

各種 PC クラスタの性能評価

Evaluations of PC Cluster Systems

正 廣安 知之 (同志社大工) 正 三木 光範 (同志社大工)
学 谷村 勇輔 (同志社大院)

Tomoyuki HIROYASU, Doshisha University, Tatara Miyakodani 1-3, Kyo-Tanabe, Kyoto
Mitsunori Miki, Doshisha University
Yusuke TANIMURA, Graduate School of Engineering, Doshisha University

Key word: PC Cluster Benchmarks Optimization Genetic Algorithm

1 はじめに

PC クラスタは一般に使用されている PC をネットワーク接続することにより構築する分散メモリ型並列計算機システムである。近年の PC のコモディティハードウェアの性能の向上には目をみはるものがある。同時に、それらのコストも激減している。コモディティハードウェアの性能の向上と価格の下落で、手軽に高性能な PC クラスタを構築することが可能となってきた。

我々の研究グループでは、本年度、新たに 2 種類の PC クラスタシステムを導入した。一つは、高性能ネットワーク (Myrinet 2000) を有するクラスタシステムであり、もう一つは、ネットワークは通常の FastEthernet であるが、256 node という多ノードを有するシステムである。

本研究では、これらの 2 種類のアーキテクチャの異なる PC クラスタシステムに対して、最適化アプリケーションを含むいくつかのベンチマークを適用することにより、その性能の評価を行った。

その結果、アーキテクチャに適したアプリケーションが存在することがわかり、特に遺伝的アルゴリズム (Genetic Algorithm: GA) のような進化型の最適化アルゴリズムの並列化であれば、比較的低速なネットワークを持つクラスタでも、並列化効率を上げることが可能であることが明らかとなった。

2 PC クラスタ

今回比較するのは表??に示すような 2 種類のアーキテクチャを持つ PC クラスタシステムである。なお、Cambria については現在メモリを 128MB から 256MB へ増設中である。

3 ベンチマーク

本研究では、PC クラスタの性能を比較するために、数種のベンチマークを使用した。これらのベンチマークについて説明する。

表 1: PC Cluset System

システム名	Cambria	Gregor
ノード数	256 + 1	64
CPU 数	256 + 1	128
CPU	Pentiumu III	Pentium III
800MHz	1.0GHz	
メモリ	128 MB × 256	512 MB × 64
ネットワーク	FastEthernet	Myrinet 2000 FastEthernet

3.1 Linpack

Linpack はテネシー大学の Jack Dongarra らによって開発された線形代数の LU 分解を行うベンチマークである。世界のスーパーコンピュータのランキングを示すサイト (<http://www.top500.org>) ではこの Linpack の結果を用いてランキングを行っている。Atlas を用いているためにシステムのアーキテクチャに適応したプログラムが生成されるためにコンパイラなどの性能の影響が出にくい。一方で、LU 分解を行う際に使用するマトリックスのサイズは規定されていないため、できるだけ大きなマトリックスを使用することにより、各ノードでの処理量を大きくすることが可能であり、各ノードでの処理量が多くなると通信の頻度が下がるためネットワークの性能によらなくなる。ノード数が多く、メモリが大きなシステムが有利となる。

3.2 姫野ベンチ

姫野ベンチは理化学研究所の姫野氏が開発したベンチマークである。流体計算の性能評価に適したベンチマークであるといわれている。Linpack や Nas Parallel Benchmark Suite では多くの計測時間が必要であるのに対して、非常に手軽に利用できるベンチマークである。一方でキャッシュサイズが問題になるようなアプリケーションがあるので、マトリックスのサイズが大きくなるような問題を扱うようなアプリケーションとの結果は異なる場合がある。

GAPPA Bench は我々が開発している遺伝的アルゴリズム (Genetic Algorithm: GA) の並列モデルを模擬したベンチマークである。GA は生物の進化を模倣した確率的な多点探索アルゴリズムである。GA を並列で処理する手法は種々あるが、GAPPA では分散母集団モデルが採用されている。すなわち、一つの母集団を複数の母集団に分割し、各々で独立して遺伝的操作により解探索を進める。数世代後、各分割母集団の中からランダムにいくつかの個体を選択し、他母集団に移動させる。この操作が移住であり、全体としての多様性を維持し、各分割母集団が初期収束することを防いでいる。その後、各母集団ごとに遺伝的操作を再開する。

このアルゴリズムから分かる通り、データを分割することは非常に容易であり、並列処理に伴う通信量も極めて少ない。PC クラスタなどの分散並列機に適したモデルである。よって、多数の CPU を持つがネットワークは高速で無いようなアーキテクチャに適した最適化アルゴリズムであるといえる。

GAPPA ベンチでは実際には解析は行わず、評価部分では乱数を発生させているのみである。この1度の評価の際に発生させる乱数の回数で評価部の計算コストを調整する。使用 CPU 数、総個体数、世代数、評価部での計算コストのパラメータにより任意に変動できるように設計されている。

3.4 カットオフのある粒子シミュレーション

情報処理学会が毎年行っている並列処理シンポジウムでは、同時にソフトウェアコンテスト (PDC) が開催されている。2001 年の PSC では、決められた並列計算機においてプログラム能力を競う本部門に加えて、用いるプラットフォームやソフトウェアに制限のない自由部門も開催された。提出済のソフトウェアを使用することで、アーキテクチャのベンチマークとして使用することができる。2001 年度の PSC の内容は、カットオフのある粒子シミュレーションで、 n 個の質点の質量および初期速度、位置が与えられ、ある時間ステップ毎の質点の位置を求めていき、最終的な質点の位置を求めるといったプログラムであった。

4 ベンチマーク結果

ここでは紙面の制約上、GAPPA Bench とカットオフのある粒子シミュレーションによる結果を示す。

4.1 GAPPA Bench

図 1 に GAPPA Bench の結果を示す。これは 100bit 長の遺伝子を 1024 個体使用し、交叉率 0.8、突然変異率 0.01、移住率 0.3、移住間隔 5 のパラメータで行ったものである。なお結果は 10 試行の各試行の最高目的関数値の平均を示しており、load パラメータの大きなものほど評価の際に時間を必要とする。また、分割母集団数は使用する

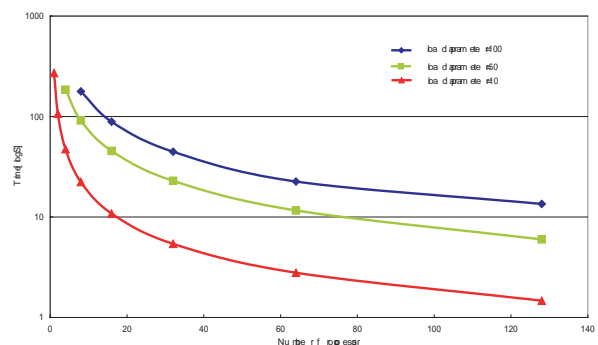


図 1: Result of GAPPA

この結果からもわかるように、Cambria は低速なネットワークのみを有しているにもかかわらず、使用するノード数が増加しても処理速度を向上させることが可能である。この傾向は、計算負荷が高いときだけでなく、計算負荷が低い際も同様である。

これは GA の分割母集団モデルが分散メモリ型の並列計算機に適しているからで、並列計算機の利用が普及する今後は、最適化の分野において GA はますます重要となる。

4.2 カットオフのある粒子シミュレーション

PSC の規定に従って計算を行った。その結果、丸山氏の作成したプログラムにより Gregor は 64node(128CPU) で 16.69[s]、Cambria は 10node(10CPU) で 253.5[s] であった。

Cambria は 256CPU のマシンであるが、10 ノード以上では性能が出なかった。

これは、本シミュレーションが頻繁にネットワーク通信を必要とするものであり、高速なネットワークを有しているアーキテクチャが高い性能を発揮したものである。

5 おわりに

並列化効率を上げるためには、教科書にある通り、ノード間での通信をできるだけ行わないことが重要となる。

本研究では 2 種の PC クラスタの性能評価を行った。Cambria は多ノードであるが低速なネットワークのみを有しているマシンである。通信量が多量に発生するベンチマークではまったく性能が上がらなかったが、GA のベンチマークでは高い性能を発揮した。この結果から GA のような最適化アルゴリズムを用いることにより低速なネットワークを持つ PC クラスタを十分利用できるといえる。

なお、本研究は文科省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行った。

