

1. 序論

近年では、計算機がなくてはならない存在になっており、大容量の記憶デバイスの登場や、ネットワークの高速化などにより、大量のデータを高速に処理することが求められている。それらを処理するには、膨大な時間がかかる。高速な処理を行うためには、複数のプロセッサを持つ並列計算機 (Parallel Computer) が用いられる。しかし、一般的に並列計算機は高価であるために、容易に用いることはできない。そこで、複数の計算機をネットワーク接続して仮想的並列計算機とする手法が注目されている。本研究では、MPI [Message Passing Interface] ¹⁾を用いて仮想並列計算を行い、その実用性を検証する。

2. 研究内容

本研究では、膨大な行列計算を1つのコンピュータで計算する場合と、MPIを用いて複数のコンピュータに計算する場合と分けて、処理時間にどれだけ差があるかを計測する。MPIとは Message Passing Interface の略であり、分散メモリ型の並列計算機において、複数のプロセッサ間で、データのやりとりをするために用いる、メッセージ通信操作の仕様標準である。

計算方法は、8個の正方行列 $A_1 A_2 \dots A_8$ の行列積

$\prod_{k=1}^8 A_k$ を4台のコンピュータを用いて計算し、その計

算時間を測定する。このとき、行列のサイズを $10 \times 10, 100 \times 100, 500 \times 500, 1000 \times 1000$ と変化させ、サイズごとに測定を行う。ホストコンピュータがすべての行列を保持している。ホストコンピュータは、他のサブコンピュータに生成した行列を2つずつ送信する。各コンピュータにて、受け取った行列同士の積を計算し、その結果を上位コンピュータに渡し、また計算を繰り返す。最後に、ホストコンピュータに送り、それらをホストコンピュータで計算する。この結果を表示するまでの時間を計測し、この作業を1台で行ったときと、複数台で行ったときの計測時間を判定する。

3. 結果・考察

表1: 行列計算の処理時間の平均(秒= s)

CPU\行列数	10	100	500	1000
1台	0.015s	0.066s	10.87s	167.2s
4台	0.012s	0.051s	7.29s	86.3s

上記の表1は行列計算を行った時の処理時間の平均を表している。表1よりCPU数1台のときと比べると、どれも大幅に速度が向上しており、特に処理の数が大きければ大きいほど、処理時間の向上率は上がっていることが示される。これは処理の送受信や、同期の時間にかかる時間があるため、それらの影響を受けやすいからと考えられる。1台の計算時間はややばらつきがあることに対し、複数のコンピュータによる処理は安定した処理速度であった。これは、複数のコンピュータの場合よりも、1台のコンピュータにかかる負担が大きくなるためであると考えられる。

4. 結論

本研究により、並列計算環境を用いて複数のコンピュータを使って膨大なデータの処理を行うことと、1台のコンピュータでその処理を行った時を比べると、大幅に処理速度が向上することを確認出来た。この並列環境を用いれば、今まで処理時間が膨大だった問題も、大幅に速度を向上して、さらにたくさんの情報を扱うことができると考えられる。

参考文献

- 1) P パチュコ 著, 秋葉博 訳: MPI 並列プログラミング, 培風館 (2001)
- 2) 渡邊真也 著: MPI による並列プログラミングの基礎
- 3) MPICH2 on Windows Local :
<http://ums.futene.net/wiki/Paralell/4D5049434832206F6E2057696E646F7773204C6F63616C.html>