

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4490745号
(P4490745)

(45) 発行日 平成22年6月30日(2010.6.30)

(24) 登録日 平成22年4月9日(2010.4.9)

(51) Int.Cl. F I
G 0 6 F 11/20 (2006.01) G 0 6 F 11/20 3 1 0 C

請求項の数 6 (全 29 頁)

(21) 出願番号	特願2004-190818 (P2004-190818)	(73) 特許権者	000005108
(22) 出願日	平成16年6月29日(2004.6.29)		株式会社日立製作所
(65) 公開番号	特開2006-12004 (P2006-12004A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成18年1月12日(2006.1.12)	(74) 代理人	110000350
審査請求日	平成18年11月29日(2006.11.29)		ポレール特許業務法人
		(74) 代理人	100068504
			弁理士 小川 勝男
		(74) 代理人	100086656
			弁理士 田中 恭助
		(74) 代理人	100094352
			弁理士 佐々木 孝
		(72) 発明者	市川 正也
			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】 ホットスタンバイシステム

(57) 【特許請求の範囲】

【請求項1】

現用系計算機と待機系計算機でストレージを共有するホットスタンバイシステムにおける系の切替え制御方法において、

該現用系計算機から該ストレージのある論理ボリュームに対する書き込みI/O要求が発生した場合に、ストレージを構成する各論理ボリューム毎に前記論理ボリュームに対する操作の種類毎に起動するプログラムを登録したデバイススイッチテーブルを参照し、前記論理ボリュームにおける前記書き込みI/O要求に対応する書き込み処理プログラムを起動するステップと、

前記書き込み処理プログラムが処理中の該I/O要求に関する情報をメモリにI/O記録として記録し、論理ボリュームに対する書き込みI/Oを処理するステップと、

該現用系計算機で障害が発生した場合に、前記デバイススイッチテーブルを参照し、前記論理ボリュームにおけるダンプ出力処理に登録されているダンプ出力プログラムを起動するステップと、

前記ダンプ出力プログラムが、該メモリに保持された該I/O記録を該待機系計算機へ送信し、該現用系計算機内のメモリダンプを前記論理ボリュームに出力するステップと、
該待機系計算機で、該現用系計算機から送信された該I/O記録を受信するステップと

、
該待機系計算機において、該受信されたI/O記録を検証するステップと、

該検証の結果、処理中の該I/O要求が存在した場合、該ストレージに対する復旧処理を

10

20

実施するステップと、
を含むことを特徴とする系の切替え制御方法。

【請求項 2】

前記書き込み I / O 要求をメモリに前期 I / O 記録として記録する対象として前記論理ボリュームを指定するステップと、

前記対象として指定された論理ボリュームに対応する前記デバイススイッチテーブルを、前記書き込み I / O 要求をメモリに前記 I / O 記録として記録することの可能な前記書き込みプログラムを登録するステップと、

前記書き込み I / O 要求をメモリに前記 I / O 記録として記録しない除外対象として前記論理ボリュームを指定するステップと、

前記除外対象として指定された論理ボリュームに対応する前記デバイススイッチテーブルを、前記前記書き込み I / O 要求をメモリに前記 I / O 記録として記録しない前記書き込みプログラムを登録するステップと、

を含むことを特徴とする請求項 1 記載の系の切替え制御方法。

【請求項 3】

前記 I / O 記録として、アクセス対象となる該論理ボリュームの識別情報と、該論理ボリュームに対する I / O 要求のうち処理中の I / O 要求をカウントする情報を、該メモリに格納することを特徴とする請求項 1 又は 2 の系の切替え制御方法。

【請求項 4】

前記 I / O 記録は、該 I / O 要求毎にエントリを有するテーブル形式で該メモリに格納され、該現用系計算機から該待機系計算機へはネットワークを介して該テーブルの情報が送信され、

該待機系計算機は、受信された該テーブルから取得可能なエントリを取得して、該エントリの該カウント情報が所定の値かを判定し、所定の値である場合には、復旧プログラムを実行せず、所定の値で無い場合には、該エントリに対応する前記論理ボリュームには中断された書き込み I / O が存在し、該論理ボリュームの内容が不整合であると見なし、該論理ボリュームに対して復旧プログラムを実行することを特徴とする請求項 3 の系の切替え制御方法。

【請求項 5】

現用系計算機と待機系計算機でストレージを共有するホットスタンバイシステムにおける系の切替え制御方法において、

該現用系計算機から該ストレージのある論理ボリュームに対する書き込み I / O 要求が発生した場合に、ストレージを構成する各論理ボリューム毎に前記論理ボリュームに対する操作の種類毎に起動するプログラムを登録したデバイススイッチテーブルを参照し、

前記書き込み処理プログラムが該 I / O 要求の前処理として、該待機系計算機に I / O 開始の通知を送信し、論理ボリュームに対する書き込み I / O を処理するステップと、

該待機系計算機で、該 I / O 開始の通知を受信し、I / O 処理中を示す情報をメモリに記録するステップと、

該現用系計算機で、前記書き込み処理プログラムは、該 I / O 要求が完了した時、I / O 後処理として該待機系計算機に I / O 完了の通知を送信するステップと、

該現用系計算機で障害が発生した場合に、該メモリに記録された該 I / O 処理中を示す情報を検証するステップと、

該検証の結果、処理中の書き込み I / O 要求が存在する場合、該ストレージに対する復旧処理を実施するステップと、

を含むことを特徴とする系の切替え制御方法。

【請求項 6】

現用系計算機と待機系計算機でストレージを共有するホットスタンバイシステムにおける系の切替え制御方法において、

該現用系計算機から該ストレージのある論理ボリュームに対する書き込み I / O 要求が発生した時、ストレージを構成する各論理ボリューム毎に前記論理ボリュームに対する操作

10

20

30

40

50

の種類毎に起動するプログラムを登録したデバイススイッチテーブルを参照し、前記論理ボリュームにおける前記書込みI/O要求に対応する書込み処理プログラムを起動するステップと、

前記書込み処理プログラムが該I/O要求の書き込み処理中であることを示す情報を該ストレージに記録するステップと、

該現用系計算機で障害が発生した場合に、該待機系計算機で該ストレージに記録された情報を参照して検証するステップと、

該情報の検証の結果、書き込み処理中である場合、該ストレージに対する復旧処理を実施するステップと、

を含むことを特徴とする系の切替え制御方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ホットスタンバイシステムに係り、特に現用系及び待機系からアクセス可能な共有ディスクを有するホットスタンバイシステムにおける共有ディスクの整合性の検査、及び系の切替え制御に関するものである。

【背景技術】

【0002】

近年、オンラインランザクション処理の分野では、可用性向上のために、現用系及び待機系からアクセス可能な共有ディスクを有するホットスタンバイ構成のシステムが適用

20

されている。この種のホットスタンバイシステムに関して、例えば、特許文献1（特開平10-289122号公報）には、現用系と待機系にそれぞれ1台のディスクを接続し、通常運用時は現用系でデータ更新が発生した場合、ネットワークを通じて待機系に更新データを送信し、待機系側のディスクに対しても同様のデータ更新を行う旨が開示されている。そして、現用系で障害が発生した際は、待機系が新しく現用系として業務を引き継ぎ、従前の現用系が以後待機系として機能し、障害から復旧した後、引継ぎ後の更新データを新たな現用系から新たな待機系へ送り、共有ディスクの同期を行っている。この従来技術によれば、待機系に切替えた時点で、両ディスク内のデータの同期を改めて取らずに業務の引継ぎ処理を行えるため、業務再開までの時間を短縮できる。

30

【0003】

【特許文献1】特開平10-289122号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

ソフトウェアによるミラーリング制御は、ミラー間の同期化処理もソフトウェアで行うため、障害発生により同期化処理が完了せずミラー間に不整合が発生することがある。しかし、上記従来技術は、待機系に処理を引継いだ時点で整合性のチェックを行わないで高速化を図っている。このため、ミラー間に不整合が発生している可能性があるにも拘わらず、それを検出する手段がないため、不整合な状態にデータを上書きしてしまう恐れがある。

40

【0005】

本発明の目的は、ホットスタンバイシステムにおいて現用系計算機で障害が発生した場合、待機系計算機側で共有ディスクの整合性を保ちつつ、処理を高速に引継ぐことにある。

【課題を解決するための手段】

【0006】

本発明は、現用系計算機と待機系計算機で、共有ディスクのようなストレージを共有するホットスタンバイシステムにおいて、現用系計算機は、ストレージに対する書き込みI/O要求が発生した場合に、処理中のI/O要求に関する情報を記憶部、例えばメモリ又

50

は上記ストレージの一部に、I/O記録として記録する処理と、現用系計算機で発生する障害を監視し、障害が発生した場合に、待機系計算機に通知する処理とを行い、待機系計算機は、現用系計算機から障害発生の通知を受けた場合、記憶部から取得されたI/O記録を検証する処理と、検証の結果処理中の該I/O要求が存在する場合、ストレージに対する復旧のための処理とを行うホットスタンバイシステムである。

【0007】

好ましい第一の例による系の切替え制御方法によれば、現用系計算機では、ストレージに対する書き込みI/O要求が発生した場合に、処理中のI/O要求に関する情報をメモリにI/O記録として記録し、I/O要求の処理が完了した場合に、メモリから対応するI/O記録を削除すると共に、現用系計算機で障害が発生した場合に、メモリに保持されたI/O記録を待機系計算機へ送信し、待機系計算機では、現用系計算機から送信されたI/O記録を受信し、受信されたI/O記録を検証するステップと、検証の結果、処理中のI/O要求が存在した場合、ストレージに対する復旧処理を実施するステップとを有する。

10

【0008】

好ましい第二の例による系の切替え制御方法によれば、現用系計算機は、ストレージに対する書き込みI/O要求が発生した場合に、I/O要求の前処理として、待機系計算機にI/O開始の通知を送信し、待機系計算機は、このI/O開始の通知を受信して、I/O処理中を示す情報をメモリに記録し、現用系計算機でI/O要求が完了した時、I/O後処理として待機系計算機へI/O完了の通知を送信し、かつ、現用系計算機で障害が発生した場合に、メモリに記録されたI/O処理中を示す情報を検証し、検証の結果、処理中の書き込みI/O要求が存在する場合、ストレージに対する復旧処理を実施する処理を行うものである。

20

【0009】

好ましい第三の例による系の切替え制御方法によれば、現用系計算機でストレージに対する書き込みI/O要求が発生した時、I/O要求の書き込み処理中であることを示す情報をストレージの一部に記録し、現用系計算機で障害が発生した時に、待機系計算機でストレージの一部に記録された情報を参照して検証し、この検証の結果、I/O要求の書き込み処理中である場合、ストレージに対する復旧処理を実施する処理を行うものである。

【0010】

好ましい例では、ストレージは、夫々固有の識別情報が付与された、論理的に定義された1又は複数の論理ボリュームを有しており、前記I/O記録として、アクセス対象となる論理ボリュームの識別情報と、論理ボリュームに対するI/O要求のうち処理中のI/O要求をカウントする情報をメモリ又はストレージの一部に格納する。

30

一例では、前記I/O記録は、I/O要求毎にエントリを有するテーブル形式で格納され、現用系計算機から待機系計算機へはネットワークを介してテーブルの情報が送信される。

【発明の効果】**【0011】**

本発明によれば、共有ディスクへの書き込みI/O処理が障害発生により中断してディスクに不整合の疑いがある場合には、待機系計算機がこれを検出して診断処理や修復処理を行い、その後の処理を引き継ぐことができる。さらに、待機系計算機における診断や修復の対象となるディスクを不整合の疑いがあるディスクに絞るため、不要な処理を省略でき、復旧時間を短縮できる。

40

【発明を実施するための最良の形態】**【0012】**

以下、図面を参照して本発明に係るいくつかの実施形態について説明する。

【0013】

・第一の実施形態

まず、図1乃至図14を参照して第一の実施例について説明する。

50

【 0 0 1 4 】

図 1 は、第一の実施形態におけるホットスタンバイシステムの概略構成図である。このシステムが適用される計算機システムは、現用系計算機（以下、単に現用系という）1 0 0 0、待機系計算機（以下、単に待機系という）2 0 0 0、両系によって共通的にアクセスされる共有ディスク 3 0 0 0、およびこれらを接続するネットワーク 4 0 0 0 から構成される。

現用系 1 0 0 0 と待機系 2 0 0 0 は、同様の構成を成し、夫々、メモリ 1 4 0 0、2 4 0 0、プロセッサ 1 4 5 0、2 4 5 0、入出力プロセッサ（以下、I O P : Input Output Processor）1 5 0 0、2 5 0 0、通信アダプタ 1 5 5 0、2 5 5 0、およびディスク制御装置 1 6 0 0、2 6 0 0 を備えて構成される。

10

【 0 0 1 5 】

共有ディスク 3 0 0 0 は、1 台あるいは複数台の論理ボリュームから構成される。各論理ボリュームは、それぞれ一意に識別可能な論理ボリューム ID を持つ。

ネットワーク 4 0 0 0 は、現用系 1 0 0 0 および待機系 2 0 0 0 を所定のプロトコルに基づき接続する。なお、図 1 の計算機システムには、通常、現用系 1 0 0 0 や待機系 2 0 0 0 以外の第三の計算機や電子機器が接続されることがあるが、図示では省略されている。

【 0 0 1 6 】

現用系 1 0 0 0 と待機系 2 0 0 0 のメモリ 1 4 0 0、2 4 0 0 に格納されるソフトウェアとしては、アプリケーションプログラム 1 0 5 0、2 0 5 0、系切替え制御プログラム 1 1 0 0、2 1 0 0、整合性検査プログラム 1 1 5 0、2 1 5 0、OS (Operating System) 1 2 0 0、2 2 0 0、アクセス要求監視プログラム 1 2 5 0、2 2 5 0、アクセス要求管理テーブル 1 3 0 0、2 3 0 0、およびデバイスドライバ 1 3 5 0、2 3 5 0 が含まれる。

20

アプリケーションプログラム 1 0 5 0、2 0 5 0 は、1 つまたは複数存在し、トランザクション処理などの業務を行い、共有ディスク 3 0 0 0 に対して I / O を発行する。系切替え制御プログラム 1 1 0 0、2 1 0 0 は、自系の障害を検出し、ホットスタンバイ切替えを実行するプログラムである。現用系、待機系それぞれで動作している系切替え制御プログラム 1 1 0 0、2 1 0 0 は、現用系で障害が発生した場合これを検出し、両系で連携してホットスタンバイ切替えを実行する。

【 0 0 1 7 】

整合性検査プログラム 1 1 5 0、2 1 5 0 は、共有ディスクを構成する論理ボリュームに対し、整合性のチェックを行う。チェックの結果、不整合の疑いがあるボリュームに関しては復旧プログラムを実行する。OS 1 2 0 0、2 2 0 0 は、アプリケーションプログラム 1 0 5 0、2 0 5 0 から発行される論理ボリュームへの I / O 要求に応じて、アクセス要求監視プログラム 1 2 5 0、2 2 5 0 またはデバイスドライバ 1 3 5 0、2 3 5 0 に I / O 処理を要求する。アクセス要求監視プログラム 1 2 5 0、2 2 5 0 は、共有ディスク 3 0 0 0 を構成する論理ボリュームに対する I / O 要求を監視し、要求が発行された場合はそれをアクセス要求管理テーブル 1 3 0 0、2 3 0 0 に記録する。

30

【 0 0 1 8 】

アクセス要求管理テーブル 1 3 0 0、2 3 0 0 は、アプリケーションプログラム 1 0 5 0、2 0 5 0 が共有ディスク 3 0 0 0 に対して発行した I / O 要求を登録し管理する。デバイスドライバ 1 3 5 0、2 3 5 0 は、ドライバ外部からの要求に応じて I / O 処理を実行する。

40

【 0 0 1 9 】

図 2 は、第一の実施形態における共有ディスク 3 0 0 0 の構成図である。共有ディスク 3 0 0 0 は、1 つまたは複数の論理ボリューム 3 1 0 0 ~ 3 3 0 0 から構成される。ここで、論理ボリュームとは、物理ボリューム装置の一部または全部の領域から構成された論理的なディスク装置である。物理ボリュームとは、物理的なディスク装置または OS から物理的なディスク装置として認識可能な記憶装置である。第一の実施形態における論理ディスクは、二つの物理ボリューム 3 4 0 0、3 5 0 0 を二重化して、それぞれ

50

同一のデータを書き込むミラーリング処理を施している。

【 0 0 2 0 】

以下の説明では、待機系 2 0 0 0 を構成する各部 2 0 5 0 ~ 2 6 0 0 の構成は、現用系 1 0 0 0 の各部 1 0 5 0 ~ 1 6 0 0 と同様であるため、現用系の説明をもって、待機系の説明は省略する。

【 0 0 2 1 】

図 3 は、第一の実施形態における OS 1 2 0 0 およびデバイスドライバ 1 3 5 0 の構成を示す図である。

OS 1 2 0 0 は、デバイススイッチテーブル 1 2 1 0、I/O 要求処理プログラム 1 2 2 0、障害検出プログラム 1 2 3 0、デバイススイッチテーブル書き換えプログラム 1 2 4 0 から構成される。デバイスドライバ 1 3 5 0 は、書き込み処理プログラム 1 3 6 0、ダンプ出力プログラム 1 3 7 0 から構成される。なお、上記以外にも OS やデバイスドライバとして必要なプログラム部分は存在するが、本実施の形態の説明に直接関係が無いため、説明を省略する。

10

【 0 0 2 2 】

デバイススイッチテーブル 1 2 1 0 は、共有ディスク 3 0 0 0 を構成する各論理ボリュームが持つ論理ボリューム ID と、論理ボリュームに対する操作の種類書き込み、障害発生時のダンプ出力などの組み合わせに対して、処理を実施するプログラムのアドレスを登録するテーブルである。操作の種類に関しては、本実施形態では操作の種類として書き込み処理とダンプ出力処理以外に関する説明を省略しているが、これら以外の操作があってもよい。図 3 では、論理ボリューム ID = 0 に対する書き込み処理に対応するデバイススイッチテーブルエントリ 1 2 1 1 に、デバイスドライバ 1 3 5 0 の書き込み処理プログラム 1 3 6 0 のアドレスが登録されており、同じく論理ボリューム ID = 0 に対するダンプ出力処理に対応するデバイススイッチテーブルエントリ 1 2 1 2 に、デバイスドライバ 1 3 5 0 のダンプ出力プログラム 1 3 7 0 のアドレスが登録されている。

20

【 0 0 2 3 】

I/O 要求処理プログラム 1 2 2 0 は、アプリケーションプログラム 1 0 5 0 からの I/O 要求によって呼び出され、I/O 対象に指定された論理ボリューム ID と操作の種類を元に、デバイススイッチテーブル 1 2 1 0 に登録されたアドレスのプログラムを呼び出す。障害検出プログラム 1 2 3 0 は、現用系 1 0 0 0 において処理を継続できなくなる程度の障害が発生すると、これを検出する。そして、予め登録された論理ボリュームについて、対応するデバイススイッチテーブルエントリ 1 2 1 2 を参照し、登録されたアドレスのプログラムを呼び出す。デバイススイッチテーブル書き換えプログラム 1 2 4 0 は、OS 1 2 0 0 の内部または外部プログラムからの要求によって呼び出され、デバイススイッチテーブル 1 2 1 0 内のエントリを書き換える。

30

【 0 0 2 4 】

書き込み処理プログラム 1 3 6 0 は、論理ボリュームへの書き込み要求に応じて、対応する物理ボリュームの領域に書き込む処理を行う。このとき、論理ボリュームにミラーリング処理が施されている場合は、多重化された物理ボリュームのそれぞれに同一の書き込み処理を行う。なお、本実施形態では、ミラーリング処理を書き込み処理プログラム 1 3 6 0 が行っているが、書き込み処理プログラム 1 3 6 0 以外にミラーリング処理を行うプログラムが存在してもよい。

40

【 0 0 2 5 】

ダンプ出力処理プログラム 1 3 7 0 は、OS 1 2 0 0 に障害が発生し処理続行不能に陥った場合に呼び出され、メモリ 1 4 0 0 に記録されているデータをディスク装置に書き込む処理を行う。

【 0 0 2 6 】

図 4 は、第一の実施形態におけるアクセス要求監視プログラム 1 2 5 0 の構成を示す図である。

アクセス要求監視プログラム 1 2 5 0 は、アクセス監視対象論理ボリュームの登録プログ

50

ラム 1 2 5 5、アクセス監視対象論理ボリュームの除外プログラム 1 2 6 0、アクセス監視状態の取得プログラム 1 2 6 5、アクセス要求記録プログラム 1 2 7 0、アクセス記録の送信プログラム 1 2 7 5、初期化処理プログラム 1 2 8 0 から構成される。なお、アクセス監視対象論理ボリュームの除外プログラム 1 2 6 0 およびアクセス監視状態の取得プログラム 1 2 6 5 は、無くても良い。

【 0 0 2 7 】

アクセス監視対象論理ボリュームの登録プログラム 1 2 5 5 は、指定された論理ボリュームをアクセス監視対象に追加する。追加された論理ボリュームは、アクセス要求記録プログラム 1 2 7 0 によりアクセス要求を記録可能となる。なお、アクセス監視対象論理ボリュームの登録プログラム 1 2 5 5 は、アプリケーションプログラム 1 0 5 0 からシステムコールまたはユーザコマンドにより呼出される。

10

【 0 0 2 8 】

アクセス監視対象論理ボリュームの除外プログラム 1 2 6 0 は、指定された論理ボリュームをアクセス監視対象より除外する。除外された論理ボリュームは、アクセス要求記録プログラム 1 2 7 0 によるアクセス要求の記録が不可能となる。なお、アクセス監視対象論理ボリュームの除外プログラム 1 2 6 0 は、アプリケーションプログラム 1 0 5 0 からシステムコールまたはユーザコマンドにより呼出される。

【 0 0 2 9 】

アクセス監視状態の取得プログラム 1 2 6 5 は、プログラム外部または内部からの問い合わせに応じ、論理ボリュームの現在のアクセス監視状態を返す。アクセス監視状態とは、アクセス監視対象として登録されているか否か、現在のアクセス状況などの情報を表す。アクセス監視状態の取得プログラム 1 2 6 5 は、アクセス要求監視プログラム 1 2 5 0 の内部または、アプリケーションプログラム 1 0 5 0 からシステムコールまたはユーザコマンドにより呼び出される。

20

【 0 0 3 0 】

アクセス要求記録プログラム 1 2 7 0 は、アクセス監視対象論理ボリュームへのアクセスの際に I / O 要求処理プログラム 1 2 2 0 から呼び出される。アクセス監視状態の判定プログラム 1 2 7 0 は、アクセス要求が共有ディスク 3 0 0 0 に対する書き込み処理要求である場合に、その要求をアクセス要求管理テーブル 1 3 0 0 に記録する。

【 0 0 3 1 】

アクセス記録の送信プログラム 1 2 7 5 は、アクセス要求管理テーブル 1 3 0 0 に記録された情報を、ネットワーク 4 0 0 0 を通じて待機系 2 0 0 0 の整合性検査プログラム 2 1 5 0 に送信する。アクセス要求記録の送信プログラム 1 2 7 5 は、障害発生時に系切替え制御プログラム 1 1 0 0 または障害検出プログラム 1 2 3 0 から呼び出される。

30

【 0 0 3 2 】

初期化処理プログラム 1 2 8 0 は、ダンプ出力時にアクセス記録の送信プログラム 1 2 7 5 が呼び出されるように、デバイススイッチテーブル 1 2 1 0 を書き換える処理を行う。初期化処理プログラム 1 2 8 0 は、システム立ち上げ時に系切替え制御プログラム 1 1 0 0 から呼び出される。

【 0 0 3 3 】

図 5 は、第一の実施形態におけるアクセス要求管理テーブル 1 3 0 0 を構成するアクセス要求管理テーブルエントリ 1 3 1 0 の構成図である。アクセス要求管理テーブルエントリ 1 3 1 0 は、アクセス対象の論理ボリュームを識別する論理ボリューム ID 1 3 2 0、共有ディスクへの書き込み要求の中で処理中のものをカウントする書き込み要求カウンタ 1 3 3 0 から構成される。アクセス要求管理テーブルエントリ 1 3 1 0 は、アクセス要求記録プログラム 1 2 7 0 により更新される。なお、図 5 において、斜線部分は他のフィールドを格納してもよい。

40

【 0 0 3 4 】

以下、図 6 ~ 図 1 1 を参照して、第一の実施形態におけるアクセス要求監視プログラム 1 2 5 0 を構成する各プログラムの処理の流れを説明する。

50

図6は、第一の実施形態におけるアクセス監視対象論理ボリュームの登録プログラム1255の処理の流れを示す図である。

アクセス監視対象論理ボリュームの登録プログラム1255は、現用系1000および待機系2000のシステム立ち上げ時にシステムコールまたはコマンドにより、登録対象の論理ボリュームIDをパラメータとして呼び出す。次に、デバイススイッチテーブル書き換えプログラム1240を呼び出し、登録対象の論理ボリュームへの書き込み処理に対応するデバイススイッチテーブルエントリ1211にアクセス要求記録プログラム1270を登録する。これにより、登録対象の論理ボリュームへの書き込みI/O要求が発生すると、アクセス要求記録プログラム1270が呼び出される(S5000)。

【0035】

次に、ステップS5000で行ったデバイススイッチテーブルエントリ1211の書き換えが成功したかどうか判定する。ここで、この書き換え処理の結果は、書き換えを実施するプログラムからの戻り値として表現される。成功の戻り値が返っていればステップS5030へ進み、失敗の戻り値が返っていればステップS5020へ進む(S5010)。

【0036】

ステップS5010で失敗の戻り値が返った場合、呼び出し元プログラムにデバイススイッチテーブル書き換え失敗を表す戻り値を返し、このプログラムの処理を終了する(S5020)。一方、ステップS5010で成功の戻り値が返った場合、アクセス要求管理テーブル1300にアクセス要求管理テーブルエントリ1310を1つ追加する。この時、そのエントリの論理ボリュームID1320には登録対象論理ボリュームの論理ボリュームIDをセットし、書き込み要求カウンタ1330には「0」をセットする(S5030)。

最後に、処理が成功したことを表す戻り値を呼び出し元のプログラムに返し、このプログラムを終了する(S5040)。

【0037】

図7は、第一の実施形態におけるアクセス監視対象論理ボリュームの除外プログラム1260の処理の流れを示す図である。

アクセス監視対象論理ボリュームの除外プログラム1260は、システムコールまたはコマンドにより、登録対象の論理ボリュームIDをパラメータとして呼び出される。

まず、パラメータとして渡された論理ボリュームIDに基づき、アクセス要求管理テーブル1300から論理ボリュームIDに対応するアクセス要求管理テーブルエントリ1310を検索する(S5100)。

【0038】

この検索の結果、該当するエントリが見つかったかどうか判定する(S5110)。見つかった場合にはステップS5130に進み、見つからなかった場合にはステップS5120に進む。ステップS5110で、該当するエントリが見つからなかった場合、該当エントリ無しを表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する(S5120)。一方、該当するエントリが見つかった場合には、そのエントリの書き込み要求カウンタ1330が「0」であるかどうか判定する。「0」である場合はステップS5150に進み、「0」でない場合はステップS5140に進む(S5130)。

【0039】

ステップS5130で、書き込み要求カウンタ1330が「0」でない場合、指定された論理ボリュームはI/O処理中である旨を表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する(S5140)。一方、書き込み要求カウンタ1330が「0」である場合、デバイススイッチテーブル書き換えプログラム1240を呼び出し、指定された論理ボリュームへの書き込み処理に対応するデバイススイッチテーブルエントリ1211に書き込み処理プログラム1360を登録する。これにより、その論理ボリュームへの書き込みI/O要求が発生すると、書き込み処理プログラム1360が呼び出される(S5150)。

10

20

30

40

50

【 0 0 4 0 】

次に、ステップ S 5 1 5 0 で行ったデバイススイッチテーブルエントリ 1 2 1 1 の書き換えが成功したかどうか判定する (S 5 1 6 0)。ここで、この書き換え処理の結果は、書き換えを実施するプログラムからの戻り値として表現される。成功の戻り値が返っていればステップ S 5 1 8 0 へ進み、失敗の戻り値が返っていればステップ S 5 1 7 0 へ進む。ここで、失敗の戻り値が返った場合、デバイススイッチテーブル書き換え失敗を表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 1 7 0)。一方、成功の戻り値が返った場合には、ステップ S 5 1 0 0 で検索したアクセス要求管理テーブルエントリ 1 3 1 0 をアクセス要求管理テーブル 1 3 0 0 から削除する (S 5 1 8 0)。

10

最後に、このプログラムが成功した旨を表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 1 9 0)。

【 0 0 4 1 】

図 8 は、第一の実施形態におけるアクセス監視状態の取得プログラム 1 2 6 5 の処理の流れを示す図である。

アクセス監視状態の取得プログラム 1 2 6 5 は、取得対象の論理ボリューム ID をパラメータとして、システムコールまたはコマンドにより呼び出されるか、アクセス監視プログラム 1 2 5 0 内の他のプログラムから内部的に呼び出される。なお、論理ボリューム ID の他にも、監視状態を示す情報を格納するバッファのアドレス等の追加パラメータを伴ってもよい。

20

【 0 0 4 2 】

まず、パラメータとして渡された論理ボリューム ID に基づき、アクセス要求管理テーブル 1 3 0 0 からその論理ボリューム ID に対応するアクセス要求管理テーブルエントリ 1 3 1 0 を検索する (S 5 2 0 0)。

次に、この検索の結果、該当するエントリが見つかったかどうか判定する (S 5 2 1 0)。該当するエントリが見つかった場合はステップ S 5 2 3 0 に進み、見つからなかった場合はステップ S 5 1 2 0 に進む。該当するエントリが見つからなかった場合、該当エントリ無しを表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 2 2 0)。一方、該当するエントリが見つかった場合には、該当エントリの情報を呼び出し元に返す (S 5 2 3 0)。このとき、追加パラメータでバッファのアドレスを渡されていれば、そのバッファに格納してもよい。

30

次に、このプログラムの処理が成功した旨を表す戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 2 4 0)。

【 0 0 4 3 】

図 9 は、第一の実施の形態におけるアクセス要求記録プログラム 1 2 7 0 の処理の流れを示す図である。

アクセス要求の記録プログラム 1 2 7 0 は、アクセス監視対象の論理ボリューム対し I / O 要求が発生した場合に、I / O 要求処理プログラム 1 2 2 0 から論理ボリューム ID や I / O に用いるデータなどをパラメータとして呼び出される。

【 0 0 4 4 】

40

まず、パラメータとして渡された論理ボリューム ID に基づき、アクセス要求管理テーブル 1 3 0 0 からその論理ボリューム ID に対応するアクセス要求管理テーブルエントリ 1 3 1 0 を検索する (S 5 3 0 0)。検索の結果、該当するエントリが見つかったかどうか判定する (S 5 3 1 0)。該当するエントリが見つかった場合はステップ S 5 3 2 0 に進み、見つからなかった場合にはステップ S 5 3 3 0 に進む。なお、このステップは無くても良い。

【 0 0 4 5 】

ステップ S 5 3 1 0 で該当するエントリが見つかった場合、ステップ S 5 3 0 0 で検索したアクセス要求管理テーブルエントリ 1 3 1 0 の書き込み要求カウンタ 1 3 3 0 の値に「 1 」を加える (S 5 3 2 0)。

50

【 0 0 4 6 】

次に、書き込み処理プログラム 1 3 6 0 を呼び出し、実際の I / O 処理を実施する。このとき、I / O 要求処理プログラム 1 2 2 0 がこのプログラムに渡したパラメータをそのまま渡す。書き込み処理プログラム 1 3 6 0 は、実際の I / O を行い、I / O 完了した時点でその処理を終了し戻ってくる (S 5 3 3 0)。このとき、対象の論理ボリュームに対応する物理ボリュームそれぞれに書き込みを実施する。

【 0 0 4 7 】

次に、ステップ S 5 3 0 0 で実施した検索で該当するエン트리があるかどうか判定する (S 5 3 4 0)。該当するエン트리がある場合はステップ S 5 3 5 0 に進み、無い場合にはステップ S 5 3 6 0 に進む。なお、このステップは必ずしも必要ない。

ステップ S 5 3 4 0 で、該当するエン트리がある場合は、そのエントリの書き込み要求カウンタ 1 3 3 0 の値を「 1 」減じる。この結果、書き込み要求カウンタ 1 3 3 0 の値が「 0 」になれば、論理ボリュームへの I / O 処理はすべて終了したことを意味する (S 5 3 5 0)。

最後に、デバイスドライバから返された戻り値を呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 3 6 0)。

【 0 0 4 8 】

図 1 0 は、第一の実施の形態におけるアクセス記録の送信プログラム 1 2 7 5 の処理の流れを示す図である。アクセス記録の送信プログラム 1 2 7 5 は、障害発生を契機として、障害検出プログラム 1 2 3 0 から呼び出される。

【 0 0 4 9 】

まず、このプログラム内部で用いる内部カウンタの値を「 0 」にセットする (S 5 4 0 0)。次に、待機系 2 0 0 0 の系切替え制御プログラム 2 1 0 0 に対し、アクセス要求管理テーブル 1 3 0 0 の内容を、ネットワーク 4 0 0 0 を経由して送信する (S 5 4 1 0)。このとき、I O P 1 5 0 0 を経由して、通信アダプタ 1 5 5 0 を用いて通信を行う。なお、この処理は外部プログラムが行っても良い。次に、内部カウンタの値に「 1 」を加える (S 5 4 2 0)。

【 0 0 5 0 】

次に、内部カウンタの値が規定の値に達したかを判定する (S 5 4 3 0)。規定の値に達した場合はステップ S 5 4 4 0 に進み、達していない場合にはステップ S 5 4 1 0 に進む。

この判定で、内部カウンタの値が規定の値に達した場合には、ダンプ出力処理プログラム 1 3 7 0 を呼び出す。ダンプ出力処理プログラム 1 3 7 0 は、ダンプ出力処理を実施し、処理の結果を戻り値として返す (S 5 4 4 0)。ダンプ出力処理プログラム 1 3 7 0 の戻り値を、このプログラムの呼び出し元プログラムに返し、このプログラムの処理を終了する (S 5 4 5 0)。

【 0 0 5 1 】

図 1 1 は、第一の実施形態における初期化処理プログラム 1 2 8 0 の処理の流れを示す図である。初期化処理プログラム 1 2 8 0 は、システム立ち上げ時に系切替え制御プログラム 1 1 0 0 からダンプ出力対象となる論理ボリューム I D をパラメータとして呼び出される。なお、これ以外に追加パラメータがあってもよい。

【 0 0 5 2 】

まず、O S 1 2 0 0 のデバイススイッチテーブル書き換えプログラム 1 2 4 0 を呼び出し、指定された論理ボリュームへのダンプ出力処理に対応するデバイススイッチテーブル 1 2 1 2 にアクセス記録の送信プログラム 1 2 7 5 のアドレスを登録する。これにより、障害が発生してその論理ボリュームへのダンプ出力要求が発生すると、障害検出プログラム 1 2 3 0 はアクセス記録の送信プログラム 1 2 7 5 を呼び出す (S 5 5 0 0)。

【 0 0 5 3 】

次に、ステップ S 5 5 0 0 の処理が成功したかどうか判定する (S 5 5 1 0)。成功した場合はステップ S 5 5 3 0 に進み、失敗した場合にはステップ S 5 5 2 0 へ進む。

10

20

30

40

50

即ち、ステップS 5 5 1 0で書き換えに失敗した場合は、デバイススイッチテーブル書き換えに失敗した旨を表す戻り値を呼び出し、元プログラムに返して、このプログラムの処理を終了する(S 5 5 2 0)。一方、書き換えに成功したと判定された場合には、このプログラムの処理が成功した旨を表す戻り値を呼び出して元プログラムに返し、このプログラムの処理を終了する(S 5 5 3 0)。

【 0 0 5 4 】

図12は、第一の実施形態における系切替え制御プログラム1100の処理の流れを示す図である。以下、待機系2000の系切替え制御プログラム2100について記述する。系切替え制御プログラム2100は、現用系1000の障害を監視し、障害が発生した場合は現用系1000で行っていた業務の処理を待機系2000に引き継がせる処理を行う。

10

【 0 0 5 5 】

まず、現用系1000のアクセス記録の送信プログラム1275から送信されたアクセス要求管理テーブル1300の内容を受信する(S 5 6 0 0)。次に、ステップS 5 6 0 0での受信処理が成功したかどうか判定する(S 5 6 1 0)。受信が成功した場合はステップS 5 6 3 0に進み、受信が失敗した場合にはステップS 5 6 2 0に進む。

【 0 0 5 6 】

ステップS 5 6 1 0で、受信失敗と判定された場合は、受信失敗に対応する処理を行う(S 5 6 2 0)。なお、受信が失敗した状態で業務引継ぎ処理を行っても良い。一方、受信成功と判定された場合には、待機系2000の整合性検査プログラム2150を呼び出し、整合性の検査処理を行う(S 5 6 3 0)。このとき、ステップS 5 6 0 0で受信したアクセス要求管理テーブル1300の内容をパラメータとして渡す。なお、アクセス要求管理テーブル1300以外の追加パラメータがあってもよい。

20

最後に、業務引継ぎ処理を行い、待機系2000は新現用系として業務を引き継ぐ(S 5 6 4 0)。

【 0 0 5 7 】

図13は、第一の実施形態において、待機系2000における整合性検査プログラム2100の処理の流れを示す図である。整合性検査プログラム2100は、系切替え処理の中で、系切替え制御プログラム2100から呼び出される。このとき、パラメータとして、現用系1000のアクセス記録の送信プログラム1275から送信されたアクセス要求管理テーブル1300の内容を受け取る。なお、これ以外に追加パラメータがあってもよい。

30

【 0 0 5 8 】

まず、パラメータとして受け取ったアクセス要求管理テーブル1300の中から、取得可能なアクセス要求管理テーブルエントリ1310が存在するかどうか判定する。この判定を全エントリに対して検索して行う(S 5 7 0 0)。該当するエントリが存在する場合は、ステップS 5 7 1 0に進み、存在しない場合にはこのプログラムの処理を終了する。ステップS 5 7 0 0で、取得可能なアクセス要求管理テーブルエントリ1310が存在する場合は、エントリを1つ取得する(S 5 7 1 0)。そして、その取得したアクセス要求管理テーブルエントリ1310の書き込み要求カウンタ1330の値が「0」かどうか判定する(S 5 7 2 0)。

40

【 0 0 5 9 】

判定の結果、「0」である場合は、そのエントリに対応する論理ボリュームは書き込みI/Oが全て完了した状態であることを意味し、復旧プログラム実行の必要がない。このため復旧プログラムを実行せずに、ステップS 5 7 0 0に進む。一方、「0」でない場合は、そのエントリに対応する論理ボリュームは処理中で中断された書き込みI/Oが存在し、論理ボリューム内が不整合である可能性があることを意味している。そこでステップS 5 7 3 0に進み、論理ボリュームの復旧プログラムを実行する。

【 0 0 6 0 】

ステップS 5 7 2 0での判定の結果、書き込み要求カウンタ1330の値が「0」でな

50

いと判定された場合は、対応する論理ボリュームに対して論理ボリュームの復旧プログラムを実行する。これにより、その論理ボリューム内に不整合の疑いがある場合は物理ボリュームの一部または全部を他の物理ボリュームにコピーし、不整合な状態を修復する。修復プログラム実行後はS 5 7 0 0に進む(S 5 7 3 0)。

【 0 0 6 1 】

図 1 4 は、第一の実施形態におけるホットスタンバイ処理の全体の流れを示す図である。

まず、現用系 1 0 0 0 のシステム立ち上げ時の処理が実行される。

現用系 1 0 0 0 は、システム立ち上げ時に初期化処理プログラム 1 2 8 0 を呼び出し、実行する(S 5 8 0 0)。これにより、デバイススイッチテーブルエントリ 1 2 1 2 を書き換え、ダンプ出力時にアクセス要求記録の送信プログラム 1 2 7 5 が呼び出されるように設定する。

10

【 0 0 6 2 】

次に、アクセス管理対象論理ボリュームの登録プログラム 1 2 5 5 を呼び出し、実行する(S 5 8 0 5)。これにより、デバイススイッチテーブルエントリ 1 2 1 1 を書き換え、論理ボリュームへの書き込み I / O 要求発生時にアクセス要求記録プログラム 1 2 7 0 を呼び出すように設定する。なお、アクセス監視対象とするべき論理ボリュームが複数ある場合は、個々に対してこのステップの処理を行う。

【 0 0 6 3 】

続いて、現用系 1 0 0 0 の通常運用時の処理が実行される。

20

アプリケーションプログラム 1 0 5 0 が論理ボリュームに対し書き込みの I / O 要求を発行すると、OS 1 2 0 0 の I / O 要求処理プログラム 1 2 2 0 が呼び出され、実行される(S 5 8 1 0)。I / O 要求処理プログラム 1 2 2 0 は、I / O 対象となる論理ボリュームと、操作の内容に基づき、デバイススイッチテーブルエントリ 1 2 1 1 の内容を参照する。

【 0 0 6 4 】

次に、I / O 要求処理プログラム 1 2 2 0 は、その参照したデバイススイッチテーブルエントリ 1 2 1 1 に登録されているアクセス要求記録プログラム 1 2 7 5 を呼び出し、実行する(S 5 8 1 5)。アクセス要求記録プログラム 1 2 7 5 は、下記(1)～(3)の処理を行う。

30

I / O 対象の論理ボリュームに対応するアクセス要求管理テーブルエントリ 1 3 1 0 の書き込み要求カウンタ 1 3 3 0 に「1」を加える。

(2) デバイスドライバ 1 3 5 0 の書き込み処理プログラム 1 3 6 0 を呼び出して実際の I / O 処理を行う。このとき、対象となる論理ボリュームに対応する物理ボリュームそれぞれに同一の書き込みを実施する。

(3) 上記(1)で操作した書き込み要求カウンタ 1 3 3 0 の値を「1」減じた後、呼び出し元のアプリケーションプログラム 1 0 5 0 に I / O 完了を報告する。

このステップ S 5 8 1 0 と S 5 8 1 5 は、現用系 1 0 0 0 で障害が発生するまで繰り返えし、実行される。

【 0 0 6 5 】

40

ここで、現用系 1 0 0 0 で障害が発生した場合には、系の切替え処理が実行される。

即ち、現用系 1 0 0 0 で障害が発生すると、この障害を障害検出プログラム 1 2 3 0 が検出する。障害検出プログラム 1 2 3 0 は予め登録された論理ボリュームへのダンプ出力処理に対応するデバイススイッチテーブルエントリ 1 2 1 2 を参照し、そのエントリに登録されたプログラムのアドレスを取得する(S 5 8 2 0)。

【 0 0 6 6 】

次に、障害検出プログラム 1 2 3 0 は、ステップ S 5 8 2 0 で取得したアドレスのプログラムを呼び出す。これにより、アクセス要求記録の送信プログラム 1 2 7 5 を呼び出し、実行する(S 5 8 2 5)。アクセス要求記録の送信プログラム 1 2 7 5 は、アクセス要求管理テーブル 1 3 0 0 の内容を待機系 2 0 0 0 の系切替え制御プログラム 2 1 0 0 に送

50

信する。

【 0 0 6 7 】

次に、障害検出プログラム 1 2 3 0 は、デバイスドライバ 1 3 5 0 のダンプ出力処理プログラム 1 3 7 0 を呼び出し、実行する (S 5 8 3 0)。これにより、システム内のメモリダンプが論理ボリュームに出力される。ダンプ出力が終了すると、現用系 1 0 0 0 は障害を復旧し、新たな待機系として立ち上がる。このとき、後述するステップ S 5 8 5 0 以降の処理を行う (S 5 8 3 5)。

【 0 0 6 8 】

これに対して、待機系 2 0 0 0 ではシステム立ち上げ処理が実行される。待機系 2 0 0 0 では、システム立ち上げ時に初期化プログラム 2 2 8 0 を呼び出して実行する (S 5 8 5 0)。これにより、OS 2 2 0 0 のデバイススイッチテーブルエントリ 2 2 1 2 を書き換え、ダンプ出力時にアクセス要求記録の送信プログラム 2 2 7 5 が呼び出されるよう設定する。

10

【 0 0 6 9 】

次に、アクセス管理対象論理ボリュームの登録プログラム 2 2 5 5 を呼び出し、実行する (S 5 8 5 5)。これにより、OS 2 2 0 0 のデバイススイッチテーブルエントリ 2 2 1 1 を書き換え、論理ボリュームへの書き込み I / O 要求発生時にアクセス要求記録プログラム 2 2 7 0 を呼び出すよう設定する。なお、アクセス監視対象とすべき論理ボリュームが複数ある場合には、夫々の論理ボリュームに対してこのステップの処理を行う。以上の処理を終えると、待機系 2 0 0 0 の系切替え制御プログラム 2 1 0 0 は現用系 1 0 0 0 のアクセス記録の送信プログラム 1 2 7 5 からアクセス要求管理テーブル 1 3 0 0 の内容が送信されるのを待つ。

20

【 0 0 7 0 】

続いて、障害時における待機系 2 0 0 0 系切替え処理が実行される。待機系 2 0 0 0 の系切替え制御プログラム 2 1 0 0 は、現用系 1 0 0 0 から送信されたアクセス要求管理テーブル 1 3 0 0 の内容を受信する (S 5 8 6 0)。そして、受信したアクセス要求管理テーブル 1 3 0 0 の内容をパラメータとして整合性検査プログラム 2 1 5 0 を呼び出し、実行する (S 5 8 6 5)。即ち、整合性検査プログラム 2 1 5 0 は、パラメータとして渡されたアクセス要求管理テーブル 1 3 0 0 の内容に基づき、論理ボリュームの整合性をチェックする。このとき、途中で中断された書き込み I / O が存在する場合は、その論理ボリュームに対して復旧プログラムを実行する。復旧プログラムでは、中断された I / O の対象となった領域を他の物理ボリュームとの間でコピーし、不整合な状態を修復する。なお、論理ボリュームに対応する領域全てをコピーしても良い。以上の処理が終了すると、系切替え制御プログラム 2 1 0 0 は、業務引継ぎ処理を行い、新たな現用系として業務を再開する (S 5 8 7 0)。

30

【 0 0 7 1 】

・第二の実施形態

次に、図 1 5 ~ 図 2 2 を参照して第二の実施形態について説明する。上述した第一の実施形態との相違点は、現用系のアプリケーションが共有ディスクへの書き込み I / O 要求を発行する度に、I / O 実施前と完了時に現用系のアクセス要求監視プログラムが待機系のアクセス要求監視プログラムに通知し、待機系側のアクセス要求監視プログラムがアクセス要求管理テーブルに記録する点である。

40

【 0 0 7 2 】

図 1 5 は、第二の実施形態における計算機システムの構成を示す図である。系切替え制御プログラム 6 0 0 0、6 2 0 0、整合性検査プログラム 6 0 5 0、6 2 5 0、アクセス要求監視プログラム 6 1 0 0、6 3 0 0 を除き、その構成は第一の実施の形態と同じであるため、同じ部分について説明は以降省略する。また、以下の説明では、待機系 2 0 0 0 を構成する 6 2 0 0 ~ 6 3 0 0 の構成は、現用系 1 0 0 0 の 6 0 0 0 ~ 6 1 0 0 と同様であるため、現用系 1 0 0 0 の説明をもって代用する。

【 0 0 7 3 】

50

図16は、第二の実施形態におけるアクセス要求監視プログラム6100の構成を示す図である。

アクセス要求監視プログラム6100は、アクセス監視対象論理ボリュームの登録プログラム1255、アクセス監視対象論理ボリュームの除外プログラム1260、アクセス監視状態の取得プログラム1265、アクセス要求記録プログラム6110、アクセス要求記録の受信プログラム6120から構成される。なお、アクセス監視対象論理ボリュームの除外プログラム1255およびアクセス監視状態の取得プログラム1265は、無くても良い。また、プログラム1255～1265の構成および動作は第一の実施形態と同じであるため、説明は省略する。

【0074】

図17は、第二の実施の形態における、現用系1000で書き込みI/O要求が発生した時に、待機系へ送信されるアクセス要求通知パケット6400の構成を示す図である。アクセス要求通知パケット6400は、通知の内容を示す通知の種類6410とアクセス対象の論理ボリュームID6420を含んで構成される。ここで、通知の種類6410は、待機系への通知の種類を示す値を格納し、「書き込みI/O開始」を示す値または「書き込みI/O完了」を示す値のどちらかを示す。また、論理ボリュームID6420は、書き込みI/Oの対象となった論理ボリュームを識別するための論理ボリュームIDを格納する。

【0075】

現用系1000で書き込みI/O要求発生が発生し、アクセス要求記録プログラム6110が呼び出されると、アクセス要求記録プログラム6110はアクセス要求通知パケット6400を作成し、待機系2000のアクセス記録の受信プログラム6320に送信する。

【0076】

図18は、第二の実施形態におけるアクセス要求記録プログラム6110の処理の流れを示す図である。アクセス要求記録プログラム6110は、第一の実施の形態と同じく、共有ディスク3000への書き込みI/O要求発生を契機として、I/O要求処理プログラム1220から論理ボリュームIDやI/Oに用いるデータなどをパラメータとして呼び出される。

【0077】

まず、書き込みI/O開始を示すアクセス要求通知パケット6400を作成する(S7000)。このとき、通知の種類6410には「書き込みI/O開始」を示す値を格納し、論理ボリュームID6420にはI/O対象となった論理ボリュームの論理ボリュームIDを格納する。

【0078】

次に、作成されたアクセス要求通知パケット6400を、ネットワーク4000を經由して待機系のアクセス記録の受信プログラム6320へ送信する(S7010)。この場合、IOP1500を經由して通信アダプタ1550を用いて通信を行うが、この処理は外部プログラムが行っても良い。次に、書き込み処理プログラム1360を呼び出し、実際のI/O処理を実施する(S7020)。このとき、I/O要求処理プログラム1220がこのプログラムに渡したパラメータをそのまま渡す。書き込み処理プログラム1360は、実際のI/Oを行い、I/O完了した時点でその処理を終了し戻ってくる。

【0079】

次に、書き込みI/O完了を示すアクセス要求通知パケット6400を作成する(S7030)。この場合、通知の種類6410には「書き込みI/O完了」を示す値を格納し、論理ボリュームID6420にはI/O対象となった論理ボリュームの論理ボリュームIDを格納する。

そして、上記のように作成されたアクセス要求通知パケット6400を、ネットワーク4000を經由して待機系のアクセス記録の受信プログラム6320に送信する(S7040)。ここで、IOP1500を經由して通信アダプタ1550を用いて通信を行うが、

10

20

30

40

50

この処理は外部プログラムが行っても良い。

【0080】

最後に、書き込み処理プログラム1360から返された戻り値を呼び出し、元プログラムに返して、このプログラムの処理を終了する(S7050)。

【0081】

図19は、アクセス要求の受信プログラム6120の処理の流れを示す図である。アクセス要求の受信プログラム6120は、通常運用時の待機系2000上で常にアクセス要求通知パケット6400を待ち受けており、アクセス要求通知パケット6400の到着を契機として処理を行う。以降は、待機系2000のアクセス記録の受信プログラム6320について説明する。

【0082】

まず、現用系1000のアクセス要求記録プログラム6110から送信されたアクセス要求通知パケット6400を受信する(S7100)。次に、ステップS7100で受信したアクセス要求通知パケット6400の論理ボリュームID6420の内容に基づき、アクセス要求管理テーブルエントリ2310を検索する(S7110)。

【0083】

この検索の結果、該当するエントリが見つかったかどうか判定する(S7120)。該当エントリが見つかった場合はステップS7130に進み、見つからなかった場合には操作すべきエントリが見つからなかったと判断し、ステップS7100に進む。ステップS7120の判定で該当エントリが見つかった場合、ステップS7100で受信したアクセス要求通知パケット6400の通知の種類6410の内容を判定する(S7130)。この内容が「書き込みI/O開始」を示している場合は、ステップS7110で検索したアクセス要求管理テーブルエントリ2310の書き込み要求カウンタ2330の値に「1」を加え、ステップS7100に進む(S7140)。

【0084】

一方、ステップS7130の判定の結果、通知の種類6410が「書き込みI/O完了」を示している場合には、ステップS7110で検索したアクセス要求管理テーブルエントリ2310の書き込み要求カウンタ2330の値を「1」減じ、ステップS7100に進む(S7150)。

【0085】

図19は、第二の実施形態における障害発生時における系切替え制御プログラム6000の処理の流れを示す図である。

待機系2000の系切替え制御プログラム6000は、通常運用中は常に現用系1000の障害を監視しており、障害発生を検出すると系切替えの処理を実施する。以降は、待機系2000の系切替え制御プログラム6200について説明する。

【0086】

まず、系切替え制御プログラム6200は、現用系1000での障害発生を検出する(S7200)。このとき、障害を検出する手段としては、ALIVEメッセージなど既存の技術による。次に、整合性検査プログラム6250を呼び出し、共有ディスク3000を構成する論理ボリュームに関して、整合性の検査を実施する(S7210)。このとき、論理ボリュームの論理ボリュームIDの集合をパラメータとして整合性検査プログラム6250に渡す。整合性検査プログラム6250は、検査が終了した後、プログラムの処理を終えて戻ってくる。

【0087】

整合性検査プログラム6250の処理が終了すると、待機系2000に業務引継ぎ処理を行わせる。このステップの処理が終了すると、待機系2000は新現用系として業務を再開する(S7220)。

【0088】

図21は、第二の実施形態における待機系2000における整合性検査プログラム6250の処理の流れを示す図である。整合性検査プログラム6250は、系切替え処理の中

10

20

30

40

50

で、系切替え制御プログラム6200から呼び出される。このとき、パラメータとしてアクセス監視対象の論理ボリュームIDの列を受け取る。なお、これ以外に追加パラメータがあってもよい。

【0089】

まず、論理ボリュームIDの列から、取得可能な論理ボリュームIDが存在するかどうか判定する(S7300)。存在する場合は、ステップS7310に進み、存在しない場合にはこのプログラムの処理を終了する。

取得可能な論理ボリュームIDが存在する場合は、1つのエントリを取得し、そのIDに対応するアクセス要求管理テーブルエントリ2310を検索する(S7310)。このとき、取得した論理ボリュームIDはパラメータとして受け取ったIDの列から削除する。

10

【0090】

次に、取得したアクセス要求管理テーブルエントリ2310の書き込み要求カウンタ2330の値が「0」かを判定する(S7320)。「0」である場合は、そのエントリに対応する論理ボリュームは書き込みI/Oが全て完了した状態であることを意味し、復旧プログラム実行の必要がない。このため復旧プログラムを実行せずステップS7300に進む。一方、「0」でない場合には、そのエントリに対応する論理ボリュームは処理中で中断された書き込みI/Oが存在し、論理ボリューム内が不整合である可能性があることを意味している。この時には、対応する論理ボリュームに対して論理ボリュームの復旧プログラムを実行する(S7330)。これにより、その論理ボリューム内に不整合がある場合には復旧プログラムにより修復される。復旧プログラムの実行後はS7300に進む。

20

【0091】

図22は、第二の実施形態におけるホットスタンバイ処理の全体的な流れを示す図である。

【0092】

まず、現用系1000のシステム立ち上げ時の処理が実行される。現用系1000は、システム立ち上げ時にアクセス管理対象論理ボリュームの登録プログラム1255を呼び出し、実行する。これにより、OS1200のデバイススイッチテーブルエントリ1211を書き換え、論理ボリュームへの書き込みI/O要求発生時にアクセス要求記録プログラム6110を呼び出すよう設定する。なお、アクセス監視対象とするべき論理ボリュームが複数ある場合は、夫々の論理ボリュームに対してこの処理を行う(S7400)。

30

【0093】

続いて、現用系1000の通常運用時の処理が実行される。アプリケーションプログラム1050が論理ボリュームに対して書き込みのI/O要求を発行すると、OS1200のI/O要求処理プログラム1220が呼び出され、実行される。I/O要求処理プログラム1220は、I/O対象となる論理ボリュームと、操作の内容に基づき、デバイススイッチテーブルエントリ1211の内容を参照する(S7405)。

【0094】

40

次に、I/O要求処理プログラム1220は、その参照したデバイススイッチテーブルエントリ1211に登録されているアクセス要求記録プログラム6110を呼び出し、実行する。アクセス要求記録プログラム6110は、下記(1)~(3)の処理を行う。なお、下記(2)の処理は、I/Oの対象となる論理ボリュームに対応する物理ボリュームそれぞれに対して行う。

待機系2000のアクセス要求通知受信プログラム6120に「書き込みI/O開始」を示すアクセス要求通知パケット6400を送信し、

デバイスドライバ1350の書き込み処理プログラム1360を呼び出して実際のI/O処理を行い、

待機系2000のアクセス要求通知受信プログラム6120に「書き込みI/O完了」を

50

示すアクセス要求通知パケット6400を送信した後、呼び出し元のアプリケーションプログラム1050にI/O完了を報告する。

このステップS7405とS7410は、現用系1000で障害が発生するまで繰り返し、実行される(S7410)。

【0095】

この状態で現用系1000に障害が発生した場合には、系切替え処理が実行される。即ち、現用系1000で障害が発生し、障害検出プログラム1230がこれを検出すると、障害検出プログラム1230は予め登録された論理ボリュームへのダンプ出力処理に対応するデバイススイッチテーブルエントリ1212を参照し、当該エントリに登録されたプログラムのアドレスを取得する(S7415)。

10

【0096】

次に、障害検出プログラム1230は、ステップS7415で取得したアドレスのプログラムを呼び出す。これにより、デバイスドライバ1300のダンプ出力処理プログラム1320を呼び出し実行する。ダンプ出力処理プログラム1320は、システム内のメモリダンプを論理ボリュームに出力する。ダンプ出力が終了すると、現用系1000は障害を復旧し、新たな待機系として立ち上がる。このとき、後述するステップS7450以降の処理を行う(S7420)。

【0097】

続いて、待機系2000ではシステム立ち上げ処理が実行される。待機系2000は、システム立ち上げ時にアクセス管理対象論理ボリュームの登録プログラム2255を呼び出し、実行する。これにより、OS2200のデバイススイッチテーブルエントリ2211を書き換え、論理ボリュームへの書き込みI/O要求発生時にアクセス要求記録プログラム6310を呼び出すように設定する。なお、アクセス監視対象とすべき論理ボリュームが複数ある場合は、夫々の論理ボリュームに対してこのステップの処理を行う(S7450)。

20

【0098】

以上の処理を終えると、アクセス要求通知受信プログラム6320は、現用系1000からのアクセス要求通知パケット6400を待ち受ける。現用系1000で書き込みI/O要求が発生すると、アクセス要求記録プログラム6110からアクセス要求通知パケット6400が送信される。そして、アクセス要求通知受信プログラム6320は受信したアクセス要求通知パケット6400に応じた処理を実施する(S7455)。

30

【0099】

続いて、障害時における待機系2000系切替え処理が行われる。待機系2000の系切替え制御プログラム6200は、現用系1000の障害発生を検出すると、整合性検査プログラム6250を呼び出し、論理ボリュームの整合性を検査する(S7460)。

【0100】

次に、系切替え制御プログラム6200から呼び出された整合性検査プログラム6250は、パラメータとして渡された論理ボリュームの列に基づき、論理ボリュームの整合性をチェックする。このとき、途中で中断された書き込みI/Oが存在する場合は、復旧プログラムを実行する。復旧プログラムは、この復旧プログラムでは、中断されたI/Oの対象となった領域を他の物理ボリュームとの間でコピーし、不整合な状態を修復する。なお、論理ボリュームに対応する領域全てをコピーしても良い(S7465)。
以上の処理が終了すると、系切替え制御プログラム6200は、業務引継ぎ処理を行い、新たな現用系として業務を再開する(S7470)。

40

【0101】

・第三の実施形態

次に、図23～図29を参照して第三の実施形態について説明する。

第一および第二の実施形態との相違点は、現用系から共有ディスクへの書き込みI/Oを実施する際に、I/O処理中であるボリューム上のアドレスをその論理ボリューム上に記

50

録し、I/O完了時にこの記録を削除する点である。これにより、障害発生により書き込みI/O中が強制中断された場合は、ボリューム上に処理中のアドレスの記録が残るため、不整合な状態を検出可能となる。

【0102】

図23は、第三の実施の形態における計算機システムの構成を示す図である。整合性チェックプログラム8000、8200、デバイスドライバ8100、8300、共有ディスク8400以外の構成については、第二の実施形態と同じであるため、以降の説明は省略する。また、以下の説明では、待機系2000を構成する8200、8300の構成は現用系1000の8000、8100と同様であるため、現用系1000の説明をもって代用する。

10

【0103】

図24は、第三の実施の形態における共有ディスク8400の構成を示す図である。共有ディスク8400は複数の論理ボリューム8410、8420、8430から構成されている。各論理ボリュームの領域は、それぞれ物理ボリューム8450、8460に対応付けられている。このとき、論理ボリュームは、物理ボリュームに一对一で割り当てても良いし、物理ボリュームの領域を分割して一部のみを論理ボリュームに対応付け、複数の論理ボリュームで一つの物理ボリュームを共有しても良い。さらに、分割した領域とは別に領域が設けられ、書き込みI/O要求を記録するアクセス要求管理テーブル8455、8465が格納される。デバイスドライバ8100、8300は、書き込みI/O実施の際にアクセス要求管理テーブル8455、8460を更新する。なお、アクセス要求管理テーブルは、論理ボリュームに対応付けた領域それぞれの中に存在してもよい。各論理ボリュームには、固有の論理ボリュームIDが付与され、それらのIDにより論理ボリュームが認識される。なお、通常、論理ボリュームを複数個有するのが望ましいが1つでもよい。

20

【0104】

図25は、上述のアクセス要求管理テーブル8455（同様の構成であるので、符号8465は省略）を構成するアクセス要求管理テーブルエン트리8440の構成を示す図である。アクセス要求管理テーブル8455は、1つまたは複数のアクセス要求管理テーブルエン트리8440の集合により構成され、各エン트리8440は、書き込み対象アドレス8445を含む。なお、斜線部分は他のフィールドを含んでもよい。

30

【0105】

書き込み対象アドレス8445は、デバイスドライバ8100が書き込みI/Oを行う対象となる領域のアドレスを格納する。例えば、論理ボリューム8410に対する書き込みI/Oを実施する場合は、デバイスドライバ8100がI/Oの対象となるアドレスを書き込み対象アドレス8445に格納したアクセス要求管理テーブルエン트리8440をアクセス要求管理テーブル8455、8465に確保した後、実際のI/O処理を行う。そして、実際のI/O処理が完了した後、I/O処理前に確保したエントリーを削除する。なお、エントリーを削除せず、このI/Oが完了したことを示す値を格納してもよい。また、書き込み処理は、対象論理ボリュームに対象となった物理ボリューム全てに対して実施する。このため、アクセス要求管理テーブル8465に対しても同様の処理を行う。

40

【0106】

図26は、第三の実施の形態におけるデバイスドライバ8100の構成を示す図である。ダンプ出力処理プログラム1370の構成および動作は第一および第二の実施の形態と同じであるが、書き込み処理プログラム8110の動作が異なる。

【0107】

図27は、第三の実施の形態における書き込み処理プログラム8110の処理の流れを示す図である。書き込み処理プログラム8110は、論理ボリュームへの書き込みI/O要求発生時に、I/O処理プログラム1220により呼び出される。なお、図27の処理は、対象論理ボリュームに対応する物理ボリュームのそれぞれに対し行う。

【0108】

50

まず、対象となる論理ボリューム 8410 に対応する物理ボリューム 8450, 8460 のアクセス要求管理テーブル 8455, 8465 に、アクセス要求管理テーブルエントリ 8440 を 1 つ追加する。このとき、書き込み対象アドレス 8450 に、書き込み I/O を行う論理ボリューム上のアドレスを格納する (S9000)。次に、論理ボリューム 8410 に対して実際の書き込み I/O 処理を実施する (S9010)。

最後に、ステップ S9000 で追加したアクセス要求管理テーブルエントリ 8440 を削除する。このとき、このエントリ 8440 の書き込み対象アドレス 8445 が格納されていた領域には、空白を示す値を格納する (S9020)。なお、エントリを削除せず、この I/O が完了したことを示す値を格納してもよい。

【0109】

図 28 は、第三の実施形態における待機系 2000 の整合性検査プログラム 8200 の処理の流れを示す図である。整合性検査プログラム 8200 は、系切替え処理の中で、系切替え制御プログラム 6200 から呼び出される。このとき、パラメータとしてアクセス監視対象の論理ボリューム ID の列を受け取る。なお、これ以外に追加パラメータがあってもよい。

【0110】

まず、論理ボリューム ID の列から、取得可能な論理ボリューム ID が存在するかどうか判定する (S9100)。この判定で、取得可能な論理ボリューム ID が存在する場合は、1 つの論理ボリューム ID を取得する。このとき、取得した論理ボリューム ID はパラメータとして受け取った ID の列から削除する。そして、取得した論理ボリューム ID を元に、対応する論理ボリューム 8410 のアクセス要求管理テーブル 8455, 8465 を読み込む (S9110)。一方、この判定の結果、取得可能な論理ボリューム ID が存在しなければ、このプログラムの処理を終了する。

【0111】

次に、ステップ S9110 で読み込んだアクセス要求管理テーブル 8455, 8465 の中に書き込み対象アドレス 8445 に空白を表す値以外が格納されたものが在るかどうか判定する。1 つも存在しない場合は、そのエントリに対応する論理ボリュームは書き込み I/O が全て完了した状態であることを意味し、復旧プログラム実行の必要がない。このため復旧プログラムを実行せずステップ S9100 に進む。一方、1 つでも存在する場合は、そのエントリに対応する論理ボリュームは処理中で中断された書き込み I/O が存在し、論理ボリューム内が不整合である可能性があることを意味している。そこでステップ S9130 に進み、論理ボリュームの復旧プログラムを実行する (S9120)。復旧プログラムは、その論理ボリュームに対応する物理ボリュームの一部または全部を他の物理ボリュームにコピーし、不整合な状態を修復する。

【0112】

ステップ S9130 では、対応する論理ボリュームに対して論理ボリュームの復旧プログラムを実行する。これにより、その論理ボリューム内に不整合がある場合は復旧プログラムにより修復される (S9130)。復旧プログラムの実行後は S7300 に進み、上記のステップを繰り返す。

【0113】

図 29 は、第三の実施の形態におけるホットスタンバイ処理の流れを示す図である。まず、現用系 1000 のシステム立ち上げ処理を行う (S9200)。この起動は通常のやり方と同じである。

続いて、現用系 1000 の通常運用時の処理が実行される。

即ち、アプリケーションプログラム 1050 が論理ボリュームに対し書き込みの I/O 要求を発行すると、OS 1200 の I/O 要求処理プログラム 1220 が呼び出され、実行される。I/O 要求処理プログラム 1220 は、I/O 対象となる論理ボリュームと、操作の内容及び、デバイススイッチテーブルエントリ 1211 の内容を参照する (S9205)。

【0114】

次に、I/O要求処理プログラム1220は、上記参照されたデバイススイッチテーブルエントリ1211に登録されている書き込み処理プログラム8110を呼び出し実行する。書き込み処理プログラム8110は、書き込み対象論理ボリュームのアクセス要求管理テーブル8455、8465にアクセス要求管理テーブルエントリ8440を1つ追加し、実際のI/O処理を行う。更に、上記の追加したアクセス要求管理テーブルエントリ8440を削除した後、呼び出し元のアプリケーションプログラム1050にI/O完了を報告する。上記ステップS9205とS9210は、現用系1000で障害が発生するまで繰り返えされる(S9210)。

【0115】

この状態で、現用系1000に障害が発生した場合には、系切替え処理が実行される。即ち、現用系1000で障害が発生し、障害検出プログラム1230がこれを検出すると、障害検出プログラム1230は予め登録された論理ボリュームへのダンプ出力処理に対応するデバイススイッチテーブルエントリ1212を参照し、そのエントリに登録されたプログラムのアドレスを取得する(S5720)。

10

【0116】

次に、障害検出プログラム1230は、ステップS5720で取得したアドレスのプログラムを呼び出す。これにより、デバイスドライバ8100のダンプ出力処理プログラム1320を呼び出し、実行する。ダンプ出力処理プログラム1320は、システム内のメモリダンプを論理ボリュームに出力する(S9215)。

ダンプ出力が終了すると、現用系1000は障害を復旧し、新たな待機系として立ち上がる。このとき、後述するステップS9250以降の処理を行う(S9220)。

20

【0117】

続いて、待機系2000のシステム立ち上げ処理が実行される。

この処理では、まず、待機系2000のシステム立ち上げ処理が行われる(S9250)。システム立ち上げが終了すると、系切替え制御プログラム6200は現用系1000の障害を監視するステップS9255。

【0118】

続いて、障害時における待機系2000系切替え処理が実行される。

待機系2000の系切替え制御プログラム6200は、現用系1000の障害発生を検出すると、整合性検査プログラム8200を呼び出し、論理ボリュームの整合性を検査する(S9255)。

30

【0119】

次に、系切替え制御プログラム6200から呼び出された整合性検査プログラム8200は、パラメータとして渡された論理ボリュームの列に基づき、論理ボリューム内のアクセス要求管理テーブルを読み出して整合性のチェックを行う。このとき、途中で中断された書き込みI/Oが存在する場合は、その論理ボリュームに対し復旧プログラムを実行する(S9260)。復旧プログラムは、その論理ボリュームに対応する物理ボリュームの一部または全部を他の物理ボリュームにコピーし、不整合な状態を修復する。

以上の処理が終了すると、系切替え制御プログラム8200は、業務引継ぎ処理を行い、新たな現用系として業務を再開する(S9265)。

40

【0120】

以上により、現用系1000に障害が発生した場合、待機系2000が業務を引き継ぐ前に共有ディスク3000の整合性をチェックし、不整合の状態であればこれを修復することができる。

【図面の簡単な説明】

【0121】

【図1】本発明の第一の実施形態によるホットスタンバイシステムの概略構成図である。

【図2】第一の実施形態における共有ディスク3000の構成を示す図である。

【図3】第一の実施形態におけるOS1200とデバイスドライバ1300の構成および相関を示す図である。

50

【図 4】第一の実施形態におけるアクセス要求監視プログラム 1 2 5 0 の構成を示す図である。

【図 5】第一の実施形態におけるアクセス要求管理テーブル 1 3 0 0 のエントリの構成を示す図である。

【図 6】第一の実施形態におけるアクセス管理対象論理ボリュームの登録プログラム 1 2 5 5 の処理の流れを示す図である。

【図 7】第一の実施形態におけるアクセス管理対象論理ボリュームの除外プログラム 1 2 6 0 の処理の流れを示す図である。

【図 8】第一の実施形態におけるアクセス要求監視状態の取得プログラム 1 2 6 5 の処理の流れを示す図である。

10

【図 9】第一の実施形態におけるアクセス要求記録プログラム 1 2 7 0 の処理の流れを示す図である。

【図 10】第一の実施形態におけるアクセス記録の送信プログラム 1 2 7 5 の処理の流れを示す図である。

【図 11】第一の実施形態における初期化プログラム 1 2 8 0 の処理の流れを示す図である。

【図 12】第一の実施形態における系切替え制御プログラム 1 1 0 0 の処理の流れを示す図である。

【図 13】第一の実施形態における整合性検査プログラム 1 1 5 0 の処理の流れを示す図である。

20

【図 14】第一の実施形態におけるホットスタンバイ処理の流れを示す図である。

【図 15】第二の実施形態によるホットスタンバイシステムの概略構成図である。

【図 16】第二の実施形態におけるアクセス要求監視プログラム 6 1 0 0 の構成を示す図である。

【図 17】第二の実施形態におけるアクセス要求通知パケット 6 4 0 0 の構成を示す図である。

【図 18】第二の実施形態におけるアクセス要求記録プログラム 6 1 1 0 の処理の流れを示す図である。

【図 19】第二の実施形態におけるアクセス要求の受信プログラム 6 1 2 0 の処理の流れを示す図である。

30

【図 20】第二の実施形態における系切替え制御プログラム 6 0 0 0 の処理の流れを示す図である。

【図 21】第二の実施形態における整合性検査プログラム 6 2 5 0 の処理の流れを示す図である。

【図 22】第二の実施形態におけるホットスタンバイ処理の流れを示す図である。

【図 23】第三の実施形態によるホットスタンバイシステムの概略構成図である。

【図 24】第三の実施形態における共有ディスク 8 4 0 0 の構成を示す図である。

【図 25】第三の実施形態におけるアクセス要求管理テーブルのエントリ 8 4 4 0 の構成を示す図である。

【図 26】第三の実施形態におけるデバイスドライバ 8 1 0 0 の構成を示す図である。

40

【図 27】第三の実施形態における書き込み処理プログラム 8 1 1 0 の処理の流れを示す図である。

【図 28】第三の実施形態における整合性検査プログラム 8 2 0 0 の処理の流れを示す図である。

【図 29】第三の実施形態におけるホットスタンバイ処理の流れを示す図である。

【符号の説明】

【 0 1 2 2 】

1 0 0 0 : 現用系計算機、 2 0 0 0 : 待機系計算機、

1 0 5 0、2 0 5 0 : アプリケーションプログラム、

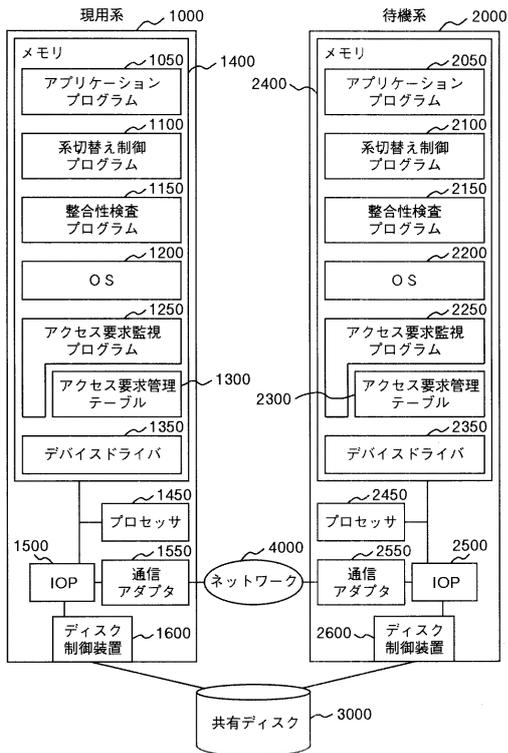
1 1 0 0、2 1 0 0 : 系切替え制御プログラム、

50

1 1 5 0、2 1 5 0 : O S、 1 2 0 0、2 2 0 0 : ア ク セ ス 要 求 制 御 プ ロ グ ラ ム
 、 1 3 0 0、2 3 0 0 : デ バ イ ス ド ラ イ バ、 1 3 5 0、2 3 5 0 : プ ロ セ ッ サ、
 1 4 0 0、2 4 0 0 : メ モ リ、 1 4 5 0、2 4 5 0 : プ ロ セ ッ サ、
 1 5 0 0、2 5 0 0 : 入 出 力 プ ロ セ ッ サ (I O P)、
 1 5 5 0、2 5 5 0 : デ ィ ス ク 制 御 装 置、 1 6 0 0、2 6 0 0 : 通 信 ア ダ プ タ、
 3 0 0 0 : 共 有 デ ィ ス ク、 4 0 0 0 : ネ ャ ト ワ ー ク

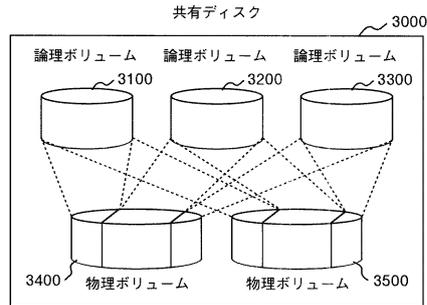
【 図 1 】

図 1



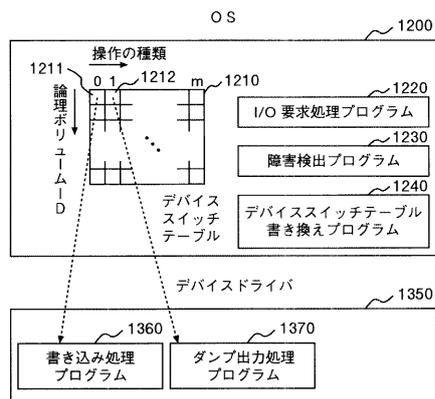
【 図 2 】

図 2



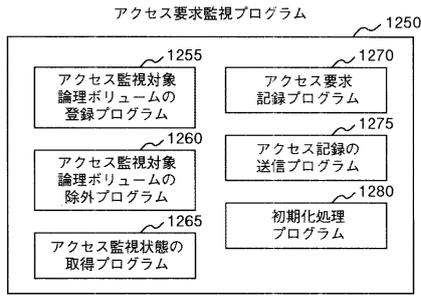
【 図 3 】

図 3



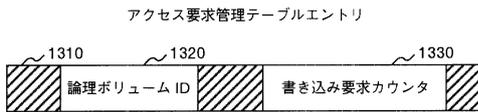
【図4】

図 4



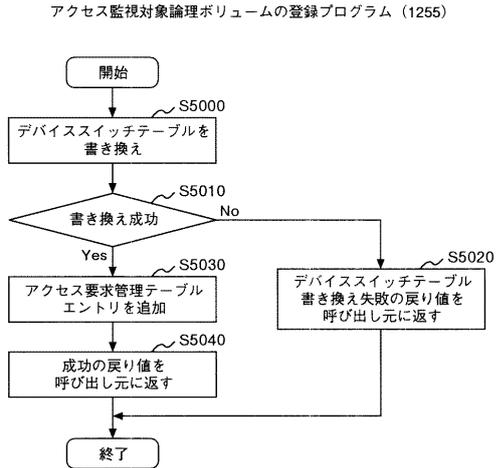
【図5】

図 5



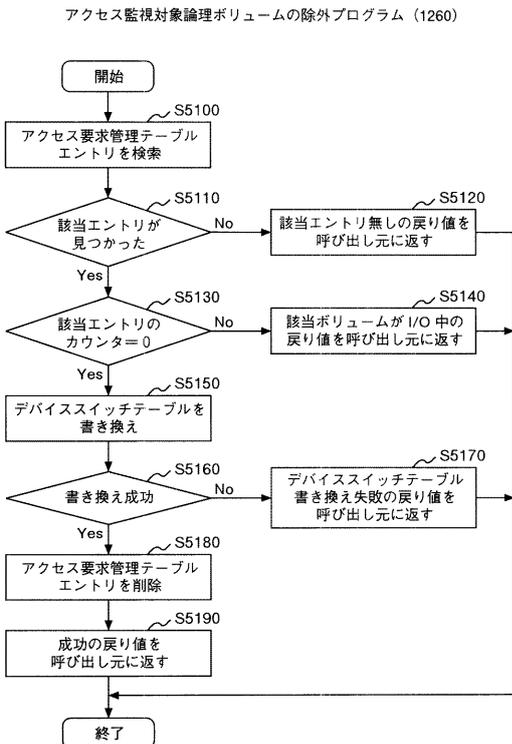
【図6】

図 6



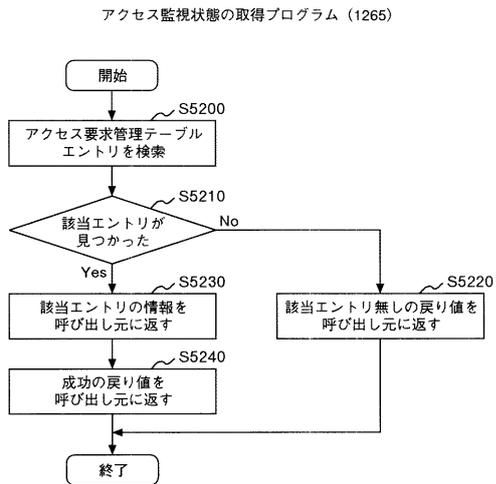
【図7】

図 7



【図8】

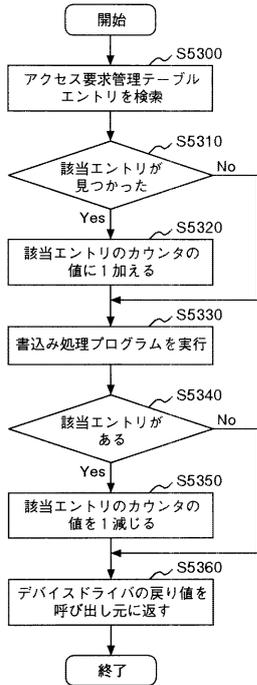
図 8



【図 9】

図 9

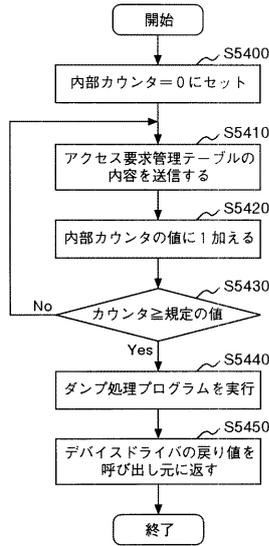
アクセス要求記録プログラム (1270)



【図 10】

図 10

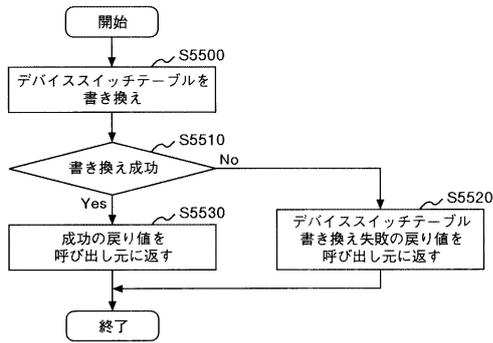
アクセス記録の送信プログラム (1275)



【図 11】

図 11

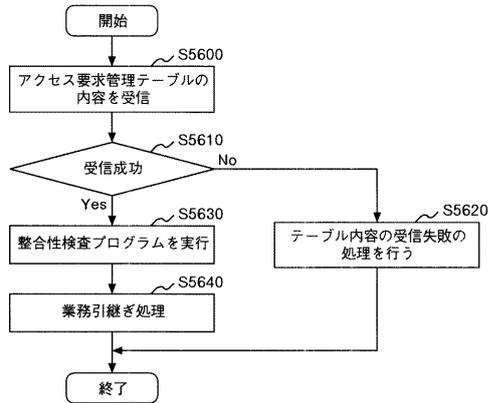
初期化処理プログラム (1280)



【図 12】

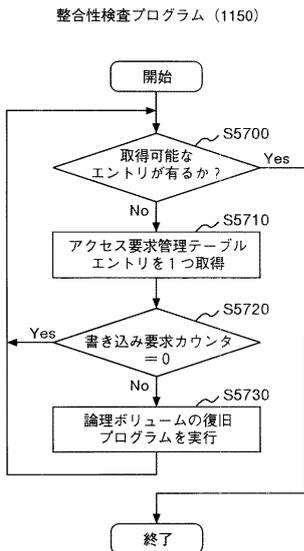
図 12

系切り替え制御プログラム (1100)



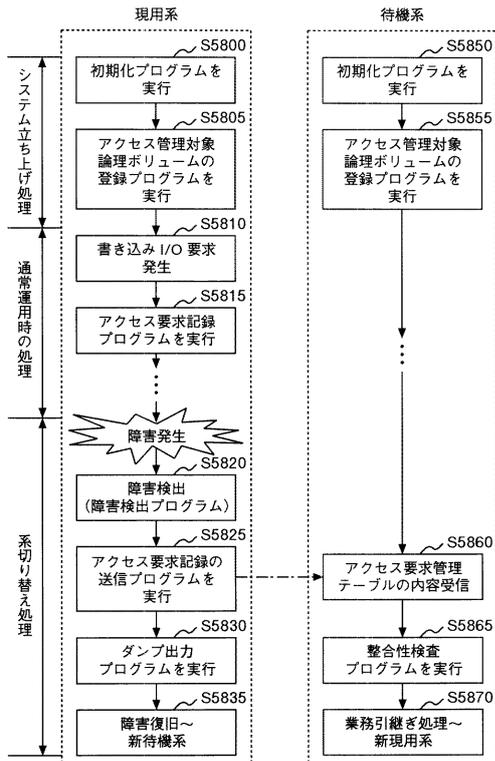
【図 13】

図 13



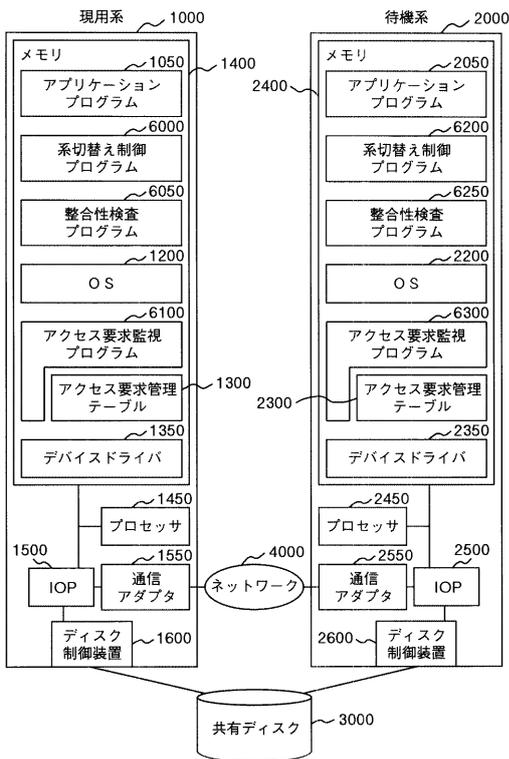
【図 14】

図 14



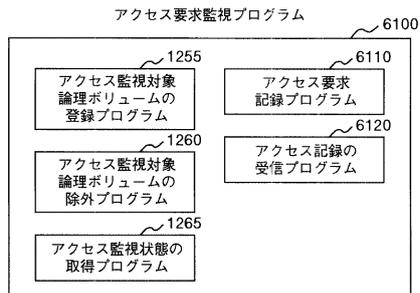
【図 15】

図 15



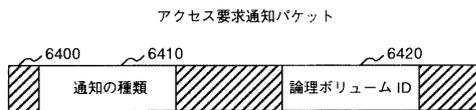
【図 16】

図 16



【図 17】

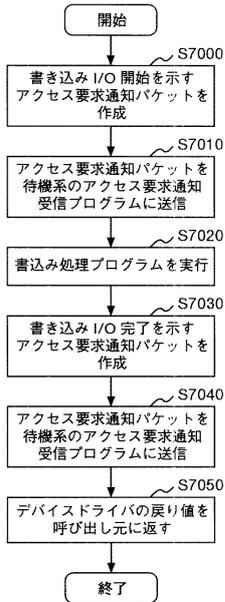
図 17



【図18】

図 18

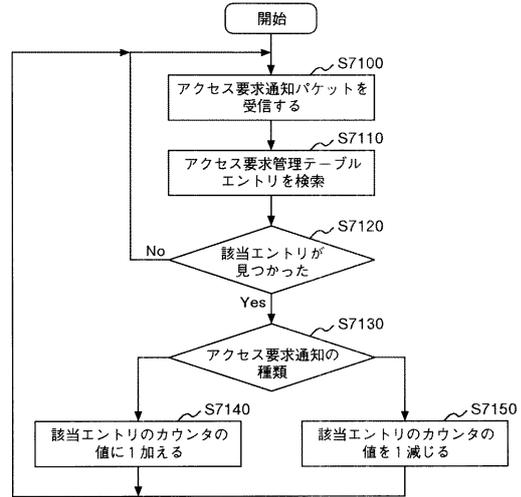
アクセス要求記録プログラム (6110)



【図19】

図 19

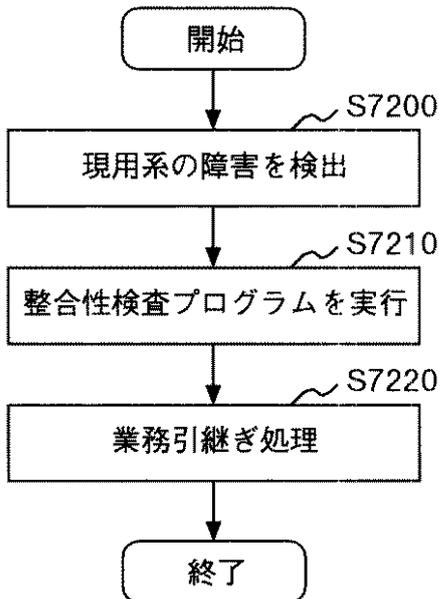
アクセス要求の受信プログラム (6120)



【図20】

図 20

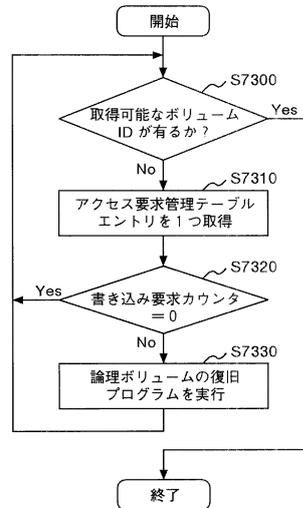
系切り替え制御プログラム (6000)



【図21】

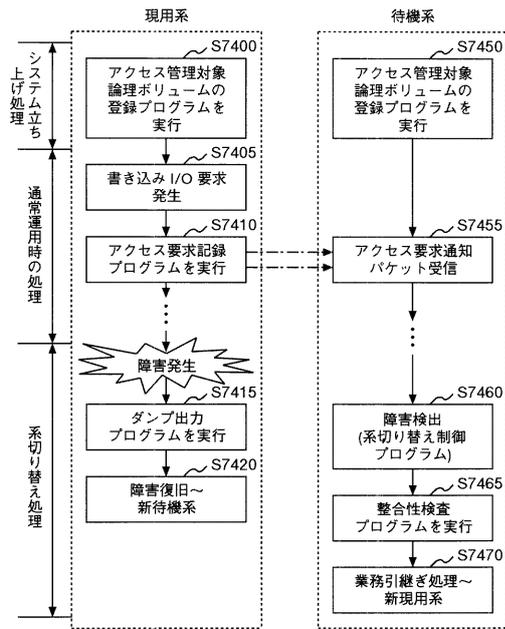
図 21

整合性検査プログラム (6250)



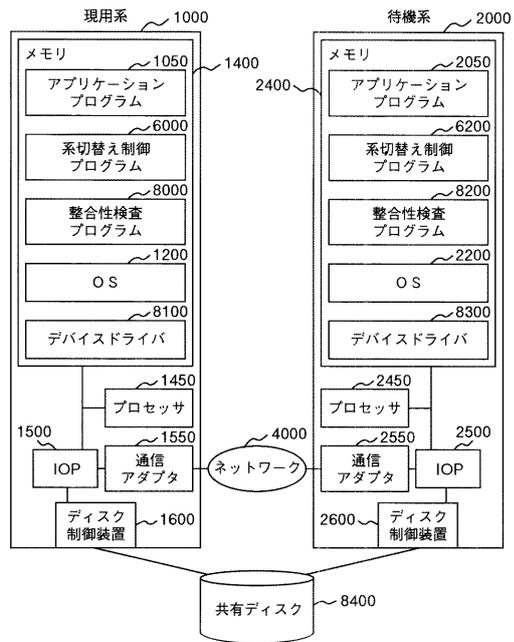
【図 2 2】

図 2 2



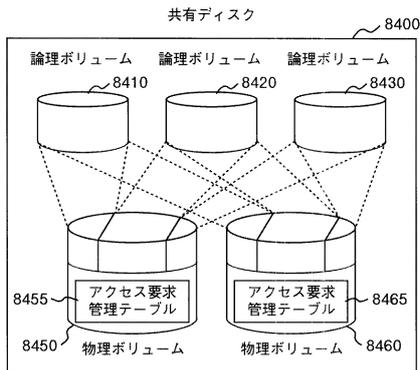
【図 2 3】

図 2 3



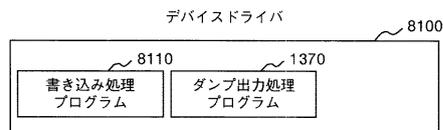
【図 2 4】

図 2 4



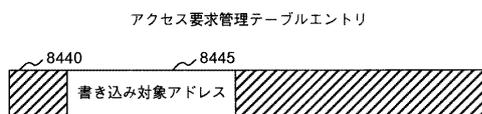
【図 2 6】

図 2 6



【図 2 5】

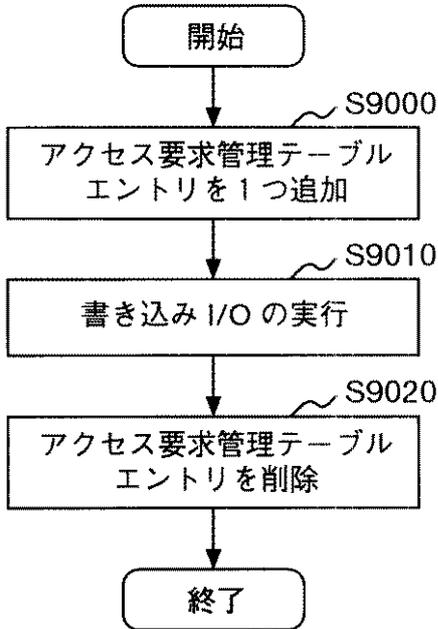
図 2 5



【図 27】

図 27

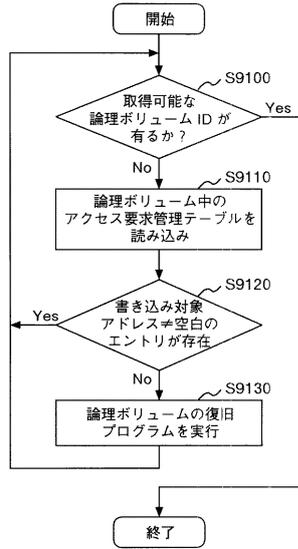
書き込み処理プログラム (8110)



【図 28】

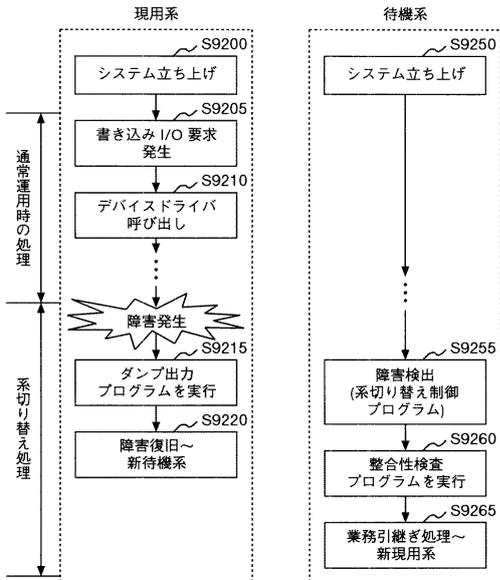
図 28

整合性検査プログラム (8200)



【図 29】

図 29



フロントページの続き

(72)発明者 真矢 讓

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

(72)発明者 三瓶 英智

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

審査官 久保 正典

(56)参考文献 特開平08-212095(JP,A)

特開平07-036721(JP,A)

特開平02-077943(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F11/16-11/20