

クラスタリングシステムのジョブスケジューリング

5Aa-5

栗原素子 森田眞司 竹中安生
(株)日立製作所 ソフトウェア開発本部

1. はじめに

近年マイクロプロセッサの性能の飛躍的な向上やネットワーク技術の進歩によって分散コンピューティング環境が一般的なものとなってきた。それとともに分散したコンピューティング資源をいかに有効に活用するかが新たな課題となった。

我々は、クラスタという概念を導入しネットワークに接続されたワークステーションを一つの計算機システムのように扱うことを可能にした。さらに、CPU 使用率とメモリ使用率を中心とするリソース管理を行い、個々のワークステーションを最大限に活用するジョブの負荷バランシングシステムを構築した。

本稿ではその成果について述べる。

2. 問題点

本節では分散コンピューティングにおける問題点について述べる。

2.1 負荷のアンバランス

分散コンピューティング環境では負荷が特定のマシンに偏る現象が起りがちである。一般にユーザはどのマシンでプログラム等を実行させるのが最適であるのかを把握しつつ作業を行うことはない。したがってより処理能力の高いマシンを選択しがちであり、その結果特定のマシンに負荷が集中してしまうネットワーク全体の処理効率が低くなってしまふ。このような状況を避けるためには、ネットワークに接続されたマシンの負荷を個々に調べ把握しなければならない。この方法は非常に煩雑である。

2.2 グルーピング

ネットワークにはプラットフォームの異なる様々なワークステーションが接続されており、それぞれ固有の名称を持ち、またその特性も異なっている。ユーザは処理の内容によって特定種類のワークステーションを使用するし、また、所属する部署により主に使用するワークステーションが決まっていることもある。このような場合、特性や使用環境など同じ特性を持つ複数のワークステーションを一つのワークステーションのように扱えることが望ましい。しかしながら現状では、各ユーザがそれぞれワークステーションごとにその特性や使用環境を記憶して使用しなければならず非常に効率が悪い。

3. 機能

本節では本システムにより上述した問題をどのように解決したかを述べる。

3.1 負荷分散

負荷のアンバランスを解決するために、ネットワーク上のコンピュータ間での処理の負荷を均一にすることを可能にした。本システムではネットワーク上のワークステーションのクライアントから一定の更新間隔で CPU 使用率とメモリ使用率をサーバに送信する。サーバは処理を行う余力が最も大きいワークステーションを決定し、ジョブを実行する。ジョブを実行するには以下のステップを経る。

1. 負荷順のワークステーションのリストを問い合わせる。
2. 指定されたクラスタを構成すること、ジョブを実行する際に必要とするメモリ使用率を満たしていることなどすべてのジョブの実行に関する条件を満たすワークステーションを上記の負荷順リストから選択する。

Job Scheduling on Clustering System

Motoko Kurihara, Shinji Morita,

Yasuo Takenaka

Hitachi, Ltd. Software Development Center

3. 選択したワークステーションにジョブを転送し、実行する。

3.2 クラスタ

グルーピングの問題を解決するために複数のワークステーションを一つのワークステーションのように扱うことを可能にした。ここでグルーピングされた複数のワークステーションがクラスタである。一つのワークステーションは複数のクラスタに所属することもできる。これによってユーザが所属している組織の構造に対応したグルーピングがネットワーク上でも実現できる。また、プラットフォームごとに一つのクラスタを構成することも可能である。

4. 評価

4.1 評価方法

通常の開発作業を行っている5台のワークステーション上で5つのジョブを10秒間隔で投入し、評価を行った。本システムを使用してどのワークステーションでジョブを実行するかが自動的に決定される場合と、本システムを使用しないで5台のワークステーションで一つずつジョブを実行する場合を設定した。なお、本システムの負荷更新間隔は5秒とした。

4.2 測定結果

評価に用いたジョブを以下に示す。

- クイックソート
- ガウス消去法によるN元連立一次方程式

クイックソートでは2,000,000のランダムに並んだ実数の昇順への並べ換えを行い、ガウス消去法では500元連立一次方程式の解を求めた。表1に測定結果を示す。表の値はジョブの平均実行時間で単位は秒である。

表1 測定結果

	本システムなし	本システム使用
qsort	27.2	15.6
Gauss	64.8	60.8

4.3 評価

いずれも本システムを用いた場合の方が実行時間が短い。

5. 考察

負荷分散の効果を決定する要素として以下が挙げられる。

- ジョブの投入間隔
- 負荷情報の更新間隔
- ジョブの実行時間

ジョブの投入間隔が負荷の更新間隔より短く、かつ、ジョブの実行時間が負荷の更新間隔よりも長い場合、負荷の更新間隔だけジョブが特定のホストに集中しシステム全体の処理能力が低下する可能性がある。

この問題を避け、更に効率的なジョブのスケジューリングを行うためには、(1)ジョブを実行した場合そのワークステーションの負荷に一定の値を加えて次のワークステーションの選択を行う。(2)ジョブが終了した時点でサーバにそのワークステーションの最新負荷を送信する。(3)ジョブの実行時間の履歴を用いる等が考えられる。

また、ジョブを実行するワークステーションを選択するためのオーバーヘッドについても検討する必要がある。