

1 M-9

同一機種分散システムにおける負荷分散方式

今井 功, 佐藤 文明, 中川路 哲男, 水野 忠則

三菱電機(株) 情報電子研究所

1 はじめに

分散処理システムでは、負荷をネットワーク上の各ノードに最適に分散することにより、高いパフォーマンスを得ることができる。

従来の負荷分散システムの方式としては、例えば、処理アクセス要求群を一つのスケジューラに転送し、そのスケジューラが処理を行っていないデータ処理装置に対して、転送するといった方法が提案されているが、この様な集中管理方式では、スケジューラに多数のアクセス要求が発行されると、スケジューリングの処理が全体のボトルネックになる可能性がある。また他の方法としては、資源をアクセスする側が各ノードの負荷状況を問い合わせ、最も負荷の軽いノードに処理を依頼する方式があるが、問い合わせの頻度が高くなると問い合わせのメッセージの処理のオーバーヘッドが問題になってくるなど、いずれも必ずしも最適とはいえない。

本稿では、同一機種によって構成された分散システムにおいて、負荷分散を行なうための一方式について提案する。

2 システムの概要

我々が提案する計算機ネットワーク環境を図1に示す。本システムは、各種資源を管理するリソース・マネージャと、この資源に対するアクセス要求を発行する計算機と、これらを接続する伝送経路によって構成される。

リソース・マネージャに資源アクセス要求が到達されると、まずそれはキューに蓄積される。キューには、あらかじめ閾値を設定しておき、キュー長が閾値を越えた場合、それ以後のアクセス要求はネットワークシステム内にある他のリソース・マネージャに転送され、代替処理を行なうことによって負荷の分散をはかる。転送されるアクセス要求は、各ノードから転送されるアクセス要求と自ノードから転送するアクセス要求をもとに、負荷の軽いリソース・マネージャを一つもしくは複数予測し、その中でランダムに送れるようにする。

また、資源の状態を監視することによって、処理が不可能となるアクセス要求を、キューに蓄積される前に他のノードに転送する。そして、キューの長さ、アクセス要求の処理率、アクセス要求の転送コストによって閾値を一定間隔で見直して、負荷の変動にもなる適用的な

分散構想を実現する。

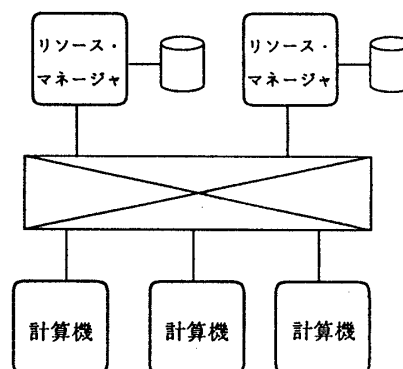


図1 計算機ネットワーク

3 システムの構成

本稿で提案したリソース・マネージャの内部構成を図2に示す。本リソース・マネージャは、以下のコンポーネントから構成されている。

- (1) 通信制御モジュール
- (2) アクセス要求受付モジュール
- (3) アクセス要求分析モジュール
- (4) 閾値設定モジュール
- (5) キュー管理モジュール
- (6) アクセス要求処理モジュール
- (7) 資源入出力モジュール
- (8) アクセス要求転送モジュール
- (9) 資源転送モジュール
- (10) 転送先設定モジュール
- (11) 資源監視モジュール

本リソース・マネージャの特徴は、動的に変化する閾値を持つところにある。以下に、閾値設定モジュール内で閾値を設定するためのアルゴリズムについて示す。

【STEP 1】初期閾値、初期キュー長、初期アクセス要求処理率、初期アクセス転送率を設定する。

【STEP 2】一定時間経過後、新キュー長、新アクセス要求処理率、アクセス要求転送率をアクセス要求分析モジュールより受信する。

【STEP 3】アクセス要求処理率が低下しているかどうかの判定を行なう。処理率が低下している時はステップ4を行ない、そうでない時はステップ5を行なう。

【STEP 4】キュー長が増加しており転送率が増加していない場合は、閾値を下げる。キュー長と同時に転送率も増加している場合、もしくは転送率は増加しているがキュー長が低下している場合は、閾値は変更しない。変更した閾値をキュー管理モジュールに送信した後、ステップ2に移る。

【STEP 5】キュー長が減少している場合や転送率が増加している場合は閾値を上げ、そうでない場合は閾値は変更しない。変更した閾値をキュー管理モジュールに送信した後、ステップ2に移る。

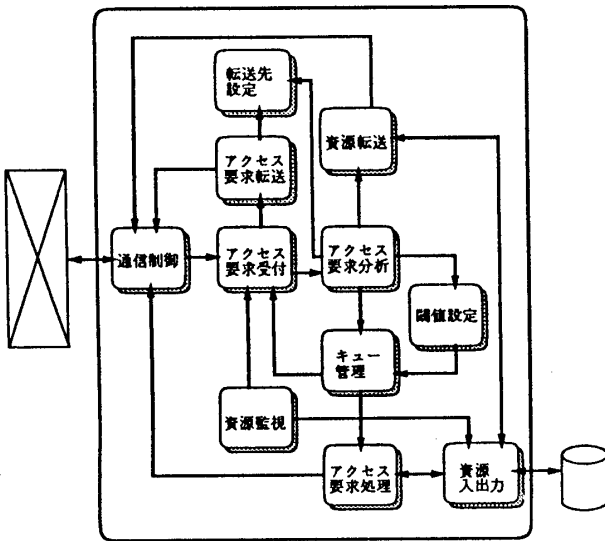


図2 リソース・マネージャの内部構成図

4 実験例

図1及び図2に基づいて、本リソース・マネージャへの処理依頼の流れを説明する。

まず各計算機から資源アクセス要求をリソース・マネージャに送信されると、リソース・マネージャでは通信制御モジュールが受信し、その資源アクセス要求をアクセス要求受付モジュールに転送する。アクセス要求受付モジュールでは、キュー管理モジュールに対してはまだ受付可能かどうか、資源監視モジュールに対して要求する資源が使用可能かどうかを問い合わせ、その結果がいずれも可能であれば資源アクセス要求はアクセス要求分析モジュールに転送され、その結果のいずれかでも不可能であれば、資源アクセス要求転送モジュールに送られる。

資源アクセス要求がアクセス要求分析モジュールに転送されると、そこで資源の要求元、資源のアクセス頻度、キュー長、資源アクセス要求の処理率、資源アクセス要求の転送率、資源アクセス要求の転送コストが計算される。アクセス要求転送モジュールに転送された資源アクセス要求は、転送先設定モジュールに問い合わせ、代替

資源管理モジュールに対して資源アクセス要求を転送する。転送先設定モジュールはアクセス要求分析モジュールから受け取った各ノードからの資源アクセス要求情報と自ノードからの転送情報をもとに、代替リソース・マネージャを決定する。

アクセス要求分析モジュールでの計算後、資源アクセス要求はキュー管理モジュールに含まれるキューに転送される。閾値管理モジュールでは、アクセス要求分析モジュールで分析された統計情報をもとに、キューの長さの閾値を決定し、アクセス要求分析モジュールを通過した資源アクセス要求をキューに格納し、アクセス要求処理モジュールにキューの先頭のものから転送する。また、アクセス要求受付モジュールからアクセス要求の受付が可能かどうかの問い合わせに対して、閾値管理モジュールから設定された閾値と、現在のキューの長さを比較して受付可能かどうかを応答する。

アクセス要求処理モジュールは資源入出力モジュールを使って実際の資源との入出力を行ない、資源アクセス要求の処理を行なって、その応答を通信制御モジュールを経由して応答する。資源転送モジュールは、アクセス要求分析モジュールが資源アクセスの頻度が著しく高いと判断したノードに対して、資源転送の可能性を問い合わせ、可能であれば資源入出力モジュールを使って資源を転送する。資源監視モジュールは常に資源の状態を監視しており、アクセス要求受付モジュールからの資源使用の問い合わせに対し応答する。管理している資源がダウンした時や復旧した時は、その応答を資源入出力モジュールから受け取る。

5 おわりに

本稿では、計算機ネットワーク内に分散された資源を、効率良く利用するための管理方式として、動的に変化する閾値を持つリソース・マネージャによって負荷分散を行なう方式について報告した。

これによると、計算機ネットワークにおいて分散して配置された資源をアクセスする際に、資源に対する負荷が分散されるように資源に対するアクセス要求を制御することが可能となり、応答性や全体の資源の利用効率を向上させることができる。また、資源を管理する機構にボトルネックが生じないため、多くの資源アクセス要求が生じた場合にも対応することできるといったことが期待できる。

今後は、既存の負荷分散方式と、本稿で提案した分散システムとの、処理応答時間について比較評価を行なう予定である。

参考文献

- [1] DEREK LEAGER, et al : " Adaptive Load Sharing in Homogeneous Distributed System ", IEEE TRANSACTION ON SOFTWARE ENGINEERING, VOL. SE-12, NO.5, MAY 1986.