4L−12

ミッドレンジコンピュータでのOLTP高信頼化手法

西　敦子　野崎正治　滝本秀明　井口宗樹
(株)東芝　府中工場

1．はじめに

情報システムの利用範囲の拡大とともに、信頼性向上が大きな課題となっている。
ミニコンベースのオンラインシステムでは、価格性能比の面からも信頼性向上の手法としてホットスタンバイシステムを構築するのが主流である。
ミッドレンジコンピュータＤＳ６５００シリーズのトランザクション管理システムでは、ホットスタンバイシステムにおいて障害発生時の高速な運転継続の実現をおこない他の計算機との差別化を図っている。

2．ホットスタンバイ制御の開発方針

高信頼性のオンラインシステムを構築するための手法の１つとして系の多重化がある。多重化の制御には、両系で全く同じ処理を行い片系障害時、dead time=0で待機系が引き継ぐ方法と待機系が稼動系を監視しながら稼動系障害時最小のdead　time　で運転を継続する方法がある。

本論文は、ＤＳ６５００のトランザクション処理システムで高信頼性オンラインシステムを構築するために後者の方式を採用し、オンラインホットスタンバイ制御を実現したので報告する。

［ホットスタンバイ方式とした理由］

●系切り替え高速化によりdead timeを極力小さくし、端末と連携して無停止運転を実現する。
●待機系でバッチ処理、プログラム開発等のバックグラウンド処理ができる。

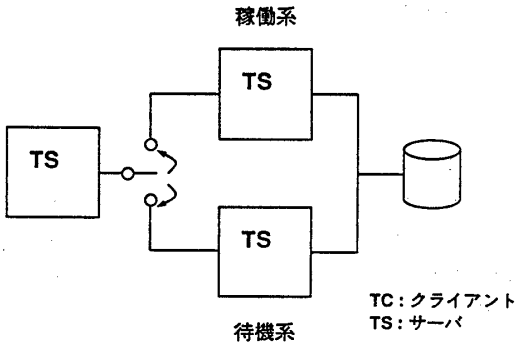

図１　ホットスタンバイ構成

3．ホットスタンバイシステムの構成と機能

ホットスタンバイシステムの構成を図１に示す。ホットスタンバイシステムは次の要素より構成される。

●共有メモリで結合された複合系システム
●複合系システムの両系からアクセスできるマルチコールディスク。
●マルチコールディスク上に配置されたシステムファイル（システムステータスログファイル、ジャーナルファイル）。

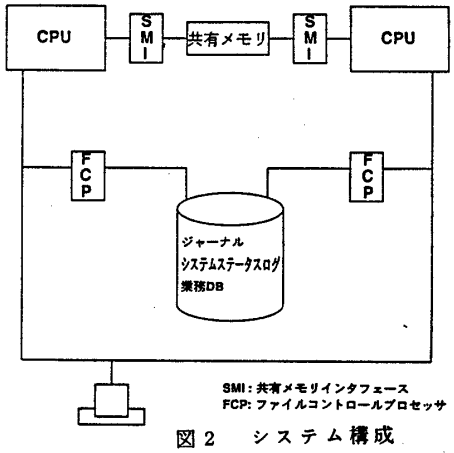

図２　システム構成

待機系システムでは、予めオンラインシステム運転に必要となる資源の生成を行なった後、稼動系システムの監視を開始する。稼動系の障害を検出すると待機系システムではマルチコールボリュームの切り替えを行ない、端末等の組み込みやシステムステータスログを用いた運転状態のリカバリ、ジャーナルを用いたデータベース（ＤＢと略す）とメッセージのリカバリを行ないオンラインシステムの再開を行なう。

4．切り替え制御の高速化

高速切り替え処理のポイントは

●共有メモリを用いた高速な障害診断
●ＤＢリカバリの高速化
●マルチプロセッサを用いた処理の高速化

待機システムに高速に制御を移すための両系運転監視には、共有メモリを利用して実現する。計算機障害は共有メモリインターフェースが検出し待機システムに緊急割り込みを発生することにより実現する。待機予システムでは、オンライン運転に必要となるディスクボリュームの組み込みや端末の組み込み処理（系切り替え、初期化処理）を並列して処理した後、システムステータスログからの運転状態リカバリ、ジャーナルデータによるＤＢとメッセージのリカバリを行いシステムを再開する。最もリカバリ時間のかかるＤＢについては、オンラインで使用するＤＢは新規作成や削除が発生しないのに着目し、障害によるボリュームの

High Reliabirity Methods for OLTP on middle range Computers.
Atsuko Nishi. Masaharu Nozaki
Hideaki Takimoto. Muneki Iguchi.
TOSHIBA CORPORATION.

チェック処理を行なわずジャーナルによるロールフォワードカバリのみを行なうこととした。ロールフォワード処理によるリカバリはデータバッファ上のデータの復旧のみで完了とするためリカバリ処理は高速におこなうことができる。更に、リカバリ処理は端末へのメッセージのリカバリとＤＢのリカバリを並列で処理することにより、高速化を図っている。

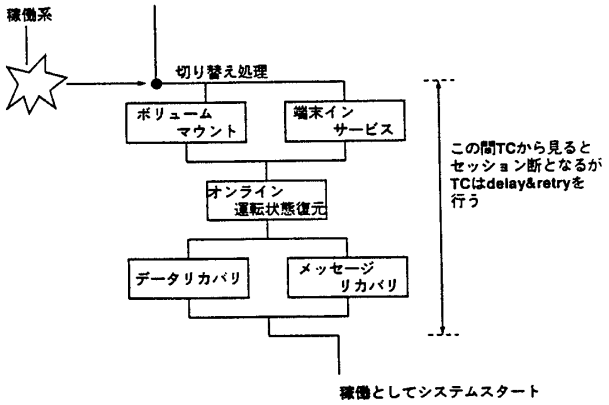

図３　系切り替え処理の流れ

５．両系切り替え制御

稼動システムから待機システムへのシステムの切り替えは、システム障害により自動的に運転を変更する場合と、システムの保守等によりオペレータの指示により運転を変更する場合の２種類がある。

障害は、表１のレベルでトランザクション処理が検出を行いシステム障害のときのみ予備機への運転の切り替えを行なう。トランザクションの運転監視によりトランザクションの停滞を検出した場合には、稼動システムは自分からシステム停止する。更にこのトランザクションの停滞監視機能を予備システムで定期的に監視し不具合を検出すると待機システムから稼動システムを強制的に停止させる。

表１　障害処理

障害の種類	検出方法	障害処理
ＤＢ　通信 端末	ＯＳ	障害資源のみ閉塞して運転を継続する。 フォールバック運転
業務プログラム	割り込み処理 LAPSE, CPU時間監視	トランザクションキャンセル 必要なら業務を閉塞する。
システム （ﾃﾞｰﾀ破壊 ｽﾄｰﾙ等、ﾊｰﾄﾞ）	ＯＳ ﾊｰﾄﾞｳｪｱ	システムダウンを発生させ系を切り替える。
OLTPｼｽﾃﾑﾌｧｲﾙ	ＯＬＴＰ	オンラインシステム運転継続できない場合稼働系が停止し自動的に系が切り替わる。
トランザクション処理停滞	ＯＬＴＰ停滞監視処理	ﾄﾗﾝｻﾞｸｼｮﾝの系内滞留時間が異常に増えた場合自系を停止し系を切り替える。
ＯＬＴＰ停滞監視処理の異常	待機系から稼動系周期監視	待機系から稼働系を強制的に停止させ、系の切り替えを行う。

オペレータの指示により切り替え処理を行なう時には、両系が決して同時に稼動システムとならないために（両系とも稼動中のときは）、系間通信処理でかならず自系の動作を相手系に通知するとともに相手系の許可を得てから運転切り替えを行なうという切り替え打診処理を行ない、複合系システムに矛盾を発生させない。

６．単体系システムとホットスタンバイシステム

単体系システムでは、次にあげる無停止運用機能を実現し稼動率99.5%を実現している。

●高信頼性ディスク
ディスクの２重化機能とｺﾝﾄﾛｰﾗからﾃﾞｨｽｸへの経路の２重化と動的経路選択。片系障害時の無停止ﾒﾝﾃﾅﾝｽ機能

●プロセッサ縮退機能
最大４台までのﾏﾙﾁﾌﾟﾛｾｯｻで構成でき、ﾌﾟﾛｾｯｻが故障すれば自動的に切り離して縮退運転を行なう。

●オンライン構成制御
オンライン運転中に端末、ＤＢ、業務プログラムの追加、削除、変更などの構成制御を行なうことができる

●運転自動化機能
システムの障害要因となるｵﾍﾟﾚｰﾀの操作を業務ﾌﾟﾛｸﾞﾗﾑが代行して行える自動運転支援環境

更に、ホットスタンバイシステムを利用することにより99.999%の稼動率となる。単体系システムからホットスタンバイシステムへの移行に際しては、オンライン業務プログラムは何も変更する必要はない。

６．おわりに

本論文では、稼動／待機で構成される２台系のホットスタンバイシステムの実現を報告した。

参考文献

[1]森　他:ﾐﾆｺﾝﾋﾟｭｰﾀでのｼﾞｬｰﾅﾙﾘｶﾊﾞﾘ高速化手法(1)
[2]松井　他:ﾐﾆｺﾝﾋﾟｭｰﾀでのｼﾞｬｰﾅﾙﾘｶﾊﾞﾘ高速化手法(2)
[3]鈴木　他:ﾐﾆｺﾝﾋﾟｭｰﾀでのｼﾞｬｰﾅﾙﾘｶﾊﾞﾘ高速化手法(3)
[4]松井　他:ﾐﾆｺﾝﾋﾟｭｰﾀでのｼﾞｬｰﾅﾙﾘｶﾊﾞﾘ高速化手法(4)
[5]広兼　他:ﾐﾆｺﾝﾋﾟｭｰﾀでのｼﾞｬｰﾅﾙﾘｶﾊﾞﾘ高速化手法(5)
以上情報処理第４１回全国大会(1990)
[6]滝本　他:高速ＯＬＴＰ環境を実現した新技術
東芝ﾚﾋﾞｭｰ11月号
[7]香川　他:ｵﾝﾗｲﾝｼｽﾃﾑのﾒｯｾｰｼﾞﾘｶﾊﾞﾘ手法
情報処理第４２回全国大会(1990)