JAIST Repository
https://dspace.jaist.ac.jp/

Title	ワークステーションクラスタによる高速並列処理に関する研究
Author(s)	奥野，弘之
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1123
Rights	
Description	Supervisor:堀口 進，情報科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

ワークステーションクラスタによる 高速並列処理に関する研究

奥野 弘之

北陸先端科学技術大学院大学 情報科学研究科

1998 年 2 月 13 日

キーワード： Workstation Cluster,Performance Evaluation,Message Passing Library, Parallel Processing.

1 はじめに

ネットワークで互いに接続されたワークステーションにPVM(Parallel Virtural Machine)[1] や MPI(Message Passing Interface)[2] などに代表されるメッセージ通信ライブラリを用いて，ワークステーションクラスタを構築し，並列処理を行わせる研究が盛んになってきた．これによって専用の並列計算機を用意するのに比べて手軽に並列処理環境を入手することが可能となったが，ワークステーションクラスタではメッセージの通信に既存のネットワークを利用しているため，10Base-T のような低速なネットワークでは処理時間全体に占めるメッセージ通信時間の割合が大きく，投入台数に見合う性能の向上が得られないという問題がある．

本研究ではスイッチングハブを用いたワークステーションクラスタにおけるメッセージライブラリの通信性能についてベンチマークを行い，その性能を詳しく検討する．次に幾つかのアプリケーションを実行し，低速なネットワークを利用したワークステーションクラスタで有効な並列処理の方法を提案する．

2 ワークステーションクラスタの性能向上に関する研究

ワークステーションクラスタの実行性能を向上させるためにハードゥエア面で取り組まれている方法としては，ワークステーションを接続するネットワークを従来の 10Base-T ハブから 100Base-T ハブ,FDDI,ATM, 商用ギガビット LAN の Myrinet などへ置き換え，

Copyright © 1998 by Okuno Hiroyuki

物理的な通信速度の向上を達成するものや，10-BaseT スイッチハブの利用による通信帯域の拡張が挙げられる．

しかし，FDDI,ATM,Myrinet などは導入に対するコストが大きく容易に入手することが難しい．比較的入手が容易となった 100Base-T ハブや 10Base-T スイッチハブでは大きな通信速度の向上が得られない問題がある．

ソフトウェア面で取り組まれている方法は，ワークステーションクラスタで用いる PVM や MPI などのメッセージ通信ライブラリに対して機能拡張を行い，通信手続きのためのオーバーヘッドを削減して実行性能を向上させる研究が多い．メッセージ通信におけるオーバーヘッドを減少を主目的とし，PVM に対しては，lightweight process を用い，SMP(Symetric Multi Processors) で動作する LPVM や，PVM のプログラム実行単位であるタスクをスレッド単位で動かすことを可能にした TPVM(Thread-oriented PVM)，商用の高速なメッセージ通信ライブラリである pardi を用いる pardi-PVM などがある．その他には PVM を利用してワークステーションクラスタにおいて分散共有メモリシステムを提供する Adsmith が提案されている。

これらの手法は LPVM のように動作環境が制限されていたり，商用であるため容易に入手が可能でないもの，インタフェースが PVM や MPI から大きく変わってしまい，今までに開発されたソフトゥエア資産を利用できないという問題がある．

3　ワークステーションクラスタにおけるメッセージ通信

ワークステーションクラスタで用いられる PVM と MPI のメッセージ通信性能を評価するためにワークステーションクラスタを構築し，メッセージ通信についてのベンチマーク測定を行った．

ベンチマークの結果，PVM はメッセージ通信に TCP/IP を用いるとメッセージ通信性能がよいこと，MPI は PVM に比べて全体的に通信性能がよいことが分かった．

メッセージ通信性能の評価結果より，PVM のブロードキャストにおける通信速度を向上させる，スイッチングハブを利用した階層型ブロードキャストを提案した．その通信速度の性能の評価を行った結果，8 台以上で 256KB 以上のデータをブロードキャストすると通信速度が向上することが確認された．また，この手法を改善する選択型ブロードキャストについて提案した．

4　ワークステーションクラスタの並列処理性能

ワークステーションの動作台数の増加に伴い処理性能の向上が得られるアプリケーションを提案するために，ワークステーションクラスタでソーティング，N クィーン問題，TSP(Travelling Salesman Problem) を実行し，その処理性能の評価とメッセージ通信性能の影響について検討を行った．

ソーティングでは大きなデータのメッセージ通信が行われるため，メッセージ通信時間が処理時間全体のうちの多くを占め，性能向上が得られないことが分かった．

N クイーン問題ではメッセージ通信を伴わずに並列化が行えるため，ワークステーション 1 台あたりへの負荷分散が均等であれば動作台数の増加に伴う線形的な速度の向上を得られることが分かった．

TSP ではプログラムの実行を分散型，マスタースレーブ型，協調型の 3 種類の方法で行い，それぞれの処理性能について評価した．

また，TSP において発生するメッセージ通信は，メッセージ通信性能のベンチマークで得られた結果より，ワークステーションクラスタの実行時間に殆んど影響せず，ワークステーションクラスタに有効な処理であることが分かった．

5 まとめ

本研究ではワークステーションクラスタに適した処理方法を提案するため，ワークステーションクラスタにおけるメッセージ通信性能の評価を行った．通信性能の評価結果と各種アプリケーションの実行によって，10Base-T スイッチングハブで接続されるネットワークでは大量のデータを通信する処理はワークステーションクラスタには適していないことが分かった．N クィーン問題の実行において，メッセージ通信を伴わず，負荷分散が均等であれば線形的な実行性能の向上が認められることを確認した．そして TSP の実行で小さなデータのメッセージ通信を行うアプリケーションは動作台数に見合うだけの性能向上がワークステーションクラスタで得られることより PVM を利用した変数共有システムを提案した．

今後の課題としては提案した階層型ブロードキャストの改善，PVM による変数共有システムの実装とその性能の評価などが挙げられる．

参考文献

[1] V.S.Sunderam.PVM:A Framework for Parallel Distributed Computing.*Concurrency: Practice and Experience*,2(4):315-339,1990.

[2] MPI Forum, “MPI: A Message-Passing Interface Standard”, *International Journal of Supercomputer Application Vol. 8 No. 3/4*, 1994