分散処理技術を用いた高信頼通信システムの一構成法

7D－5

藤長 昌彦　加藤 聰彦　鈴木 健二
国際電信電話株式会社 研究所

1　はじめに

筆者らはこれまで、クライアントサーバ・モデルに基づくRPC等の分散処理技術を用いて複数の計算機を結合し、その上にメッセージ通信処理システムやINの付加サービス処理システム等の通信システムを構築することを検討してきた[1, 2]。このような通信システム(以下、分散型通信システムと呼ぶ)では、複数のサーバが相互に協調して、メッセージ交換やIN付加サービスの実行等のひとまとまりの通信処理を行なう。

分散型通信システムの信頼性を高めるためには、ソフトウェアのバグやハードウェアの故障等の原因で一部のサーバに障害が発生した場合にその回復を試み、回復できない場合にはそのサーバをシステムから切り離して、残りのサーバによりシステム全体としての動作を継続する必要がある。本稿では、分散型通信システムを高信頼化することを目的に、障害に対処するための特別なサーバや、サーバのレプリケーション等の分散処理技術を用いて、障害回復機能やシステムの動的な再構築機能を実現するための一構成法について述べる。

2　構成法の概要

高信頼な分散型通信システムを構成する方法の概要を以下に示す。

- ここでは、ひとまとまりの通信処理を行なう複数のサーバを抽象化して通信処理サーバと呼ぶ。図1に示すように、通信処理サーバのレプリケーションを行ない、通信処理サーバに加えてログサーバ[3]とシステム監視サーバを導入する。
- ログサーバは、二重化ディスク等を利用して、障害発生時にもデータを紛失することなく、安全に保持する機能を提供する。通信処理サーバはその処理の区切りにおいて障害からの回復に必要な、受信データや処理の進行状況等の情報をログサーバに書き込む。障害発生の後に再起動されると、通信処理サーバは、ログサーバに保持された情報を読み出し、最新の処理の区切りにおける内部状態を回復して処理を再開する。
- ログサーバの障害に対処するために、ログサーバ自体のレプリケーションも行なう。通信処理サーバは、ログサーバへの書き込み時にその障害を検出すると、他のログサーバに切り替えて処理を続行する。
- システム監視サーバは、通信処理サーバとログサーバの動作を監視し、サーバの障害が発生した場合にはその再起動を行なう。ハードウェア故障等、自動的には回復できない障害が発生した場合には、システム管理者に通知するとともに、障害部分をシステムから切り離してシステムの再構築を行なう。

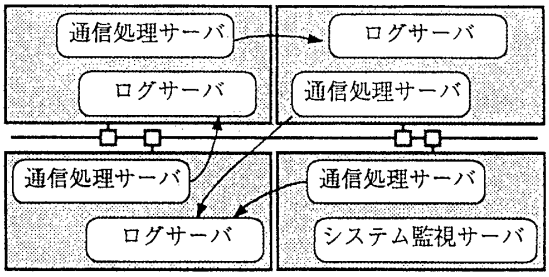


図1: 高信頼な分散型通信システムの構成

3　障害時における回復の手順

通信処理サーバは、通常動作時、レプリケートされたログサーバの内ひとつを選択し、障害回復に必要な情報を記録する。そのログサーバが利用不能になった時には他のログサーバを選び、選択したログサーバのサーバ名と時刻の組(以下、切替え情報と呼ぶ)を利用可能なすべてのログサーバに書き込む。

通信処理サーバに障害が発生して再起動された場合には、以下の手順により、障害が発生した時点で使用していたログサーバを決定し、そのログサーバに保持された情報を読み出して処理を再開する(図2参照)。

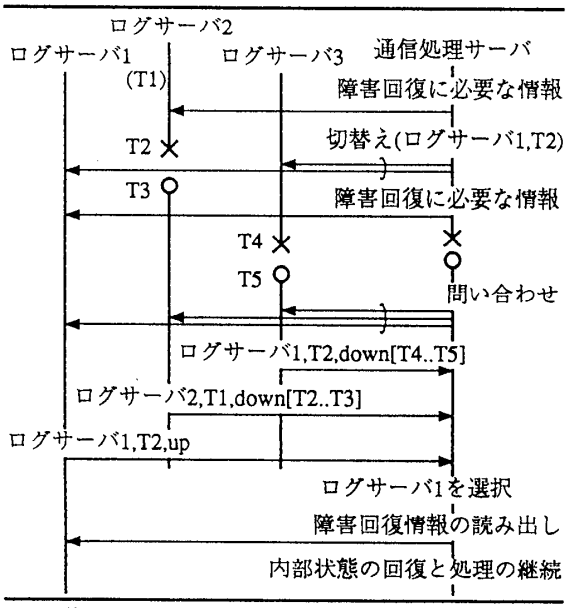
(1) 分散型通信システム内に導入されているすべてのログサーバに対して、その通信処理サーバに関する最新の切替え情報を問い合わせる。

(2) ログサーバは、保持している最新の切替え情報と、その切替え情報に記された時刻以降にログサーバ自身に障害が発生していた場合には、障害の発生時刻及び回復した時刻を通信処理サーバに通知する。

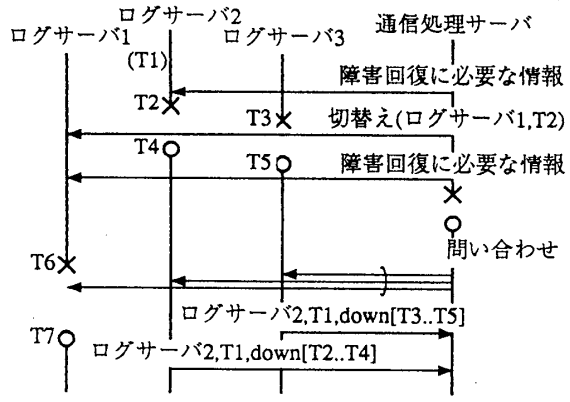
(3) 図2(a)に示すようにすべてのログサーバから応答があった場合には、その中の最新の切替え情報により、障害発生時に使用していたログサーバ(図2(a)ではログサーバ1)を知る。

(4) 問合せに応じない(障害中の)ログサーバがある場合には、得られた切替え情報の内最新のもの以降に、問合せに応じたすべてのログサーバが障害中であった時間帯があるか確認する。もしあれば、図2(b)に示すように、障害発生時に使用していたログサーバを決定することができない。図2(b)では、問合せに応じたログサーバ2及びログサーバ3の切替え情報はログサーバ2を示しているが、時刻T1以降に両者共障害中であった時間帯[T3..T4]が存在する。この

An Approach to Build High Reliable Communication System using Distributed Processing Technologies
Masahiko FUJINAGA, Toshihiko KATO and Kenji SUZUKI
KDD R & D Laboratories



(a) 最後に使用していたログサーバを決定できる場合の例



(b) 最後に使用していたログサーバを決定できない場合の例 注) ×：障害の発生, ○：再起動, Tn：時刻

図 2: 通信処理サーバの回復手順の例

期間中にログサーバの切替えを行なった可能性があるため、障害中のログサーバ 1 の回復を待つ必要がある。問合せに応じたログサーバがすべて障害中であった時間帯がなければ、最新の切替え情報によりログサーバを決定する。

(5) 通信処理サーバは、障害発生時に使用していたログサーバにアクセスして保持された情報を読み出し、内部状態を回復して処理を再開する。そのログサーバが障害中である場合には回復を待つ。

ログサーバの障害回復では、最新の書き込み情報を検索し、書き込みの途中で障害が発生していた場合にはそれを廃棄した上で、再起動した時刻を記録する。

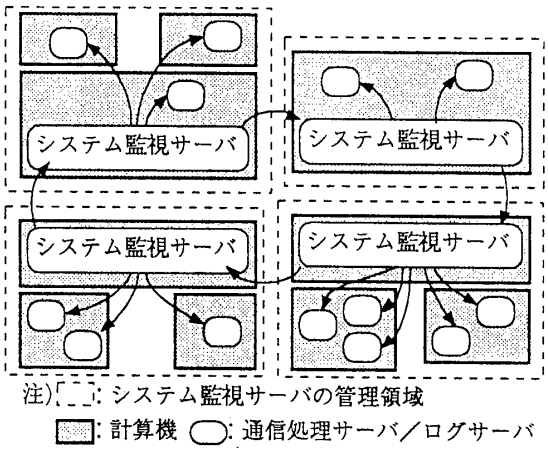


図 3: システム監視サーバによるシステム状態の監視

4 システム状態の監視と動的な再構築

システム監視サーバは、図 3に示すように、割り当てられた計算機上で動作すべき通信処理サーバとログサーバ、及び他のシステム監視サーバ に対して定期的にヘルスチェックを行なう。計算機の障害を検知した時には一定時間その再起動を待つ。再起動されない場合その計算機をシステムから切り離し、システム管理者に通知すると共に、残りの計算機による縮退運転を行なう。障害中となった計算機上のサーバについては、可能であれば他の計算機上でその処理を肩代りするサーバを起動する。

サーバの障害を検知した場合にはそのサーバを再起動する。ログサーバを再起動した場合、その障害が発生した時刻も通知する。計算機や通信処理サーバ、ログサーバの新規追加あるいはハードウェア障害からの回復の通知は、オペレータコマンドによりシステム監視サーバに通知する。

5 むすび

本稿では、分散処理技術を用いて構築する分散型の通信システムについて、その高信頼化のための構成法について検討した。本構成法は、通信処理を行なうサーバに加えて、障害回復に必要な情報を安全に保持するサーバとシステムの状態を監視するサーバを導入し、これらのサーバのレプリケーションを行なうことにより信頼性を高めることを特徴としており、障害からの回復機能と、回復不能な場合の動的な再構築機能を容易に実現できる。最後に、日頃御指導戴く KDD 研究所 浦野所長、真家次長に感謝する。

参考文献

[1] 加藤, 藤長, 鈴木：分散処理技術を用いた通信システムの構築に関する一考察, 情報処理学会第 43 回全国大会, No. 7T-1 (1991).

[2] Fujinaga, M., Kato, T. and Suzuki, K.: Implementing IN Functional Entities on top of Distributed Operating System, in *Proceedings of the XIV International Switching Symposium*, Vol. 1, pp. 268 – 272 (1992).

[3] 藤長, 加藤：高信頼アプリケーションのための汎用ログサーバ, 電子情報通信学会秋季全国大会, No. D-115 (1990).