

UNIXにおける 5G-1 ホットスタンバイシステムの構築

石田 太門* , 香川 弘一*
* (株) 東芝 府中工場

1. はじめに

オンライントランザクション処理 (OLTP) システムは企業の基幹業務をはじめ様々な分野で利用されている。しかし、システムダウンによるサービスの停止が発生した場合は、重大な損害を発生させることとなる。したがって、システムダウンを防ぎ、連続稼働性を保証する機能が重要視される。UNIX計算機はハードウェア的な信頼性の向上に加え、ソフトウェア的にも High Availability技術によって、基幹業務のサーバとしても利用されてきている。

当社のUNIX*計算機 (UXシリーズ) 向けの複合系サポートソフトウェア (以下「HAソフト」) を使用し、当社独自のTPモニターであるTX/ATPSにホットスタンバイシステムを実現したので報告する。

本システムの特長を以下に示す。

- クライアントはサーバのシステム障害を意識する必要がない
- 単体システムからホットスタンバイシステムへの移行でサーバ側・クライアント側双方の業務処理プログラムの変更が不要
- 相互スタンバイシステムの構築
- 障害からの高速な回復

2. 要素技術

サーバ障害をクライアントへ隠蔽するために必要な要素技術を以下に述べる。

(1) メッセージリカバリ処理

トランザクションメッセージは、ジャーナルファイル (JNF) 中のメッセージジャーナルまたは、物理ディスク上のファイルキュー (MRF, SMRF) 中に格納される。障害復旧処理では、それぞれの情報からメッセージのリカバリ処理を行う。またファイルキューの制御部の情報もファイルに記録されているため、システム障害発生後も高速にファイルキューの復元処理を行うことができる。ただしこの制御部の書き出しは、即座に行われないため、システム障害発生時にファイルキューの

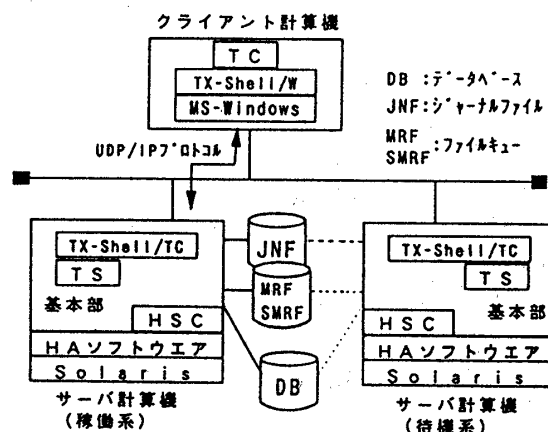
制御情報が失われる可能性がある。この場合は、タイムスタンプの比較によりファイル中の情報が正当性を判定し、不整合の場合は格納されているトランザクションメッセージから制御部の復元を行う。

(2) メッセージ通番制御

各宛先毎にメッセージ通番器を持ち、通番をTXメッセージに付加し、メッセージの送受信やリカバリ、ファイルキュー上での順序性を保証する。特に、分散トランザクション制御では、ネットワークをまたがったTXメッセージの重送や脱送を防止する。

(3) 回線再接続処理

サーバとの接続時にサーバのホットスタンバイシステムの有無を確認する。ネットワーク回線の切断が発生した際に、サーバがホットスタンバイ構成であれば、回線の再接続を試行し、メッセージの再送受信を行う。この処理はクライアントプログラム開発のためのAPIに組み込まれているために、開発者およびユーザはサーバの障害を意識する必要がない。



第1図 全体構成図

3. 処理の流れ

●クライアントへのサーバ障害の隠蔽

前節で述べたように、TX/ATPSでは端末への入出力メッセージのリカバリ処理を行うことができる。ホットスタンバイシステム構成時には、サーバ計算機の障害をクライアントに隠蔽し、処理を継続しているかのように見せるための処理は以下に行われ

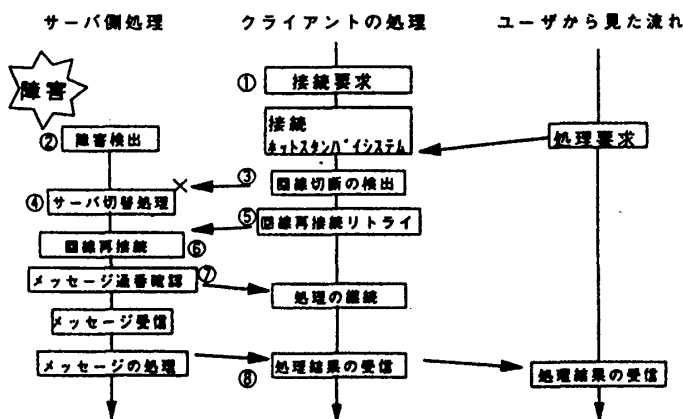
An Implementation of Hot-Standby System for OLTP of UNIX.

Tamonn Ishida, Koichi Kagawa, TOSHIBA CORP.

る。

- ①：クライアントはサーバとの接続時にサーバがホットスタンバイシステムかどうかの情報を得る。
- ②：待機系計算機は稼働系計算機の状態をHeatbeatにより監視しており、稼働系計算機に障害が発生した場合、障害の検出を行う。
- ③：クライアントはメッセージを送信するが、サーバ計算機は障害（切り替え処理中）のために応答を返さない。クライアントは回線（サーバ）障害の発生を認識する。
- ④：障害検出後、待機系計算機は切り替え処理を開始し、OLTPシステムおよびサーバプログラムの起動を行う。切り替えの終了まで、ネットワーク回線（IPアドレス）は停止されている。
- ⑤：クライアント側では回線の再接続を試みる。
- ⑥：サーバの切り替え終了後、ネットワーク回線が使用可能となり、サーバクライアント間の回線が再確立される。
- ⑦：回線の再確立の際にサーバは、クライアントからのメッセージ通番情報を受け取り、メッセージの順序性を損なうことのないよう処理を再開する。
- ⑧：クライアントは、サーバから送られたメッセージに従って、処理を続ける。

第2図にサーバ側、クライアント側およびユーザから見た処理の流れを示す。縦軸方向の矢印は時間の流れを示し、横方向の矢印はメッセージの流れを示す。



第2図 処理の流れ

●障害からの高速な回復

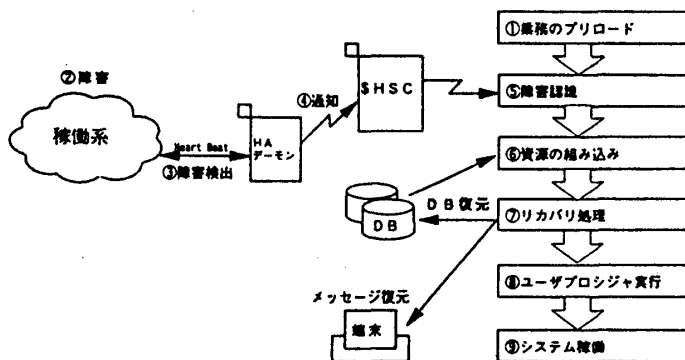
HAソフトだけで構成したホットスタンバイシステムでは、稼働系の障害を認識した後に初めて、待機系でOLTPシステムが起動される。障害回復時間短縮のために、待機系でのOLTPシステムの起動処理の

うち、待機中に実行可能な処理をあらかじめ行っておく。この方式によって、系の切り換え時には共有資源（ディスクなど）の確保などの処理のみでOLTPシステムを再開することができる。

待機系でのTX/ATPS起動・障害時の切り替え処理とをコントロールするために、ホットスタンバイシステムをコントロールするプロセス（\$HSC）を作成し、待機系ではTX/ATPSをいわば起動途中の状態にとどめておくようにする。

障害発生時のリカバリ処理の流れを第3図に示す。

また、リカバリ処理対象のファイルはユーザのデータベースなど最少限とし、OLTPシステムの起動に必要なシステムファイルは同じものを両計算機に配置することで、リカバリ処理時間の減少を行う



第3図 リカバリ処理の流れ

4. まとめ

アプリケーションの開発者やユーザにシステムの障害を意識させる必要はない。High Availability 技術によってシステムの耐障害性を向上させることができるが、障害発生時の処理をOLTPシステムに持たせることで、開発もより容易に行うことができる。さらに、アプリケーションユーザも通常の処理のみに意識を集中させることができる。

5. 参考文献

- [1] 香川、滝本、関田：分散OLTPシステムにおけるデフォルト・オンライン制御、情報処理学会第49回全国大会（1994）

*UNIXはX/Openカンパニーリミテッドがライセンスしている米国ならびに他の国における登録商標です。