

W/S保守情報管理のための分散データ管理システムの構築

7H-1

橋元 直 川邊恵久 今田正卓 守屋康正

富士ゼロックス システム技術研究所

1. はじめに

一人一台以上のワークステーション(以下WSと略す)を有する研究/開発環境において、そのネットワーク上で、異機種接続や、異なるバージョンのOSの使用例が増加している。しかし、このような環境下において、個々の保守情報を集中管理化した場合、下記のような不具合が考えられる。

- 保守情報を一括管理しているWSがダウンしている場合、その保守情報にアクセスできない
- 一斉的なソフトウェアのversion up等において、保守情報を保持しているWSに負荷が集中する
- ディスククラッシュにより全てのデータを失う危険性がある

これらの問題に対処するための一手法として、保守情報を複数のWSに分散して保持するようにすれば、全てのデータを失う危険性や、不必要なリモートアクセスが減少すると考えられる。こうした背景から、保守情報を複数のWSに分散して保持させ、多数のユーザによる分散管理を可能とするシステムを開発したのでここに報告する。

2. システムの構成

本システムは、図1のように複数のワークステーションが接続されたネットワーク上で、クライアント・サーバモデルを採用し、データを分散管理することを目的としている。

図1においてA~Dは、通信機能を有しネットワークによって接続されたノード(WS)を表す。この図は、B,C,D上でサーバを起動し、C上でクライアントを起動した状態を表している。サーバは、クライアントからの要求の受け付けを行ない、データの保持、参照、更新、同時アクセスの制御を行なう。クライアントは、ユーザとサーバの間のインターフェイスの役割を果たす。具体的にはユーザからのコマンドの解釈、サーバに対する要求の発行、データの授受、及びそのデータの処理を行なう。

本システムが管理するデータは、ネットワークに接続されたWSの、管理者、OSのバージョン、購入先等の保守情報である。データベースは各サーバ上に分散される。データ分散の方式¹⁾としては各種の方式が考えられるが、ここでは、データ更新時の手順が簡単になるという点から、

- データベースを互いに重複させずに分割
- 分散型の非重複ディレクトリ方式

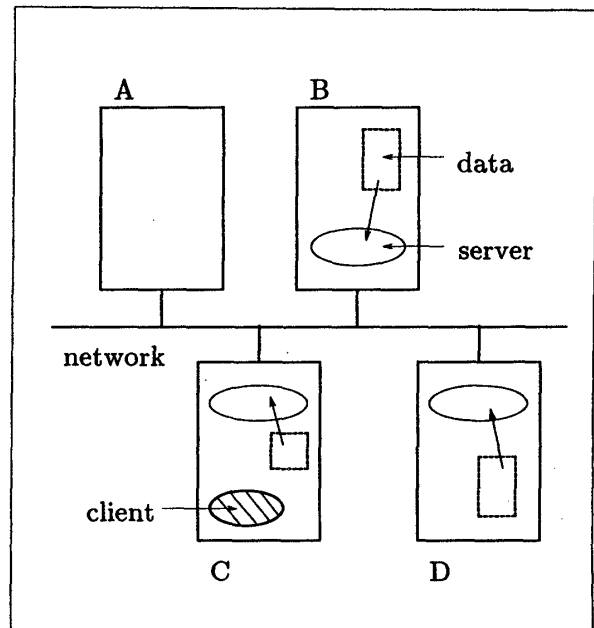


図1 HPSのシステムモデル

といった形式を取っている。

3. 分散データベースとしての対応

ここでは本システムが、分散したデータをどのようにして管理するのかを述べる。

3.1 問合せ処理

サーバは、クライアントからの同報通信機能を用いた問合せに対する応答機能を持つ。このため、クライアントは明確なサーバの存在を意識することなくデータの要求を行うことが可能になり、結果的に不可視性が実現されている。

3.2 名前管理による重複チェック機構

サーバは、名前(key)とデータ(entity)の組合せによりデータの管理を行なう。前述した問合せ処理に起因して、名前はシステム内で重複しないようにする必要がある。これを実現するために、新しいデータをシステムに登録する時点で、システム内に同一の名前が存在しないことを同報

通信を用いて確認する機構を設けた。すでにその名前が存在する場合は、登録を中断し重複を通知する。

3.3 アクセスの排他制御

複数のクライアントが同時に同一のデータの更新を行なうと、データベースの一貫性が保てなくなる。そこでサーバは、クライアント間のデータの更新に対して排他制御を行なう必要がある。本システムでは、flockシステムコールを採用し、更新の前でファイルのロック、アンロックを行なうことで排他制御を実現している。ロックを用いる排他制御では、複数の共有資源にアクセスする場合デッドロックが起こり得るが、本システムでは複数ファイルの同時アクセスは行えないため、デッドロックは起こらない。

3.4 アクセス権の設定

本システムでは、将来的な利用形態に対処するため、サーバが有する保守情報にアクセスできるユーザのリストを、各サーバごとに作成し保持することにより、保守情報に対するアクセスコントロールを行なっている。

3.5 通信のプロトコル

本システムでは、クライアントとサーバの間で、メッセージを授受することで、処理が進む。例えば、データの更新を行なう処理では、クライアント、サーバ間でのメッセージの流れは図2のようになる。ここで使用しているメッセージの種類及びその意味を次表に示す。

送り手	メッセージ	意味	通信の種類
client	search	データの検索	同報通信
	require	データの要求	一対一通信
	send_data	転送開始	一対一通信
server	administor	データ保持	一対一通信
	send_data	転送開始	一対一通信
	complete	処理終了	一対一通信

表1 データ更新処理で使うメッセージ

図2で示すように、クライアントはまず、同報通信で全てのサーバにメッセージを送り、一つのサーバがそれに答える。このようにして、サーバに対する問合せが行われ、その後の一対一通信により、データの更新が行われる。

4. 結果

ネットワーク上に分散する保守情報を管理するシステムを実装し、本システムが論理的に動作することを確認した。また、不可視性を実現することにより、データの所在を意識することなくアクセスすることが可能になった。

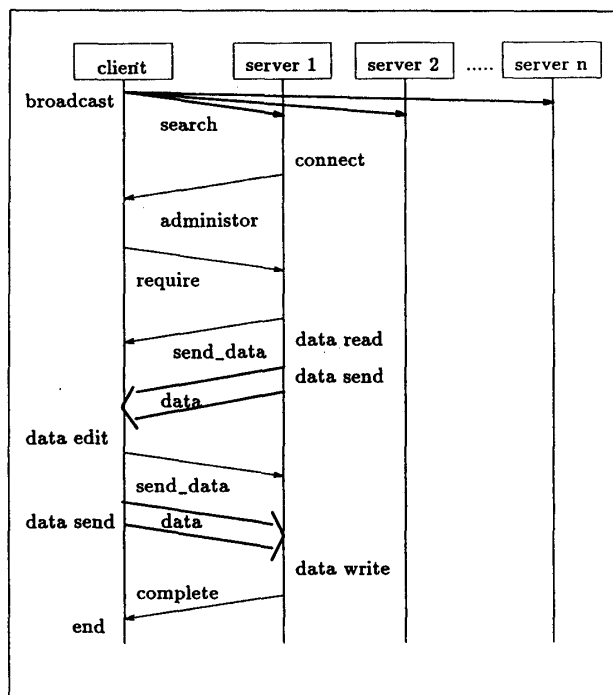


図2 データ更新のためのプロトコル

本システムの実装により、保守情報の分散管理に対して、必要な尺度となる

- 危険の分散
- 負荷の分散
- アクセスのオーバーヘッド

等を定量的に把握/評価するためのテストベッドができたと考えている。

5. 今後の課題

今後は、

- 実装したシステムの有効性について、定量的な評価を行う。
- 装置の故障やシステムの異常に起因するデータの消失に対する対策として、データを重複して保持する機構を追加する。
- 新たにデータを登録する場合、同報通信によって名前の一貫性を確認するが、この方法では一意性の保証が完全ではないため、新たなプロトコルを考える。

といった課題を解決していくつもりである。

[参考文献]

- 1) "ネットワークシステムに関する調査(ネットワーク専門委員会報告書)"、日本電子工業振興協会、1987年3月
- 2) 松下、"コンピュータ・ネットワーク"、培風館、1983
- 3) "File Servers for Network-Based Distributed Systems", ACM Computing surveys, Vol.16, No.4, pp.353-398, 1984
- 4) "Sun Workstation Network Programming Manual", Sun Microsystems, Inc., 1988