

超並列プロセッサ向けルーティング方式の定量的検討

佐藤寿倫 高橋真史 藤井洋重 吉田尊

株式会社東芝 研究開発センター

5M-4

1. はじめに

ネットワークを介してメッセージの通信を行う分散メモリ型並列プロセッサでは、小レイテンシで高スループットなルーティング方式が要求される。小レイテンシのルーティング方式にワームホールルーティングがあるが、(1)デッドロックの発生、(2)低いスループットという問題がある。

今回私達は、デッドロック回避とスループットの向上を同時に満足するルーティング方式を検討し、シミュレーションによりその効果を確認した。

2. ルーティング方式

2.1 ワームホールルーティング

ワームホールルーティングでは、メッセージをフリットと呼ばれる数バイトのデータに分割し、中継プロセッサではフリットだけが蓄えられる。メッセージの先頭フリットにだけ宛先の情報があり、中継プロセッサは先頭フリットを受信すると、中継ルートを決定し転送する。メッセージの転送されている間は、その中継ルートは占有される。

ワームホールルーティングは小レイテンシのルーティング方式であるが、(1)デッドロックの発生、(2)低いスループットという問題がある。

Dallyらはヴァーチャルチャネルを用いて上記の問題の解決を図っている[1-2]。しかしこれらは、2つの問題を同時に解決していない。また、堀江らは構造化バッファプールを用いて、同時に2つの問題を解決しているが、各プロセッサにネットワーク規模に応じた数のバッファが必要になる[3]。

2.2 チャネルの多重化

ここでは、デッドロック回避のルーティングの下で、チャネルを多重化することにより、上記2つの問題を同時に解決することを試みる。

他のメッセージによるブロックを回避できれば、システム全体のスループットは向上する。そのためには、ブロックの回避のためのチャネルを増設すればよい[2]。これをデッドロック回避ルーティングに適応させる。

図1に、4つのプロセッサからなる1次元トラスネットワークに本方式を適用した場合を示す。

デッドロックを回避するためには、チャネル依存グラフに閉路ができないようにルーティングを行う[1]。そのために、

各物理チャネルを2つの論理チャネルとして用いる。したがって、各物理チャネルには2組のバッファが必要になる。

スループットを向上させるために、さらに各論理チャネルを多重化する。多重化する論理チャネルの数だけ、バッファを用意すればよい。図1では各論理チャネルをさらに2重化した場合を示してあり、1つの物理チャネルを4つの論理チャネルとして用いることになる。

このバッファの数はネットワークの規模に独立であり、超並列プロセッサ向けの方法である。

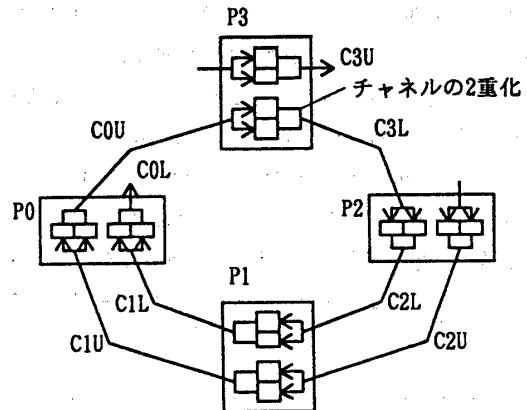


図1 チャネルの2重化

3. シミュレーションモデル

図2にシミュレーションに用いたプロセッサのモデルを示す。ネットワークは2次元トラスを想定しており、まずY方向でルーティングを行い、つづいてX方向でルーティングする。バッファはFIFOになっており、容量は1フリットを単位とする。FIFOの容量とチャネルの多重度を可変にしてシミュレーションを行う。

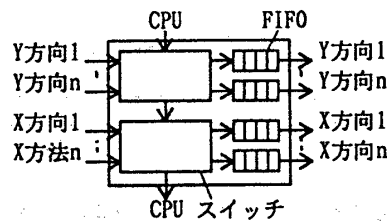


図2 プロセッサモデル

チャネルの割当てには、乱数で選んだり、最も過去に利用されたものを選んだりする方法が考えられるが、今回は簡単のために、通信距離で使い分けることにした。すなわちチャネルをN重化した場合には、各次元の最大距離をN等分して割当てることにした。たとえば図1のような4プロセッサの1次元

トラスでチャンネルを2重化した場合では、通信距離が2より大きいかわかで、割当てるチャンネルを選択する。

4. シミュレーション結果

チャンネルを2重、4重、8重化した場合のシミュレーションを、以下の条件で行った。

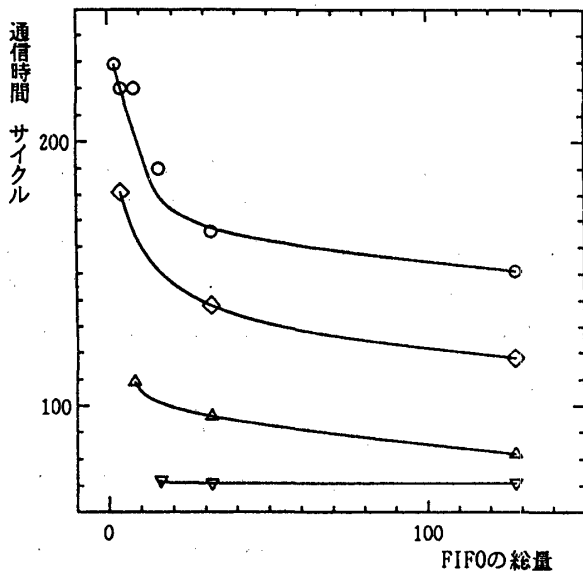
1. 8×8のトラス上で単方向通信。
2. 1メッセージは3フリット。
3. 通信の終了は、全メッセージの通信の終了時。

シミュレーション結果の通信時間は、隣接プロセッサ間で1フリットの通信を行うに必要な時間を1サイクルとして示した。

4.1 ランダム通信

宛先を一樣乱数で決め、各プロセッサが10個のメッセージを連続して送信する。同様の試行を30回行い、平均を求めた。図3にFIFOの容量を変えてランダム通信を行った結果を示す。横軸は、1物理チャンネル当りのFIFOの総量である。

チャンネルの多重度を増すにしたがって、通信時間は短縮している。FIFOの総量が同じ場合では、各FIFOの容量を増すよりも、チャンネルの多重度を増した方が効果的である。また、チャンネルの多重度が大きな場合では、FIFOの容量が小さくても、通信時間が短縮される。チャンネルを多重化した効果により、ブロックによるスループットの低下が改善されたためである。



○: 多重化なし ◇: 2重化 △: 4重化 ▽: 8重化

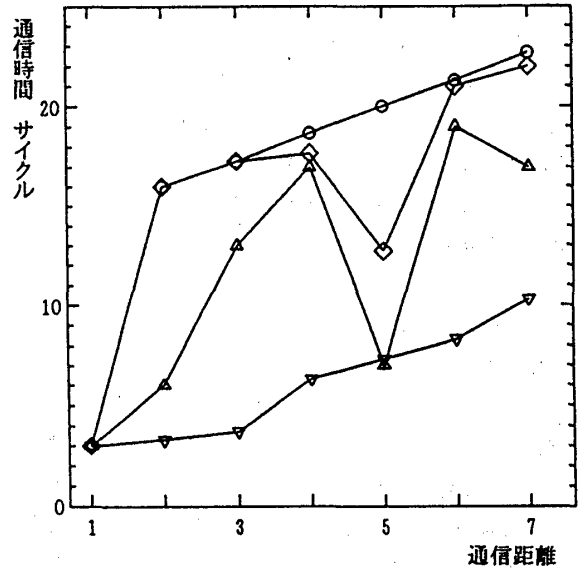
図3 ランダム通信のシミュレーション結果

4.2 等距離一斉通信

全プロセッサが1次元方向に一定距離だけ離れたプロセッサに通信を行うシミュレーションを行った。これを、等距離一斉通信と呼ぶことにする。各論理チャンネルのFIFOの容量を1とし、各プロセッサが1つのメッセージを通信するときの、等距

離一斉通信をシミュレーションした。図4に結果を示す。

等距離一斉通信の場合は、8重化するとかなりの効果が得られるが、2重化ではほとんど効果がない。また2重化、4重化の場合、通信距離が長くなっているにもかかわらず、通信時間が短くなる場合がある。いずれも、現在の割当法では、チャンネルが有効に利用されない場合が多いためである。チャンネルの割当法を改良すれば、より一層の効果が期待できる。



○: 多重化なし ◇: 2重化 △: 4重化 ▽: 8重化

図4 等距離一斉通信のシミュレーション結果

5. まとめ

ワームホールルーティングにおいて、デッドロックを回避しつつ、スループットを向上させることを検討した。チャンネルを多重化することで、スループットの向上を図った。

シミュレーションの結果、多重化をしないモデルと比較すると、概ね多重化の効果が現れた。また、FIFOの総量が同じ場合は、各FIFOの容量よりも、チャンネルの多重度を増す方が性能は向上した。

現在、効率のよいチャンネルの割当法を検討中である。

参考文献

[1] W.J.Dally et al., "Deadlock-Free Message Routing in Multiprocessor Interconnection Networks," IEEE Trans. Comput., vol. C-36, pp. 547-553, 1987.
 [2] W.J.Dally, "Virtual-Channel Flow Control," IEEE Trans. Parallel and Distributed Sys., vol. 3, No. 2, pp. 194-205, 1992.
 [3] 堀江 他, 「並列計算機AP1000における相互結合網のルーティング方式」, 通信学会論文誌, Vol. J75-D-1, No. 8, pp. 600-606, 1992.