

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-18556
(P2012-18556A)

(43) 公開日 平成24年1月26日(2012.1.26)

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

F I
G06F 11/20 310B

テーマコード(参考)
5B034

審査請求 未請求 請求項の数 16 O L (全 28 頁)

(21) 出願番号 特願2010-155596(P2010-155596)
(22) 出願日 平成22年7月8日(2010.7.8)

(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 100075513
弁理士 後藤 政喜
(74) 代理人 100114236
弁理士 藤井 正弘
(74) 代理人 100120260
弁理士 飯田 雅昭
(72) 発明者 爲重 貴志
神奈川県川崎市麻生区王禅寺1099番地
株式会社日立製作所システム開発研究所
内

最終頁に続く

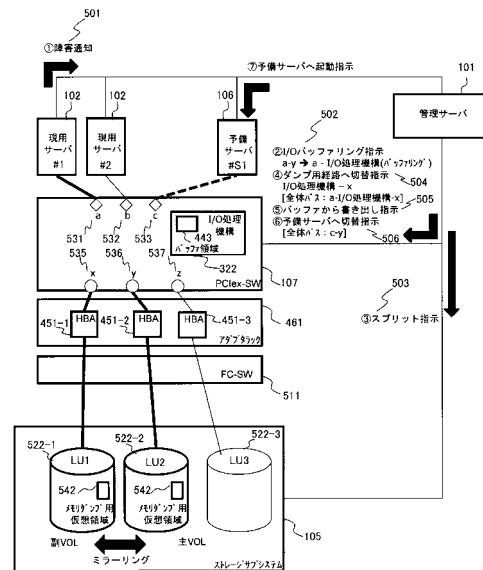
(54) 【発明の名称】 計算機システム及び計算機システムの系切替制御方法

(57) 【要約】

【課題】OSの種別に係らずメモリダンプを取得しながら系切替を高速で行うことを目的とする。

【解決手段】第1の計算機及び第2の計算機とストレージ装置との間で、第1の計算機からのI/O出力を一時的に格納するバッファと、バッファの内容をストレージ装置に出力する制御部と、を備えたI/O処理部を備えて、管理計算機は系切替の契機で、第1の計算機のI/O出力をI/O処理部のバッファへ格納し、ミラーボリュームの第1の記憶部と第2の記憶部を分離させ、バッファと第2の記憶部とを接続し、第2の計算機と第1の記憶部とを接続させ、バッファの内容を第2の記憶部に出力し、第2の計算機を第1の記憶部から起動させる。

【選択図】図5



【特許請求の範囲】

【請求項 1】

プロセッサとメモリ及び I / O インターフェースを備える第 1 の計算機と、
 プロセッサとメモリ及び I / O インターフェースを備える第 2 の計算機と、
 前記第 1 の計算機と第 2 の計算機からアクセス可能なストレージ装置と、
 ネットワークを介して前記第 1 の計算機と第 2 の計算機に接続されて、所定のタイミン
 グで前記第 1 の計算機を、前記第 2 の計算機に引き継ぐ系切替を行う管理計算機と、を備
 え、前記第 1 の計算機が、所定の条件となったときに前記メモリの内容を前記ストレ
 ージ装置に書き込む I / O 出力を送信する計算機システムにおいて、

前記ストレージ装置は、

10

前記第 1 の計算機がアクセスする第 1 の記憶部と、前記第 1 の記憶部のミラーリングを
 行う第 2 の記憶部と、を有し、

前記第 1 の計算機及び第 2 の計算機と前記ストレージ装置との間で、前記 I / O 出力を
 一時的に格納するバッファと、前記バッファの内容を前記ストレージ装置に出力する制御
 部と、を備えた I / O 処理部と、

前記 I / O 処理部と前記第 1 の計算機及び第 2 の計算機が前記ストレージ装置をアクセ
 スする経路を切り替えるスイッチ部と、を備え、

前記管理計算機は、

前記所定のタイミングとなったときに、前記第 1 の計算機の前記 I / O 出力を前記バッ
 ファへ格納する指令を前記前記 I / O 処理部に送信するバッファリング指示部と、

20

前記第 1 の記憶部と第 2 の記憶部を分離する指令を前記ストレージ装置に送信するスト
 レージ制御部と、

前記バッファと前記第 2 の記憶部とを接続し、前記第 2 の計算機と前記第 1 の記憶部と
 を接続する指令を前記スイッチ部に送信する経路切替部と、

前記バッファの内容を前記第 2 の記憶部に出力する指令を前記 I / O 処理部へ送信する
 書き出し指示部と、

前記第 2 の計算機を前記第 1 の記憶部から起動させる系切替部と、を有することを特徴
 とする計算機システム。

【請求項 2】

請求項 1 に記載の計算機システムであって、

30

前記管理計算機は、

前記第 1 の計算機に障害が発生したことを検知する障害検知部をさらに有し、

前記障害を検知したときを前記所定のタイミングとして前記系切替を行うことを特徴と
 する計算機システム。

【請求項 3】

請求項 1 に記載の計算機システムであって、

前記第 1 の計算機が前記 I / O 出力を出力したことを検知する監視部をさらに有し、

前記管理計算機は、

前記監視部が、前記第 1 の計算機が前記 I / O 出力を検知したときを、前記所定のタイ
 ミングとして前記系切替を行うことを特徴とする計算機システム。

40

【請求項 4】

請求項 1 に記載の計算機システムであって、

前記ストレージ制御部は、

前記第 1 の記憶部へ前記 I / O 出力が完了した後に、当該第 1 の記憶部を予め設定した
 保守用のグループへ移動させることを特徴とする計算機システム。

【請求項 5】

請求項 1 に記載の計算機システムであって、

前記ストレージ制御部は、

前記第 2 の計算機がアクセスする前記第 2 の記憶部のミラーリングを行う第 3 の記憶部
 を設定することを特徴とする計算機システム。

50

【請求項 6】

請求項 1 に記載の計算機システムであって、
前記スイッチ部は、
前記第 1 の計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、前記第 2 の計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、を制御する I / O スwitch であることを特徴とする計算機システム。

【請求項 7】

請求項 1 に記載の計算機システムであって、
物理計算機を仮想化する仮想化部をさらに有し、
前記仮想化部は、
前記第 1 の計算機として、仮想プロセッサと仮想メモリ及び仮想 I / O インターフェースとを備えた第 1 の仮想計算機を割り当て、
前記第 2 の計算機として、仮想プロセッサと仮想メモリ及び仮想 I / O インターフェースとを備えた第 2 の仮想計算機を割り当て、
前記スイッチ部として、前記第 1 の仮想計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、前記第 2 の仮想計算機の仮想 I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、を制御し、
前記第 1 の計算機は、所定の条件となったときに前記仮想メモリの内容を出力するメモリダンプ部を有し、
前記メモリダンプ部は、
前記所定の条件となったときに前記仮想メモリの内容を前記ストレージ装置に書き込む I / O 出力を前記仮想 I / O インターフェースに送信することを特徴とする計算機システム。

10

20

【請求項 8】

請求項 1 に記載の計算機システムであって、
前記経路切替部は、
前記第 1 の計算機の I / O インターフェースと前記バッファとを接続し、前記バッファと前記第 2 の記憶部とを接続する指令を前記スイッチ部へ送信し、
前記第 2 の計算機の I / O インターフェースと前記第 1 の記憶部とを接続する指令を前記スイッチ部に送信することを特徴とする計算機システム。

30

【請求項 9】

プロセッサとメモリ及び I / O インターフェースを備える第 1 の計算機と、プロセッサとメモリ及び I / O インターフェースを備える第 2 の計算機と、前記第 1 の計算機と第 2 の計算機からアクセス可能なストレージ装置と、ネットワークを介して前記第 1 の計算機と第 2 の計算機に接続されて、所定のタイミングで前記第 1 の計算機を前記第 2 の計算機に引き継ぐ系切替を行う管理計算機と、を備えて、前記第 1 の計算機が、所定の条件となったときに前記メモリの内容を前記ストレージ装置に書き込む I / O 出力を送信する計算機システムの系切替制御方法において、

40

前記計算機システムは、
前記第 1 の計算機及び第 2 の計算機と前記ストレージ装置との間で、前記 I / O 出力を一時的に格納するバッファと、前記バッファの内容を前記ストレージ装置に出力する制御部と、を備えた I / O 処理部と、

前記 I / O 処理部と前記第 1 の計算機及び第 2 の計算機が前記ストレージ装置にアクセスする経路を切り替えるスイッチ部と、を備え、

前記系切替制御方法は、

前記管理計算機が、前記第 1 の計算機がアクセスする第 1 の記憶部と、前記第 1 の記憶部のミラーリングを行う第 2 の記憶部とを前記ストレージ装置に設定する第 1 のステップと、

前記管理計算機が、前記所定のタイミングとなったときに、前記第 1 の計算機の前記 I

50

／＼出力を前記バッファへ格納する指令を前記前記 I / O 処理部に送信する第 2 のステップと、

前記管理計算機が、前記第 1 の記憶部と第 2 の記憶部を分離する指令を前記ストレージ装置に送信する第 3 のステップと、

前記管理計算機が、前記バッファと前記第 2 の記憶部とを接続し、前記第 2 の計算機と前記第 1 の記憶部とを接続する指令を前記スイッチ部に送信する第 4 のステップと、

前記管理計算機が、前記バッファの内容を前記第 2 の記憶部に出力する指令を前記 I / O 処理部へ送信する第 5 のステップと、

前記管理計算機が、前記第 2 の計算機を前記第 1 の記憶部から起動させる第 6 のステップと、を含むことを特徴とする計算機システムの系切替制御方法。

10

【請求項 10】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前記第 2 のステップは、

前記管理計算機が、前記第 1 の計算機に障害が発生したことを検知するステップをさらに有し、

前記障害を検知したときを前記所定のタイミングとして前記系切替を行うことを特徴とする計算機システムの系切替制御方法。

【請求項 11】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前記計算機システムは、

前記第 1 の計算機が前記 I / O 出力を出力したことを検知する監視部をさらに有し、

前記第 2 のステップは、

前記管理計算機が、前記第 1 の計算機が前記 I / O 出力を前記監視部で検知したときを、前記所定のタイミングとして前記系切替を行うことを特徴とする計算機システムの系切替制御方法。

20

【請求項 12】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前記管理計算機は、前記第 1 の記憶部へ前記 I / O 出力が完了した後に、当該第 1 の記憶部を予め設定した保守用のグループへ移動させる指令を送信する第 7 のステップを、さらに含むことを特徴とする計算機システムの系切替制御方法。

30

【請求項 13】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前記第 6 のステップは、

前記管理計算機が、前記第 2 の計算機がアクセスする前記第 2 の記憶部のミラーリングを行う第 3 の記憶部を設定する指令を前記ストレージ装置に送信するステップを含むことを特徴とする計算機システムの系切替制御方法。

【請求項 14】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前記スイッチ部は、

前記第 1 の計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、前記第 2 の計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、を制御する I / O スwitch であることを特徴とする計算機システムの系切替制御方法。

40

【請求項 15】

請求項 9 に記載の計算機システムの系切替制御方法であって、

物理計算機を仮想化する仮想化部をさらに有し、

前記仮想化部は、

前記第 1 の計算機として、仮想プロセッサと仮想メモリ及び仮想 I / O インターフェースとを備えた第 1 の仮想計算機を割り当て、

前記第 2 の計算機として、仮想プロセッサと仮想メモリ及び仮想 I / O インターフェー

50

スとを備えた第 2 の仮想計算機を割り当て、

前記スイッチ部として、前記第 1 の仮想計算機の I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、前記第 2 の仮想計算機の仮想 I / O インターフェースと前記ストレージ装置とを接続する I / O デバイスの経路と、を制御し、

前記第 1 の計算機は、所定の条件となったときに前記仮想メモリの内容を入力するメモリダンプ部を有し、

前記メモリダンプ部は、

前記所定の条件となったときに前記仮想メモリの内容を前記ストレージ装置に書き込む I / O 出力を前記仮想 I / O インターフェースに送信することを特徴とする計算機システムの系切替制御方法。

10

【請求項 16】

請求項 9 に記載の計算機システムの系切替制御方法であって、

前第 4 のステップは、

前記管理計算機が、前記第 1 の計算機の I / O インターフェースと前記バッファとを接続し、前記バッファと前記第 2 の記憶部とを接続する指令を前記スイッチ部へ送信し、

前記管理計算機が、前記第 2 の計算機の I / O インターフェースと前記第 1 の記憶部とを接続する指令を前記スイッチ部に送信することを特徴とする計算機システムの系切替制御方法。

【発明の詳細な説明】

【技術分野】

20

【0001】

本発明は、障害の発生した計算機を切り替えるコールドスタンバイシステムに関し、特に系切替を高速化することによる可用性を向上させる技術に関する。

【背景技術】

【0002】

計算機システムにおいて、障害が発生した計算機の OS が出力するメモリダンプは、障害の原因を特定する上で有益な情報である。また、障害が発生した計算機システムを早期に復旧させ、業務を再開することは計算機システムにとって重要である。例えば、コールドスタンバイシステムにおいて、系切替時に障害解析用のメモリダンプを取得する方法が提案されている。現用系におけるメモリダンプ出力が完了した後、予備系へ LU (Logical Unit) を接続して系切替を実施するが、メモリダンプの採取と系切替がシーケンシャルであるため切替までに時間を要する。そのため、メモリダンプを採取しつつ、障害発生後に速やかに予備系で業務を再開させる、迅速なシステム復旧の実現が望まれている。また、OS によってはメモリダンプ用の領域をブートボリュームに持つ必要があり、メモリダンプ用の領域を分離出来ない。

30

【0003】

また、障害発生時のメモリダンプを高速化する技術としては特許文献 1 が知られている。

【先行技術文献】

【特許文献】

40

【0004】

【特許文献 1】特開 2007 - 257486

【発明の概要】

【発明が解決しようとする課題】

【0005】

従来のコールドスタンバイシステムでは、メモリダンプの出力が完了するのを待って、系切替を行うか、一部の OS は対応していないブートボリュームとメモリダンプ出力先となる LU を分離するシステム構成にせざるを得なかった。

【0006】

また、上記特許文献 1 では、メモリを二重化することで、系切替を実施する際にメモリ

50

の内容を保存出来るシステム構成となっている。ただし、特許文献1ではメモリダンプを取得する計算機が同一のため、系切替時にメモリダンプを採取できない、という問題があった。

【0007】

そこで、本発明は上記問題点に鑑みてなされたもので、OSの種別に係わらずメモリダンプを取得しながら系切替を高速で行うことを目的とする。

【課題を解決するための手段】

【0008】

本発明は、プロセッサとメモリ及びI/Oインターフェースを備える第1の計算機と、プロセッサとメモリ及びI/Oインターフェースを備える第2の計算機と、前記第1の計算機と第2の計算機からアクセス可能なストレージ装置と、ネットワークを介して前記第1の計算機と第2の計算機に接続されて、所定のタイミングで前記第1の計算機を、前記第2の計算機に引き継ぐ系切替を行う管理計算機と、を備え、前記第1の計算機が、所定の条件となったときに前記メモリの内容を前記ストレージ装置に書き込むI/O出力を送信するメモリダンプ部を有する計算機システムにおいて、前記ストレージ装置は、前記第1の計算機がアクセスする第1の記憶部と、前記第1の記憶部のミラーリングを行う第2の記憶部と、を有し、前記第1の計算機及び第2の計算機と前記ストレージ装置との間で、前記I/O出力を一時的に格納するバッファと、前記バッファの内容を前記ストレージ装置に出力する制御部と、を備えたI/O処理部と、前記I/O処理部と前記第1の計算機及び第2の計算機が前記ストレージ装置をアクセスする経路を切り替えるスイッチ部と、を備え、前記管理計算機は、前記所定のタイミングとなったときに、前記第1の計算機の前記I/O出力を前記バッファへ格納する指令を前記前記I/O処理部に送信するバッファリング指示部と、前記第1の記憶部と第2の記憶部を分離する指令を前記ストレージ装置に送信するストレージ制御部と、前記バッファと前記第2の記憶部とを接続し、前記第2の計算機と前記第1の記憶部とを接続する指令を前記スイッチ部に送信する経路切替部と、前記バッファの内容を前記第2の記憶部に出力する指令を前記I/O処理部へ送信する書き出し指示部と、前記第2の計算機を前記第1の記憶部から起動させる系切替部と、を有する。

【発明の効果】

【0009】

したがって、本発明は、障害等の所定のタイミングで、現用系の第1の計算機からのI/O出力の収集をOSの種類にかかわらず確実にいながらも、予備系の第2の計算機への系切替を迅速に行うことが可能となる。特に、ミラーボリュームの第1の記憶部と第2の記憶部をスプリットした後は、第1の計算機のI/O出力と第2の計算機への系切替を並列的に行うことで、I/O出力の完了を待たずに系切替を開始できるので、コールドスタンバイによるフェイルオーバの高速化を図ることができる。

【図面の簡単な説明】

【0010】

【図1】本発明の第1の実施形態を示し、系切替を行う計算機システムの一例を示すブロック図である。

【図2】本発明の第1の実施形態を示し、管理サーバの構成を示すブロック図である。

【図3】本発明の第1の実施形態を示し、現用系のサーバ102または予備系のサーバ106の構成を示すブロック図である。

【図4】本発明の第1の実施形態を示し、PCIe x - SW107及びアダプタの構成を示すブロック図である。

【図5】本発明の第1の実施形態を示し、PCIe x - SW107を主体とするフェイルオーバの概略を示すブロック図である。

【図6】本発明の第1の実施形態を示し、サーバ管理テーブル221を示す説明図である。

【図7】本発明の第1の実施形態を示し、LUマッピング管理テーブル222を示す説明

10

20

30

40

50

図である。

【図 8】本発明の第 1 の実施形態を示し、LU 管理テーブル 2 2 3 を示す説明図である。

【図 9】本発明の第 1 の実施形態を示し、PCIe x - SW 1 0 7 の I / O 処理機構 3 2 2 内の I / O バッファ管理テーブル 4 1 1 を示す説明図である。

【図 1 0】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 の制御部 1 1 0 で行われる処理の一例を示すフローチャートである。

【図 1 1】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 の I / O バッファリング指示部 2 1 1 で行われる処理の一例を示すフローチャートである。

【図 1 2】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 の経路切替部 2 1 3 で行われる処理の一例を示すフローチャートである。

【図 1 3】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 の I / O バッファ書出し指示部 2 1 4 で行われる処理の一例を示すフローチャートである。

【図 1 4】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 の N + M 切替指示部 2 1 5 で行われる処理の一例を示すフローチャートである。

【図 1 5】本発明の第 1 の実施形態を示し、I / O 処理機構 3 2 2 の I / O バッファリング制御部 4 0 1 で行われる処理の一例を示すフローチャートである。

【図 1 6】本発明の第 1 の実施形態を示し、管理サーバ 1 0 1 が管理する業務及び S L A 管理テーブル 2 2 4 の一例を示す説明図である。

【図 1 7】本発明の第 2 の実施形態を示し、第 2 の実施形態を示すサーバ 1 0 2 (または 1 0 6) のブロック図である。

【図 1 8】本発明の第 2 の実施形態の処理の概要を示すブロック図である。

【図 1 9】本発明の第 3 の実施形態を示し、PCIe x - SW 1 0 7 を主体とするフェイルオーバーの概略を示すブロック図である。

【図 2 0】本発明の第 1 の実施形態を示し、メモリダンプの書き込みが完了した LU 1 を、予め設定した保守用の領域へ退避させる例を示すブロック図である。

【発明を実施するための形態】

【0 0 1 1】

以下、本発明の一実施形態を添付図面に基づいて説明する。

【0 0 1 2】

図 1 は、本発明の第 1 の実施形態を示し、系切替を行う計算機システムの一例を示すブロック図である。

【0 0 1 3】

管理サーバ 1 0 1 は、NW - SW (管理用ネットワークスイッチ) 1 0 3 を介して、NW - SW 1 0 3 の管理インタフェース(管理 I / F) 1 1 3、NW - SW (業務用ネットワークスイッチ) 1 0 4 の管理インタフェース 1 1 4 へ接続されており、管理サーバ 1 0 1 から各 NW - SW の VLAN (Virtual LAN) を設定することが可能である。

【0 0 1 4】

NW - SW 1 0 3 は、管理用のネットワークを構成し、現用系のサーバ 1 0 2 や予備系のサーバ 1 0 6 に対して、OS やアプリケーションの配布や電源制御等の運用管理をするためのネットワークである。NW - SW 1 0 4 は、業務用のネットワークを構成し、サーバ 1 0 2、1 0 6 上で実行される業務用アプリケーションが使用するネットワークである。なお、NW - SW 1 0 4 は、WAN 等に接続されて計算機システムの外部のクライアント計算機と通信を行う。

【0 0 1 5】

管理サーバ 1 0 1 は、FC - SW (ファイバーチャネル・スイッチ) 5 1 1 を介してストレージサブシステム 1 0 5 に接続される。管理サーバ 1 0 1 は、ストレージサブシステム 1 0 5 内の N 個の LU (Logical Unit) 1 ~ LU n を管理する。

【0 0 1 6】

管理サーバ 1 0 1 上では、サーバ 1 0 2、1 0 6 を管理する制御部 1 1 0 が実行され、管理テーブル群 1 1 1 を参照および更新する。管理テーブル群 1 1 1 は制御部 1 1 0 によ

10

20

30

40

50

って所定の周期などで更新される。

【0017】

管理対象となるサーバ102は、N+Mコールドスタンバイを提供するシステムにおける現用系サーバであり、同様に予備系である物理サーバ106とともに、PCIex-SW107とI/Oデバイス(図中、HBA)を介して、NW-SW103および104に接続される。PCIex-SW107には、PCI Express規格のI/Oデバイス(NIC(Network Interface Card)、HBA(Host Bus Adapter)、CNA(Converged Network Adapter)、といったI/Oアダプタ)が接続されている。一般的に、PCIex-SW107は、マザーボード(またはサーバブレード)より外へPCI Expressのバスを延長し、さらに多数のPCI-Expressデバイスを接続すること

10

20

30

40

50

【0018】

本実施形態の計算機システムでは、PCIex-SW107内の通信経路を切り替えることで、N+Mコールドスタンバイシステムを実現する。N+Mコールドスタンバイシステムでは、現用系のサーバ102に障害が発生すると、当該サーバ102の業務を予備系のサーバ106に引き継ぐ系切替が管理サーバ101によって実施される。系切替の際、障害が発生した瞬間から特定のI/O出力として出力される現用系サーバ102のメモリダンプを漏れなく収集し、かつ、障害発生から間を置かず、障害が発生した現用系サーバ102で稼働していた業務システムを予備系のサーバ106へフェイルオーバーさせる。これにより、収集したメモリダンプから障害原因を特定しつつ、業務システムは再起動程度の寸断で動き続けることが可能になる。

【0019】

また、管理サーバ101はPCIex-SW107の管理インターフェース1070に接続され、サーバ102、106とI/Oデバイスの接続関係を管理する。

【0020】

また、サーバ102、106は、PCIex-SW107に接続されたI/Oデバイス(図中HBA)を介してストレージサブシステム105のLU1~LU_nにアクセスする。また、ディスクインターフェース203は、管理サーバ101の内蔵ディスクやストレージサブシステム105のインターフェースである。また、現用系のサーバ102は図中#1~#3で識別し、予備系のサーバ106は、図中#S1、#S2で識別する。

【0021】

図2は、管理サーバ101の構成を示すブロック図である。管理サーバ101は、演算を処理するCPU(Central Processing Unit)201、CPU201で演算するプログラムや、プログラムの実行に伴うデータを格納するメモリ202、プログラムやデータを格納するストレージ装置とのディスクインタフェース203、IPネットワークを介した通信のためのネットワークインタフェース204から構成される。

【0022】

図2では、ネットワークインタフェース204及びディスクインタフェース203を、それぞれ代表して一つずつ示しているが、各々が複数あるものとする。たとえば、管理用ネットワークのNW-SW103と業務用ネットワークのNW-SW104との接続は、各々異なるネットワークインタフェース204を用いる。

【0023】

メモリ202には、制御部110および管理テーブル群111が格納されている。制御部110は、障害検出部210、I/Oバッファリング指示部211(図11参照)、ストレージ制御部212、経路切替部213(図12参照)、I/Oバッファ書出し指示部214(図13参照)、及びN+M切替指示部215(図14参照)を有する。

【 0 0 2 4 】

障害検出部 2 1 0 は、サーバ 1 0 2、1 0 6 の障害を検知し、障害を検知したときには N + M 切替指示部 2 1 5 により後述のサーバ管理テーブル 2 2 1 を参照して上述の系切替を実施する。なお、障害の検知及びフェイルオーバーについては公知または周知の技術を適用すればよいので、本実施形態では詳述しない。

【 0 0 2 5 】

ストレージ制御部 2 1 2 は後述の L U 管理テーブル 2 2 3 を用いてストレージサブシステム 1 0 5 の L U 1 ~ L U n を管理する。

【 0 0 2 6 】

管理テーブル群 1 1 1 は、サーバ管理テーブル 2 2 1 (図 6 参照)、L U マッピング管理テーブル 2 2 2 (図 7 参照)、L U 管理テーブル 2 2 3 (図 8 参照)、業務及び S L A (Service Level Agreement) 管理テーブル 2 2 4 (図 1 6 参照) を有する。

10

【 0 0 2 7 】

各テーブルの情報収集は O S (図示省略) の標準インタフェースや情報収集用プログラムを使用した自動収集でも良いし、手動で利用者 (または管理者) に入力させても良い。ただし、規則や方針といった情報のうち物理的要件や法律の要請で限界値が決定されるもの以外は、利用者に予め入力させる必要があり、入力用のインタフェースを備えてもよい。また、利用者の方針によって、限界値に至らない運用をする場合も同様に条件を入力するインタフェースを備えてもよい。

【 0 0 2 8 】

管理サーバ 1 0 1 の種別については、物理サーバ、ブレードサーバ、仮想化されたサーバ、論理分割または物理分割されたサーバなどのいずれであっても良く、いずれを使った場合も本発明の効果を得ることが出来る。

20

【 0 0 2 9 】

図 3 は、現用系のサーバ 1 0 2 または予備系のサーバ 1 0 6 の構成を示すブロック図である。現用系のサーバ 1 0 2 と予備系のサーバ 1 0 6 の構成が一致する必要性は必ずしもない。ただし、構成が一致する場合、N + M コールドスタンバイにて切替を実施した場合に、問題が発生しにくい。これは、N + M コールドスタンバイによる切替動作が、O S にとっては再起動と同じように見えることに起因する。この効果は、本願でも有効である。以下では、現用系のサーバ 1 0 2 と予備系のサーバ 1 0 6 が同一の構成である場合について説明する。

30

【 0 0 3 0 】

サーバ 1 0 2、1 0 6 は、演算を処理する C P U 3 0 1、C P U 3 0 1 で演算するプログラムや、プログラムの実行に伴いデータを格納するメモリ 3 0 2、プログラムやデータを格納するストレージ装置とのディスクインタフェース 3 0 4、I P ネットワークを介して通信を行うためのネットワークインタフェース 3 0 3、電源制御や各インタフェースの制御を行う B M C (Basement Management Controller) 3 0 5、P C I e x - S W に接続するための P C I - E x p r e s s インタフェース 3 0 6 を有する。

【 0 0 3 1 】

メモリ 3 0 2 上の O S 3 1 1 が C P U 3 0 1 によって実行され、サーバ 1 0 2 または 1 0 6 内のデバイス及びタスクの管理を行っている。O S 3 1 1 の下で、業務を提供するアプリケーション 3 2 1 や監視プログラム 3 2 2 などが動作する。監視プログラム 3 2 2 はサーバ 1 0 2、1 0 6 の障害を検知し、管理サーバ 1 0 1 に通知する。O S 3 1 1 は、所定の条件でメモリ 3 0 2 の内容を、ストレージサブシステム 1 0 5 に書き込むメモリダンプを出力するメモリダンプ部 3 1 1 0 を有する。なお、O S 3 1 1 がメモリダンプ部 3 1 1 0 を機能させる所定の条件は、システム障害の発生時や、所定のコマンドの受け付け時などである。

40

【 0 0 3 2 】

図 3 では、ネットワークインタフェース 3 0 3、ディスクインタフェース 3 0 4 および

50

PCI-Express インタフェース 306 を、それぞれ代表して一つずつ示しているが、各々が複数あるものとする。たとえば、管理用ネットワークの NW-SW103 と業務用ネットワークの NW-SW104 との接続は、各々異なるネットワークインタフェース 303 を用いる。あるいは、サーバ 102、106 は、図 1 のように PCIex インタフェースを介して接続された NIC を経由して NW-SW103 と業務用ネットワークの NW-SW104 に接続してもよい。

【0033】

現用系のサーバ 102 に障害が発生しておらず N+M 切替が発生していない場合、予備系のサーバ 106 のメモリ 302 上では OS 311 や他のプログラムは動作していない。ただし、情報収集や障害が発生していないかをチェックするプログラムが所定の周期などで実行されることはある。

10

【0034】

図 4 は、PCIex-SW107 を中心に、現用系のサーバ 102、予備系のサーバ 106 と、PCI-Express のアダプタ 451-1~451-5 (NIC、HBA、CNA などの I/O デバイス) およびそれらを格納したアダプタラック 461 やアダプタ 451 との接続構成を示している。なお、以下ではアダプタ 451-1~451-5 の総称をアダプタ 451 とする。

【0035】

PCIex-SW107 は、現用系のサーバ 102 および予備系のサーバ 106 と、PCIex インタフェース 306 を介して接続されている。また、PCIex-SW107 は、複数の PCI-express アダプタ 451 に接続されている。アダプタ 451 は、アダプタラック 461 に収められていても良いし、アダプタ 451 が直接、PCIex-SW107 に接続されていても良い。

20

【0036】

PCIex-SW107 は、I/O 処理機構 322 を備え、現用系のサーバ 102 または予備系のサーバ 106 がアダプタ 451 に接続される際に、I/O 処理機構 322 を経由するパスと経由しないパスを持つ。本実施形態では、現用系のサーバ 102 のメモリダンプを漏れなく取得する機構の動作には、I/O 処理機構 322 がメモリダンプを一時的に保持するバッファ領域 443 と、バッファ領域 443 を制御する制御部 441 ならびに管理テーブル群 442 を備える。管理テーブル群 442 は、制御部 441 によって所定の周期、あるいは管理サーバ 101 からの構成変更の指令などに応じて更新される。

30

【0037】

制御部 441 は、アダプタ (I/O デバイス) 451 と現用系のサーバ 102 及び予備系のサーバ 106 の接続を制御し、バッファ領域 443 へのアクセスを制御する I/O バッファリング制御部 401 から構成されている (図 15 参照)。

【0038】

管理テーブル群 442 は、I/O バッファリング管理テーブル 411 から構成されている (図 9 参照)。

【0039】

また、PCIex-SW107 は、後述するように、サーバ 102、106 に接続されるポート (上流ポート) と、アダプタ 451-1~451-5 に接続されるポート (下流ポート) を備える。制御部 441 は、上流ポートと下流ポートの接続関係を変更することで、サーバ 102、106 に割り当てるアダプタ 451-1~451-5 を変更することができる。なお、図示の例では、アダプタ 451-1~451-5 が 5 つの場合を示しているが、図 1 に示す NIC、HBA のように、多数のアダプタ 451 を備えることができる。また、本実施形態では、アダプタ 451-1~451-3 が HBA で構成された例を示す。

40

【0040】

図 5 は、PCIex-SW107 を主体とするフェイルオーバの概略を示すブロック図である。図 5 の例は、現用系のサーバ 102 (以下、現用系サーバ #1) で障害が発生し

50

て、現用系サーバ# 1のメモリダンプを行いながら、予備系のサーバ106（以下、予備系サーバ# S1）に系切替を行う例を示している。

【0041】

前提条件としては、現用系サーバ# 1は、PCIex - SW107のポートa531に接続され、予備系サーバ# S1はポートc533に接続される。また、PCIex - SW107を介して現用系サーバ# 1に割り当てられたストレージサブシステム105の記憶領域は、LU2（522 - 2）がポートy536に接続されて主ボリュームとして機能する。LU2にはOSのブートイメージ、業務アプリケーション等が格納される。また、LU1（522 - 1）はLU2の副ボリュームとして設定され、ミラーボリュームが構成される。ポートy536にはHBAで構成されたアダプタ451 - 2が接続され、FC - SW511を介してLU2に接続される。また、ポートy535にはHBAで構成されたアダプタ451 - 1が接続される。

10

【0042】

現用系サーバ# 1がミラーボリュームの主ボリュームであるLU2にデータを書き込むと、ストレージサブシステム105のミラーリング機能によって、LU2の内容が副ボリュームLU1に複製される。

【0043】

PCIex - SW107は、ポートa531とポートy536を接続し、現用系サーバ# 1からHBAで構成されたアダプタ451 - 2を介して主ボリュームのLU2にアクセスする。LU2に書き込まれたデータは、ストレージサブシステム105によってLU1

20

【0044】

管理サーバ101は、（1）現用系サーバ# 1（または他の現用系のサーバ102）から送られてくる障害通知501を受信した契機で、（2）I/O処理機構322へI/Oバッファリング指示を出し、ポートa531とポートy536が接続されていた構成から、ポートa531とI/O処理機構322を接続する。そして、I/O処理機構内のバッファ領域443へ、障害が発生した現用系サーバ# 1のI/O（メモリダンプ）を蓄積可能な構成へ変更する（502）。

30

【0045】

障害が発生した現用系サーバ# 1は、障害発生と同時にメモリダンプを出力（送信）しており、メモリダンプの一部は既に主VOLであるLU2（522 - 2）のメモリダンプ用仮想領域542へ出力されている。本実施形態では、LU2（522 - 2）を副ボリュームLU1とミラー構成とすることで、既に出力されたメモリダンプを漏らすことなく副ボリュームであるLU1にもコピーしておく。そして、I/O処理機構322は、現用系のサーバ102からのメモリダンプをバッファ領域443に蓄積する。I/O処理機構523はバッファ領域443にバッファリングしたメモリダンプを続けて書き込むことで、全てのメモリダンプのデータを回収することが可能になる。

【0046】

（3）管理サーバ101のストレージ制御部212が、主ボリュームのLU2と副ボリュームのLU1のミラーリングをスプリットする指示を出す（503）。なお、ストレージ制御部212はスプリット前に、強制的にミラーリングの同期をとるよう指示を出しても良い。強制的にミラー同期処理を入れる場合、同期処理が完了してからスプリットを実行する。次に、ストレージ制御部212スプリットした副ボリュームのLU1を主ボリュームに変更するよう指示を出す。これにより、障害発生と同時に主ボリュームのLU2のメモリダンプ用仮想領域542に書き込まれたメモリダンプを持つLU1、2が2つ作成されたことになる。どちらも、サーバ102または106に接続し、再起動することで業務を再開することが出来、また、メモリダンプを引き続いて書き込んでも漏れなくメモリダンプを採取することが可能である。

40

50

【 0 0 4 7 】

ここで、予備系のサーバ106に接続して業務を再開するLU1と、副ボリュームとして、ある別のLU_n（第3の記憶部）をミラー構成のペアとすることで、再度、障害が発生しても、本発明の効果を得つつ、別のシステムに高速に切替えることが可能になる。

【 0 0 4 8 】

（4）経路切替部213（図12参照）が、I/O処理機構322と先の2つの主ボリュームのLU1を接続する（504）。すなわち、I/O処理機構523のバッファ領域443とポートx535を接続し、HBA451-1を介してLU1に接続する。このとき、元々、副ボリュームであったLU1を選択し、メモリダンプを書き出す先として選択しても良いし、予備系のサーバ106に接続するようにしても良い。LU1をメモリダンプの書き出し先として選択すると、残ったLU2（最初から主ボリュームで、元々業務を提供していたLU522-2）は予備系のサーバ106（#S1）に接続することになる。この構成をとるメリットは、HBA451-2が切替前後で変わらないことである。これにより、予備系サーバ#S1を稼働させて業務を提供するOSやミドルウェアをはじめとするソフトウェア群からは、現用系サーバ#1から予備系サーバ#S1に代わっただけ（サーバ部分（主にCPUやメモリ）のみが代わっただけ）のようになるため、切替後の稼働に悪影響を及ぼしにくい。悪影響には、起動しない、だけでなく、起動後にデバイスが変わったとOSが認識することによるデバイスドライバの再組み込みや、再組み込みによるOS設定情報の破棄（再設定が必要になる）を回避することが出来る。しかし、HBA451-2が他のHBAに変わること特に業務継続に支障がないことが分かっていたり、対策を実施している場合、どちらのLU1、2を使っても良い。例えば、本実施形態ではI/O処理機構322とPCIex-SW107のポートx535を接続してバッファ領域443の内容を書き込む場合を詳述する。

10

20

【 0 0 4 9 】

この場合、障害が発生した現用系サーバ#1は、PCIex-SW107のポートa531と接続されているため、I/O処理機構322を介して、元々副ボリュームとしてペアを組んでいたLU2に接続されることになる。

【 0 0 5 0 】

（5）I/Oバッファ書出し指示部214（図13参照）が、I/O処理機構322へバッファ領域443に蓄積しているメモリダンプを書き出すよう指示を出す（505）。これにより、LU1のメモリダンプ用仮想領域542にバッファリングされた後のデータがバッファ領域443から書き加えられていく。

30

【 0 0 5 1 】

このようにして、障害発生と同時に書き出されるメモリダンプのデータを漏らすことなく、LU1に格納することが可能になる。

【 0 0 5 2 】

（6）N+M切替指示部215（図14）が、PCIex-SW107にLU2と予備系サーバ#S1を接続するよう指示する。具体的には、PCIex-SW107のポートc533とポートy536を接続する（506）。

40

【 0 0 5 3 】

上記のようにして、ブート用LU2とメモリダンプ用仮想領域542が同じLUまたはひとつのボリュームにしかメモリダンプ用仮想領域542の存在を許さない種類のOSでも、メモリダンプを採取しつつ、予備系サーバ#S1への切替と再起動を実施することが可能になる。

【 0 0 5 4 】

上記の（4）、（5）と（6）は並行して処理が実行されても良く、並行して実施することで予備系のサーバ106での起動開始を早められ、更なる高速切替を実現できる。

【 0 0 5 5 】

また、メモリダンプの書き込みが完了したLU1は、保守用の領域へ退避させたり、アクセス制限するなどして保護することで、操作ミスによるメモリダンプを採取したLU1

50

の喪失を防ぐことができ、本実施形態の効果を更に高めることが可能である。この例については、図 20 に後述する。

【0056】

図 6 は、サーバ管理テーブル 221 を示す説明図である。サーバ管理テーブル 221 は管理サーバ 101 の制御部 110 で管理される。

【0057】

カラム 601 には、サーバ 102、106 の識別子を格納しており、本識別子によって各サーバ 102、106 を一意に識別する。カラム 601 へ格納するデータは、本テーブルで使用される各カラムのいずれか、または複数カラムを組み合わせたものを指定することで入力を省略することが出来る。また、識別子は昇順などで管理サーバ 101 等が自動的に割り振っても良い。

10

【0058】

カラム 602 には、UUI D (U n i v e r s a l U n i q u e I D e n t i f i e r) が格納されている。UUI D は、重複しないように形式が規定された識別子である。そのため、各サーバ 102、106 に対応して、UUI D を保持することにより、確実なユニーク性を保証する識別子となる。ただし、カラム 601 には、システム管理者がサーバを識別する識別子を設定すれば良く、また管理する対象となるサーバ 102、106 間で重複することがなければ問題ないため、UUI D を使うことが望ましいものの必須とはならない。例えば、カラム 601 のサーバ識別子には、MAC アドレス、WWN (W o r l d W i d e N a m e) などを用いても良い。

20

【0059】

カラム 603 には、サーバの種別として、現用系サーバか予備系サーバかを格納している。また、系切替時にはどのサーバからの切替を受け付けたかも格納しても良い。

【0060】

カラム 604 には、サーバ 102、106 のステータスが格納されており、問題がなければ正常、障害が発生していれば障害を、それぞれ表すステータスが格納されている。障害発生時には、メモリダンプを書き出し中などの情報を格納しても良い。

【0061】

カラム 605 (カラム 621 ~ カラム 623) は、アダプタ 451 に関する情報を格納している。カラム 621 には、アダプタ 451 のデバイス種別を格納している。HBA (H o s t B u s A d a p t o r) や NIC や CNA (C o n v e r g e d N e t w o r k A d a p t e r) などが格納される。カラム 622 には、HBA の識別子である WWN、NIC の識別子である MAC アドレスが格納されている。

30

【0062】

カラム 606 には、現用系のサーバ 102 や予備系のサーバ 106 がアダプタ 451 を介して接続している NW - SW 103、104 や FC - SW 511 に関する情報が格納されている。種別や接続ポートおよびセキュリティ設定情報が格納されている。

【0063】

カラム 607 には、サーバのモデルを格納している。インフラに関する情報であり、性能や構成可能なシステム限界を知ることが出来る情報である。また、構成が同じか否かを判別することが出来る情報である。

40

【0064】

カラム 608 は、サーバの構成を格納している。プロセッサのアーキテクチャ、シャーシやスロットなどの物理位置情報、特徴機能 (ブレード間 SMP : S y m m e t r i c M u l t i - P r o c e s s i n g 、 H A 構成などの有無) を格納している。

【0065】

カラム 609 には、サーバの性能情報を格納している。

【0066】

図 7 は、LU マッピング管理テーブル 222 を示す説明図である。LU マッピング管理テーブル 222 は、管理サーバ 101 の制御部 110 で管理され、LU 522 とアダプタ

50

451とサーバ102、106との接続関係を格納している。

【0067】

カラム701には、ストレージサブシステム105内のLUの識別子を格納しており、本識別子によって各LUを一意に識別する。

【0068】

カラム702(カラム721~カラム722)には、アダプタ451に関する情報を格納している。カラム721には、デバイス種別を格納している。HBA(Host Bus Adaptor)やNICやCNA(Converged Network Adaptor)などが格納される。カラム722には、HBAの識別子であるWWN、NICの識別子であるMACアドレスが格納されている。

10

【0069】

カラム703には、PCIex-SW情報を格納している。PCIex-SW107のどのポートとポートが接続関係にあるか、また、I/O処理機構322との接続関係を格納している。

【0070】

図8は、LU管理テーブル223を示す説明図である。LU管理テーブル223は、管理サーバ101の制御部110で管理され、LUの種別やミラーリングの有無、ミラーのペア、ステータスを管理している。

【0071】

カラム801には、LUの識別子を格納しており、本識別子によって各LUを一意に識別する。

20

【0072】

カラム802には、LU種別を格納している。主ボリュームか副ボリュームか、といったミラーリングの主従関係を示す情報などが格納されている。

【0073】

カラム803には、ミラーリングを組んでいるペアとなる副ボリュームのLUを格納している。

【0074】

カラム804には、LUのステータスを格納している。ミラーリング状態、スプリット中、副ボリュームから主ボリュームへ変更中、ミラーリングする予定である予約、などを格納している。

30

【0075】

図9は、PCIex-SW107のI/O処理機構322内のI/Oバッファリング管理テーブル411を示す説明図である。I/Oバッファリング管理テーブル411は、制御部441によって管理され、バッファ領域443が接続されているサーバ102やアダプタ451および、バッファ領域443のステータスを管理している。

【0076】

カラム901は、I/Oバッファの識別子を格納しており、本識別子によって各バッファ領域443を一意に識別する。この識別子は、制御部441が予め設定した識別子を用いることができる。

40

【0077】

カラム902は、サーバ102、106の識別子を格納しており、本サーバ識別子によって各サーバを一意に識別する。サーバ識別子は管理サーバ101のサーバ管理テーブル221から取得した値を用いることができる。

【0078】

カラム903(カラム921~カラム922)には、アダプタ451に関する情報を格納している。カラム921には、デバイス種別を格納している。HBA(Host Bus Adaptor)やNICやCNA(Converged Network Adaptor)などが格納される。カラム922には、HBAの識別子であるWWN、NICの識別子であるMACアドレスが格納されている。アダプタ451に関する情報は管理サ

50

サーバ101のサーバ管理テーブル221から取得した値を用いることができる。あるいは、制御部441がアダプタ451をアクセスした値を用いてもよい。

【0079】

カラム904には、バッファ領域443のステータスを格納している。バッファ要求受付、データをバッファ中、バッファしたデータを書き出し中、などが格納される。

【0080】

カラム905には、バッファ領域443の使用ステータスが格納されている。使用中なのか未使用なのか、また使用している場合は使用している容量、エラー情報などである。また、予約する容量や優先順位に関する情報を格納し、バッファ領域443の容量を超えるデータをバッファするよう要求されたときに、どのバッファ領域のデータを救済するかを判定することが可能になる。

10

【0081】

カラム902やカラム903に格納されているアダプタ、デバイス、サーバはPCIe x-SW107のポート番号またはスロット番号で置き換えられる情報が格納されても良い。

【0082】

さらに、I/Oバッファリング管理テーブル411にはバッファ領域443でバッファリングに失敗した場合の対処を格納するカラムを設けても良い。例えば、再送要求を現用系のサーバ102に出す、失敗通知を管理サーバ101へ通知する、などである。また、管理サーバ101は、別のLUにつながったアダプタ451を障害が発生した現用系のサーバ102へ通知し、別LUへメモリ302の内容を書き出すようにしても良い。それにより、あふれたデータを救済することが可能になる。

20

【0083】

図10は、管理サーバ101の制御部110で行われる処理の一例を示すフローチャートである。この処理は、管理サーバ101がサーバ102、106から障害通知501を受信したときに起動される。なお、障害通知501は、サーバ102、106のBMC305やOS311等が障害を検知したときに管理サーバ101へ送信する。なお、以下では、現用系サーバ、LUの識別子を図5に示した値を用いる。

【0084】

ステップ1001で、障害検出部210が障害通知501により障害を検出する。障害を検出した場合、ステップ1002へ進む。

30

【0085】

ステップ1002で、I/Oバッファリング指示部211が、I/O処理機構322へ障害が発生した現用系サーバ#1のI/O出力(メモリダンプ)をバッファするよう指示し、ステップ1003へ進む。

【0086】

ステップ1003で、ストレージ制御部212が、ストレージサブシステム105に対して現用系サーバ#1が使用している主ボリュームLU2へミラーリングの同期処理を指示し、ステップ1004へ進む。

【0087】

ステップ1004で、ストレージ制御部212が、ストレージサブシステム105へLU2のミラーリング構成のスプリットを指示し、ステップ1005へ進む。このとき、スプリットした後に、必要に応じてペアであった副ボリュームのLU1を主ボリューム化する。また、別の副ボリュームであるLUを用意しておき、元のLU(予備系のサーバ106と接続して業務を再開するLU)とペアを組み、ミラーリング構成を再構成しても良い。

40

【0088】

ステップ1005で、経路切替部213が、I/O処理機構322とアダプタ451(メモリダンプ出力用のLU1に接続されているデバイス)と接続するよう指示し、ステップ1006へ進む。

50

【0089】

ステップ1006で、I/Oバッファ書出し指示部214がI/O処理機構322に対してバッファ領域443へ蓄積したメモリダンプのデータをステップ1005で設定したLU1に書き出すよう指示し、ステップ1007へ進む。

【0090】

ステップ1007で、N+M切替指示部215が、PCIex-SW107に予備系サーバ#S1に、障害が発生した現用系サーバ#1が使用していたアダプタ451(LU2)を接続するよう指示し、ステップ1008へ進む。

【0091】

ステップ1008で、予備系サーバ#S1を起動するよう指示し、処理を完了する。

10

【0092】

上記処理により、図5で示したように、現用系サーバ#1から障害通知501を受信すると、管理サーバ101はPCIex-SW107に対してバッファ領域443で現用系サーバ#1からのI/O出力を格納する指令を送信する。次に、管理サーバ101はストレージサブシステム105に対して現用系サーバ#1が利用しているLU2についてミラーリングの同期指示を送信し、主ボリュームのLU2と副ボリュームのLU1を同期させる。その後、管理サーバ101はストレージサブシステム105のミラーボリュームにスプリットを指示し、ミラーリングのペアを分離する指令を送信する。次に、管理サーバ101は、ミラーリングのペアを解除した一方のLU1にバッファ領域443の内容を書き込むようPCIex-SW107の制御部441に指令する。さらに、管理サーバ101は、ミラーリングのペアを解除した他方のLU2を主ボリュームとし、予備系サーバ#S1に接続するようPCIex-SW107に対して指令する。その後、管理サーバ101は予備系サーバ#S1に起動を指令してフェイルオーバを完了する。

20

【0093】

以上により、障害が発生した現用系サーバ#1のメモリダンプの収集をOSの種類にかかわらず確実にいながらも、予備系サーバ#S1への系切替を迅速に行うことが可能となるのである。特に、ミラーボリュームLU1, LU2をスプリットした後は、障害が発生した現用系サーバ#1のメモリダンプと予備系サーバ#S1への系切替を並列的に行うことで、メモリダンプの完了を待たずに系切替を開始できるので、フェイルオーバの高速化を図ることができる。

30

【0094】

図11は、管理サーバ101のI/Oバッファリング指示部211で行われる処理の一例を示すフローチャートである。この処理は、図10のステップ1002で行われる処理である。

【0095】

ステップ1101で、I/Oバッファリング指示部211は、サーバ管理テーブル221を参照し、ステップ1102へ進む。

【0096】

ステップ1102で、I/Oバッファリング指示部211は、障害通知501とサーバ管理テーブル221から障害が発生した現用系サーバ#1に接続されたアダプタ451とPCIex-SW107の接続ポートを特定し、ステップ1103へ進む。

40

【0097】

ステップ1103で、I/Oバッファリング指示部211は、I/O処理機構322に対して、ステップ1004で特定したPCIex-SW107の接続ポートとI/O処理機構322のバッファ領域443とを接続するよう指示し、ステップ1104へ進む。

【0098】

ステップ1104で、I/Oバッファリング指示部211は、I/O処理機構322に対して、当該現用系サーバ#1からのI/O出力をバッファするよう指示し、ステップ1105へ進む。

【0099】

50

ステップ 1 1 0 5 で、I / O バッファリング指示部 2 1 1 は、I / O バッファリング管理テーブル 4 1 1 を更新し、処理を完了する。

【 0 1 0 0 】

上記処理により、障害が発生した現用系サーバ # 1 からの I / O 出力は、P C I e x - S W 1 0 7 のバッファ領域 4 4 3 に格納される。

【 0 1 0 1 】

図 1 2 は、管理サーバ 1 0 1 の経路切替部 2 1 3 で行われる処理の一例を示すフローチャートである。この処理は、図 1 0 のステップ 1 0 0 5 で行われる処理である。

【 0 1 0 2 】

ステップ 1 2 0 1 で、経路切替部 2 1 3 は、L U 管理テーブル 2 2 3 を参照し、障害が発生した現用系サーバ # 1 に割り当てられた L U とペアの関係にある L U 1 を特定し、ステップ 1 2 0 2 へ進む。

10

【 0 1 0 3 】

ステップ 1 2 0 2 で、経路切替部 2 1 3 は、L U マッピング管理テーブル 2 2 2 を参照し、障害が発生した現用系サーバ # 1 に割り当てられた L U とポートの関係を特定してステップ 1 2 0 3 へ進む。

【 0 1 0 4 】

ステップ 1 2 0 3 で、経路切替部 2 1 3 は、I / O 処理機構 3 2 2 のバッファ領域 4 4 3 と、メモリダンプ出力用 L U 1 (スプリットした元々副ボリュームであった L U) とを接続するよう指示し、処理を完了する。

20

【 0 1 0 5 】

以上の処理により、バッファ領域 4 4 3 に副ボリュームの L U 1 が接続され、バッファ領域 4 4 3 の内容を L U 1 に書き込むことができる。

【 0 1 0 6 】

図 1 3 は、管理サーバ 1 0 1 の I / O バッファ書出し指示部 2 1 4 で行われる処理の一例を示すフローチャートである。この処理は、図 1 0 のステップ 1 0 0 6 で行われる処理である。

【 0 1 0 7 】

ステップ 1 3 0 1 で、I / O バッファ書出し指示部 2 1 4 は、I / O 処理機構 3 2 2 に対してバッファ領域 4 4 3 へ蓄積した I / O データを書き出すよう指示し、ステップ 1 3 0 2 へ進む。

30

【 0 1 0 8 】

ステップ 1 3 0 2 で、I / O バッファ書出し指示部 2 1 4 は、書き出しを指令したバッファ領域 4 4 3 について I / O バッファリング管理テーブル 4 1 1 を更新し、処理を完了する。

【 0 1 0 9 】

上記処理により、P C I e x - S W 1 0 7 のバッファ領域 4 4 3 に格納されたメモリダンプが、スプリットによりペアが解除された L U 1 に書き込まれる。

【 0 1 1 0 】

図 1 4 は、管理サーバ 1 0 1 の N + M 切替指示部 2 1 5 で行われる処理の一例を示すフローチャートである。この処理は、図 1 0 のステップ 1 0 0 7 で行われる処理である。

40

【 0 1 1 1 】

ステップ 1 4 0 1 で、N + M 切替指示部 2 1 5 は、サーバ管理テーブル 2 2 1 を参照し、障害が発生した現用系サーバ # 1 と、引き継ぎ先の予備系サーバ # S 1 を特定してステップ 1 4 0 2 へ進む。

【 0 1 1 2 】

ステップ 1 4 0 2 で、N + M 切替指示部 2 1 5 は、ステップ 1 4 0 1 で特定した予備系サーバ # S 1 と、障害が発生した現用系サーバ # 1 が使用していたアダプタ 4 5 1 を接続するよう、P C I e x - S W 1 0 7 に指示し、ステップ 1 4 0 3 へ進む。

【 0 1 1 3 】

50

ステップ 1 4 0 3 で、N + M 切替指示部 2 1 5 は、予備系サーバ # S 1 に接続した L U 2 について、L U 管理テーブル 2 2 3 を更新し、ステップ 1 4 0 4 へ進む。

【 0 1 1 4 】

ステップ 1 4 0 4 で、N + M 切替指示部 2 1 5 は、予備系サーバ # S 1 に接続した L U 2 について、L U マッピング管理テーブル 2 2 2 を更新し、ステップ 1 4 0 5 へ進む。

【 0 1 1 5 】

ステップ 1 4 0 5 で、N + M 切替指示部 2 1 5 は、障害が発生した現用系サーバ # 1 と、引き継ぎ先の予備系サーバ # S 1 についてサーバ管理テーブル 2 2 1 を更新し、処理を完了する。

【 0 1 1 6 】

上記処理により、障害が発生した現用系サーバ # 1 の L U 2 が、予備系サーバ # S 1 に引き継がれる。

【 0 1 1 7 】

図 1 5 は、I / O 処理機構 3 2 2 の I / O バッファリング制御部 4 0 1 で行われる処理の一例を示すフローチャートである。この処理は、図 1 1 のステップ 1 1 0 4 で行われる処理である。

【 0 1 1 8 】

ステップ 1 5 0 1 で、I / O バッファリング制御部 4 0 1 は、I / O バッファリング管理テーブル 4 1 1 を参照し、メモリダンプの書き込み先となるバッファ領域 4 4 3 を特定してステップ 1 5 0 2 へ進む。

【 0 1 1 9 】

ステップ 1 5 0 2 で、障害が発生した現用系サーバ # 1 と I / O 処理機構 3 2 2 およびバッファ領域 4 4 3 を接続されるのを待って、ステップ 1 5 0 3 へ進む。

【 0 1 2 0 】

ステップ 1 5 0 3 で、I / O バッファリング制御部 4 0 1 は、当該バッファ領域 4 4 3 へ当該現用系サーバ # 1 からの I / O データをバッファリングし、処理を完了する。

【 0 1 2 1 】

図 1 6 は、管理サーバ 1 0 1 が管理する業務及び S L A 管理テーブル 2 2 4 の一例を示す説明図である。業務及び S L A 管理テーブル 2 2 4 は、現用系サーバ 1 0 2 が提供する業務毎にどのような業務およびソフトウェアで、どのような設定がされていて、どのような Service Level を、どの程度満たす必要があるか、それぞれの優先順位付け、といった情報を管理している。

【 0 1 2 2 】

カラム 1 6 0 1 には、業務識別子を格納しており、本識別子によって業務を一意に識別する。

【 0 1 2 3 】

カラム 1 6 0 2 には、U U I D が格納されている。カラム 1 6 0 1 に格納されている業務識別子の候補であり、広範囲に渡ったサーバ管理には非常に有効である。ただし、カラム 1 6 0 1 には、システム管理者がサーバを識別する識別子を使用すれば良く、また管理する対象となるサーバ間で重複することがなければ問題ないため、U U I D を使うことが望ましいものの必須とはならない。例えば、カラム 1 6 0 1 のサーバ識別子には、業務設定情報（カラム 1 6 0 4 へ格納）を用いても良い。

【 0 1 2 4 】

カラム 1 6 0 3 は、業務種別を格納しており、使用するアプリケーションやミドルウェアといった業務を特定するソフトウェアに関する情報が格納されている。業務で使用する論理的な I P アドレスや I D、パスワード、ディスクイメージ、業務で使用するポート番号などが格納されている。ディスクイメージは、設定前後の業務が現用系のサーバ 1 0 2 上の O S へ配信されたシステムディスクのディスクイメージを指す。カラム 1 6 0 4 へ格納するディスクイメージに関する情報は、データディスクを含めても良い。

【 0 1 2 5 】

10

20

30

40

50

コラム 1605 は、優先順位や S L A の内容を格納しており、それぞれの業務間の優先順位やそれぞれの業務が必要とする要件が格納されている。これにより、どの業務が優先的に救済される必要があり、メモリダンプ採取が必要か否か、また N + M 切替高速が必要か否か、を設定することが出来る。本発明では、バッファ領域 443 をどのように使うかが重要なポイントであり、これにより最も本発明の効果を得ることが出来る運用を決めることが可能になる。

【0126】

管理サーバ 101 は、業務及び S L A 管理テーブル 224 で、S L A 1605 がメモリダンプ不要であれば、上記図 5 に示した処理を行わずに、フェイルオーバを実施すればよい。

10

【0127】

図 20 は、メモリダンプの書き込みが完了した L U 1 を、予め設定した保守用の領域へ退避させる例を示すブロック図である。管理サーバ 101 は、メモリダンプの書き込みが完了した L U 1 を、予備系サーバ # S 1 が使用するホストグループ 1 (550) から分離して、予め設定した保守用グループ 551 に変更し、アクセスを制限する。

【0128】

以上により、障害が発生した現用系サーバ # 1 のメモリダンプを、OS の種類にかかわらず、確実に L U 1 に収集し、保守用グループ 551 に移動させることで、メモリダンプの内容を誤って消去するなどの誤操作を防止することができる。

20

【0129】

< 第 2 実施形態 >

図 17 は、第 2 の実施形態を示すサーバ 102 (または 106) のブロック図である。第 2 実施形態は、前記第 1 実施形態の I / O 処理機構 322 を、仮想化機構 1711 に組み込んだものである。図 17 では、サーバ 102、仮想化機構 1711 および仮想サーバ 1712 の構成を示す。サーバ 102 の物理的な計算機資源を仮想化機構 1711 が仮想化し、複数の仮想サーバ 1712 を提供している。なお、仮想化機構 1711 としては、V M M (Virtual Machine Monitor) やハイパーバイザで構成することができる。

【0130】

メモリ 302 には、物理的な計算機資源を仮想化するサーバ仮想化技術を提供する仮想化機構 1711 が配備され、仮想サーバ 1712 を提供する。また、仮想化機構 1711 は、制御用インタフェースとして仮想化機構管理用インタフェース 1721 を備えている。

30

【0131】

仮想化機構 1711 は、サーバ 102 (ブレードサーバでも良い) の物理的な計算機資源を仮想化し、仮想サーバ 1712 を構成する。仮想サーバ 1712 は、仮想 CPU 1731、仮想メモリ 1732、仮想ネットワークインタフェース 1733、仮想ディスクインタフェース 1734、仮想 P C I e x インタフェース 1735 から構成されている。仮想メモリ 1732 には、OS 1741 が配備され仮想サーバ 1712 内の仮想デバイス群を管理している。また、OS 1741 上では、業務アプリケーション 1742 が実行されている。OS 1741 上で稼働する管理プログラム 1743 によって、障害検知や OS 電源制御、インベントリ管理などが提供されている。仮想化機構 1711 は、物理計算機資源と仮想計算機資源の対応付けを管理しており、物理計算機資源と仮想計算機資源の対応付けの生成や解除を行うことが出来る。また、どの仮想サーバ 1712 がサーバ 102 の計算機資源を、どれくらい割り当てられ、また、使用しているかといった構成情報および稼働履歴を保持している。なお、OS 1741 は、前記第 1 実施形態と同様に、所定の条件で仮想メモリ 1732 の内容を出力するメモリダンプ部 17410 を有する。

40

【0132】

仮想化機構管理用インタフェース 1721 は、管理サーバ 101 と通信をするためのインタフェースであり、仮想化機構 1711 から管理サーバ 101 へ情報を通知したり、管理サーバ 101 から仮想化機構 1711 へ指示を送るときに使われる。また、ユーザが直

50

接、使用することも可能である。

【0133】

仮想化機構1711には、I/O処理機構322が内包され、例えば、仮想PCIE xインタフェース1735と物理PCIE xインタフェース306の接続に関わる。仮想サーバ1712の障害発生時に、仮想メモリ1732のダンプを取得しつつ、他の仮想サーバ(同じ物理サーバ上または別の物理サーバ上)で業務を再開させるフェイルオーバを実施する。

【0134】

本第2実施形態では、サーバ102とストレージサブシステム105の接続について、前記第1実施形態に示したPCIE x - SW107を使用してもよいが、PCIE x - SW107の内部で経路を切り替えることなく、仮想化機構1711で複数の仮想サーバ1712とLUの接続関係を切り替えることができる。

10

【0135】

このため、本第2実施形態では、サーバ102は、仮想サーバ1712が使用するストレージサブシステム105のLUの経路数に応じて複数のディスクインターフェース304-1、304-2を備えるものとする。以下の説明では、サーバ102のディスクインターフェース304-1、304-2がFC-SW511(図1参照)を介してストレージサブシステム105のLU2(及びLU1)に接続された例を示す。

【0136】

図18は、第2の実施形態の処理の概要を示すブロック図である。図18において、仮想サーバ#VS1(1712-1)が現用系サーバとして稼動し、仮想サーバ#VS1に障害が発生したときに、仮想サーバ#VS1のメモリダンプを収集しながら、予備系として機能する仮想サーバ#VS2(1712-2)へ処理を引き継ぐ例を示す。

20

【0137】

現用系の仮想サーバ#VS1は、前記第1実施形態の図5と同様に、LU1を主ボリュームとし、LU2副ボリュームとするミラーボリュームに対してアクセスする。

【0138】

仮想化機構1711は、仮想サーバ#VS1の仮想メモリの監視と、ストレージサブシステム105のメモリダンプ用仮想領域542への仮想サーバ#VS1からの書き込みの監視と、仮想サーバ#VS1等のOS1741のシステム領域(メモリダンプ用プログラム)の読み込みの監視と、OS1741のメモリダンプ用プログラムを呼び出すシステムコールの監視と、仮想サーバ#VS1の障害発生の監視を行う。この他、仮想化機構1711は、予備系の仮想サーバ#VS2への計算機資源の割り当てなどを管理する。なお、管理サーバ101は、仮想化機構1711の仮想化機構管理用インターフェース1721を介して指令を行う。

30

【0139】

仮想サーバ#VS1に障害が発生すると、仮想化機構1711は管理サーバ101に対して障害通知を送信する(S1)。管理サーバ101は、仮想化機構1711に対して仮想サーバ#VS1のI/O出力をバッファ領域443に格納する指令を送信する(S2)。

40

【0140】

仮想化機構1711は、現用系の仮想サーバ#VS1の仮想ディスクインターフェース1734の接続先を、I/O処理機構322のバッファ領域443に切り替える(S3)。これにより、障害が発生した仮想サーバ#VS1は、仮想メモリ1732の内容をI/O処理機構322のバッファ領域443に格納する。

【0141】

次に、管理サーバ101は、ストレージサブシステム105に対して、仮想サーバ#VS1に接続されているLU1、LU2をスプリットする指令を送信する(S3)。

【0142】

次に、管理サーバ101は、仮想化機構1711に対して、バッファ領域443の内容

50

を副ボリュームであったLU1に書き込むよう経路を切り替える指令を送信する(S4)。仮想化機構1711は、バッファ領域443の接続先をLU1に接続されたディスクインターフェース304-2に切り替える。これにより、仮想化機構1711はバッファ領域443の内容をLU1に書き込む。

【0143】

管理サーバ101は、仮想化機構1711に対して予備系の仮想サーバ#VS2を割り当て、LU2を仮想サーバ#VS2に切り替える指令を送信する(S6)。仮想化機構1711は、管理サーバ101からの指令に基づいて仮想サーバ#VS2に計算機資源を割り当て、仮想ディスクインターフェース1734の接続先をLU1に設定されたディスクインターフェース304-1に設定する。

10

【0144】

管理サーバ101は、仮想化機構1711に対して予備系の仮想サーバ#VS2を起動する指令を送信する(S7)。仮想化機構1711は、計算機資源とディスクインターフェース304-1を割り当てた仮想サーバ#VS2を起動して、LU2のOS1741及び業務アプリケーション1742を実行することで、現用系の仮想サーバ#VS1の処理を引き継ぐことができる。

【0145】

以上のように、現用系の仮想サーバ#VS1に障害が発生した場合にも、OSの種類にかかわらず、メモリダンプとフェイルオーバを並列的に行って、系切替を高速化することができる。

20

【0146】

<第3実施形態>

図19は、第3の実施形態を示し、PCIex-SW107を主体とするフェイルオーバの概略を示すブロック図である。第3の実施形態では、ストレージサブシステム105に、メモリダンプ用仮想領域542への書き込みを監視する管理及び監視インターフェース600を配備して、現用系サーバ#1(102)がメモリダンプを開始したことを契機にして、フェイルオーバとメモリダンプのバッファリングを実行するものである。その他の構成は、前記第1実施形態と同様である。

【0147】

管理及び監視インターフェース600は、現用系サーバ#1がアクセスする主ボリュームとしてのLU1について、メモリダンプ用仮想領域542への書き込みを監視する。メモリダンプ用仮想領域542への書き込みが開始されると、管理及び監視インターフェース600は、管理サーバ101に現用系サーバ#1のメモリダンプが発生したことを通知する。

30

【0148】

管理サーバ101は、メモリダンプの発生を検知すると、前記第1実施形態と同様にして、現用系サーバ#1から予備系サーバ#S1へのフェイルオーバと、現用系サーバ#1のメモリダンプを並列的に実行する。

【0149】

ここで、管理及び監視インターフェース600は、メモリダンプ用仮想領域542への書き込みを監視し、また、OS311のシステム領域(メモリダンプ用プログラム)の著見込みを監視する。

40

【0150】

メモリダンプ用仮想領域542への書き込みの検知は、管理及び監視インターフェース600が、ストレージサブシステム105内の特定の領域(ブロック)からメモリダンプ用の書き込みの有無を検知する。メモリダンプ用仮想領域542の位置を特定するために、前以てメモリダンプ用の特定ファイルにサンプルデータを書き込む、または、疑似障害を用いてプログラムを起動しメモリダンプ用のデータを書き込ませる、などして領域を特定しても良い。

【0151】

50

なお、管理及び監視インターフェースは、ストレージサブシステム 105 の他に、図示の 601、602 のように FC - SW 511 またはアダプタラック 461 に設けることができる。この場合、管理及び監視インターフェース 601、602 は I/O 出力をスワッピングするなど監視し、宛先と内容からメモリダンプの開始を検知する。

【0152】

以上のように、第 1 ~ 第 3 の実施形態によれば、現用系サーバ # 1 のメモリダンプを一時的に蓄積するバッファ領域 443 を備えた I/O 処理機構 322 と、メモリダンプの経路をミラーボリュームの主ボリューム (LU) から副ボリューム (LU2) へ切り替える経路切替部として PCIex - SW 107 または仮想化機構 1711 に備える。そして、管理サーバ 101 がミラーボリュームをスプリットした後に、予備系サーバ # S1 を主ボリューム (LU1) で起動させることで、系切替とメモリダンプを並列的に実行する。これにより、メモリダンプの完了を待たずに系切替を開始できるので、OS の種類にかかわらずメモリダンプを確実に収集しながらも系切替の高速化を図ることができる。

10

【0153】

なお、上記各実施形態では、ストレージサブシステム 105 の LU でミラーボリュームを構成した例を示したが、物理的なディスク装置でミラーボリュームを構成してもよい。

【0154】

また、上記各実施形態では、FC - SW 511 と NW - SW 103、104 で SAN と IP ネットワークを分離する例を示したが、IP - SAN 等を用いてひとつのネットワークとしてもよい。

20

【産業上の利用可能性】

【0155】

以上のように、本発明はコールドスタンバイを用いて系切替を行う計算機システムや I/O スイッチあるいは仮想化機構に適用することができる。

【符号の説明】

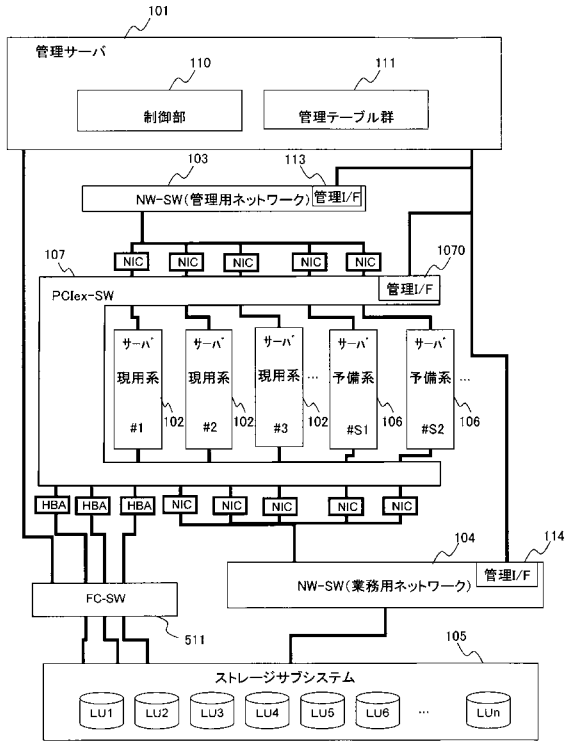
【0156】

- 101 管理サーバ
- 102 サーバ
- 105 ストレージサブシステム
- 107 PCIex - SW
- 110 制御部
- 210 障害検出部
- 211 I/O バッファリング指示部
- 212 ストレージ制御部
- 213 経路切替部
- 214 I/O バッファ書き出し指示部
- 215 N + M 切替指示部
- 221 サーバ管理テーブル
- 222 LU マッピング管理テーブル
- 223 LU 管理テーブル
- 224 業務及び SLA 管理テーブル
- 322 I/O 処理機構
- 401 I/O バッファリング制御部
- 411 I/O バッファリング管理テーブル
- 441 制御部

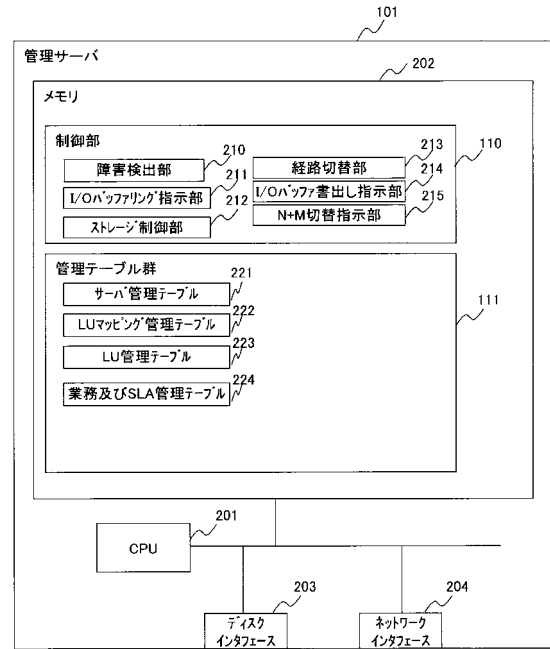
30

40

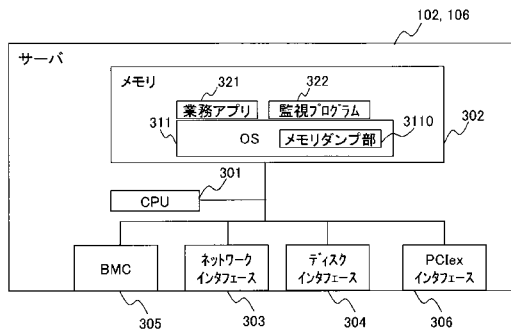
【 図 1 】



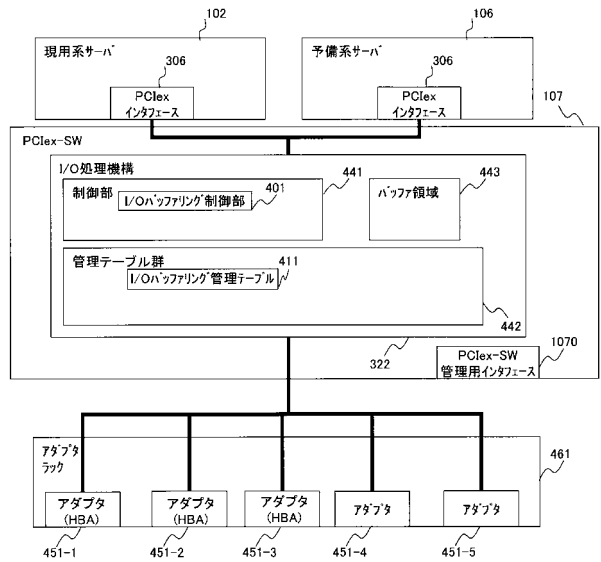
【 図 2 】



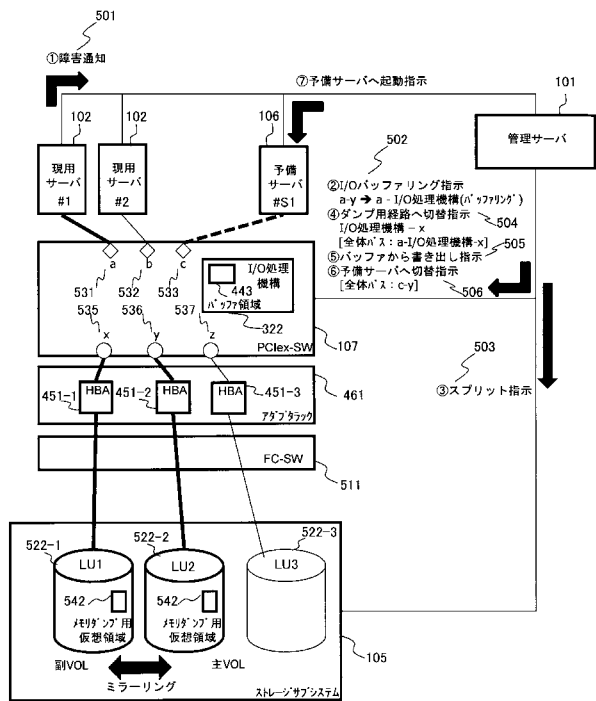
【 図 3 】



【 図 4 】



【図5】



【図6】

サーバ識別子	サーバ種別	サーバ状態	アダプタ種別	アダプタWWN	スイッチ情報	サーバ種別	構成	性能情報
サーバ1	現用	正常	HBA1	WWN1	FCSW1, port1, SAN Security	A社	IA32	CPU:3GHz/core:4 memory:12GB HDD(SATA):120GB NIC:1Gbps, HBA:3Gbps
サーバ2	現用	正常	HBA2	WWN2	FCSW2, port1, SAN Security	A社	IA32	CPU:3GHz/core:4 memory:4GB HDD(SAS):120GB NIC:1Gbps, HBA:3Gbps
サーバ3	現用	正常	HBA5	WWN5	FCSW1, port3, SAN Security	A社	Chassis1 SMP:No	CPU:2GHz/core:2 memory:4GB NIC:1Gbps, HBA:3Gbps
サーバ4	現用	正常	HBA6	WWN6	FCSW2, port3, SAN Security	A社	Chassis2 SMP:No	CPU:2GHz/core:2 memory:4GB NIC:1Gbps, HBA:3Gbps
サーバ5	現用	正常	HBA7	WWN7	FCSW4, port4, SAN Security	B社	IA32	CPU:1.6GHz/core:4 HDD(SATA):80GB memory:4GB NIC:1Gbps, HBA:3Gbps
サーバ6	予備	正常	HBA8	WWN8	FCSW1, port3, SAN Security	B社	Chassis4 SMP:No	CPU:1.6GHz/core:4 HDD(SATA):80GB memory:4GB NIC:1Gbps, HBA:3Gbps
サーバ7	予備	正常	HBA9	WWN9	FCSW1, port3, SAN Security	B社	Chassis4 SMP:No	CPU:1.6GHz/core:4 HDD(SATA):80GB memory:4GB NIC:1Gbps, HBA:3Gbps

【図7】

LU識別子	アダプタ種別	アダプタWWN	PClex-SW情報	サーバ識別子
LU1	HBA1	WWN1	port1, I/O処理機構, portA	サーバ1
LU2	HBA2	WWN2	port2, portB	サーバ2
LU3	HBA3	WWN3	port3, portC	サーバ3
LU4	HBA4	WWN4	port4, portD	サーバ4
LU5	HBA5	WWN5	port5, portE	サーバ5
LU6	HBA6	WWN6	port6, portF	サーバ6
LU7	HBA7	WWN7	port7, portG	サーバ7
---	HBA8	WWN8	port8, portH	サーバ8
---	HBA9	WWN9	port9, portI	サーバ9

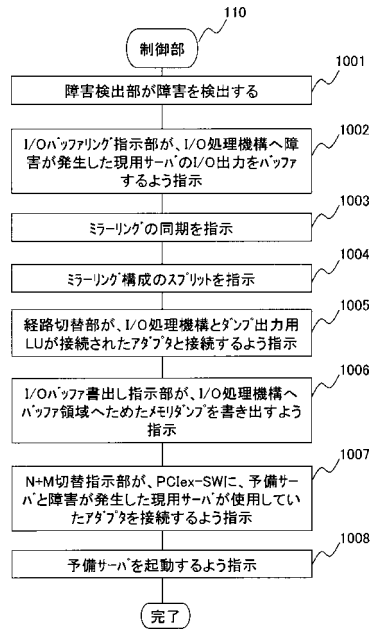
【図8】

LU識別子	LU種別	ミラーリング先LU	ステータス
LU1	主	---	アクティブ(LU11)
LU2	主	LU12	ミラーリング
LU3	主	LU13	ミラーリング
LU4	主	LU14	ミラーリング
LU5	主	LU15	ミラーリング
LU6	主	LU16	ミラーリング
LU7	主	LU17	ミラーリング
LU8	主	---	---
LU9	主	---	---
LU10	主	---	---
LU11	副	---	副→主
LU12	副	---	---
LU13	副	---	---
LU14	副	---	---
LU15	副	---	---
LU16	副	---	---
LU17	副	---	---
LU18	副	---	ミラーリング予約(LU1)

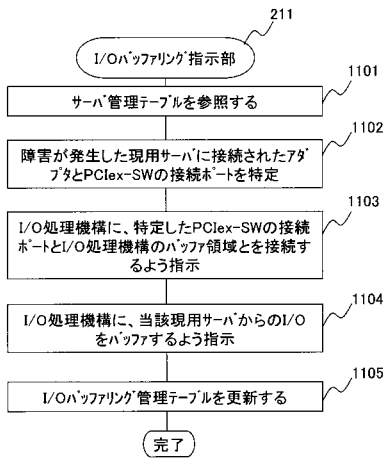
【 図 9 】

I/Oハッパ識別子	サーバ識別子	アダプタ デバイス	アダプタ WWN	サーバ要求受付・ハッパ中	ハッパ使用時-95
951	サーバ1	HBA1	WWN1	---	---
		HBA2	WWN2	---	---
952	サーバ2	HBA3	WWN3	---	---
		HBA4	WWN4	---	---
953	サーバ3	HBA5	WWN5	---	---
954	サーバ4	HBA6	WWN6	---	---
955	サーバ5	HBA7	WWN7	---	---
956	サーバ6	HBA8	WWN8	---	---
957	サーバ7	HBA9	WWN9	---	---
	

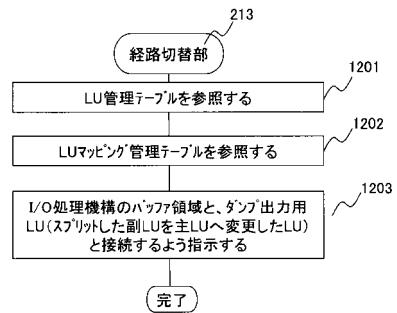
【 図 1 0 】



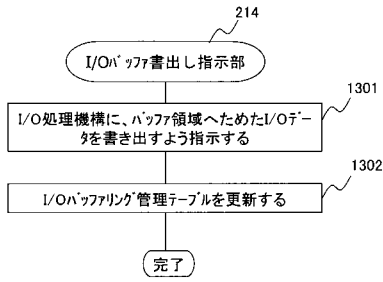
【 図 1 1 】



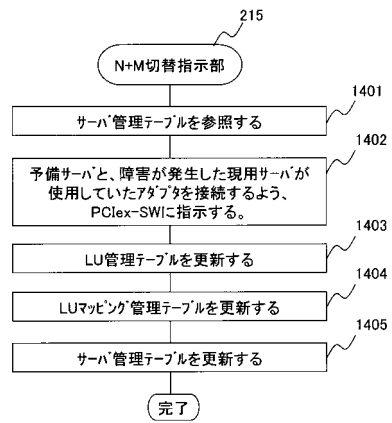
【 図 1 2 】



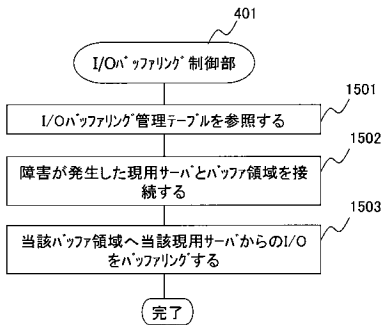
【 図 1 3 】



【 図 1 4 】



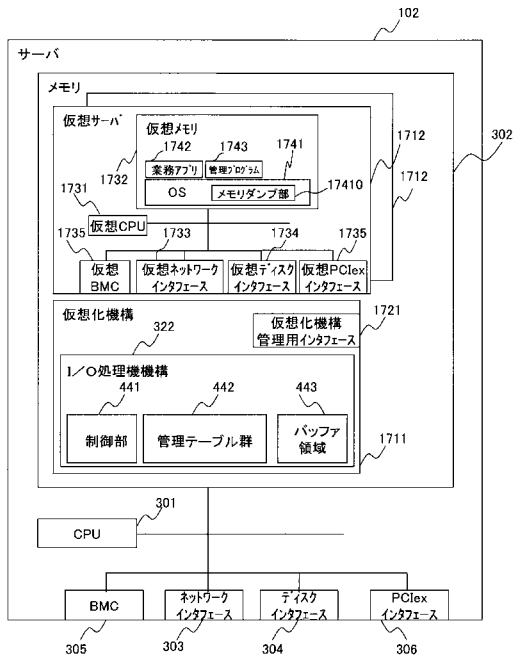
【 図 1 5 】



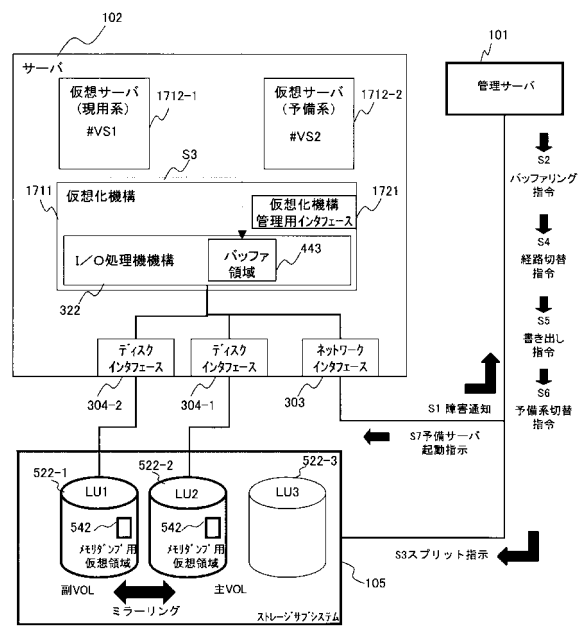
【 図 1 6 】

業務識別子	UUID	業務種別	業務設定情報	優先順位 SLA
業務1	401	777717 A	ID 431, Password 441 DiskImage 451 PortNumber 461	優先順位: 1 SLA: バックアップ 採取必須 N+M切替高速
業務2	402	777717 B	ID 432, Password 442 DiskImage 452 PortNumber 462	優先順位: 4 SLA: N+M高速切替 バッキングとれればとる
業務3	403	777717 B	ID 433, Password 443 DiskImage 453 PortNumber 463	優先順位: 2 SLA: バッキング 必須
業務4	404	777717 B	ID 434, Password 444 DiskImage 454 PortNumber 464	優先順位: 3 SLA: バッキング 必須
業務5	405	777717 C	ID 435, Password 445 DiskImage 455 PortNumber 465	優先順位: 5 SLA: N+M高速切替 バッキングとらない
...

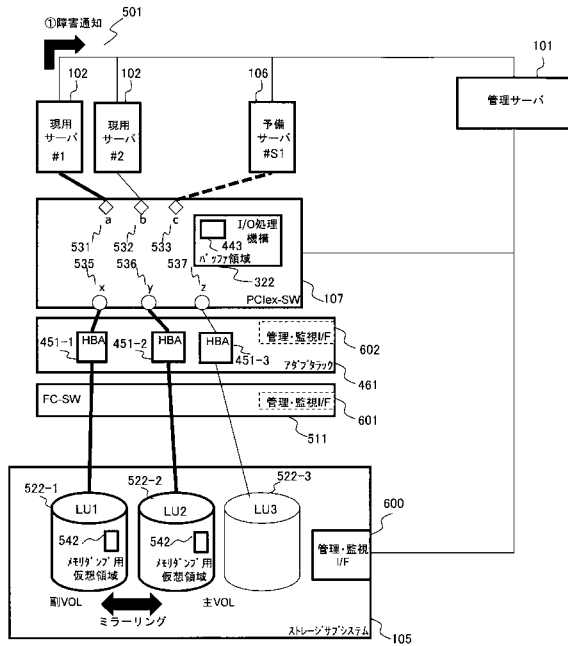
【図 17】



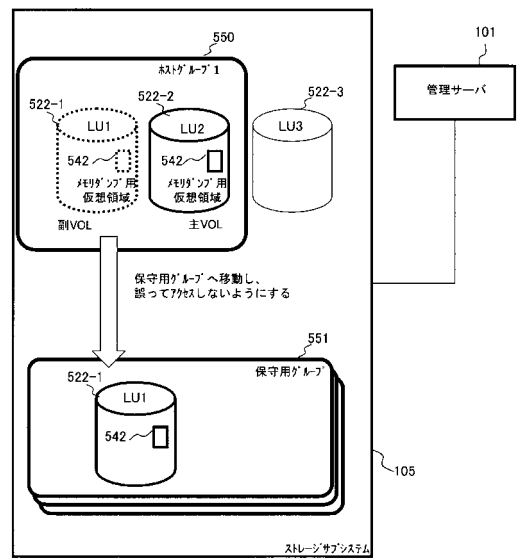
【図 18】



【図 19】



【図 20】



フロントページの続き

(72)発明者 高本 良史

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア事業部内

(72)発明者 寺村 健

神奈川県川崎市麻生区王禅寺1099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