

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3877519号
(P3877519)

(45) 発行日 平成19年2月7日(2007.2.7)

(24) 登録日 平成18年11月10日(2006.11.10)

(51) Int. Cl.

G06F 11/20 (2006.01)

F I

G06F 11/20 310C

請求項の数 7 (全 12 頁)

(21) 出願番号	特願2000-381623 (P2000-381623)	(73) 特許権者	000005108
(22) 出願日	平成12年12月15日(2000.12.15)		株式会社日立製作所
(65) 公開番号	特開2002-183088 (P2002-183088A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成14年6月28日(2002.6.28)	(73) 特許権者	000233055
審査請求日	平成15年8月12日(2003.8.12)		日立ソフトウェアエンジニアリング株式会社
			神奈川県横浜市鶴見区末広町一丁目1番43
		(74) 代理人	100083552
			弁理士 秋田 収喜
		(72) 発明者	岩本 孝寿
			神奈川県横浜市中区尾上町6丁目81番地
			日立ソフトウェアエンジニアリング株式会社内

最終頁に続く

(54) 【発明の名称】 システム回復方法およびその実施計算機システム並びにその処理プログラムを記録した記録媒体

(57) 【特許請求の範囲】

【請求項1】

実行システムで障害が発生した場合に業務処理を待機システムに切り替えて続行するシステム回復方法において、

実行システムで行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機システムに転送するステップと、

前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させる追跡処理を実施するステップと、

業務処理を実行中の実行システムの稼働状態を監視し、実行システムの障害を検知した場合に、前記追跡処理の行われた入出力バッファを用いて前記業務処理を待機システムで続行させるステップとを有し、

前記ログ情報として待機システムに転送される参照履歴は、実行システムの入出力バッファに存在していないデータへの参照処理の履歴を示すものであることを特徴とするシステム回復方法。

【請求項2】

実行システム及び待機システムで共用される記憶装置に前記ログ情報を出力するステップと、前記追跡処理の行われたログ情報と実行システムから転送されたログ情報とが連続していない場合に、その間のログ情報を前記記憶装置から読み出し、その読み出したログ情報及び前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させるステップとを有することを特徴とする

請求項 1 に記載されたシステム回復方法。

【請求項 3】

実行システムで障害が発生した場合に業務処理を待機システムに切り替えて続行する計算機システムにおいて、

実行システムで行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機システムに転送するログ転送処理部と、

前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させる追跡処理を実施する追跡処理部と、

業務処理を実行中の実行システムの稼動状態を監視し、実行システムの障害を検知した場合に、前記追跡処理の行われた入出力バッファを用いて前記業務処理を待機システムで続行させるモニタ処理部とを備え、

前記ログ情報として待機システムに転送される参照履歴は、実行システムの入出力バッファに存在していないデータへの参照処理の履歴を示すものであることを特徴とする計算機システム。

【請求項 4】

実行システムで障害が発生した場合に業務処理を待機システムに切り替えて続行するシステム回復方法を実現させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体において、

前記プログラムは、実行システムで行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機システムに転送するステップと、

前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させる追跡処理を実施するステップと、

業務処理を実行中の実行システムの稼動状態を監視し、実行システムの障害を検知した場合に、前記追跡処理の行われた入出力バッファを用いて前記業務処理を待機システムで続行させるステップとを有し、

前記ログ情報として待機システムに転送される参照履歴は、実行システムの入出力バッファに存在していないデータへの参照処理の履歴を示すものであるシステム回復方法を実現させるためのプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項 5】

実行システム及び待機システムで共用される記憶装置に前記ログ情報を出力するステップと、前記追跡処理の行われたログ情報と実行システムから転送されたログ情報とが連続していない場合に、その間のログ情報を前記記憶装置から読み出し、その読み出したログ情報及び前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させるステップとを有するシステム回復方法を実現させるためのプログラムを記録したことを特徴とする請求項 4 に記載されたコンピュータ読み取り可能な記録媒体。

【請求項 6】

実行システムで障害が発生した場合に業務処理を待機システムに切り替えて続行するシステム回復方法において、

実行システムで行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機システムに転送するステップと、

前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させる追跡処理を実施するステップとを有し、

前記ログ情報として待機システムに転送される参照履歴は、実行システムの入出力バッファに存在していないデータへの参照処理の履歴を示すものであることを特徴とするシステム回復方法。

【請求項 7】

実行システム及び待機システムで共用される記憶装置に前記ログ情報を出力するステップと、前記追跡処理の行われたログ情報と実行システムから転送されたログ情報とが連続

10

20

30

40

50

していない場合に、その間のログ情報を前記記憶装置から読み出し、その読み出したログ情報及び前記転送されたログ情報に従って、待機システム内の入出力バッファの内容を実行システム内の入出力バッファの内容に一致させるステップとを有することを特徴とする請求項6に記載されたシステム回復方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はシステム停止に伴う高速な回復処理を必要とするオンライン処理システムに関し、特に多量の更新トランザクションが発生するオンラインデータベースシステム等のオンライン処理システムに適用して有効な技術に関するものである。

10

【0002】

【従来の技術】

従来の一般的なオンラインシステムの回復方法では、実行オンラインシステムが障害によって停止した場合の回復処理の為に、システムの回復に必要な履歴情報であるログ情報を実行オンラインシステムで外部記憶装置上のログファイルに記憶し、実行オンラインシステムに障害が発生すると、待機オンラインシステムがそのログ情報を読み出してシステムの回復に必要な処理を実行していた。

【0003】

この高速化の技術として、例えば特開昭62-57030号公報に記載のものが知られており、その概要は、ホストコンピュータ間で共用する外部記憶装置上のログ情報を、障害の発生前から待機オンラインシステムで読み込み、待機オンラインシステムが実行オンラインシステムの停止以前にその処理を追跡することによって障害が発生した場合のログ情報の読み込み量を減らすものである。

20

【0004】

また、特開平2-77943号公報に記載の様に、ログ情報を実行オンラインシステムでホストコンピュータ間で共用する外部記憶装置上のログファイルに記憶すると共に、ホストコンピュータ間で共用する拡張記憶装置にも同時に記憶し、実行オンラインシステムに障害が発生した後の待機オンラインシステムでのシステム回復処理においては、この拡張記憶装置からログ情報を読み込むことで、外部記憶装置からのログ情報の読み込みを無くす方法が知られている。

30

【0005】

更に、特開平10-49418号公報に記載の様に、実行オンラインシステムのログ情報を通信により待機オンラインシステムに転送して、障害による切り替えが発生する前に待機オンラインシステムで追跡処理を実施することにより、障害が発生した後の回復を高速化する方法が知られている。

【0006】

【発明が解決しようとする課題】

しかしながら、特開昭62-57030号公報に記載の技術は、障害による待機オンラインシステムへの切り替えが発生した後に、チェックポイント以降のログ情報を外部記憶装置上のログファイルから入力する必要があり、チェックポイントの間隔によっては膨大なログ情報を読み込むことになって高速なシステム回復の実現を妨げる大きな要因となっていた。一方、障害による切り替えが発生した後の読み込むログ情報の量を減らす為にチェックポイントの間隔を狭めると、実行オンラインシステムでのオーバーヘッドを高めてしまう問題がある。

40

【0007】

また、特開平2-77943号公報に記載の技術は、ログ情報の読み込みは高速化しているものの、特開昭62-57030号公報に記載の技術と同様に、障害による待機オンラインシステムへの切り替えが発生した後に、チェックポイント以降のログ情報を読み込む必要があり、チェックポイントの間隔によっては膨大なログ情報を読み込むことになって高速なシステム回復の実現を妨げる大きな要因となっていた。一方、障害による切り替え

50

が発生した後の読み込むログ情報の量を減らす為にチェックポイントの間隔を狭めると、実行オンラインシステムでのオーバヘッドを高めてしまう問題がある。

【0008】

また、特開平10-49418号公報に記載の技術は、実行オンラインシステムのログ情報を通信により待機オンラインシステムへ転送して、障害による切り替えが発生する前に待機オンラインシステムでの追跡処理を実施しているが、更新履歴のログ情報のみを用いて追跡処理を行っている為、実行オンラインシステムで行われたインデクスへの参照等の参照処理の結果が待機オンラインシステムのメモリ上に反映されておらず、障害による切り替えが発生した場合にインデクス検索等の参照処理の効率が低下するという問題がある。更にログファイルやデータベース等の外部記憶装置を共用していない為、外部記憶装置が2倍必要となるという問題があり、また待機オンラインシステム側での障害等により一旦冗長構成が崩れると、冗長構成に戻す際の整合化の為にトランザクションの実行を一時中断する必要がある為、24時間365日連続運転が実現できないという問題がある。

10

【0009】

本発明の目的は上記問題を解決し、実行オンラインシステムの障害発生時に待機オンラインシステムへの切り替えを高速に行うことが可能な技術を提供することにある。

【0010】

本発明の他の目的は待機オンラインシステム内の入出力バッファの内容を実行オンラインシステム内の入出力バッファの内容に一致させる為のログ情報の転送負荷を軽減することが可能な技術を提供することにある。

20

【0011】

本発明の他の目的は待機オンラインシステムが障害や保守の後に再稼働した場合に、実行オンラインシステムでの業務処理の実行に影響を与えることなくホットスタンバイ状態の再確立を行うことが可能な技術を提供することにある。

【0012】

【課題を解決するための手段】

本発明は、実行オンラインシステムで障害が発生した場合に業務処理を待機オンラインシステムに切り替えて続行するオンライン処理システムにおいて、実行オンラインシステムで障害が発生した場合に、実行オンラインシステムの入出力バッファの内容に予め一致させておいた待機オンラインシステムの入出力バッファを用いて待機オンラインシステムで業務処理を続行させるものである。

30

【0013】

本発明では、実行オンラインシステムの稼働中に実行オンラインシステムで行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機オンラインシステムに転送し、前記ログ情報を受け取った待機オンラインシステムでは、その転送されたログ情報に従って、実行オンラインシステムの入出力バッファで行われた参照処理及び更新処理に相当する処理を待機オンラインシステムの入出力バッファに対して行い、待機オンラインシステム内の入出力バッファの内容を実行オンラインシステム内の入出力バッファの内容に一致させる追跡処理を実施する。

【0014】

また、業務処理を実行中の実行オンラインシステムの稼働状態を監視し、実行オンラインシステムの障害を検知した場合には、前記追跡処理の行われた入出力バッファを用いて前記業務処理を待機オンラインシステムで続行させる。

40

【0015】

前記の様に本発明によれば、障害による待機オンラインシステムへの切り替えが発生した後の外部記憶装置上のログファイルからのログ情報の入力を無くしつつ、実行オンラインシステムでのオーバヘッドが少ないオンラインシステムの回復方法を実現することができる。

【0016】

以上の様に本発明のオンライン処理システムによれば、実行オンラインシステムで障害が

50

発生した場合に、実行オンラインシステムの入出力バッファの内容に予め一致させておいた待機オンラインシステムの入出力バッファを用いて待機オンラインシステムで業務処理を続行させるので、実行オンラインシステムの障害発生時に待機オンラインシステムへの切り替えを高速に行うことが可能である。

【 0 0 1 7 】

【 発明の実施の形態 】

以下に実行オンラインシステムで障害が発生した場合に業務処理を待機オンラインシステムに切り替えて続行する一実施形態のオンライン処理システムについて説明する。

【 0 0 1 8 】

図 1 は本実施形態のオンライン処理システムの概略構成を示す図である。図 1 に示す様に本実施形態のホストコンピュータ 1 0 は、モニタ処理部 1 1 と、ログ出力処理部 1 5 と、ログ転送処理部 1 6 とを有している。

10

【 0 0 1 9 】

モニタ処理部 1 1 は、モニタ処理部 2 1 との間で相互監視の為の制御電文を交換し合うことにより相手システムの稼働状態を監視する処理部である。ログ出力処理部 1 5 は、ログ I / O バッファ 1 4 に格納されているログ情報を実行オンラインシステム 1 2 及び待機オンラインシステム 2 2 で共用される記憶装置に出力する処理部である。

【 0 0 2 0 】

ログ転送処理部 1 6 は、実行オンラインシステム 1 2 で行われた参照処理の履歴を示す参照履歴と更新処理の履歴を示す更新履歴とを表すログ情報を待機オンラインシステム 2 2 に転送する処理部である。

20

【 0 0 2 1 】

ホストコンピュータ 1 0 をモニタ処理部 1 1、ログ出力処理部 1 5 及びログ転送処理部 1 6 として機能させる為のプログラムは、C D - R O M 等の記録媒体に記録され磁気ディスク等に格納された後、メモリにロードされて実行されるものとする。なお前記プログラムを記録する記録媒体は C D - R O M 以外の他の記録媒体でも良い。

【 0 0 2 2 】

ホストコンピュータ 2 0 は、モニタ処理部 2 1 と、追跡処理部 2 7 とを有している。モニタ処理部 2 1 は、モニタ処理部 1 1 との間で相互監視の為の制御電文を交換し合うことにより、業務処理を実行中の実行オンラインシステム 1 2 の稼働状態を監視し、実行オンラインシステム 1 2 の障害を検知した場合に、追跡処理の行われたデータベース I / O バッファ 2 3 を用いて前記業務処理を待機オンラインシステム 2 2 で続行させる処理部である。

30

【 0 0 2 3 】

追跡処理部 2 7 は、前記転送されたログ情報に従って、待機オンラインシステム 2 2 内のデータベース I / O バッファ 2 3 の内容を実行オンラインシステム 1 2 内のデータベース I / O バッファ 1 3 の内容に一致させる追跡処理を実施する処理部である。

【 0 0 2 4 】

ホストコンピュータ 2 0 をモニタ処理部 2 1 及び追跡処理部 2 7 として機能させる為のプログラムは、C D - R O M 等の記録媒体に記録され磁気ディスク等に格納された後、メモリにロードされて実行されるものとする。なお前記プログラムを記録する記録媒体は C D - R O M 以外の他の記録媒体でも良い。

40

【 0 0 2 5 】

本実施形態のオンライン処理システムは、実行オンライン側のホストコンピュータ 1 0 と、実行オンライン側のモニタ処理部 1 1 と、実行オンライン側の実行オンラインシステム 1 2 (例えばデータベース管理システム)と、待機オンライン側のホストコンピュータ 2 0 と、待機オンライン側のモニタ処理部 2 1 と、待機オンライン側の待機オンラインシステム 2 2 (例えばデータベース管理システム)とを有している。

【 0 0 2 6 】

また、実行オンライン側の実行オンラインシステム 1 2 と待機オンライン側の待機オンラ

50

インシステム 2 2 とで共用する不揮発な記憶装置（一般には磁気ディスク装置）上にはログ情報 3 1 を格納するログファイル 3 0 やデータベース 4 0 を有している。

【 0 0 2 7 】

更に、本実施形態のオンライン処理システムは、実行オンラインシステム 1 2 がデータベース 4 0 とのレコードの入出力に使用するデータベース I / O バッファ 1 3 と、実行オンラインシステム 1 2 がログファイル 3 0 とのログ情報の入出力に使用するログ I / O バッファ 1 4 と、待機オンラインシステム 2 2 がデータベース 4 0 とのレコードの入出力に使用するデータベース I / O バッファ 2 3 と、待機オンラインシステム 2 2 がログファイル 3 0 とのログ情報の入出力に使用するログ I / O バッファ 2 4 とを有している。

【 0 0 2 8 】

また、実行オンラインシステム 1 2 がログ I / O バッファ 1 4 に格納されているログ情報をログファイル 3 0 に出力するログ出力処理部 1 5 と、実行オンラインシステム 1 2 がログ I / O バッファ 1 4 に格納されているログ情報を待機オンラインシステム 2 2 のログ情報受信バッファ 2 5 に転送するログ転送処理部 1 6 と、転送されたログ情報に従って、実行オンラインシステム 1 2 での業務処理と並行して待機システムでの追跡処理を実施する追跡処理部 2 7 と、モニタ処理部 1 1 とモニタ処理部 2 1 間で相互監視の為の制御電文（A l i v e 電文）を交換しあう為の通信媒体 5 0 と、実行オンラインシステム 1 2 から待機オンラインシステム 2 2 へのログ情報の転送の為の通信媒体 5 1 と、待機オンラインシステム 2 2 でログファイル 3 0 上のログ情報 3 1 を入力する為のログ I / O バッファ 2 4 とを有している。

【 0 0 2 9 】

ここで、通信媒体 5 0 と通信媒体 5 1 は物理的に一つでも良いが、ログ情報の転送トラフィックが高くなった場合の制御信号の電送遅延による誤動作を防ぐ為に、本実施形態では別々の通信媒体としている。

【 0 0 3 0 】

更に、データベース I / O バッファ 1 3、ログ I / O バッファ 1 4、データベース I / O バッファ 2 3、ログ I / O バッファ 2 4 やログ情報受信バッファ 2 5 は、それぞれ一つであっても良いが、性能・信頼性を確保する為に、それぞれ複数面のバッファを持ってバッファリングを行う。

【 0 0 3 1 】

また図 1 では実行オンラインシステム 1 2 にログ出力処理部 1 5 及びログ転送処理部 1 6 を示し、待機オンラインシステム 2 2 には追跡処理部 2 7 を示しているが、実行オンラインシステム 1 2 と待機オンラインシステム 2 2 とでは、実装している機能に違いはなく、実行系になるか待機系になるかによって振る舞いが変わるだけであるものとする。

【 0 0 3 2 】

よって、ホストコンピュータ 1 0 で障害が発生してトランザクションの実行権がホストコンピュータ 2 0 に切り替わり、待機オンラインシステム 2 2 がトランザクションサービスを開始したら、待機オンラインシステム 2 2 が実行系になり、ホストコンピュータ 1 0 の障害復旧後は、実行オンラインシステム 1 2 が待機系になる。

【 0 0 3 3 】

図 2 は本実施形態の実行オンラインシステム 1 2 及び待機オンラインシステム 2 2 の処理手順を示すフローチャートである。図 2 に示す様に本実施形態の実行オンラインシステム 1 2 は、起動されるとまず初期設定処理を行う（ステップ 1 2 2）。

【 0 0 3 4 】

この初期設定処理では、処理プログラムのロード、各種の定義情報や実行パラメタの取り込み、仮想記憶上の制御テーブルの作成、データベースのオープン、トランザクション実行空間（実行プロセスとも呼ぶ）の起動等を行い、更に、ログファイルの終端にあるログ情報の検出・記憶を行う。ここでは、更に待機オンラインシステム 2 2 とのログ情報転送に関わるバッファの確保・ページ固定・バッファの位置情報の交換等を行う。オンラインシステムでは、この他に端末との通信セッションの確立や切り替え準備等も行われるが、

10

20

30

40

50

本実施形態の範囲外である為、図 1 及び図 2 には示されていない。

【 0 0 3 5 】

一方、待機オンラインシステム 2 2 でも、待機システムとしての同様な初期設定を行う (ステップ 2 2 2)。この時点で、モニタ処理部 1 1 とモニタ処理部 2 1 による相互監視状態が開始される。

【 0 0 3 6 】

相互監視状態が開始されると、実行オンラインシステム 1 2 では業務トランザクションの処理を行う (ステップ 1 2 3)。

【 0 0 3 7 】

参照処理または更新処理を行う業務トランザクションで取得されたログ情報 1 2 4 が待機オンラインシステム 2 2 に転送されると、待機オンラインシステム 2 2 では、ログ情報 1 2 4 に従ってメモリ中でトランザクション状態やデータベースのレコードの参照及び更新状態を追跡する (ステップ 2 2 3)。このとき、ログファイル 3 0 やデータベース 4 0 は実行オンラインシステム 1 2 から更新される為、待機オンラインシステム 2 2 では、これらの外部記憶装置への書き込みは行わずに、データベースのインデックスの参照状態やレコードの更新状態の追跡さえも、メモリ中のデータベース I / O バッファ 2 3 上でだけ実施する。

10

【 0 0 3 8 】

実行オンラインシステム 1 2 に障害が発生すると (ステップ 1 2 5)、モニタ処理部 1 1 またはモニタ処理部 2 1 が障害を検知して、業務トランザクションの実行権を待機オンラインシステム 2 2 に切り替える (ステップ 1 2 6)。

20

【 0 0 3 9 】

実行オンラインシステム 1 2 だけの障害に留まっている場合は、モニタ処理部 1 1 が障害を検知してモニタ処理部 2 1 に通知する。ホストコンピュータ 1 0 全体に障害が及んでモニタ処理部 1 1 さえ正常に稼働できない場合は、モニタ処理部 1 1 からモニタ処理部 2 1 への制御電文 (A l i v e 電文) が途切れる為、モニタ処理部 2 1 が自発的に相手システムである実行オンラインシステム 1 2 の障害を検知する。

【 0 0 4 0 】

待機オンラインシステム 2 2 への切り替えが発生すると、未処理のログ情報 1 2 4 の追跡の完了を待ち合わせた後 (ステップ 2 2 4)、新たな業務トランザクションのサービスを開始する (ステップ 2 2 5) と共に、並行して未完了トランザクションのロールバック (後退復帰) を行う (ステップ 2 2 6)。

30

【 0 0 4 1 】

図 3 は本実施形態の業務トランザクション処理の処理手順を示すフローチャートである。図 3 を用いて、図 2 のステップ 1 2 3 の業務トランザクション処理を説明する。

【 0 0 4 2 】

トランザクションが開始されると、トランザクションの開始を示すログをログ I / O バッファ 1 4 にバッファリングする (ステップ 1 2 3 1)。次に、データベース I / O バッファ 1 3 上でレコードの参照や更新を行う (ステップ 1 2 3 2) と共に、レコードの参照ログや更新ログをログ I / O バッファ 1 4 にバッファリングする (ステップ 1 2 3 3)。一つのトランザクションでのデータベースレコードの参照や更新が完了すると、トランザクション終了ログをログ I / O バッファ 1 4 にバッファリングし (ステップ 1 2 3 4)、ログファイル 3 0 へ未出力のログ情報を強制的に出力する (ステップ 1 2 3 5)。

40

【 0 0 4 3 】

ここで、ステップ 1 2 3 2 において、データベース I / O バッファ 1 3 に存在していないデータへの参照が行われた場合に、ステップ 1 2 3 3 でその参照ログをログ I / O バッファ 1 4 にバッファリングするものとし、ログ情報の出力や転送に必要な負荷を軽減させても良い。

【 0 0 4 4 】

図 4 は本実施形態のログ情報のバッファリング処理の処理手順を示すフローチャートであ

50

る。図4を用いて、図3のステップ1231、ステップ1233及びステップ1234のログ情報のバッファリング処理を説明する。

【0045】

まず、現在のバッファリング先となっているログI/Oバッファに空きがあるか調べる(ステップ12311)。空きがあれば、該ログI/Oバッファにログ情報を格納する(ステップ12315)。

【0046】

空きがなければ、他のログI/Oバッファに空きがあるか調べる(ステップ12312)。ここで、空きが見つければ、そのログI/Oバッファを新たなバッファリング先として設定して(ステップ12314)、ログ情報を格納する(ステップ12315)。

10

【0047】

空きのログI/Oバッファが一つもない場合は、空きができるのを待ち合わせる(ステップ12313)。ここで、空きのログI/Oバッファが一つも無かった場合に、動的に新たなログI/Oバッファを確保する方法もあるが、メモリ不足を引き起こして障害の引き金になる可能性がある為、本実施形態では採用しない。

【0048】

図5は本実施形態の未出力ログの強制出力処理の処理手順を示すフローチャートである。図5を用いて、図3のステップ1235における未出力ログの強制出力処理を説明する。

【0049】

まず、現在のバッファリング先となっているログI/Oバッファを、“空き無し”の状態にして該ログI/Oバッファへの新たなバッファリングを抑止する(ステップ12351)。

20

【0050】

次に、ログファイル30へまだ出力していないログI/Oバッファを順次出力する(ステップ12352)。この出力は外部記憶装置へのI/O動作が完了するまで制御が戻らない同期書き込み方式でも、I/O動作の完了を待たない非同期書き込み方式でも良いが、本実施形態では、待機オンラインシステム22へのログ情報の転送処理が実行オンラインシステム12のトランザクションへ与える影響を無くす為に、非同期書き込み方式を採用する。

【0051】

ログファイル30への書き込みの完了を待っている間に、前記ステップ12352で対象にしたログI/Oバッファの内容を、通信媒体51を介して待機オンラインシステム22のログ情報受信バッファ25に直接書き込む(ステップ12353)。この書き込み位置等の情報は、前記のステップ122で示した初期設定時点及び前記のステップ123で示した前回の書き込み時の返信情報で把握しておくものとする。

30

【0052】

ここで、待機オンラインシステム22が稼動していない場合は、前記ステップ12353は失敗するが、実行オンラインシステム12側では成功したものとして扱う。この不整合状態は、待機オンラインシステム22に切り替える際に、その時点の最新のログ情報受信バッファ25のログまでの差分をログファイル30から読み込んで追い付かせることで解消させる。この処理によって、シーソーゲームの様に、実行系と待機系が頻繁に切り替わった場合であっても自動的に追従することができる。

40

【0053】

次に、前記ステップ12352のI/O動作の完了を待ち合わせる(ステップ12354)。前記ステップ12352及びステップ12353の両方が完了したログI/Oバッファを空きバッファとする(ステップ12355)。

【0054】

図6は本実施形態の追跡処理の処理手順を示すフローチャートである。図6を用いて、図2のステップ223の追跡処理を説明する。

【0055】

50

まず、待機オンラインシステム 22 の初期設定処理 222 時点で記憶したログファイルの終端にあるログ情報と、ログ情報受信バッファ 25 に送られてきたログ情報を比べる（ステップ 22301）。

【0056】

ログ情報が不連続（具体的には、ログファイルの世代番号及びログレコードの集合体であるログブロックの通番からなる番号が不連続であり、途中のブロックが抜けている状態）であれば、ログファイル 30 からログ情報 31 を入力して、ログ情報受信バッファ 25 の時点まで追いつかせる（ステップ 22302）。ここでの追いつき処理の具体的な方法は、後述のステップ 22303～ステップ 22308 と同様である。

【0057】

次に、ログ情報受信バッファ 25 に格納された個々のログ情報を順次調べ、トランザクションの開始ログや終了ログの様にトランザクションの状態の変更を記録したログであれば（ステップ 22303）、メモリ中のトランザクション毎の管理情報を更新する（ステップ 22304）。

【0058】

データベースのレコードの参照ログまたは更新ログであれば（ステップ 22305）、データベース I/O バッファ 23 に該当するページがあるかを調べ（ステップ 22306）、データベース I/O バッファ 23 に該レコードのページがなければデータベース 40 からデータベース I/O バッファ 23 に該レコードのページを読み込む（ステップ 22307）。次に、該ログが更新ログの場合にはその内容に従ってデータベース I/O バッファ 23 上でレコードを更新する（ステップ 22308）。

【0059】

これらステップ 22303～ステップ 22308 の処理を、ログ情報受信バッファ 25 にある全てのログ情報に対して繰り返す（ステップ 22309）。

【0060】

次に、モニタ処理部 11 またはモニタ処理部 21 から障害の検知が通知されているかを確認し、自システムが待機システムのままかどうかを調べる（ステップ 22310）。待機システムのままであれば、ログ情報の受信を待つ（ステップ 22313）、前記ステップ 22303～ステップ 22308 の処理を繰り返す。一方、モニタ処理部 11 またはモニタ処理部 21 による障害の検知により実行システムへの切り替えが指示されていれば、実行システムとして業務トランザクション処理を実行する。

【0061】

以上説明した様に本実施形態のオンライン処理システムによれば、実行オンラインシステムで障害が発生した場合に、実行オンラインシステムの入出力バッファの内容に予め一致させておいた待機オンラインシステムの入出力バッファを用いて待機オンラインシステムで業務処理を続行させるので、実行オンラインシステムの障害発生時に待機オンラインシステムへの切り替えを高速に行うことが可能である。

【0062】

また本実施形態のオンライン処理システムによれば、実行オンラインシステムの入出力バッファに存在していないデータへの参照処理が行われた場合にその参照履歴をログ情報として待機オンラインシステムに転送するので、待機オンラインシステム内の入出力バッファの内容を実行オンラインシステム内の入出力バッファの内容に一致させる為のログ情報の転送負荷を軽減することが可能である。

【0063】

また本実施形態のオンライン処理システムによれば、追跡処理の行われたログ情報と実行オンラインシステムから転送されたログ情報とが連続していない場合に、その間のログ情報を記憶装置から読み出して、待機オンラインシステム内の入出力バッファの追付き処理を行うので、待機オンラインシステムが障害や保守の後に再稼働した場合に、実行オンラインシステムでの業務処理の実行に影響を与えることなくホットスタンバイ状態の再確立を行うことが可能である。

10

20

30

40

50

【0064】

【発明の効果】

本発明によれば実行オンラインシステムで障害が発生した場合に、実行オンラインシステムの入出力バッファの内容に予め一致させておいた待機オンラインシステムの入出力バッファを用いて待機オンラインシステムで業務処理を続行させるので、実行オンラインシステムの障害発生時に待機オンラインシステムへの切り替えを高速に行うことが可能である。

【図面の簡単な説明】

【図1】本実施形態のオンライン処理システムの概略構成を示す図である。

【図2】本実施形態の実行オンラインシステム12及び待機オンラインシステム22の処理手順を示すフローチャートである。

10

【図3】本実施形態の業務トランザクション処理の処理手順を示すフローチャートである。

【図4】本実施形態のログ情報のバッファリング処理の処理手順を示すフローチャートである。

【図5】本実施形態の未出力ログの強制出力処理の処理手順を示すフローチャートである。

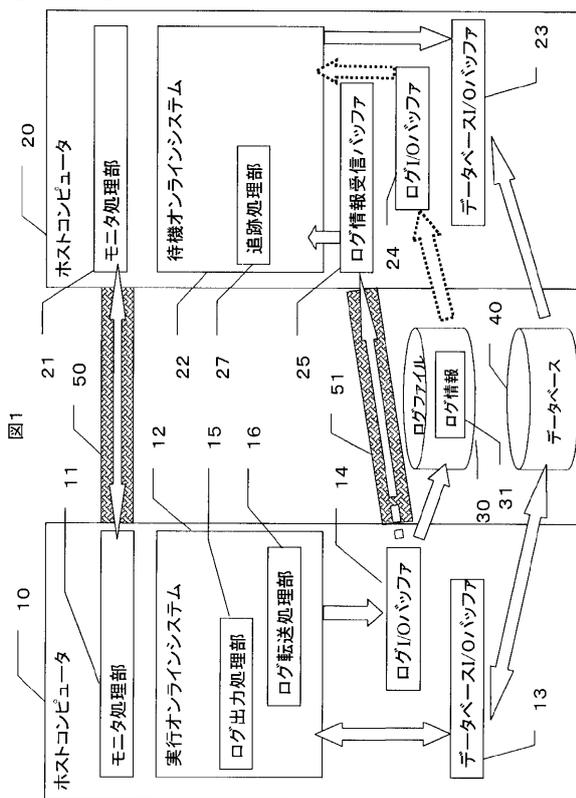
【図6】本実施形態の追跡処理の処理手順を示すフローチャートである。

【符号の説明】

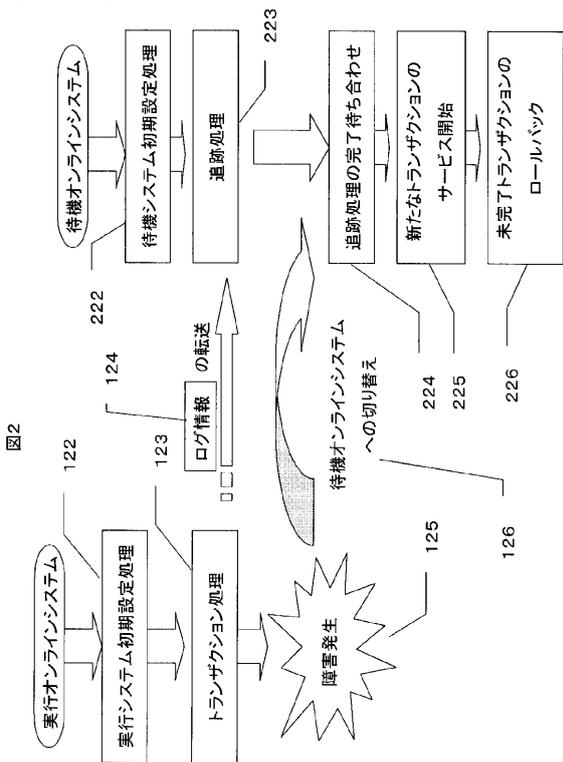
10...ホストコンピュータ、12...実行オンラインシステム、13...データベースI/Oバッファ、14...ログI/Oバッファ、20...ホストコンピュータ、22...待機オンラインシステム、23...データベースI/Oバッファ、24...ログI/Oバッファ、25...ログ情報受信バッファ、30...ログファイル、31...ログ情報、40...データベース、50及び51...通信媒体、11...モニタ処理部、15...ログ出力処理部、16...ログ転送処理部、21...モニタ処理部、27...追跡処理部、28...ログ情報受信バッファ、29...ログI/Oバッファ、32...データベース、41...データベース、52...通信媒体、17...ログ出力処理部、18...ログ転送処理部、23...データベースI/Oバッファ、24...ログI/Oバッファ、25...ログ情報受信バッファ、30...ログファイル、31...ログ情報、40...データベース、50及び51...通信媒体、11...モニタ処理部、15...ログ出力処理部、16...ログ転送処理部、21...モニタ処理部、27...追跡処理部、124...ログ情報。

20

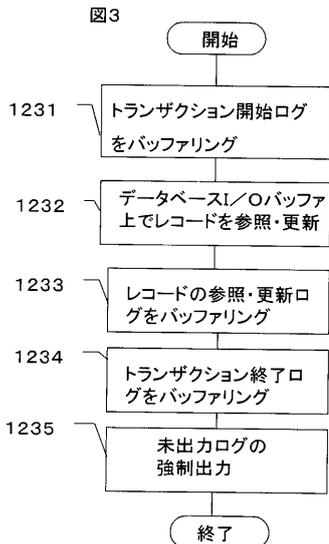
【図1】



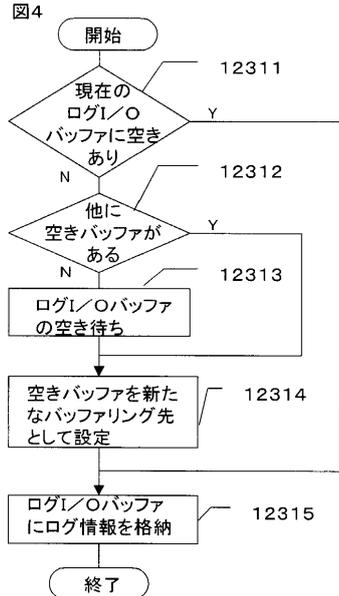
【図2】



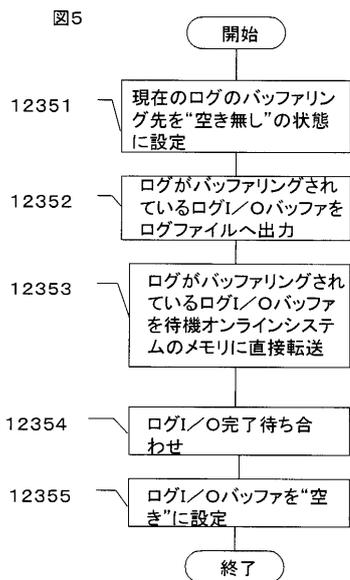
【 図 3 】



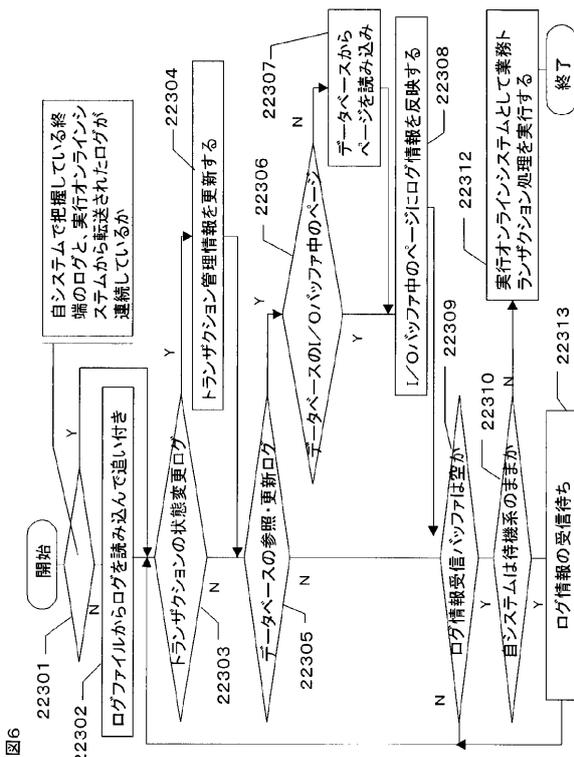
【 図 4 】



【 図 5 】



【 図 6 】



フロントページの続き

審査官 久保 正典

(56)参考文献 特開平05 - 216697 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 11/20