

トーラス型並列計算機の通信操作の性能評価*

1N-2

牧野 浩之 渋谷 進†

茨城大学工学部情報工学科‡

1 はじめに

ハードウェア的に逐次処理の速度が限界に近付いており、そこで逐次処理に代わって並列処理が求められている。並列処理の通信操作にかかる時間を知ることは並列度を決定する際に重要なデータとなり、ユーザが並列プログラミングを行う際の助けとなる。

そこで本研究では、並列計算機 AP1000 を用いて通信操作の評価を行った。基本的な通信操作である対一通信性能を測定し、セルの位置と通信速度の関係を調査した。また、集合通信であるブロードキャストや分配・収集、更に通信操作が多いトータルエクスチェンジの通信性能を測定した。

通信時間に関する近似式を導き出すことにより、AP1000 を用いて並列プログラムを組む際に、効率的な並列度や通信時間の予測が可能となる。

2 準備

2.1 並列計算機アーキテクチャ

本研究では、8×2個のセルから構成される並列計算機 AP1000 を用いて測定した(図1)。AP1000 は、T-Net・B-Net・S-Net のネットワークから成り立つ。

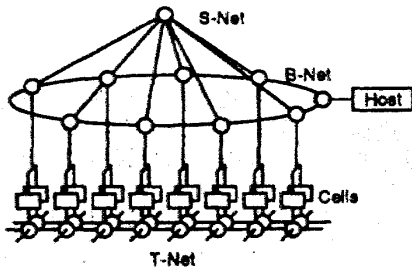


図1: AP1000 のアーキテクチャ

2.2 通信操作・測定について

本研究で測定した通信操作は、対一通信・ブロードキャスト・収集・分配・トータルエクスチェンジである。トータルエクスチェンジは行列変換に似ており、 N 個のセルがそれぞれ N 個のデータを持ち、各セルが i 番目のデータを i のセルに送信する。最終的には i 番目のセルが通信前に各セルが持っていた i 番目のデータを持つことになる。

*Performance evaluation of communication operations on a torus-connected parallel computer

†Hiroyuki Makino, Susumu Shibusawa

‡Department of Computer & Information Sciences, Faculty of Engineering, Ibaraki University

また、メッセージサイズを変化させる・使用するセル数を変化させる・合計メッセージサイズを固定する方法で、それぞれ10回測定した。その中で通信時間が最小のデータを採用した[2]。

2.3 セルの指定方法

単純に送信セルと受信セルを指定して通信する方法を、絶対セル指定と呼ぶ。一方、通信路を1方向に限定した上で、送信セルから見た受信セルの距離を指定して通信する方法を、相対セル指定と呼ぶ。

3 測定結果および考察

3.1 対一通信

対一通信について、ホストから各セルへの通信時間の結果を図2に示す。

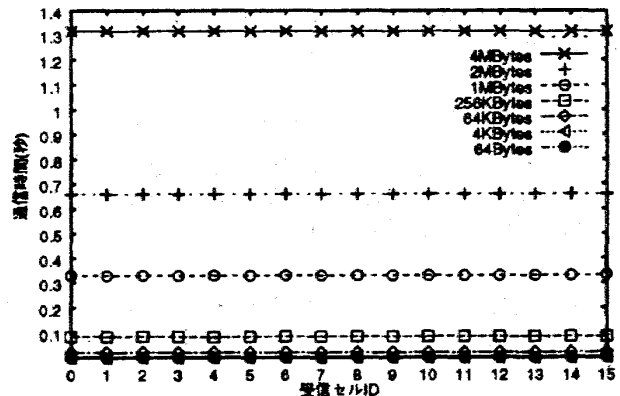


図2: ホスト-セル間通信の測定結果

セル相互通信の結果も、図2と同様に通信距離に依存しない結果が得られた。

本通信操作は、ホスト-セル間通信はB-Netを用いて、セル相互通信はT-Netを用いた。トーラス結合における通信時間は距離に比例するとされていた。しかし本結果は、T-Netを用いたセル間通信は通過する距離には比例せず、ほぼ一定の値を示した。

AP1000では、トーラス結合上にセルが存在せず、メッセージコントローラ経由でトーラス結合に接続されている。これにより、不要なメッセージはセルへ送られない。このため、セルへの入出力に伴うオーバーヘッドが削減されている。

本結果から、対一通信の通信時間 t (秒) は、セルの位置に依存しない近似式

$$t = a_m m + t_0 \quad (1)$$

で表すことが出来る[3]。但し a_m は1バイト当たりの通

信時間 (秒/Byte)、 m はメッセージサイズ (Bytes)、 t_0 はスタートアップ時間 (秒) を表す。

3.2 集合通信

3.2.1 メッセージサイズを変化させた場合

使用セル数を 16 セルに固定し、通信するメッセージサイズを変化させた場合の通信時間の結果を図 3 に点で示す。これらの通信時間の近似式を

$$t = a_m m + t_0 \quad (2)$$

で表し、図 3 に線で示す。

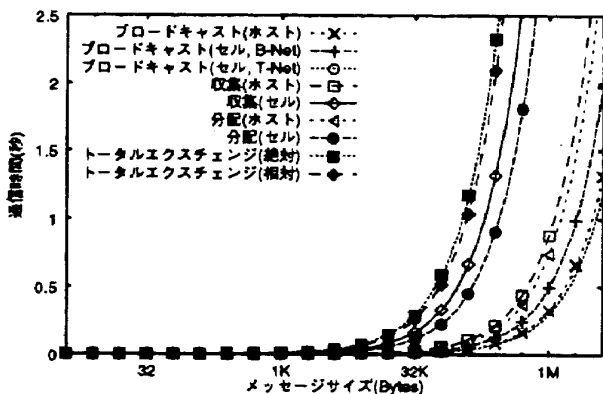


図 3: メッセージサイズと通信時間

全ての通信操作において、通信時間はメッセージサイズに比例していることが分かる。

3.2.2 使用するセル数を変化させた場合

メッセージサイズを 64KBytes に固定し、使用するセル数を変化させた場合の通信時間の結果を図 4 に点で示す。これらの通信時間を

$$t = a_n n + t_0 \quad (3)$$

の近似式で表し、図 4 に線で示す。但し、 a_n は 1 セル当たりの通信時間 (秒/セル)、 n は使用セル数 (セル) である。

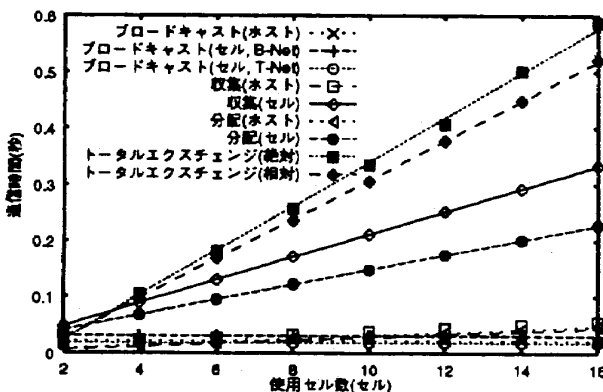


図 4: 使用セル数と通信時間

収集・分配・トータルエクスチェンジの各操作は、通信時間はセル数に比例している。一方、複数のセルがネットワーク上にあるメッセージを同時に受信することが出来るので、ブロードキャストはセル数に関係なくほぼ一定の値を示している。本操作は、使用するセル数に依存

しない。

3.2.3 合計メッセージサイズを固定させた場合

トータルエクスチェンジについて、合計メッセージサイズ m を固定させて、使用するセル数を変化させたときの通信時間を図 5 に点で示す。図 5 では、4 種類の m について示している。

1 つのメッセージサイズは m/n^2 であり、各セルはメッセージを $(n-1)$ セルへ対一通信をする。よって (1) 式に代入すると、近似式は

$$t = \left(\frac{m}{n^2} a_m + t_0 \right) (n-1) \quad (4)$$

となり、図 5 に線で示す。

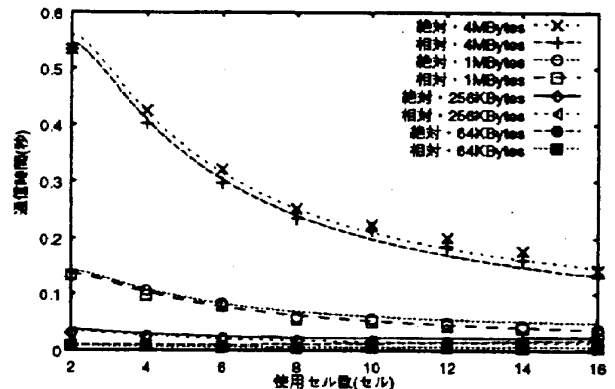


図 5: トータルエクスチェンジの測定結果

絶対セル指定の場合、2 つ以上のメッセージ間での通信路の競合が発生する。相対セル指定によって通信路の方向を 1 方向に限定することで、通信路の競合が削減できる。その結果、絶対セル指定を行うより最大で約 3 割の通信時間の短縮を図ることが出来た。

4 おわりに

本研究では、並列計算機で行われる通信操作の通信時間に関する近似式を導き出した。今後は、さらに大きな環境での測定を行い、他の並列計算機での測定も行う予定である。また、トータルエクスチェンジにおける通信路の競合について考察し、並列計算機の特性を生かした効率的な手法の開発にも取り組む予定である。

謝辞

御討論頂いた茨城大学工学部情報工学科渋沢研究室の皆様へ感謝します。

参考文献

- [1] 八代智樹, "AP1000 並列計算機の通信性能と測定とその評価," 茨城大学大学院理工学研究科情報工学専攻修士論文, 1997.
- [2] K. Hwang, Z. Xu and M. Arakawa, "Benchmark Evaluation of the IBM SP2 for parallel signal processing," IEEE Trans. on Parallel and Distib. Sys., Vol. 7 No. 5, pp. 522-536, 1996.
- [3] 古橋楠徳, "応用のための統計概論," サイエンス社, 1986.