

# F-2-3 BSP 上での並列マージソートアルゴリズム

01-1-26-043 尾原 勝博  
(情報論理工学研究室)

## 1. はじめに

本研究では、BSP (Bulk-Synchronous Parallel Machine) モデル<sup>[1]</sup>上で高速にソートを行う並列アルゴリズムを提案する。BSP モデルとは分散メモリ型の並列計算モデルであり、従来型の PRAM (Parallel Random Access Machine)<sup>[2]</sup>と異なり通信時間も考慮することができ、より現実の並列計算機に近いモデルとして注目されている。そのため現在様々な問題に対して BSP 上で動く並列アルゴリズムが求められている。

しかし、従来の PRAM のアルゴリズムは通信が考慮されておらず、これを BSP 上で実行させた場合効率良く実行できるとは限らない。従って、BSP モデル用の通信を考慮したアルゴリズムを設計する必要がある。

## 2. BSP モデル

BSP モデルは以下の 3 項目から成る。

- ① 局所メモリを持つ複数のプロセッサ
- ② プロセッサ間で 1 対 1 通信を行うネットワーク
- ③ プロセッサ間の同期機構

また、BSP モデルはプロセッサ台数  $P$ 、1 メッセージ当たりの送受信時間  $G$ 、同期時間  $L$  の 3 つのパラメータを持ち、 $N$  個のメッセージの送受信時間は  $NG+L$  時間と仮定されている。

## 3. アルゴリズム

本研究で提案するアルゴリズムは、マージソート<sup>[3]</sup>をベースとしている。マージソートとは、ソート済み配列  $A, B$  から 2 倍の長さのソート済み配列  $C$  を作るという操作を再帰的に繰り返していく  $O(M \log N)$  時間ソートである。

マージソートを BSP モデル上で実行する場合、図 1 のように各部分配列を各プロセッサに割り当て、並列にマージを行うことにより計算時間の短縮を図ることが出来る。

一般にプロセッサ台数が増加すると、内部計算時間は減少するが、通信メッセージ数および同期回数は増加し、通信時間が増加する。そこで本研究では、BSP モデルでマージソートを行ったとき最速となるプロセッサ台数を求める。

## 4. 結果および考察

本研究で提案した並列アルゴリズムは、 $P$  台のプロセッサを用いて BSP モデル上で

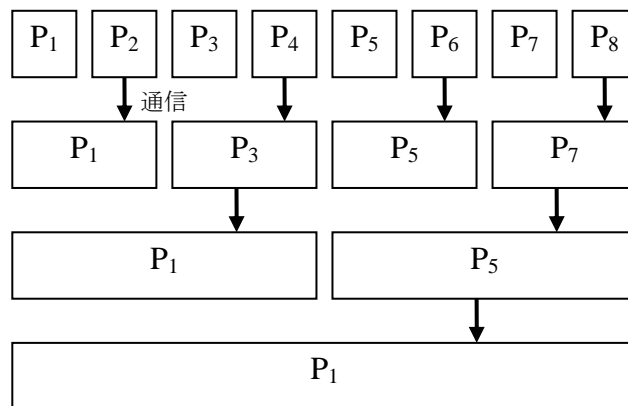


図 1 並列マージソートにおけるプロセッサへのデータの割り当て方

$$O\left(\frac{N \log N}{P}\right) + NG + L \log P$$

でソーティングを解く。このアルゴリズムは、 $G < \log N$  のときに有効なアルゴリズムとなり、

$$P = \frac{N \log N}{L}$$

のとき最適なプロセッサ数となる。

## 5. まとめ

本研究では、BSP モデル上でソーティングを行う並列アルゴリズムを提案した。このアルゴリズムは、

$$O\left(\frac{N \log N}{P}\right) + NG + L \log P$$
 時間でソーティングを

行う。また、 $P = \frac{N \log N}{L}$  のとき最適なプロセッサ数となる。

## 参考文献

- [1] L. G. Valiant: "A Bridging Model for Parallel Computation," Communication of the ACM, Vol. 33, No. 8, pp. 103-111 (1990)
- [2] Joseph JáJá: "An Introduction to Parallel Algorithms," Addison-Wesley (1992)
- [3] 石畑清: "アルゴリズムとデータ構造," 精興社, pp. 197-224 (1989)