

# PCクラスタによる大規模計算におけるネットワーク環境の影響

中央大学 学生員 立石 絢也  
中央大学 正会員 櫻山 和男

## 1. はじめに

現在、数値解析の高速化手法として並列計算が多く行われている。並列計算機は、商用高性能並列計算機と、低価格PCを組み合わせたPCクラスタが主に用いられている。商用並列計算機においては、プロセッサの高速化に伴いネットワークも高性能なものに置き換わっているが、PCクラスタにおいては、低い通信性能、100Mbps程度のFast Ethernetを用いることが多いため、プロセッサが高速化してゆくにつれて通信性能がボトルネックとなってきた。そこで本報告では、Fast Ethernetよりも高速な通信ハードウェアであるMyrinetやGigabit-Ethernetを用いたPCクラスタにより有限要素法の並列計算を行い、従来のFast Ethernetを用いた場合との並列化効率の比較検討を行った。

## 2. 並列計算機

計算機として、新情報処理開発機構(RWCP)が所有するSCore Cluster IIを使用した。SCore Cluster IIは、Pentium III 800MHzを2つ搭載する計算ノードを64台結合したPCクラスタで、転送速度12.5MB/秒のFast Ethernetに加えて、転送速度125MB/秒のGigabit-Ethernet、転送速度160MB/秒のMyrinetを使用することができる。Fast Ethernetは、データをパケットに分割して通信を行うが、大量のパケットが同時に送信されるとパケットの損失が生じ、データの再送をする必要が出てくる。Myrinetでは、インターフェイスボード上にプロトコル処理専用のプロセッサを搭載している。また、パケットのハードウェアレベルでの到着を保証しており、10 $\mu$ 秒以下の低データ遅延と高速通信を実現している。Gigabit Ethernetは、転送速度を1Gbpsに引き上げた規格のEthernetであり、Myrinetに比べ低いコストで高い性能の通信環境をつくることことができる。

## 3. 数値解析手法

### (1) 基礎方程式と離散化

無次元化された非圧縮粘性流体の運動方程式と連続の式は、それぞれ式(1)、(2)で表される。

運動方程式；

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} - \frac{1}{Re} \frac{\partial}{\partial x_j} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = 0, \quad (1)$$

連続式；

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (2)$$

ここで、 $u$ は流速、 $p$ は圧力、 $Re$ はReynolds数を表している。基礎方程式に対して空間方向の離散化にP1/P1

要素を用いた安定化有限要素法(SUPG/PSPG法)<sup>1)</sup>を、また時間方向をCrank-Nicolson法により離散化することで以下の式を得る。

$$(M + M_\delta) \frac{u_i^{n+1}}{\Delta t} + \frac{1}{2} \left\{ (K(u_j^n) + K_\delta(u_j^n)) u_i^{n+1} + \frac{1}{Re} S u_i^{n+1} \right\} - (C - C_\delta) p^{n+1} = b_i^n, \quad (3)$$

$$C^T u_i^{n+1} + M_\varepsilon \frac{u_i^{n+1}}{\Delta t} + \frac{1}{2} K_\varepsilon(u_j^n) u_i^{n+1} + C_\varepsilon p^{n+1} = d_i^n, \quad (4)$$

ここでM,K,C,Sは係数行列である。添字 $\delta, \varepsilon$ は、それぞれSUPG項、PSPG項に起因するものである。また、 $b_i^n, d_i^n$ はそれぞれ運動方程式、連続式の既知項をまとめたものである。式(3),(4)を連立方程式としてBi-CGSTAB法<sup>2)</sup>を用い計算する。係数行列には記憶容量の削減のためのElement-by-Element処理を行っている。時間方向の離散化にはCrank-Nicolson法を使用している。数値解析例として、3次元立方体キャビティ内対流問題を取り上げ、並列化性能評価を行った。

### (2) 解析条件

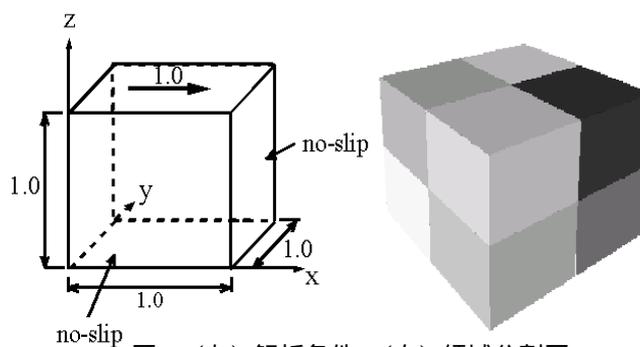


図-1 (左) 解析条件, (右) 領域分割図

解析条件として、図-1(左)のような一辺の長さが1.0の立方体の上面に一定流速を与える。境界条件は上面で $u=1.0$ 、その他の壁面上でnon-slip条件とした。有限要素分割は、各辺を32分割したMesh S(節点数35,937、要素数196,608)と、64分割したMesh L(節点数274,625、要素数1,572,864)を用いた。レイノルズ数は1000とし、微小時間増分量を0.001とした。

### 4. 並列計算法

計算領域をブロック型分割により領域分割して各ノードに割り当て、領域間の通信をMPIにより行い並列化した。1つのノードは2つのプロセッサを持っているため、各ノード1領域のみ割り当てた場合(Single)と、各ノードに2領域を割り当て共有メモリ並列化も行った

場合 (Dual) について、それぞれ Fast Ethernet, Gigabit Ethernet, Myrinet を用いて通信を行い計算した。Mesh S の場合には領域を 1~16 領域に分割した。Mesh L の場合には 16~128 領域に分割した。8 領域のときの領域分割図を図-1(右)に示す。

## 5. 並列化性能評価

Mesh S (Single) の場合と Mesh S (Dual) の場合の演算速度倍率と並列化効率を図-2, 図-3 に示す。台数が増えるにしたがって Fast Ethernet を用いたものは効率が低下しているが, Myrinet, Gigabit Ethernet を用いたものは高い効率を示している。これは、分割数を増やすにつれて隣接プロセッサ間の通信量と全プロセッサ間の通信量が増え、またパケットの損失による再送が生じて同期待ちの時間が多くなったために効率が下がったものと考えられる。Gigabit Ethernet と Myrinet との比較では Myrinet を用いた方が若干ではあるが高い性能を示している。各ノード 2 つのプロセッサを用いた場合は、1 つの場合に比べて 20% ほど効率が下がっている。これは、非対称行列の計算であるため、プロセッサ-メモリ間のデータ転送量が多くノード内の 2 つのプロセッサ間でメモリアクセスの競合を起しているためと考えられる。

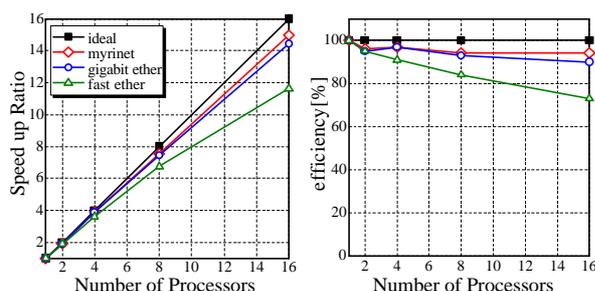


図-2 (Single) Mesh S (左) 演算速度倍率,(右) 並列化効率

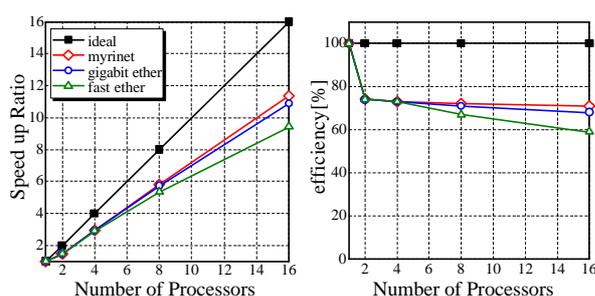


図-3 (Dual) Mesh S (左) 演算速度倍率,(右) 並列化効率

Mesh L (Single) の場合と Mesh L (Dual) の場合の演算速度倍率と並列化効率を図-4, 図-5 に示す。なお、この場合には 16 領域の計算 (各ノード 1 プロセッサで 16 ノード使用) のものを基準としている。32 分割の場合に効率が 100% を超えているのは、16 領域の計算を基準としているために、各辺を 4 分割して立方体で領域分割できる 32 領域の場合の計算が 16 領域の場合よりも高い効率の計算になっているためと考えられる。Mesh L の場合にも、Fast Ethernet では分割数が増えるにつれ

て性能が低下しているが、Gigabit Ethernet や Myrinet では高い性能を示している。この場合にも、分割数を増やすことによる隣接プロセッサ間の通信量と全プロセッサ間の通信量が増え、Fast Ethernet ではパケットの損失による再送が生じて同期待ちの時間が多くなったために効率が下がったものと考えられる。

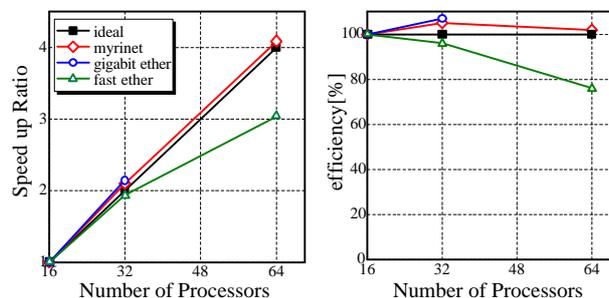


図-4 (Single) Mesh L (左) 演算速度倍率,(右) 並列化効率

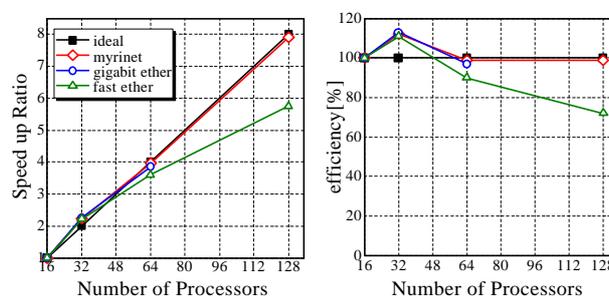


図-5 (Dual) Mesh L (左) 演算速度倍率,(右) 並列化効率

## 6. おわりに

本報告では、PC クラスタを用いた有限要素法の並列計算におけるネットワーク環境の違いが並列化性能に与える影響について調べ、以下の結論を得た。

- Fast Ethernet を用いた場合には、領域分割数を増やしたときに通信量が増えて並列化効率が下がるが、Gigabit Ethernet や Myrinet を用いた場合には、高い並列化効率を得られた。
- ノード内の 2 つのプロセッサを使用すると、メモリアクセスの競合により性能が低下することがわかった。

今後は、共有メモリ通信に OpenMP を用いた場合の並列化効率について検討をしていく予定である。

謝辞: 本研究を進めるにあたり並列計算機 SCore Cluster II の使用を許可された技術研究組合 新情報処理開発機構 (RWCP) および並列分散システムソフトウェア研究室 石川 裕様に厚く感謝いたします。

## 参考文献

- 1) T.E.Tezduyar, S.Mittal, S.E.Ray and R.Shih : "Incompressible flow computations with stabilized bilinear and linear equal-order-interpolation velocity-pressure elements" : Computer Methods in Applied Mechanics and Engineering : 95 : pp221-242 : 1992.
- 2) 福井 義成, 野寺 隆志, 久保田 光一, 戸川 隼人 : "新数値計算" : 1999