

SCore を用いた並列計算機の構築

An implementation of a parallel computer using SCore cluster system software

05528 滝澤 亮太
指導教員 内田 健

1. はじめに

分散遺伝的アルゴリズムなどで扱う実用上の応用では、一台の PC では計算資源が不足してしまい、実行が困難となるものが存在する。このような応用を実行できる PC クラスタの性能は、構成する計算ノードを接続するネットワークに依存する。本研究では、並列計算に必要なメッセージ通信用のネットワークを独立して持つ PC クラスタを構築し、その通信性能を明らかにする。

2. 提案する PC クラスタの構成

PC クラスタは、計算ノードとなる複数の PC をネットワークで接続し単一の計算資源として利用できる計算システムである。特に、SCore 型 PC クラスタ[1]は、ファイル共有などに利用される TCP(UDP)/IP 通信方式の他に、並列計算用のメッセージ通信用のプロトコル PM2 (PM/Ethernet[2])を持つ。

PC クラスタでの並列計算の効率は、このメッセージ通信をいかに効率よく行うかにかかっている。そこで、本研究では SCore 型 PC クラスタにおいてメッセージ通信と他の通信を分割し、それぞれ別々のネットワークによって構成される PC クラスタを構築する。(図 1)

図 1 における LAN_A は TCP(UDP)/IP 用のネットワークで、LAN_B は PM/Ethernet 用のネットワークである。今回、LAN_B を 1Gbps の Ethernet によって構築することで、LAN_B が無い場合に比べどの程度通信性能が向上するかを数値実験により明らかにする。

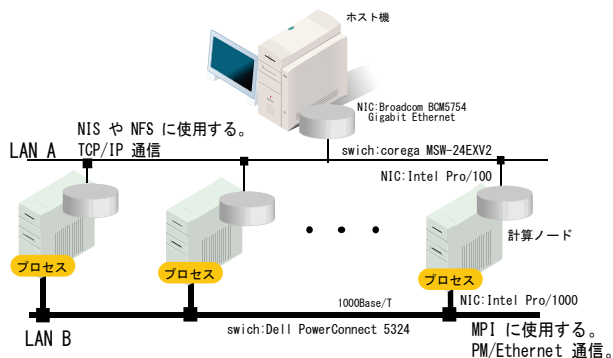


図 1. SCore 型 PC クラスタの構成

3. 実験結果

構築した SCore 型 PC クラスタの性能評価指標の一つとして、計算ノードを接続するネットワークの通信性能が重要となる。ここでは図 1 の PC クラスタ内の 2 台の計算ノード間で MPI (Message-Passing Interface) による ping-pong 通信を行い、10~5M バイトのメッセージサイズに対する通信時間より、通信性能(スループット)を算出する。各メッセージサイズについて各々 5 回通信時間を測定し、その平均値を使用する。使用した機材を図 1 の併記する。

図 2 に実験結果を示す。図 2 には図 1 の構成のもの(1000base)と図 1 の構成において LAN_B が無いもの(100base)の測定結果を用意した。提案構成によるスループット(最大)は 38.7 [M Bytes/sec]となり、1Gbps/Ethernet の理論性能値の 30%程度にとどまっている。これは、使用している計算ノードのネットワークカードが PCI 経由の接続であるためだと思われる。

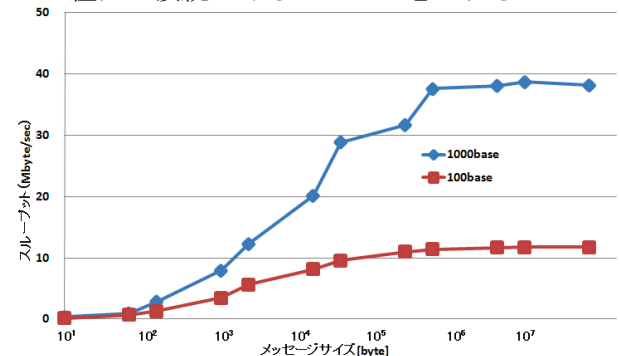


図 2. 構成した PC クラスタの通信性能

4. おわりに

本研究で構築した SCore 型 PC クラスタでは、メッセージ通信用のネットワークを増設した効果を通信性能として確認できた。しかし、増設したネットワークの理論性能値に比べ十分性能が得られていない結果であることも明らかになった。

参考文献

- 石川裕, 他, “Linux で並列計算をしよう—SCore で作るスーパーコンピューター”, 共立出版, 2002
- 住元真司, 他, ”高速通信機構 PM2 の設計と評価”, 情報処理学会論文誌, vol.41, No.HPS1, pp.80-90, 2000