

## I-A92

## PCクラスターを用いた仮想並列計算機の構築とその有効性の検討

中央大学大学院 学生員 ○須江 克章  
 日本工営株式会社 正会員 櫻庭 雅明  
 中央大学 正会員 櫻山 和男

## 1. はじめに

近年、偏微分方程式の大規模数値計算に対して計算時間の短縮と記憶容量の分散を可能にする並列計算が広く行われるようになってきている。しかし、現在の並列計算環境は商用並列計算機が依然高価なものであるため、利用者数や使用時間が制限されている状況にあり、手軽に並列計算を行うことができない。

そこで本報告では、PC(Personal Computer)8台をEthernetで接続し、通信ライブラリにMPI(Message Passing Interface)を搭載した仮想並列計算機<sup>1)</sup>を構築する。そして、要素数の異なる2種類のメッシュを用いて並列有限要素法を行い自作並列計算機の性能評価、特にHUBの性能の差異が並列計算効率に及ぼす影響について検討を行った。

## 2. 並列計算機システム

並列計算機システムはメモリの割り当て方から2つの種類に分類できる。

- 共有メモリ型並列計算機
- 分散メモリ型並列計算機

共有メモリ型並列計算機は複数のCPUが1つのメモリを共有するシステムであり、データ分散を意識したプログラミングを行う必要がない。しかし、全プロセッサのアクセスが1つのメモリに集中するため、プロセッサ数の拡張には限界がある。一方、分散メモリ型並列計算機はCPU-メモリから構成されるノードをネットワークで接続したものであり、ノードを増やしやすく拡張性が高い。しかし、プロセッサは同じノード上のメモリしかアクセスできないため、他のノード上にあるデータを利用するためにはノード間でメッセージを介して通信する必要がある。そのため、MPIやPVM(Parallel Virtual Machine)などの通信ライブラリを用いて、分散データを意識したプログラミングを行う必要がある。また、ノード間のネットワーク通信がボトルネックとなり十分な並列化効率を得られない場合がある。

自作並列計算機は各ノードをPCに置き換えることで仮想的に並列計算機として機能させるものであり、分散メモリ型並列計算機の種類に入る。PCを用いることでコストパフォーマンスが高く、構成に柔軟性があるシステムを構築することができる。

## 3. 自作並列計算機の構成

## (1) PC構成

各PCには、市販の部品を使って組み上げた自作PCを用いた。PCの構成を表-1に示す。99年2月時点でPC1台あたり約13万円で構築した。

表-1 PC仕様

|       |                                |
|-------|--------------------------------|
| CPU   | Pentium II 400MHz              |
| Cache | 512KB                          |
| RAM   | 256MB (DIMM 128MB×2)           |
| NIC   | DEC DC21x4x PCI(10/100Base-Tx) |

次に使用したソフトウェアについて説明する。一覧を表-2に示す。OSには安定性、性能、コストの面からPC-UNIXの一種であるLinux<sup>2)</sup>を選択した。通信ライブラリにはLinuxで動作するMPICHを用いた。また、ネットワークシステムとしてNFS(Network File System)を採用したことで、PC台数を意識することなくプログラムの編集、コンパイルを行うことができる。これらを使用したソフトウェアは全てGNUライセンスに基づくフリーウェアであるため、ソフトウェアについてはコストをかけることなく構築することができる。

表-2 ソフトウェア一覧

|            |              |
|------------|--------------|
| OS         | Linux-2.0.34 |
| コンパイラ      | g77,gcc      |
| 通信ライブラリ    | MPICH1.1.1   |
| ネットワークシステム | NFS          |

## (2) ネットワーク構成

PC8台をEthernetで接続し、その内の1台をGateway PCとすることにより、他のネットワークへの接続を実現している。PC間のネットワーク環境を比較するためにHUBには、10Base-TのHubと100Base-TXのHubと100Base-TXのスイッチングHUBの3種類を用意した。全体の構成図を図-1に示す。

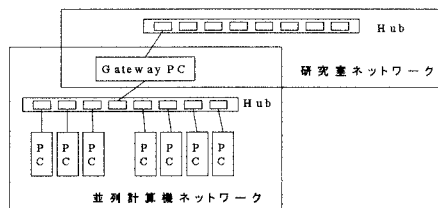


図-1 ネットワーク構成図

KeyWords: Personal Computer, 並列計算機, 有限要素法  
 〒112-8551 東京都文京区春日 1-13-27  
 TEL. 03-3817-1815 FAX. 03-3817-1803

## 4. 数値解析例

数値解析例として、安定化有限要素法<sup>3)</sup>による非圧縮粘性流れ解析を行った。仮想モデルとして、内部に構造物を含まない2平面間の一様流の解析を行った。境界条件として、両平面ともSlip条件を与えた。有限要素分割には、要素数の異なる2種類のメッシュ(MESH S:80,000, MESH L:320,000)を用いた。領域分割には節点数、要素数ともに均等にスライス分割で行った。微小時間増分を0.01とし、ステップ数を10回とした。並列処理の詳細については参考文献 4)を参照。

ネットワーク環境性能(HUBの差異)とプロセッサ数を変化させた時のCPU時間と通信時間の内訳を表-3に示す。10Base-T環境ではプロセッサ数を増やすことで100Base-TX環境と比べて効率が大幅に低下していることがわかる。その理由として、100Base-TX環境では伝送速度は100Mbps(理想速度)であるのに対し、10Base-T環境では伝送速度が10Mbps(理想速度)と遅いため、プロセッサ数が増えるにつれてデータの衝突の頻度が増え、効率が低下したと考えられる。

次に100Base-TX環境でのスイッチング機能の効果について考える。スイッチングHUBとは出力ポートが異なる場合に複数の入力をデータの衝突なしに同時に転送できる機能を有しているHUBのことであるが、今回の数値解析例ではスイッチング機能なしのHUBの方が高い並列化効率を示した。その理由についてはまだ不明なので今後の課題とする。

プロセッサ数を変化させたときの演算速度倍率を図-2、並列化効率を図-3に示す。100Base-TXスイッチング機能なしのHUBでは、MESH S, Lともにプロセッサ数が増加するに従い高い効率を示しており、大規模問題での本システムの有効性を示すことができた。

## 5. おわりに

本報告ではPCクラスターによる並列計算機を構築した。性能評価のために要素数の異なる2種類のメッシュを用いて並列有限要素法を行った結果、以下の結論を得た。

1. 10Base-Tと100Base-TX環境では大きな性能の差異が見られ、PC間のネットワークには100Base-TX環境を整える必要があるといえる。
2. 100Base-TXスイッチング機能なしのHUBでは、大規模問題に対して高い演算速度倍率と並列化効率を得ることができ、本計算機システムの有効性を示すことができた。

今後の課題として、スイッチングHUBでの並列化効率低下の理由を明らかにするとともに、性能の異なるPCが混在する場合の領域分割法を検討する。

表-3 CPU時間と通信時間の内訳

| PE                    | 100Base-TX<br>Switch Hub |                  | 100Base-TX<br>No Switch Hub |                  | 10Base-T<br>No Switch Hub |                   |
|-----------------------|--------------------------|------------------|-----------------------------|------------------|---------------------------|-------------------|
|                       | 全 CPU                    | 通信時間             | 全 CPU                       | 通信時間             | 全 CPU                     | 通信時間              |
| MESH L (総反復回数 7092 回) |                          |                  |                             |                  |                           |                   |
| 1                     | 5177.6                   | 0.0              | 5177.6                      | 0.0              | 5177.6                    | 0.0               |
| 4                     | 1470.4                   | 205.4<br>(13.9%) | 1836.6                      | 63.4<br>(4.7%)   | 1889.6                    | 612.4<br>(32.4%)  |
| 8                     | 1243.2                   | 538.5<br>(43.3%) | 832.8                       | 125.3<br>(15.0%) | 2376.1                    | 1662.4<br>(70.0%) |
| MESH S (総反復回数 2981 回) |                          |                  |                             |                  |                           |                   |
| 1                     | 565.1                    | 0.0              | 565.1                       | 0.0              | 565.1                     | 0.0               |
| 4                     | 158.3                    | 24.0<br>(15.2%)  | 147.3                       | 13.7<br>(9.3%)   | 250.9                     | 113.0<br>(45.0%)  |
| 8                     | 207.5                    | 134.3<br>(64.7%) | 104.9                       | 31.0<br>(29.5%)  | 451.5                     | 370.1<br>(82.1%)  |

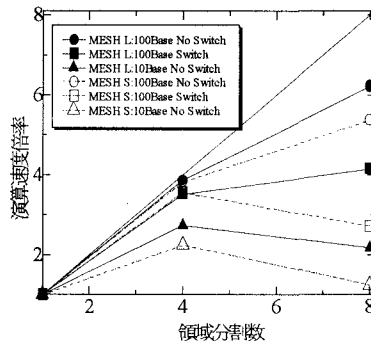


図-2 演算速度倍率図

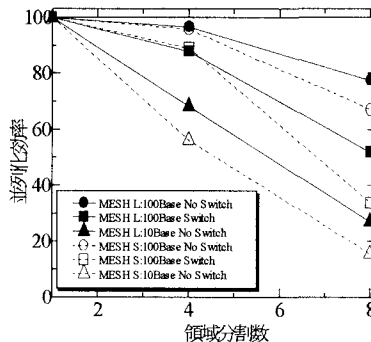


図-3 並列化効率図

## 参考文献

- 1) 塩谷隆二, 矢川元基, "自作並列計算機とその計算工学への応用" 計算工学講演会論文集 Vol.2, pp155-158, 1997.
- 2) はねひでや, やまだあきら, あべひろのぶ, "るんるんLinux": アスキー出版局.
- 3) 樫山和男, 猪股渉, "混合補間要素を用いた非圧縮粘性流れ解析のための高精度安定化有限要素法", 土木学会論文集 No.591/I-43, pp125-137, (1998).
- 4) 玉井典, 樫山和男, "非構造格子に基づく三次元非圧縮粘性流れの並列有限要素解析" 第11回数値流体力学シンポジウム論文集, pp557-558, 1997.