

分散処理の研究領域

戸田 巖（電電公社横須賀電気通信研究所）

1 まえがき

分散処理が叫ばれて約10年の歳月が流れたが、未だ十分実用的な分散処理技術が確立したとは云えない現状である。

これが当研究会が存続を許される理由の一つであろう。

本稿では、分散処理技術確立の一助として、現時点での同分野の問題点を主観的にとりまとめたものである。

2 定義

分散処理システムの議論がややもすれば噛み合わない理由の一つは、その定義が不分明なことである。

前主査元岡教授は、分散処理を“オンライン接続された複数のプロセッサから成るシステム上で、複数のプロセッサに分散されて実行されるタスクの集合から成るジョブの処理”と定義（※1）している。本定義は地域分散システムと分散形マシンの双方を含む広範なものである。

上記の定義に“システム全体の制御が非集中であること”を付加して、例えば単なるマルチプロセッサ等を分散システムから除外する定義も提案されている。

（※2） この場合システムの制御はシステムの制御情報全体のコピーを各プロセッサが夫々所有することにより又は各プロセッサが制御情報の部分を所有し、必要な情報は所有者に要求して入手することにより行われる。なお本定義は、集中的な制御を排除しているだけで、必ずしも地域分散システムのみを意味するものではない。

分散形マシンも地域分散システムも、並列処理を目標としたものであるが、前者は主として高速性の追求を目標するのに対して、後者は組織に適合したシステム構成の実現を目標としていると云えよう。

本稿では上述第二の定義、即ち非集中制の分散システムを中心に研究課題を論ずる。

3 通信網

地域分散システムでの通信網設計の問題は

- （1）プロセッサ、データベース等の配備
- （2）ノード間の通信容量の算出
- （3）ネットワークポロジの選定
- （4）通信回線・通信方式の決定等を含む。

これらの検討には、公衆網／専用線、回線交換／パケット交換、星状／網状／ループ／バストポロジ、通信速度等多数のパラメータの選定を要し、コスト、信頼性の点で最適化をはかることは大変複雑な問題である。

さらに運用時のフロー制御、ルート選定のアルゴリズムも重要である。

通信網は、光通信、ディジタル通信の普及、通信処理等網機能の高度化等、急速な変貌を遂げている。

これらの新しい状況の下で、上記の諸問題も見直していく必要がある。

分散形マシンのプロセッサ間通信路の問題は、数百のマイクロプロセッサを相互に接続する問題に加え、VLSIで実現する制約も加わって上記とは異なった技術分野を形成して居る。今後急速に進展するものと思われる。

4 プロセス間通信

プロセス間の通信は、分散システムの最も特徴的かつ基本的な機能である。

一般のプロセス間通信には、

- （1）共有アドレス空間を介しての通信
- （2）メッセージ通信

の2方法がある。前者は交信するプロセス間に共有のアドレス空間が存在する時のみ適用可能である。

狭義の分散システムの場合、全てのプロセスの共通空間は存在しないので、プロセス間通信はメッセージ通信で行うことが多い。

プロセス間通信を行うためには、交信プロセスの名前が必要となる。システム内で通用する名前を生成するに際し以下の問題を解決する必要がある。

(1) 名前の作り方

(2) 名前形成のタイミング

さらに分散システムの制御のためには、システム状態を知る必要があるが、全プロセッサから共通的に参照できるシステム状態が(定義から)存在しないので、各プロセッサは必要な状態の所有者にその都度問合わせるか、又はシステム状態のコピーを夫々のプロセッサが所有する等の対処が必要となる。いずれの場合も、通信に伴う遅延時間の存在によりシステム状態をリアルタイムに把握することが困難になる。

プロセッサ間でシステム状態を通知するためシステム内の全プロセッサに同一情報を伝達する“同報通信”を効率よく確実に実行することが特に必要となる。

5 負荷分散及び制御

プロセスの実行をどのプロセッサに分担させるかも分散システムの設計及び制御に当たって重要な問題である。

このためには、まずシステムの機能をプロセッサに割当て可能な単位に分割して表現すること、これらの実行単位相互間の制約条件を明らかにすることが必要となる。

次にこれらの実行単位を、システムの資源の制約の下で、プロセッサに割り付けていくアルゴリズムが必要となる。

これらにおいてシステム機能の表現方法、目的関数(性能、信頼性、経済性等)を最適化する割当てアルゴリズム等未解決の問題が多い。

特に現在までの研究はいずれも理論研究の域を出ず、実際の設計手法の確立が急がれる。

6 分散データベース

地域分散システムの研究において最も興味のある分野である。

(1) データの配置

データ参照、更新のトラヒック、通信及びファイルコストを与え、性能等の制約条件の下でデータの最適配置を求める問題である。

同一データを複数のノードに配置することを許すマルチプルコピーの条件の下では、可能な組み合わせの個数がノードの個数と共に急激に増大するので最適配置を求める能率のよいアルゴリズム、ヒューリスティックの発見が必要となる。

(2) ディレクトリ

分散データベースの構造を示すディレクトリの構成、及びその配置も今後の検討課題である。具体的な問題として、マルチプルコピーファイルへのアクセス要求があった場合、アクセスコストを最小とするコピーを選択する問題などがある。

(3) 並列制御

分散データベースへのアクセス要求を、データベースの一貫性を損うことなく可能な限り並列に処理する問題である。

集中データベースの場合と同様に、データ更新と読出しの順序性を保証していくことが必要で、ロック方式、タイムスタンプ方式等が使用される。

分散データベースにロック方式を適用する際には、データが複数ノードに分散配置されて居り、かつマルチプルコピーが存在するので効率よくロックを施す方法、デッドロックの検出、ファイル回復等新らしい解決が必要となる。

タイムスタンプ方式の場合は、能率、デッドロックの問題は軽減できるが、代りに各ノードのローカルロックの同期を確保する必要が生ずる。単に物理時計の精度を上げるだけでなく論理的な意味で同期をとる手法が議論されて居る。

並列度を向上させる手法としては、ロック単位の細分化、マルチコピーの配備

等が提案されている。後者の場合、コピー相互でのデータの一致性を保証する能率のよい更新アルゴリズムが必要となる。

(4) 障害回復

システム内の障害発生に際し、データベースの内容を矛盾なく復元する問題である。

障害発生時に、ファイルを特定時点の内容に復旧する方式として、ログ方式、差分ファイル方式、バックアップファイル方式、バージョン方式等が提案されている。

いずれの方式でも、集中ファイルの場合同様、事前処理及び復旧時の効率の問題が重要である。各ノードでのファイルバックアップ、マルチプルコピーファイルの復元に必要なローカルクロックの同期等が新しい問題となる。

最近提案された(※5)バージョン方式は、各ファイルレコードを更新不能のバージョンの履歴として構成する方式であり、障害回復が容易、並列処理との親和性のよい光ディスク等の追記形媒体を使用できる等の特長があるが、最新バージョンの管理等の問題が生ずる。

ファイルをどの時点まで復旧するか、ファイル障害の遡及防止等集中ファイル同様の問題を解決する必要がある。

7 信頼性

システムの信頼性設計は資源の冗長配備、障害の検出、回復の問題に帰着する。

分散システム、特にシステム制御が分散されている場合には、検出、回復をシステム内の各ノードが通信線を介して実行していく必要があり、そのアルゴリズムの開発が必要である。

信頼度と冗長性の関係も必ずしも明確でなく、今後の問題である。

災害によりノード機能全体が損われる様な場合のバックアップも別個の問題である。

8 ネットワークアーキテクチャ

現実の通信網、ノードの構成要素は、多様であり、常に進歩している。

この事態に対処するため分散システムの論理的モデルを定めるものがネットワークアーキテクチャである。

データの転送を司る下位レベルのモデルについては、ほぼ完成されつつある状態であるが、データグラム、同報通信、適応ルーティング等の取り扱いにはさらに検討を要する。

今後の重要な研究領域は、いわゆる高位レベルプロトコルであり、分散システムのプリミティブとしての機能の選定、特に本稿で論議しているプロセス間通信、データベースアクセス、ファイル、ジョブの転送、障害関係、保守・運転関係等のモデル及びプロトコルの検討が重要である。

また関連してプロトコルの形式的記述、プロトコルの検証、相互変換等も研究領域である。

9 ネットワークOS

分散システムの制御を行うオペレーティングシステムである。通常ネットワークの存在を意識させない形(Network Transparency)で、ユーザにシステム機能の提供を行うものをいう。

ネットワークOSは、前述の様に、集中的なシステム制御情報の欠如、通信の遅延の制約の下で、各種の制御を実行しなければならない点に問題がある。

ユーザにネットワークの存在を意識させる場合には、プログラミング言語に分散システム制御のための機能を付与する必要がある。プロセス間通信、プロセスの同期、並列動作の制御、障害回復処理、例外処理等の機能の表現について検討する必要がある。

10 評価、モデル、シミュレーション
分散システムの効果を何に求めるかについて意見の一致を見ていない。

たとえば経済性、信頼性、ユーザインタフェース等の効果が期待されている場合、分散システムの経済性、信頼性、ユーザの生産性等を評価する手法が必要である。

分散システムの動作の解析には、そのモデル化、ならびに効率的なシミュレーション手法の開発が必要となる。

11 機密保護

分散システムでは、地理的拡がり及び運転操作に関与する要員数の増大に伴い、機密漏洩の可能性が増大する。

反面、データが分散配置されて居り、機密防護の点から有利な面もある。

機密保護に暗号を用いる場合には、暗号鍵の管理を統一的行う手法、鍵自体を安全にユーザ、ノードに分配する方法等が研究されている。

いずれにしろこの面の研究は殆ど未着手であり、今後の検討が必要である。

12 建設・保守・運転・移行

分散システムの建設には、予め全体的に設計されたノードを建設していく場合と、分散システムを意識せずに建設されたノードを結合して分散システムとする場合の二つがある。いずれの場合にも、システムの機能・性能の確認が必要で、そのためのテストケースの設計は集中システムに比較して困難となる。

システム内の機器の保守・障害診断等も、例えば共通の保守ノードから行うことが考えられ、その効率化の問題が生じる。

運転についても、システムの運転開始・

終了、異常時の処置等も、特定のノードから自動的に行うか、各ノードの運転員が行うか等の選択が必要である。又ユーザ、構成機器、サービスの変更等の通知及び制御情報の変更の方法も確立する必要がある。

既存システムから、分散システムへ移行する場合も技術的及び運用上種々の問題を生じる。

13 まとめ

以上、分散システム技術確立のため必要と予想される問題点を列挙した。

これらの解明を通じて、いかなる条件で、分散システムが集中システムより有利となるか、分散システムに適合した応用は何か等の問題も明らかにしていく必要があるだろう。

文 献

(1) 元岡“分散処理への期待”、情報処理

vol. 20, No. 4, 1979, pp.270~ 283

(2) A. Tripathi, E. Upchurch
“An Overview of Research Directions in Distributed Processing”, Proc. of Compcon Fall 80, 1980, pp.333~ 340

(3) R. Eckhaus et al. “Issues in Distributed Processing - An Overview of Two Workshops”, Computer, Jan. 1978, pp. 22~26.

(4) J. Stankovic, A. van Dam
“Research Directions in Distributed Processing”, in “Research Directions in Software Technology” ed. by P. Wegner, 1979, pp.611~ 640.

(5) D. P. Reed “Naming and Synchronization in a Decentral-

ized Computer System", Technical Report MIT/LCS/
TR-205, MIT, Oct. 1978