卒業研究報告書

題目

HBSP モデル上での 行列積を求めるアルゴリズム

指 導 教 員

石水 隆 助教

報告者

06-1-037-0142

吉岡　健太

近畿大学理工学部情報学科

平成 22 年 2 月 5 日提出

概要

本研究では、HBSP(Heterogeneous Bilk-Synchronous Parallel)[1]モデル上で行列積を求める場合、各プロセッサへの最適なデータの割り当て方を求める。

HBSP モデルは非同期式並列計算モデルの一つであり、Williams によって提唱された BSP[2]モデルの拡張モデルである。HBSP モデルはネットワークと局所メモリを持つ性能の異なる複数のプロセッサから成り立ち、プロセッサの処理速度に応じて処理を割り当てる。HBSP モデルどの通信は通信コスト g、バリア同期時間 L といったパラメタにより抽象化されている。

目次

1　序論

1.1 並列処理

現在、地球環境のシミュレーション、天気予報で用いられる数値予報、量子力学に基づく分子設計、原始物理、宇宙物理のシミュレーション、流体の計算、経済・社会システムのシミュレーション、コンピュータグラフィックス、広域通信網、電子網等の最適設計、画像処理、大規模データベース、大規模知識ベースの検討などの様々な分野で、膨大な量の問題を高速かつ短時間で処理することが求められている。問題を高速で処理するには 2 通りの方法が挙げられる。1 つは使用するプロセッサの性能の向上である。近年ではプロセッサの発達が進み、処理速度が目覚しい向上を遂げた。しかし、これにも限界があり今以上の処理速度の著しい向上は期待できない。もう 1 つは、並列処理を用いた処理の並列化である。

並列処理とは、1 つの問題に対する処理を小さな処理に分割し、それを複数のプロセッサで並列に実行することをいう。近年、コンピュータハードウェアの値下がりによって複数のプロセッサを用いた並列計算機を構築することが可能になってきた。こういった傾向が進むにつれ並列処理による問題の処理への必要性が大幅に高くなっていくと思われる。

1.2 並列アルゴリズム

アルゴリズムとは数学、コンピューティング、言語学、あるいは関連する分野において、ある特定の目的を達成するために必要な作業の倫理や効率的手順を述べたものである。

並列処理を行うためには、プロセッサ間のデータのやり取りやメモリへのアクセス、プロセッサ間の同期等、並列特有の問題を解決せねばならない。このため、従来の逐次処理で用いられてきた逐次アルゴリズムをそのまま並列処理に用いることはできず、並列処理専用のアルゴリズム、すなわち並列アルゴリズムが必要となる。

1.3 並列計算機

複数のプロセッサを用いることで並列処理を行なうことが可能な計算機を並列計算機という。並列計算機には、全てのプロセッサが共通のメモリに対して読み書きを行い、プロセッサ間の通信はメモリを通して使用する共有メモリ型並列処理や、それぞれのプロセッサは個々に局所メモリを持ち、通信にはネットワークを使用する分散メモリ型並列処理の 2 つに大きく分けることが出来る。

その中で、共有メモリ型並列処理は同期問題やデータの送受信といった問題がメモリを共有しているため対処しやすいが、プロセッサの増減などの変化があった場合には、メモリに全てのプロセッサを接続させることが困難になってしまう。そのため、現在では局所メモリが使用できる分散メモリ型並列処理が主流となっている。

1.4 本報告書の構成

本報告書の構成を以下に述べる。 2 章では本研究で使用する HBSP モデルについて述べる。 3 章ではHBSP モデル上での行列積において、計算により各プロセッサへの最適なデータの割り当て方および計算量を述べる。 4 章では結論と今後の課題について述べる。

2　BSP モデルと HBSP モデル

2.1　BSP モデル

BSP モデルは非同期式並列計算モデルであり、以下の要素から構成される。

・局所メモリを持つ複数のプロセッサ
・プロセッサ間の 1 対 1 メッセージ通信を行う完全結合網
・プロセッサ間の同期を実現するための同期機構

図 1 に BSP モデルの概念図を示す。

次に、BSP モデルは以下のパラメタを持つ。

・p^2：プロセッサ数
・g：通信コスト
・L：バリア同期時間

BSP モデルだは基本的命令の実行については、各プロセッサは 1 単位時間に 1 個の内部命令を局所メモリ上のデータのみを用いて実行し、h 個のメッセージ、送信命令または受信命令の実行は gh 単位時間で行われると仮定されている。

また、BSP モデルのアルゴリズムは、スーパーステップと呼ばれる命令列のとして表される。スーパーステップとはプロセッサの同期を取るのに実行する命令の列であり、各スーパーステップにおいて各プロセッサはスーパーステップ開始時におけるメモリの内容にのみ内部計算命令を実行し、また送信命令を実行できるスーパーステップの全ての命令を実行すると、全てのプロセッサでバリア同期をとり、次のスーパーステップに実行が移る。実行についてはあるスーパーステップで送信されたメッセージは次のスーパーステップの開始時に受信されると仮定されている。

あるスーパーステップで各プロセッサ P_i が w_i 個の内部計算命令 h_i 個の送信命令または受信命令を実行するとき、そのときの実行時間は BSP モデルでは $O(\max_{0\leqq i<p^2}\{w_i+gh_i\}+L)$ と仮定されている。BSP アルゴリズム上での実行の概念図を図 2 に示す。

2.2　HBSP モデル

HBSP(Heterogeneous Bilk-Synchronous Parallel)モデルは BSP の拡張モデルであり、以下の要素から構成される。

・局所メモリを持つ性能の異なる複数のプロセッサ
・プロセッサ間の 1 対 1 メッセージ通信を行う相互結合網
・全体あるいは一部のプロセッサ間で同期をとるための同期機構

図 3 に HBSP モデルの概念図を示す。

次に、HBSP モデルは以下のパラメタを持つ。

・p^2：プロセッサ数
・g：通信コスト
・L：バリア同期時間
・s_i：プロセッサ Pi の速度

$$s_0 \geqq s_1 \geqq \cdots \geqq s_{p^2-1}=1$$

・s：プロセッサの速度合計 $s=\Sigma s_i$

本研究では簡単のために通信コストと各プロセッサの処理速度が比例すると仮定する。

$$g_0{:}s_0=g_1{:}s_1=g_2{:}s_2=\cdots=g_{p^2-1}{:}s_{p^2-1} \quad \cdots\cdots (1)$$

HBSP モデルでは基本的命令の実行については、プロセッサ Pi は 1 単位時間に s_i 個の内部計算命令を局所メモリ上のデータのみを用いて実行し、h 個のメッセージの送信命令または受信命令の実行は $g_i h$ 時間で行われると仮

定されている。

HBSP モデル上での並列アルゴリズムも、BSP モデル上の並列アルゴリズムと同様にスーパーステップの列として表される。

あるスーパーステップで各プロセッサ P_i が w_i 個の内部計算命令、h_i 個の送信命令または受信命令を実行するとき、その実行時間は HBSP モデルでは $O(\max_{0 \leqq i < p2}\{w_i/s_i+g_i h_i\}+L)$ と仮定されている。

HBSP アルゴリズム上での実行の概念図を図 4 に示す。

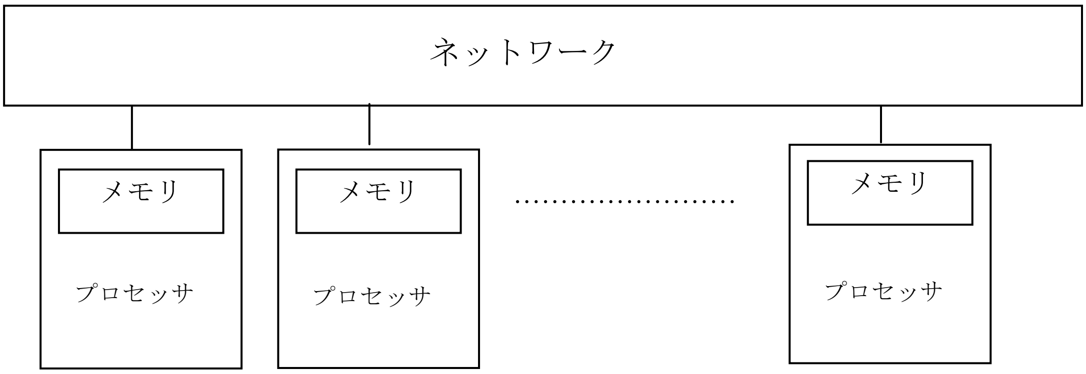

図 1　BSP モデルの概念図

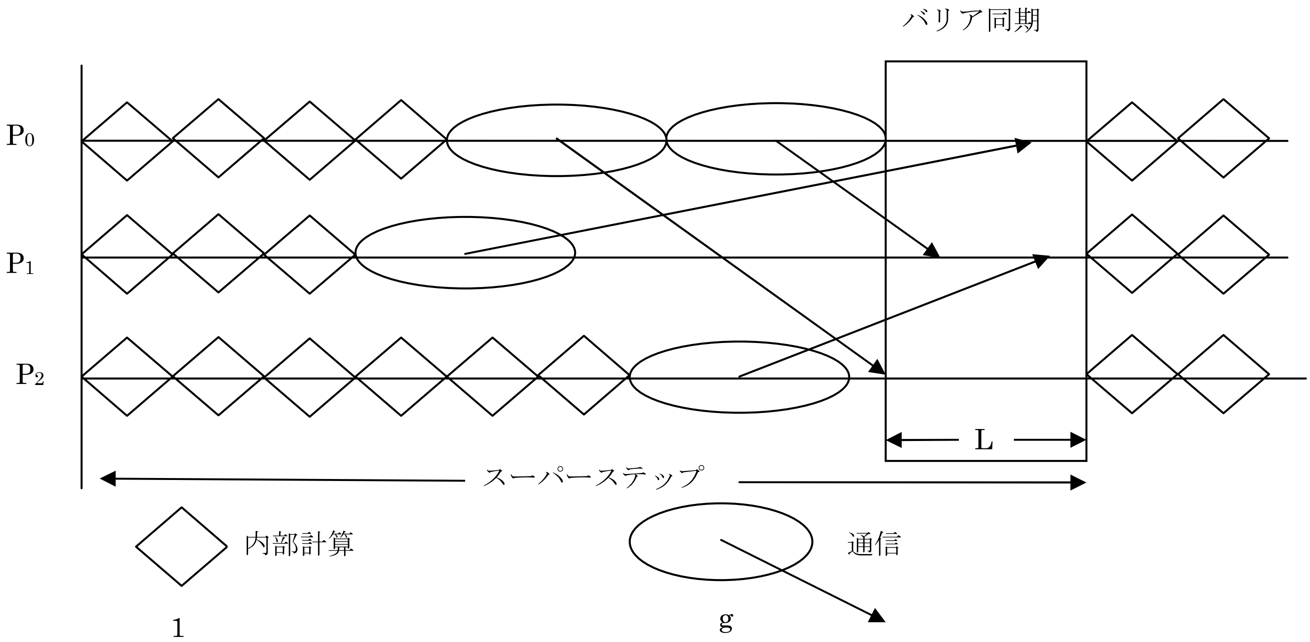

図 2　BSP アルゴリズム上での実行の概念図

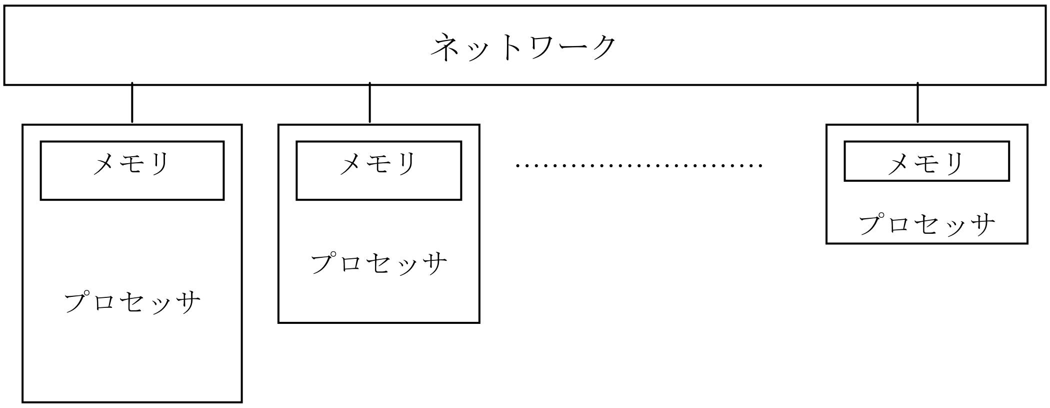

図 3　HBSP モデルの概念図

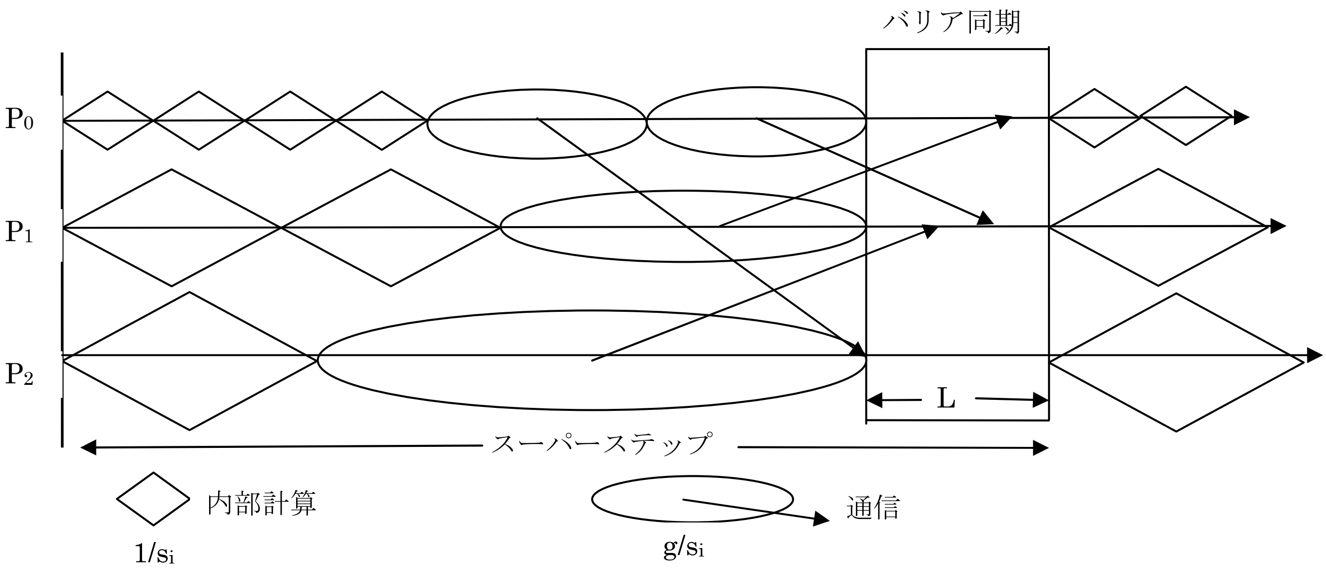

図 4　HBSP アルゴリズム上での実行の概念図

3　HBSP モデル上での行列積

本研究では HBSP モデル上で n*n の行列積を求める効率の良い並列アルゴリズムを設計する。

式(1)より、プロセッサ P_i はプロセッサ P_{p^2-1} に比べ、内部計算命令および送信命令共に s_i 倍高速に実行できる。HBSP モデル上で行列積を求める場合、n*n の行列積 C=A*B に対して、各行列A, B ,C をそれぞれ s 個の部分行列に分割し、速度 $s_i(0 \leqq i < p^2)$ であるプロセッサ p_i に対して s_i 個の部分行列を割り当てる。各プロセッサは割り当てられた部分行列の計算を担当する。

入力行列 A,B の各プロセッサへのデータの割り当て方は以下の３通りが考えられる。

α：それぞれサイズ $n/s^{0.5}*n/s^{0.5}$ 個の部分行列に分割し、各プロセッサに１つの部分行列を割り当てる
β：それぞれサイズ n/s*n 個の部分行列に分割し、各プロセッサに１つの部分行列を割り当てる
γ：それぞれサイズ n*n/s 個の部分行列に分割し、各プロセッサに１つの部分行列を割り当てる

入力行列 A,B および解行列 C がそれぞれα、β、γの３通りの分割パターンがあるため、割り当て方の組み合わせのパターンは 3^3=27 通りがある。また行列 B が送信に必要なデータ数は行列 A を縦横を入れ替えたものと等しく、解行列 C についてはβとγが向きが異なるだけなので、入力行列 A と解行列 C の組み合わせについて、通信計算量についてのみ考慮すればよい。したがって、入力行列 A の割り当て方３通りと解行列 C の割り当て方３通りにより C=A*B の組み合わせは９通りの組み合わせについて求める。

図３　部分行列の分割パターン

(証明)

以下は入力行列 A の部分行列へのデータの各プロセッサへの割り当て方および解行列 C に部分行列の計算担当範囲の各プロセッサへの割り当て方をα, β, γそれぞれ分割した場合の通信計算量を求める。

(1)A の割り当て方α、C の割り当て方αの場合

入力行列 A の部分行列を持つプロセッサが送信するデータ数は $n/s^{0.5}*n/s^{0.5}$

$s^{0.5}$ 個のプロセッサに対して送信する。

よって、全体の通信計算量は $O(g*n^2/s^{0.5}+L)$となる。

(2)A の割り当て方β、C の割り当て方αの場合

入力行列 A の部分行列を持つプロセッサが送信するデータ数は n/s*n

$s^{0.5}$ 個のプロセッサに対して送信する。

よって、全体の通信計算量は $O(g*n^2/s^{0.5}+L)$となる。

(3)A の割り当て方γ、C の割り当て方αの場合

入力行列 A の部分行列を持つプロセッサが送信するデータ数は $n/s*n/s^{0.5}$

s 個のプロセッサに対して送信する。

よって、全体の通信計算量は $O(g*n^2/s^{0.5}+L)$となる。

(4)A の割り当て方γ、C の割り当て方γの場合

入力行列 A の部分行列を持つプロセッサが送信するデータ数は n/s*n

s 個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2+L)$となる。

(5)A の割り当て方β、C の割り当て方γの場合
入力行列 A の部分行列を持つプロセッサが送信するデータ数は n*n/s
s 個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2+L)$となる。

(6)A の割り当て方α、C の割り当て方γの場合
入力行列 A の部分行列を持つプロセッサが送信するデータ数は $n/s^{0.5}*n/s^{0.5}$
s 個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2+L)$となる。

(7)A の割り当て方α、C の割り当て方βの場合
入力行列 A の部分行列を持つプロセッサが送信するデータ数は $n/s^{0.5}*n/s$
$s^{0.5}$ 個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2/s+L)$となる。

(8)A の割り当て方γ、C の割り当て方βの場合
入力行列 A の部分行列を持つプロセッサが送信するデータ数は n/s*n/s
s 個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2/s+L)$となる。

(9)A の割り当て方β、C の割り当て方βの場合
入力行列 A の部分行列を持つプロセッサが送信するデータ数は n/s*n
１個のプロセッサに対して送信する。
よって、全体の通信計算量は $O(g*n^2/s+L)$となる。

上記(1)から(9)より、解行列 C の割り当て方がαの時は $O(g*n^2/s^{0.5}+L)$、βの時は $O(g*n^2+L)$、γの時は $O(g*n^2/s+L)$となり、通信計算量は解行列 C の計算担当範囲の割り当て方に依存することが分かる。
また、行列 B のデータ送信については行列 A を縦横入れ替えたものと等しいので、解行列 C の割り当て方がαの時は $O(g*n^2/s^{0.5}+L)$、βの時は $O(g*n^2/s+L)$、γの時は $O(g*n^2+L)$となる。
よって、A と B のデータ送信を合計した場合、通信計算量が最小となる割り当て方は C=A*B に対してα＝α＊αの分割パターンにした場合であり、そのときの通信計算量は $O(g*n^2/s^{0.5}+L)$となる。
また、解行列の各要素は O(n)で計算可能で、各プロセッサは n^2/s 個の要素の計算を行うので解行列の計算は $O(n^3/s)$で行うことができる。よって、HBSP モデル上で行列積を求めるアルゴリズム計算量は $O(n^3/s+n^2/s^{0.5}+L)$となる。

4　結論・今後の課題

本研究では HBSP モデル上で効率よく行列積を求めるアルゴリズムを提案した。

HBSP モデル上で行列積を求める場合の各プロセッサへの最適なデータの割り当て方は行列積 C=A*B に対して各行列を $s^{0.5}*s^{0.5}$ 個の部分行列に分割し、プロセッサ P_i に s_i 個の部分行列を割り当てた場合である。このときの BSP モデル上で行列積を求める並列アルゴリズムの計算量は $O(n^3/s+n^2/s^{0.5}+L)$ が最適な並列アルゴリズムとなる。データを送信する最適な分割パターンは各部分行列が正方形になる場合であることが示された。

今後の課題としては、行列以外にも 2 次元配列になっているデータが存在するような問題に対するデータの分割の仕方を求めていくことが必要となる。

謝辞

本研究するにあたり、終始様々なご指導をしていただいた石水隆先生には心より感謝申し上げます。また、研究室の皆様には論文を書く上で助言などしていただき大変感謝します。

参考文献

[1] T.L Williams and R.J.Parsons,"The Heterogeneous Bulk Synchronous Parallel Model" in Workshop on Advance in Parallel and Distributed Computational Model 2000.

[2] L.G Valiant, "A Bridging Model for Parallel Computation" Communications of the ACM, Vol33, No8,pp103-111,1990.