

論文の内容の要旨

論文題目 Middleware Systems for Enabling Users to Adapt to Dynamic Changes
 in Execution Environments

(実行環境の動的な変化にユーザが適応することを可能にするためのミドルウェアシステム)

氏 名 金田 憲二

(本文) 本論文は、動的に参加・脱退する計算資源を効率的に利用するためのミドルウェアシステムについて述べる。

近年その重要度が増しつつある、PC・ワークステーションクラスタやグリッドは、動的に変化する環境である。資源共有や故障などの原因によって、アプリケーションが利用可能な資源の質は絶えず変化する。例えば、デスクトップPCなどの非占有マシンは複数の利用者によって共有されるため、その可用性が動的に変化する。数百のノードからなるシステムにおいては、マシン・ネットワークの故障が頻繁に起こり得る。このような資源の可用性が動的に変化することが、クラスタやグリッドの普及への大きな障害となっている。

本研究の目標は、利用者が自分の実行環境の変化に適応し、動的に変化する資源を有効利用することを可能にすることである。この目標に向けて、我々は、それぞれ異なる焦点を持つ二つのアプローチで取り組む。一つのアプローチは、地理的に分散した大規模環境上での並列プログラミングのための枠組みを構築することに焦点を置いている。もう一つのアプローチは、単一システムイメージを提供し、動的に変化する資源を簡便に利用可能にすることに焦点を置いている。前者のアプローチのために、我々は、Grid-enabled メッセージパッシングライブラリと、遠隔ジョブ投入のためのコマンドシェルを設計・実装した。後者のアプローチのために、単一システムイメージを提供するための仮想マシンモニタを開発した。

Grid-enabled メッセージパッシングライブラリは、Phoenixプログラミングモデルに基づく。このモデルは、プログラマが動的に物理マシンに割り当て・解放できる論理マシン識別子の集合を提供する。この論理識別子の動的割り当てによって、ライブラリは、実行時におけるノードの動的な参加・脱退を扱うことができる。さらに、このシステムは、ファイアウォールやNATなどの制限によって直接通信出来ないノード間でのメッセージ配送を扱うことができる。資源発見機構によって、静的な名前を持たないノード(例、DHCPクライアント)が、手間を必要とせずに計算に参加することも可能になっている。これらの機構を実装するために、我々のシステムは、利用可能な情報が静的に設定ファイルなどの形で与えられると仮定はせずに、分散資源発見とルーティング表の構築アルゴリズムを実行する。ルーティングアルゴリズムの性能向上のために、我々

はルーティング更新メッセージの重複転送を除去する技術を考案し、そのアルゴリズムの性能評価を3つのLANにまたがる400ノードを用いて行った。その結果、我々のアルゴリズムにおいてルーティング表の構築にかかる時間が、経路情報が静的に与えられた場合の時間の約2倍に収まることが示された。

コマンドシェルは、利用者がWAN上にちらばった数百の遠隔マシンにアクセスすることを可能にする。このシェルでは、利用者は遠隔マシンへのジョブ投入を、遠隔マシンをまたがる入出力のリダイレクションやネットワークパイプなどの機能を用いながら行える。さらに、このシェルでは、何らかの経路が存在しさえすれば、利用者はファイアウォールやNATの内側のマシンにもアクセスできる。このシステムを、我々のメッセージパッシングライブラリを用いて実装した。
仮想マシンモニタは、ネットワークでつながれた複数のマシン上に、共有メモリ型マルチプロセッサマシンを仮想的に構築する。この機能によって、分散環境を簡便に利用することが可能となる。例えば、マルチプロセッサマシン上にインストールされる既存のアプリケーション(例、SMP用のLinuxカーネル)を、コストのより低い複数のマシン上で動作させることが可能になる。このシステムは、一つ以上の仮想プロセッサを一つの実プロセッサに割り当て、さらに、その割り当てを動的に変更させることによって、動的なマシンの追加・削除を扱うことができる。我々は、予備実験として、8台の物理マシン上に仮想的に8-wayのマルチプロセッサマシンを構築し、その上で互いに独立なタスクを並列に実行した。この実験の結果は、我々のアプローチが現実的に可能であることを示している。