

PC クラスターの性能評価

- その2 ベンチマークテストの評価 -

日大生産工（学部） 佐藤隆英

日大生産工（院） 高橋大士

日大生産工 角田和彦

1. はじめに

近年、PC(Personal Computer)の性能向上やネットワークのインフラ整備によるネットワーク効率の向上により、大規模なデータを扱う事が非常に増えてきた。そこで、大量なデータを高速に処理する為に、PC を用いた分散メモリ型並列計算システムが開発されてきている¹⁾。

本論文では、その1で構築したPCクラスターの性能をベンチマークを通して、評価することを目的とする。なお、並列処理プログラムにはMPI(Message Passing Interface)を採用している²⁾。

2. 並列計算ライブラリ

PC クラスタで用いられる並列計算プログラムとしては、MPI と PVM の二つがある。

この二つは、概ね同様の処理をすることができるが、MPIの方が移植性に優れ、多数のライブラリがある為、本システムではMPIを採用している。

MPIは、新しい並列プログラミングの規格として、最も広く使われているものである。MPIは新しいプログラミングではなく、実装はCまたはFortran 77から呼びだす、並列プログラミング用ライブラリ(サブプログラムのライブラリ)である。MPIを実装する、

フリーのソフトとして、MPICHとLAMの二つがある。本システムでは、汎用性等に優れている、MPICHを採用している。

【MPICH】

MPICHは非常に移植性が高くなるように設計され、現在、多数のMPIの実装例がある、MPIインプリメンテーションである。主なMPIの関数を以下に示す³⁾。

・MPI_Init()

MPIの実行環境の初期化

・MPI_Comm_rank()

コミュニケータ内のランクを取得

・MPI_Send()

ブロッキング送信

・MPI_Recv()

ブロッキング受信

・MPI_Comm_size()

コミュニケータ内のプロセス数を取得

・MPI_Finalize()

MPIの実行環境の終了(すべてのMPI処理を終了)

3. 性能評価環境

まず、本論文では、姫野ベンチマークを主に用いてPCクラスターの性能評価を行う⁴⁾。姫野ベンチマークとは、情報基盤センター・センター長の姫野龍太郎氏が非圧縮流体解析コードの性能評価のために考えたものでポア

Performance Evaluation of the PC Cluster

-Part2 the Evaluation of Benchmark Test-

Takahide SATO, Daishi TAKAHASI and Kazuhiko KAKUDA

ソルーション方程式解法をヤコビの反復法で解く場合の主要なループの処理速度を計るものである。コードは非常に短く簡単にコンパイル・実行できるので、即座に実測速度(何 MFLOPS)を求めることができる。なお、本研究では MFLOPS だけではなく、処理時間、ループ数、評価値も合わせて計測する。なお、併せて計測した実データは、図 1 のように示される。また、計測した結果は図 2 に示す。

ベンチマークの評価のみでは、PC クラスターの性能をしっかりと評価するにあたっては不十分な為、円周率を求めるプログラムを走らせ、その処理にどの程度の処理時間を要するのかを計測し、その結果も今回の性能評価に反映させている。計測結果は図 3 に示す。

また、計測結果から、どの程度の並列化効率が得られているのかを算出した。その結果は図 4 に示す。

4. おわりに

その 1 で構成した PC クラスタに対して、MPI を導入し、ベンチマークテストを通して PC クラスタの性能評価を行ってきた。図 2 及び図 3 では、PC の台数の増加に伴い、性能が向上していることが確認できた。また、図 4 からは、並列化効率が最も低い場合でも 75%の効率が得られた。

今後の課題としては、グリッド環境にした場合にどのように結果が変わるか、また、独自のプログラムを分散処理させた時にどのような結果が得られるのかを検証したい。

```

kakuta@PC3: ~/himenos
call mpi_allreduce(wgosa,
Argument #2 of 'mpi_allreduce' is one precision at (2) but is some other precision at (1) [info=7 g77 M GLOBALS]
[kakuta@PC3 himenos]$ nohup -no 2 a.out
Sequential version array size
mimax= 129 njmax= 65 mmax= 65
Parallel version array size
mimax= 129 njmax= 65 mmax= 35
imax= 128 jsax= 64 kmax= 33
l-decomp= 1 j-decomp= 1 k-decomp= 2
Start rehearsal measurement process.
Measure the performance in 3 times.
MFLOPS: 54.822058 time(s): 0.501153 0.0030138952
Now, start the actual measurement process.
The loop will be executed in 199 times.
This will take about one minute.
Wait for a while.
Loop executed for 199 times
Gosa = 0.00168137206
MFLOPS: 57.579323 time(s): 55.913431
Score based on Pentium III 600MHz : 0.695074081
[kakuta@PC3 himenos]$

```

図 1 姫野ベンチマーク実行画面

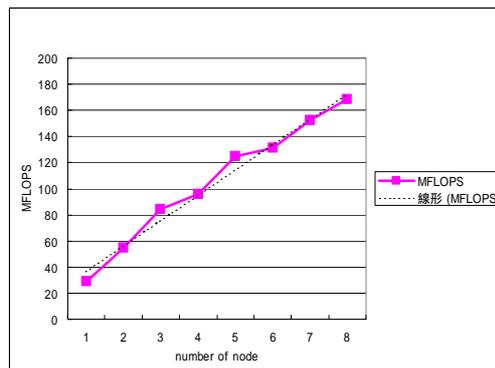


図 2 MFLOPS の計測結果

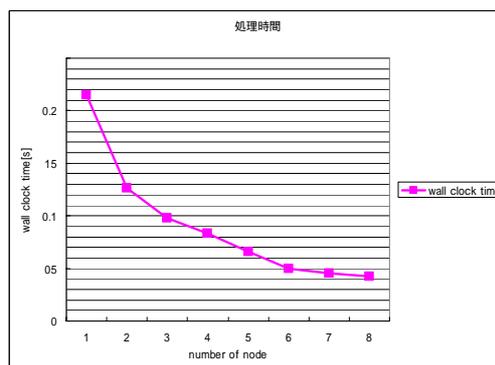


図 3 円周率プログラムの計測結果

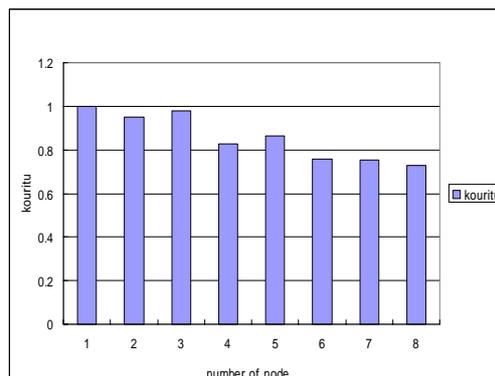


図 4 並列化効率

参考文献

- 1)角田和彦,スクーリンドサブネットワーク内での PC クラスタの構築,学術講演回,2002
- 2)P .パチェコ(秋葉博 訳),MPI 並列プログラミング,培風館,2001
- 3)http://mikilab.doshisha.ac.jp/dia/research/person/yoneda/research/2002_6_12/03-report.html
- 4)<http://accr.riken.jp/HPC/HimenoBMT/ind ex.html>