの特定領域302A、B、Cに対して、データの書き込み・読み込み動作を説明するための模式図である。同図では、コンピュータ200Aがディスク装置300Aでデータの更新が発生した場合の動作を説明する。同図において、ファイルシステム・コール252および特定領域ファイルモジュール253は、上述したデータ管理部221の1機能である。

10

【0059】

書き込み処理時には、各コンピュータ200A、B、Cの特定領域ファイルモジュール253A、B、Cは、リアルタイムにデータの更新を行うことで、データの冗長化を行う。冗長化処理中は、各コンピュータ200A、B、Cは、仮想サーバ100を構成する全てのコンピュータ200に対して、ファイルシステムの最下層のコマンドレベルでネットワーク10越しに、他のファイル更新要求を排他処理する。

【0060】

読み込み処理は、各コンピュータ200A、B、Cで逐次可能となっている。ファイル読み込み時には、各コンピュータ200A、B、Cの特定領域ファイルモジュール2531A、B、Cは、それぞれ特定領域302A、B、Cのファイルを独自のタイミングで読み取ることが可能となっている。

20

【0061】

コンピュータ200Aにおいて、アプリケーション251Aでファイルの特定領域302への書き込み要求が発生した場合には、ファイルシステム・コール252Aは、特定領域ファイルモジュール253Aに特定領域302へのファイルの書き込みを依頼する。特定領域ファイルモジュール253Aは、コンピュータ200B、Cの特定領域ファイルモジュール253B、Cに対して、ファイルとタイムスタンプを送信して、書き込みを依頼する。そして、特定領域ファイルモジュール253A、B、Cは、ディスク装置300A、B、Cの特定領域302A、B、Cに略同時に同一のファイルをそれぞれ書き込む。ここでは、アプリケーション251Aでデータの書き込み/読み込み要求が発生した場合について説明したが、外部装置20からデータの書き込み/読み込み要求が発生した場合についても同様な処理が行われる。

30

【0062】

[エラー発生時の動作・リカバリー動作]

コンピュータ200A、B、Cでエラーが発生した場合の動作、およびそのリカバリー動作を説明する。

【0063】

1. 仮想サーバが正常に動作しない場合

図8は、コンピュータ200A、B、Cのいずれかでトラブルが発生した場合の処理を説明するための図である。コンピュータ200A、B、Cのいずれかでトラブルが発生した場合は、そのコンピュータ200を切り離し、残りのコンピュータ200でRAIDシステムを継続する。

40

【0064】

図8において、例えば、コンピュータ200Aが故障した場合には、そのコンピュータ200Aを切り離し、残りのコンピュータ200B、CでRAIDシステムを継続する。

【0065】

2. 読み込み処理で特定領域302にエラーが発生した場合

図9は、コンピュータ200A、B、Cのディスク装置300A、B、Cの特定領域302A、B、Cにおいて、読み込み処理でエラーが発生した場合の処理を説明するための図である。

50

【 0 0 6 6 】

コンピュータ 2 0 0 A、B、C のいずれか 1 つの特定領域 3 0 2 A、B、C で読み込み処理中にエラーが発生した場合には、他のコンピュータ 2 0 0 の特定領域 3 0 2 で読み込み処理を継続し、以降、故障した特定領域 3 0 2 へはアクセスしない。

【 0 0 6 7 】

図 9 において、例えば、コンピュータ 2 0 0 A のアプリケーション A からの特定領域 3 0 2 A の読み込み処理でディスクエラーが発生した場合には、他のコンピュータ 2 0 0 B または 2 0 0 C の特定領域 3 0 2 A、3 0 2 B から読み込み処理を行って処理を継続する。以後、特定領域 3 0 2 A にはアクセスしない。

【 0 0 6 8 】

3 . 書き込み処理で特定領域 3 0 2 にエラーが発生した場合

図 1 0 は、コンピュータ 2 0 0 A、B、C のディスク装置 3 0 0 A、B、C の特定領域 3 0 2 A、B、C において、書き込み処理でエラーが発生した場合の処理を説明するための図である。コンピュータ 2 0 0 A、B、C のいずれか 1 つの特定領域 3 0 2 A、B、C で書き込み処理でエラーが発生した場合には、他のコンピュータ 2 0 0 の特定領域 3 0 2 で書き込み処理を継続し、以降、故障した特定領域 3 0 2 へはアクセスしない。

【 0 0 6 9 】

図 1 0 において、例えば、コンピュータ 2 0 0 A のアプリケーション A からの書き込み処理において、特定領域 3 0 2 A でディスクエラーが発生した場合には、他の特定領域 3 0 2 B、C で処理が成功している場合は、書き込み処理を継続し、以後、特定領域 3 0 2 A にはアクセスしない。

【 0 0 7 0 】

なお、上記読み込み処理および書き込み処理では、ディスク装置 3 0 0 の特定領域 3 0 2 でセクターエラーが発生したものと、アプリケーション A が処理を継続しているが、通常、物理的なディスクエラーが発生した場合には、領域を分割していても同一のディスク装置を利用している限りディスク全体に影響が発生するため、いずれシステム的なエラーが発生する可能性がある。そのため、上記 2、3 の対処法は、特定領域 3 0 2 を物理的に別のディスクに割り当てた上記図 2 の構成の場合に有効である。

【 0 0 7 1 】

4 . 特定領域のリカバリー

図 1 1 は、特定領域 3 0 2 をリカバリーする手順を説明するためのフローチャートである。特定領域 3 0 2 をリカバリーする場合には、特定領域 3 0 2 が正常なコンピュータ 2 0 0 から特定領域 3 0 2 の全エリアのブロック転送を利用する。

【 0 0 7 2 】

図 1 1 において、RAID システムへの復帰を希望するコンピュータ 2 0 0 から RAID システムが正常に動作している全てのコンピュータに復帰要求を通知する (A 1 1)。通知を受けたコンピュータ 2 0 0 は、自機の特定領域へのアクセスを一時的に禁止する (A 1 2)。通知を受けたコンピュータ 2 0 0 の 1 つは、自機の特定領域 3 0 2 の全エリアをブロック転送を利用して、RAID システムへの復帰を希望するコンピュータ 2 0 0 の特定領域 3 0 2 にコピーする (A 1 3)。この後、全てのコンピュータ 2 0 0 で RAID システムを起動する (A 1 4)。

【 0 0 7 3 】

図 1 2 は、特定領域 3 0 2 のリカバリーの具体例を示す図である。図 1 2 において、例えば、コンピュータ 2 0 0 C の特定領域 3 0 2 C でエラーが発生して、RAID システムへの復帰を希望する場合には、コンピュータ 2 0 0 C は、コンピュータ 2 0 0 A、B に復帰要求を通知する。通知を受けたコンピュータ 2 0 0 A、B は、自機の特定領域 3 0 2 A、B へのアクセスを一時的に禁止する。コンピュータ 2 0 0 A または 2 0 0 B は、自機の特定領域 3 0 2 A、B の全エリアのデータをコンピュータ 2 0 0 C の特定領域 3 0 2 C にコピーする。これにより、全てのコンピュータ 2 0 0 A、B、C で特定領域 3 0 2 A、B、C へのアクセスが可能となる。

10

20

30

40

50

【 0 0 7 4 】

なお、仮想サーバ100の全てのコンピュータ200A、B、Cの特定領域302A、B、Cが全て故障した場合（地震や火事の場合）に備えて、外部のネットワークに接続されるデータセンター500で仮想サーバ100の特定領域302のデータのバックアップを行うことにしてもよい。

【 0 0 7 5 】

図13は、データセンター500を使用して、仮想サーバ100をリカバリーする手順を説明するためのフローチャートである。図13において、まず、仮想サーバ100を構築する全てのコンピュータ200の特定領域302へのアクセスを禁止し（A21）、データセンター500は、仮想サーバ100を構成するコンピュータ200の1つの特定領域302を復元する（A22）。特定領域302を復元したコンピュータ200は、その特定領域302を他のコンピュータ200の特定領域302にコピーする（A23）。その後、全てのコンピュータ200でRAIDシステムを起動する（A24）。

10

【 0 0 7 6 】

図14は、データセンター500を使用して特定領域302をリカバリーする具体例を示す図である。図14において、まず、仮想サーバ100の全コンピュータ200A、B、Cの特定領域302A、B、Cをリカバリーする場合は、コンピュータ200A、B、Cの特定領域302A、B、Cへのアクセスを禁止し、データセンター500は、コンピュータ200Cの特定領域302Cを復元する。コンピュータ200Cは、その特定領域302Cをコンピュータ200A、Bの特定領域302A、Bにコピーする。これにより、全てのコンピュータ200A、B、Cで特定領域302A、B、Cへのアクセスが可能となる。なお、リムーバルディスクをバックアップ先に使用することにも良い。

20

【 0 0 7 7 】

[コンピュータの電源状態の遷移]

図15は、コンピュータ200の電源状態の遷移を示す図である。図15において、（1）通常起動状態（第1モード）は、使用者により電源SWがONされて、コンピュータ200の各部に電源が供給されている状態で、コンピュータ200が通常に起動している状態である。（2）スタンバイ状態（第2のモード）は、コンピュータ200を起動した状態であるが、バックアップ動作（RAID機能）を実行するための部位のみに電源が供給される省電力状態であり、バックアップ動作に必要なない部位、例えば、入力装置204および表示装置205等への電源の供給を切断した状態である。（3）電源OFF状態（第3のモード）は、ネットワーク10を介して通信を行うための部位のみに電源が供給された状態で、LANコントローラ201および電源制御装置203以外への電源の供給が切断された状態である。

30

【 0 0 7 8 】

使用者により電源SWがONされると、電源制御装置203は、電源装置208に電源ライン232からの電源の供給を開始させる。これにより、制御部202に電源が供給され、電源管理部222が起動する。電源管理部222は、電源制御装置203から入力される起動情報（電源SWのONによる起動）に基づき、電源制御装置203を介して、第1および第2スイッチ209、210をONさせる。これにより、コンピュータ200の各部に電源が供給され、通常起動状態（1）となる。

40

【 0 0 7 9 】

また、使用者により電源SWがOFFされた場合または電源OFFコマンドを受け取った場合は、電源制御装置203は、電源装置208に電源ライン232からの電源の供給を停止させる。これにより、LANコントローラ201および電源制御装置203以外への電源の供給が切断され、電源OFF状態となる。

【 0 0 8 0 】

特殊電源ONは、ネットワーク10を介して、電源OFF状態からスタンバイ状態にする機能である。コンピュータ200が電源OFF状態にある場合に、ネットワーク10を介して、特殊電源ON信号が電源制御装置203に入力されると、電源制御装置203は

50

、電源装置 208 に電源ライン 232 からの電源の供給を開始させる。これにより、制御部 202 に電源が供給され、電源管理部 222 が起動する。電源管理部 222 は、電源制御装置 203 から入力される起動情報（特殊電源 ON による起動）に基づき、電源制御装置 203 を介して、第 1 および第 2 スイッチ 209、210 を OFF させる。これにより、スタンバイ状態となり、バックアップ動作（RAID 機能）を実行するために必要な部位（制御部 202 およびディスク装置 300 等）に電源が供給され、バックアップ動作を実行するために必要のない入力装置 204 および表示装置 205 等への電源の供給が遮断される。

【0081】

特殊電源 OFF は、ネットワーク 10 を介して、スタンバイ状態から電源 OFF 状態にする機能である。コンピュータ 200 がスタンバイ状態にある場合に、ネットワーク 10 を介して特殊電源 OFF 信号が入力されると、電源管理部 222 は、電源制御装置 203 を介して、電源ライン 232 からの電源の供給を停止させる。これにより、電源 OFF 状態となる。

【0082】

[起動時の動作]

コンピュータ 200 A、B、C が電源 OFF 状態にある場合に、1 台のコンピュータ 200 のみの電源 SW を ON して、RAID システムを起動させる場合について説明する。ここでは、コンピュータ 200 A の電源 SW を ON させた場合の動作を説明する。コンピュータ 200 B、C の電源 SW を ON させた場合も同様な動作を行うのでここではその説明を省略する。

【0083】

図 16 は、コンピュータ 200 A、B、C が電源 OFF 状態にある場合に、1 台のコンピュータ 200 のみの電源 SW を ON して、RAID システムを起動させる動作を説明するためのフローチャート、図 17 は、コンピュータ 200 A、B、C が電源 OFF 状態にある場合に、1 台のコンピュータ 200 のみの電源 SW を ON して、RAID システムを起動させる動作を説明するための説明図である。

【0084】

図 16 において、コンピュータ 200 A、B、C が電源 OFF 状態にある場合に（S1）、コンピュータ 200 A において、使用者により電源 SW が ON されると（S2）、電源制御装置 203 は、電源装置 208 に電源ライン 232 からの電源の供給を開始させる（S3）。これにより、制御部 202 に電源が供給され、電源管理部 222 が起動する。電源管理部 222 は、電源制御装置 203 から入力される起動情報（電源 SW の ON による起動）に基づいて起動条件を判定する（S4）。この場合は、電源 SW の ON による起動であるので、電源管理部 222 は、電源制御装置 203 を介して、第 1 および第 2 スイッチ 209 A、210 A を ON させて、入力装置 204 および表示装置 205 に電源を供給する（S5）。そして、電源管理部 222 は、特殊電源 ON 信号をコンピュータ 200 B、C に送信する（S6）。この後、データ管理部 221 を起動させる（S7）。電源管理部 222 は、電源制御装置 203 を介して通常起動用 LED 206 を点灯させる（S8）。このようにして、通常起動状態となる（S9）。

【0085】

他方、電源 OFF 状態にあるコンピュータ 200 B、C では（S11）、電源管理装置 203 は、特殊電源 ON 信号を受信すると（S12）、電源制御装置 203 は、電源装置 208 に電源ライン 232 からの電源の供給を開始させる（S13）。これにより、制御部 202 に電源が供給され、電源管理部 222 が起動する。電源管理部 222 は、電源制御装置 203 から入力される起動情報（特殊電源 ON による起動）に基づき起動条件を判定する（S14）、この場合は、特殊電源 ON による起動であるので、電源管理部 222 は、電源制御装置 203 を介して、第 1 および第 2 スイッチ 209、210 を OFF させて、入力装置 204 および表示装置 205 に電源を供給しない（S15）。そして、データ管理部 221 を起動させる（S16）。電源管理部 222 は、電源制御装置 203 を介

10

20

30

40

50

してスタンバイ用LED207を点灯させる(S17)。このようにして、コンピュータ200B、Cはスタンバイ状態となる(S18)。この後、RAIDシステムを使用することができる。

【0086】

この後、コンピュータ200B、Cでは、電源SWがONされると(S21)、電源制御装置203は、電源SWのONを電源管理部222に通知する。電源管理部222は、電源制御装置203を介して、第1および第2スイッチ209、210をONさせて、入力装置204および表示装置205に電源を供給する(S22)。この後、電源管理部222は、電源制御装置203B、Cを介して通常起動用LED206を点灯させる(S23)。このようにして、通常起動状態となる(S24)。

10

【0087】

[電源OFF時(または電源OFFコマンドが発行された場合)の動作]

コンピュータ200A、B、Cのうち、1台のコンピュータが通常起動状態で、他のコンピュータがスタンバイ状態にある場合に、電源をOFFする動作を説明する。ここでは、コンピュータ200Aが通常起動状態で、コンピュータB、Cがスタンバイ状態にある場合に、コンピュータ200Aの電源SWがOFFされた場合(または電源OFFコマンドが発行された場合)の動作を説明する。コンピュータ200B、Cの電源SWをOFFさせた場合(または電源OFFコマンドが発行された場合)も同様な動作を行うのでここではその説明を省略する。

【0088】

図18は、コンピュータ200Aが通常起動状態で、コンピュータB、Cがスタンバイ状態にある場合に、コンピュータ200Aの電源SWがOFFされた場合または電源OFFコマンドが発行された場合の動作を説明するためのフローチャートである。図19は、コンピュータ200Aが通常起動状態で、コンピュータB、Cがスタンバイ状態にある場合に、コンピュータ200Aの電源SWがOFFされた場合または電源OFFコマンドが発行された場合の動作を説明するための説明図である。

20

【0089】

図18において、コンピュータ200Aでは、通常起動状態にある場合に(S31)、電源SWがOFF(またはソフトウェア処理により電源OFFコマンドが発行)されると(S32)、電源制御装置203は、電源SWのOFFを電源管理部222に通知する。電源管理部222は、コンピュータ200B、Cに電源状態を問い合わせ、終了条件を判定する(S33)。

30

【0090】

通常起動状態のコンピュータがある場合には、電源管理部222は、電源制御装置203を介して、第1および第2スイッチ209、210をOFFさせて、入力装置204および表示装置205への電源の供給を切断する(S34)。そして、電源管理部222は、電源制御装置203を介して、通常起動用LED206を消灯させるとともに、スタンバイ用LED207を点灯させて(S35、S36)、スタンバイ状態となる(S37)。

【0091】

他方、通常起動状態のコンピュータがない場合には、電源管理部222は、特殊電源OFF信号をコンピュータ200B、Cに送信する(S41)。そして、電源管理部222は、電源制御装置203を介して、電源装置208の電源ライン231からの電源の供給を停止させ(S42)、また、通常起動用LEDを消灯させて(S43)、電源OFF状態となる(S44)。

40

【0092】

また、スタンバイ状態にある(S51)コンピュータ200B、Cでは、コンピュータ200Aからの問い合わせに応答し(S52)、ネットワーク10を介して特殊電源OFF信号を受信すると(S53)、電源管理部222は、電源制御装置203を介して、電源ライン232からの電源の供給を停止させ(S54)、また、スタンバイ用LED20

50

7を消灯させて(S55)、電源OFF状態となる(S56)。

【0093】

以上説明したように、本実施例によれば、ネットワーク10を介して相互に接続された複数のコンピュータ200A、B、Cと、各コンピュータ200A、B、Cにそれぞれ接続されたディスク装置300A、B、Cとを備えた仮想サーバ100において、各コンピュータ200A、B、Cにそれぞれ接続されたディスク装置300A、B、Cを利用して、単一の仮想記憶領域400を構成し、ネットワーク10上に単一の仮想記憶領域400を表示することとしているので、複数台の低コストなコンピュータを使用して、仮想サーバを実現することにより、負荷を複数台のコンピュータに分散でき、安価でかつ信頼性の高いストレージシステムを提供することができる。

10

【産業上の利用可能性】

【0094】

本発明に係る仮想サーバは、低コストな構成でネットワーク上でストレージシステムを実現する場合に有用である。

【図面の簡単な説明】

【0095】

【図1】本発明の一実施例に係る仮想サーバの構成を示す図である。

【図2】仮想サーバの変形例を示す図である。

【図3】図1のコンピュータの構成を示すブロック図である。

【図4】仮想サーバを構成するコンピュータにおいて、表示装置に表示される管理装置情報の一例を示す図である。

20

【図5】外部装置において、その表示装置に表示されるネットワーク上のコンピュータリスト情報の一例を示す図である。

【図6】コンピュータが、ディスク装置の特定領域に対して、データの書き込み・読み込む動作の概略を説明するためのフローチャートである。

【図7】各コンピュータが、ディスク装置の特定領域に対して、データの書き込み・読み込む動作を説明するための模式図である。

【図8】コンピュータのいずれかでトラブルが発生した場合の処理を説明するための図である。

【図9】コンピュータのディスク装置の特定領域において、読み込み処理でエラーが発生した場合の処理を説明するための図である。

30

【図10】コンピュータのディスク装置の特定領域において、書き込み処理でエラーが発生した場合の処理を説明するための図である。

【図11】特定領域をリカバリーする手順を説明するためのフローチャートである。

【図12】特定領域のリカバリーの具体例を示す図である。

【図13】データセンターを使用して、ネットワークRAIDシステムをリカバリーする手順を説明するためのフローチャートである。

【図14】データセンターを使用して特定領域をリカバリーする具体例を示す図である。

【図15】コンピュータの電源状態の遷移を示す図である。

【図16】各コンピュータが電源OFF状態にある場合に、1台のコンピュータのみの電源SWをONして、RAIDシステムを起動させる動作を説明するためのフローチャートである。

40

【図17】各コンピュータが電源OFF状態にある場合に、1台のコンピュータのみの電源SWをONして、RAIDシステムを起動させる動作を説明するための説明図である。

【図18】コンピュータが通常起動状態で、他のコンピュータがスタンバイ状態にある場合に、当該コンピュータの電源SWがOFFされた場合(または電源OFFコマンドが発行された場合)の動作を説明するためのフローチャートである。

【図19】コンピュータが通常起動状態で、他のコンピュータがスタンバイ状態にある場合に、当該コンピュータの電源SWがOFFされた場合(または電源OFFコマンドが発行された場合)の動作を説明するための説明図である。

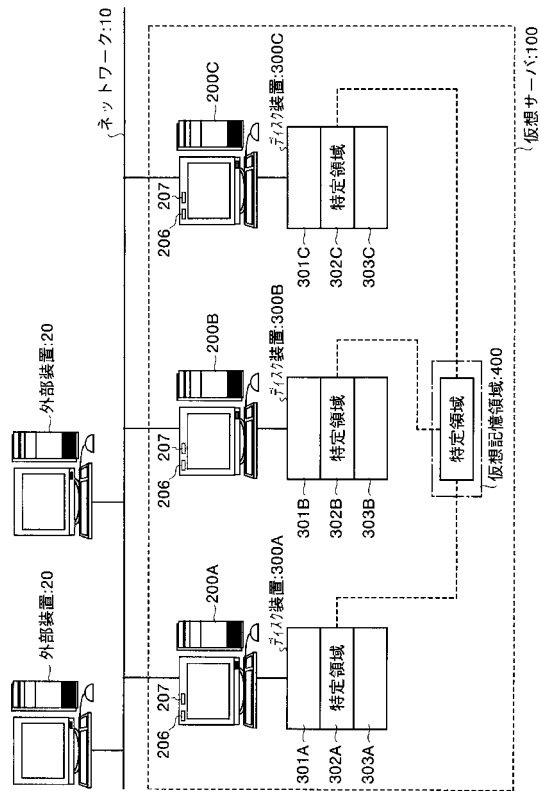
50

【符号の説明】

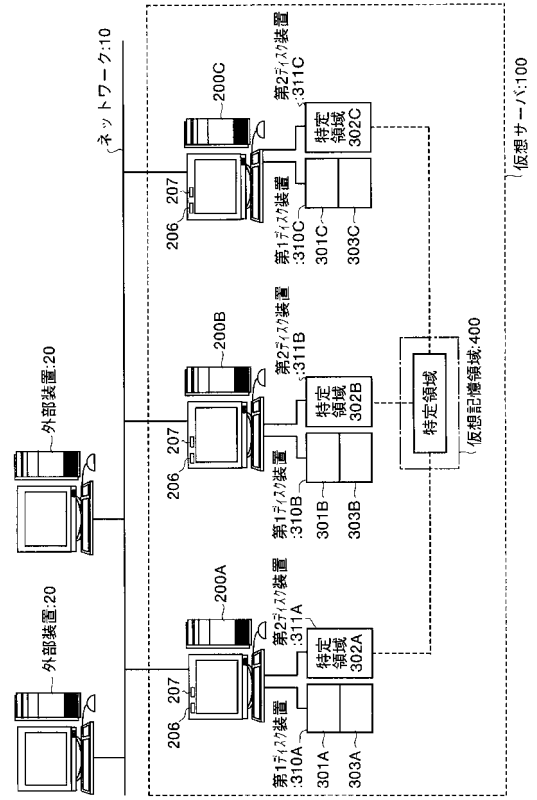
【0096】

10	ネットワーク	
20	外部装置	
100	仮想サーバ	
200	コンピュータ	
201	L A Nコントローラ	
202	制御部	
203	電源制御装置	
204	入力装置	10
205	表示装置	
206	通常起動用 L E D	
207	スタンバイ用 L E D	
208	電源装置	
209	第1スイッチ	
210	第2スイッチ	
221	データ管理部	
222	電源管理部	
231、232	電源ライン	
301	第1領域	20
302	特定領域	
303	第3領域	
310	第1ディスク装置	
311	第2ディスク装置	

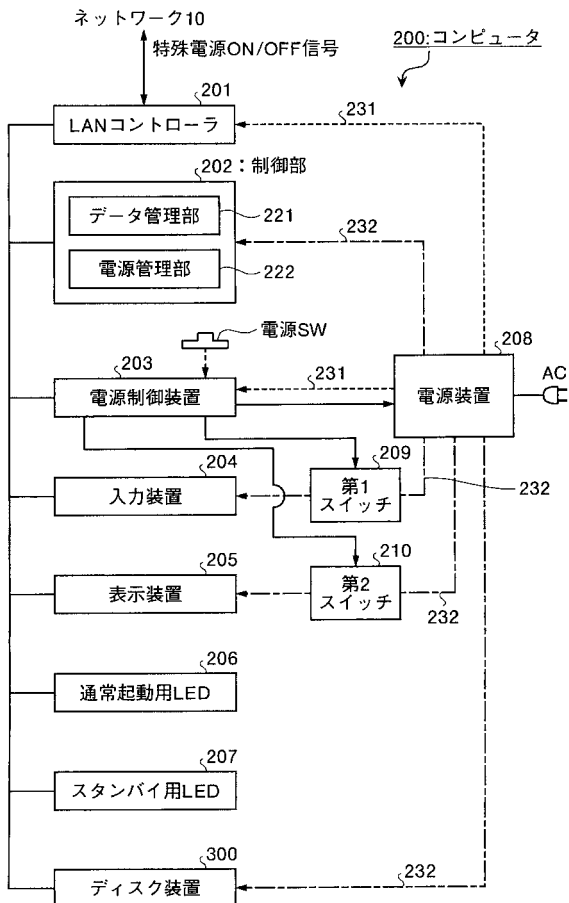
【図1】



【図2】



【図3】



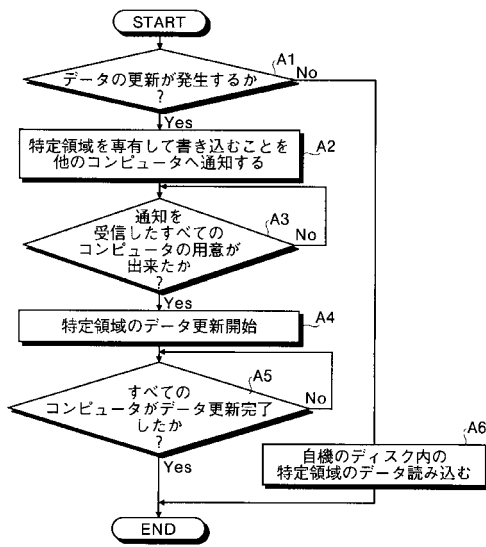
【図4】

名称	管理装置名称	空き容量
Aグループ	共通データ領域	900MB
コンピュータ1	CD-R (未装着)	
コンピュータ1	DISK	15000MB

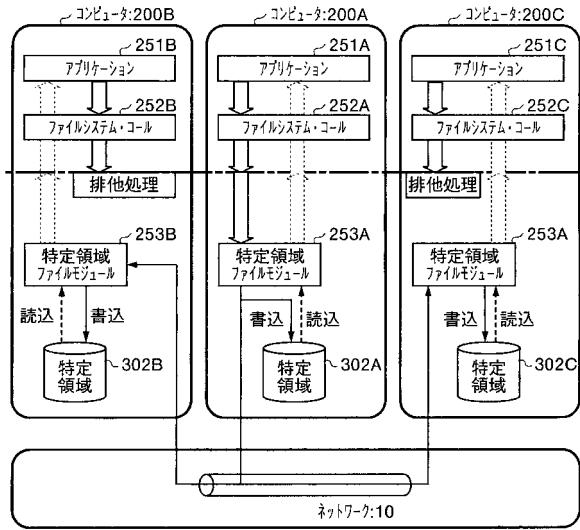
【図5】

コンピュータ名称	グループ名称	管理装置名称
コンピュータ1	Aグループ	DISK1
コンピュータ2	Aグループ	DISK2
コンピュータ3	Aグループ	DISK3
サーバA	Aグループ	共通データ領域
外部1		
外部2		

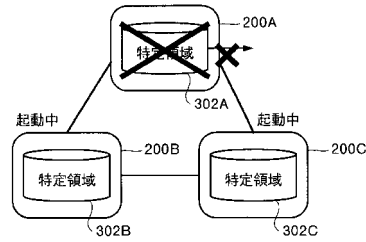
【図6】



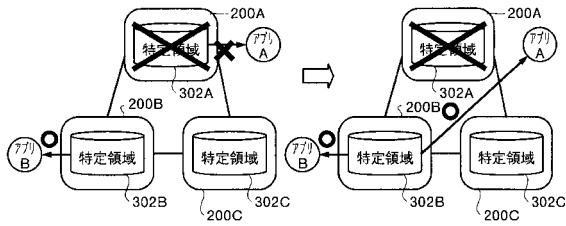
【図7】



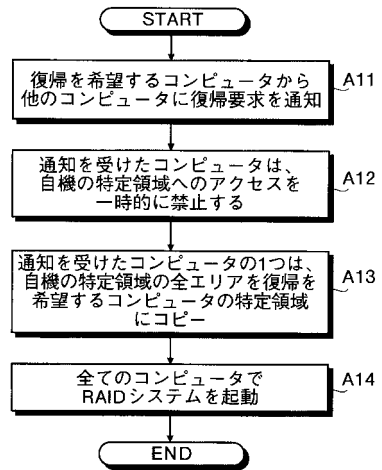
【図8】



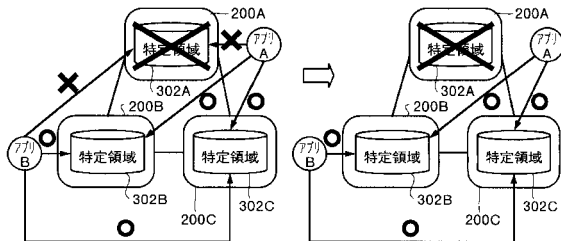
【図9】



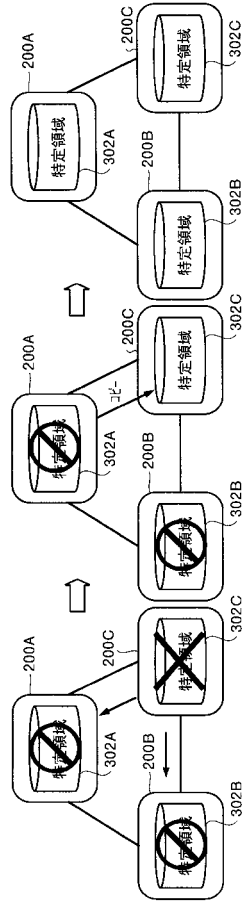
【図11】



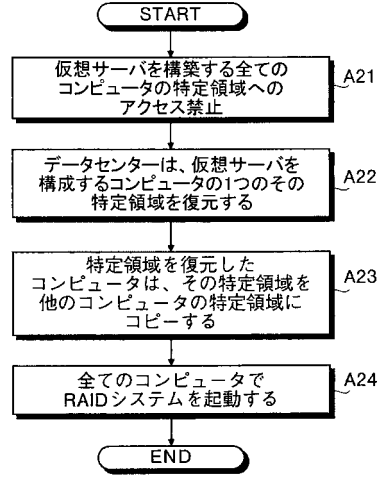
【図10】



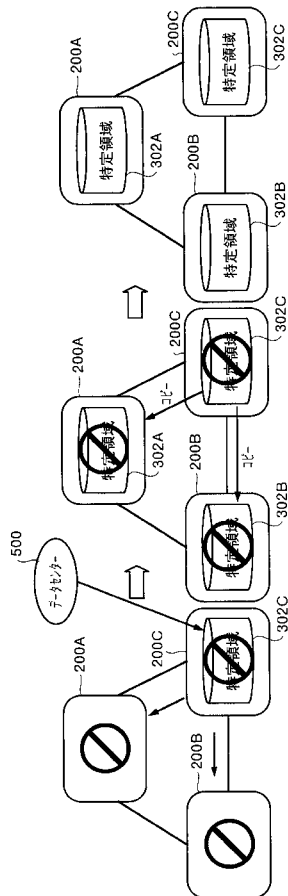
【図12】



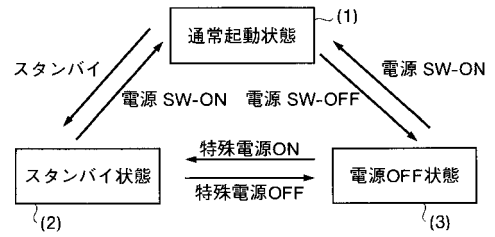
【図13】



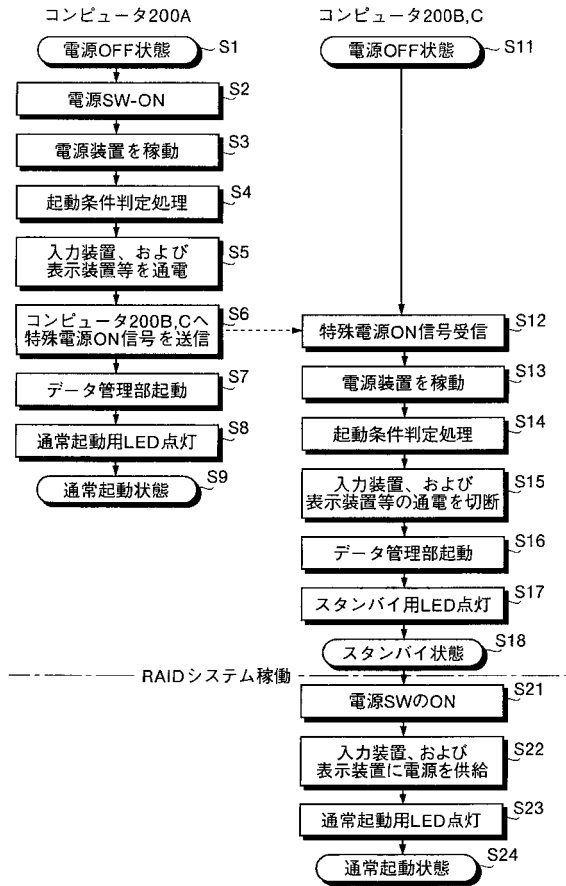
【図14】



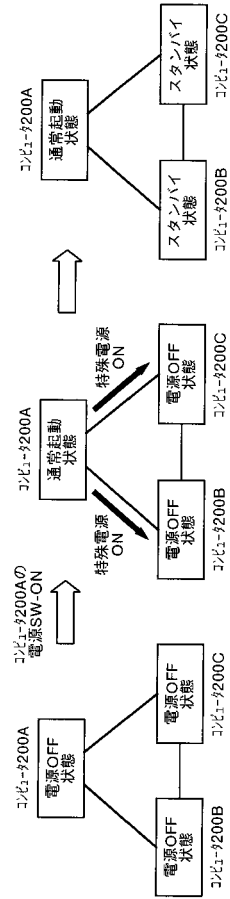
【図15】



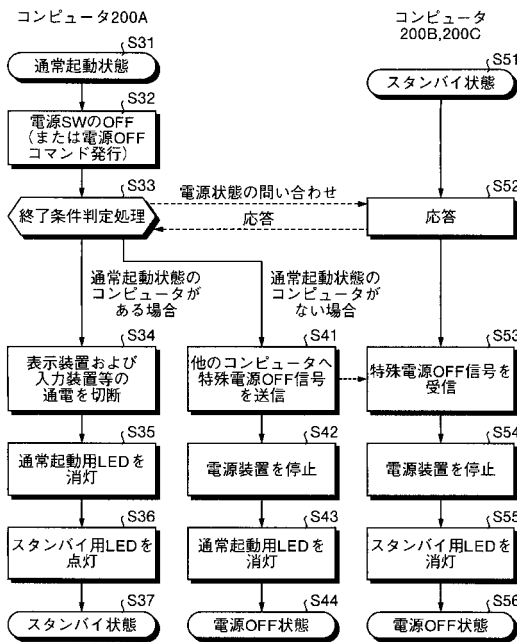
【図16】



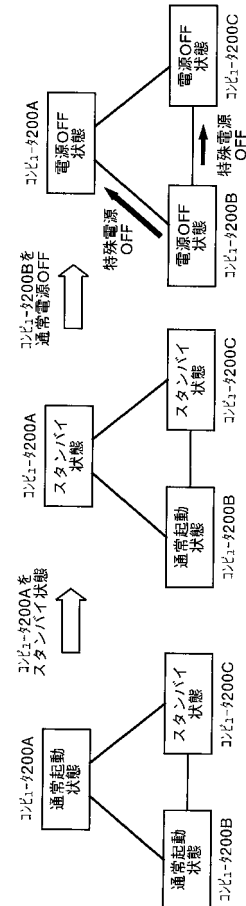
【図17】



【図18】



【図19】



フロントページの続き

(51)Int.Cl. F I
G 0 6 F 3/06 3 0 1 Z

(72)発明者 山川 正一郎
東京都江東区新砂 1 - 2 - 3 株式会社日本デジタル研究所内

審査官 田川 泰宏

(56)参考文献 特開2003-006015(JP,A)
特開平05-334006(JP,A)
特開2003-316522(JP,A)
特開2003-296039(JP,A)
特開平07-114495(JP,A)
特開2001-337789(JP,A)
特開2000-259289(JP,A)
松本 尚, NICを活用したネットワークRAID方式の提案, 情報処理学会研究報告 Vol
. 2000 No. 74, 日本, 社団法人情報処理学会, 2000年 8月 4日, 第2000巻,
p.79-84

(58)調査した分野(Int.Cl., DB名)
G 0 6 F 1 2 / 0 0
G 0 6 F 3 / 0 6