

トポロジ独立な A-NET マルチコンピュータの通信性能

2G-5

廣田守

吉永努

馬場敬信

宇都宮大学工学部*

1 はじめに

我々は、並列オブジェクト指向トータルアーキテクチャ A-NET プロジェクト [1] において A-NET マルチコンピュータを開発した。本マルチコンピュータは、種々の応用分野における様々な通信パターンの要求に応えるため、トポロジ独立な適応型ルータにより、メッセージ通信を制御する。

本研究の目的は、OS レイヤ等の上位層の影響を受けない、ハードウェアレベルでのルータの性能を評価することである。そのため、ファームウェアでメッセージ送受信を行う最小の機能だけを実装して実験を行った。なお、実験対象となるノード台数は現在実装されている 16 台とした。

2 ルータ

ルータは、ホスト用ポートと隣接ルータ間との結合用の 6 個のポートを持つ。プログラマブル通信制御装置により、メッセージの経路選択を受け付け、レシーバや送出可能なポートをプログラマブルなロジックで決定することでトポロジ独立な制御を達成している。また、出力候補ポートから利用可能なものを選択して経路を適応ルーティングする [2]。

プロトタイプは 30MHz クロックで動作し、ルータ内のピーク転送速度は 2 クロック/バイト (15MB/s) である。

3 メッセージ送受信性能評価

実験ではネットワークトポロジに 4 次元ハイパーキューブを用いた。

3.1 ルータ-ルータ間の基本通信性能

ルータ-ルータ間の基本通信性能を得るために転送距離・メッセージサイズによって、どのように遅延時間に影響があるか実験した。

この実験での遅延時間は往復転送時間であり、その時間に含まれるのは、以下の 4 つである。

- (A) ノード i からノード j までの転送時間,
- (B) ノード j がメッセージ受信から次の送信までに要する時間,
- (C) ノード j が受信メッセージをメモリに書き込む時間,
- (A') ノード j からノード i までの転送時間

結果を図 1 に示す。グラフから 1 ホップ当りの遅延時間は、往復転送で $2.0\mu s$ 、よって片道転送では、 $1.0\mu s$ である。

また、遅延時間とメッセージサイズの関係を考えると転送するメッセージサイズが 1 バイト増すごとに $0.067\mu s$ 遅延することが分かる。従って、 $C = 0.067n\mu s$ (n はバイト数) である。 B はメッセージを受信して再び送信するまでの時間でマイクロプログラムのステップ数を数え

ることにより $B = 4.8\mu s$ であることが分かる。

また A, A' について考えてみる。距離 1 の宛先ノードにメッセージの先頭が到着する時間を x とすると、1 ホップ $1.0\mu s$ であるから、距離を D とすると $x + 1.0(D - 1)\mu s$ であると考えられる。また、メッセージの先頭が到着する時間 x は転送するメッセージサイズによって、ほとんど変わらない。従って、往復転送時間 T_d は、

$$\begin{aligned} T_d &= A + B + C + A' = 2A + B + C \\ &= 2\{x + 1.0(D - 1)\} + 4.8 + 0.067n \\ &= 2x + 2D + 2.8 + 0.067n \quad (\mu s) \end{aligned}$$

と表すことができる。よって、この式より、距離 1 の時の往復時間より、 x は約 $3.5\mu s$ と計算できる。この時間に含まれているのは、メッセージヘッダの読み出し、経路選択、ポート間半二重リンクの獲得、データ転送時間である。従って、

$$T_d = 9.8 + 2D + 0.067n \quad (\mu s)$$

となる。

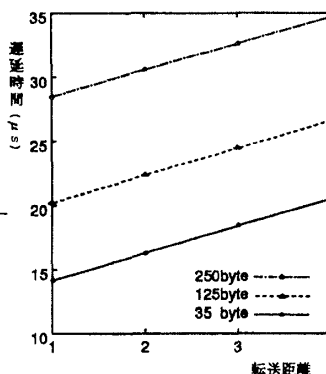


図 1: 転送距離・転送メッセージサイズと遅延時間 (往復転送)

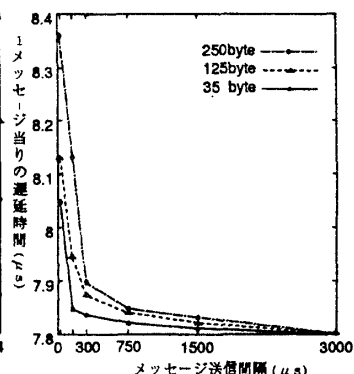


図 2: ランダム・トラフィックにおける送信間隔と遅延時間 (片道転送)

3.2 ランダム・トラフィック

次に、各々のノードが宛先をランダムに変えながら非同期にメッセージ送信する実験を行った結果について述べる。メッセージ送信間隔、メッセージサイズを変化させて実験を行ったところ図 2 のような結果を得た。縦軸は、送信ノードがメッセージを送信してから、受信ノードがメッセージの着信をハードウェア・フラグのポーリングにより検知するまでの時間を示す。

メッセージ間隔が $300\mu s$ 以上は、メッセージサイズに関わらず、ほぼ同様な値となった。これは、ネットワーク負荷が小さいためメッセージ衝突が少ないと考えられる。また、メッセージサイズが大きい方が遅延しているが、これはサイズが大きい分だけ、ネットワーク中に残っ

*Communication Performance of the Network-Topology Independent A-NET Multicomputer, Mamoru Hirota, Tsutomu Yoshinaga and Takanobu Baba, Utsunomiya University.

ている率が大きく、衝突が起きやすいと考えられる [3]。

4 ネットワークトポロジ可変性の評価

ルータがトポロジ可変であることの有効性を調べるため、図3のような通信パターンを用い、ハイパーキューブと二重リングでメッセージ送信間隔・送信回数と遅延時間の関係について実験した。ハイパーキューブと比較するのに二重リングを選んだ理由は、限られたノード数で構成でき、ハイパーキューブに埋め込まれないようなトポロジであり、また、リンク数は同じだが隣接したノードに2つの経路を持つというハイパーキューブにない特徴があるからである。

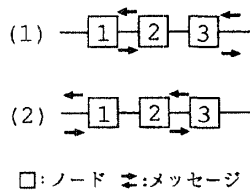


図3: 通信パターン

通信パターンは、まず、ホストから全ノードに、起動メッセージをブロードキャストして、次の(1)と(2)を繰り返す。(1)隣接ノードの一方にメッセージが向き合うように送信する。(2)もう片方の隣接ノードにも同様に送信し、2つメッセージを受信したら(1)に戻る。(図3の(1),(2)に対応)

なお、遅延時間は、起動メッセージを受信して、メッセージ送信を行った直後から、隣接ノードからの2つのメッセージを受信するまでの時間とする。

4.1 送信間隔と遅延時間の関係

メッセージ送信回数を100回に固定し、35byteのメッセージを送信間隔を変化させて実験した。結果を図4に示す。

二重リングは、送信間隔に関係なく一定である。これは二重リングが、隣接ノードに送信と受信専用の2つの経路を持っているために、ルータ-ルータ間の経路獲得のための競合が起きず、パケットがルータ上のパケットバッファ(PB)に退避されないためである。

ハイパーキューブは、送信間隔を極端に小さくした時に、大きく遅延していることが分かる。これは、隣接ノードとの経路が1つであるため、衝突も頻発し、パケットがPBに待避されることが多くなったり、待避されている時間が延びることが原因である。また、送信間隔が大きくなると規則正しく送受信が行われるようになり、遅延時間は送信間隔に関係なく一定になる。しかし、この状態での遅延時間は、二重リングのものより大きくなっている。これは、パケットがPBに常に退避されるようになり、この分だけ遅延が生じるためである。

4.2 送信回数と遅延時間の関係

メッセージ送信間隔を最小に固定し、35byteのメッセージを送信回数を変化させて実験を行った。結果を図5に示す。

ハイパーキューブの場合、最初衝突した時、大きく遅延するが、それ以降は、タイミングがずれルータ-ルータ

間の経路の優先権をめぐる衝突は減少する。しかし、送信回数が大きくなるにつれてルータ内での衝突が起き、パケットがPBに退避され、序々に遅延時間が大きくなり、定常状態となる。

二重リングでは、ルータ-ルータ間に2つの経路を持つために衝突が起こらず一定の遅延時間となる。

一般にハイパーキューブは、種々の通信パターンの埋め込み能力が高いことが知られているが、ここでの通信パターンにおいては、ハイパーキューブより二重リングの方が有利であることがわかり、トポロジ可変であることの有効性を確認できた。

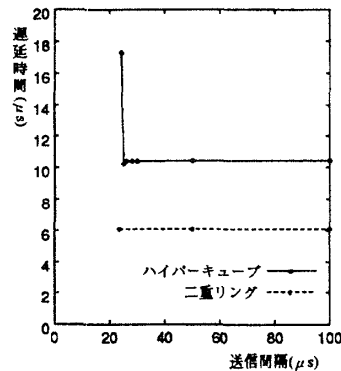


図4: メッセージ送信間隔と遅延時間

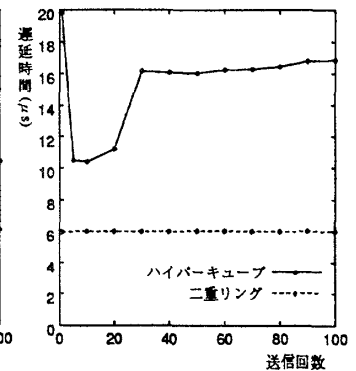


図5: メッセージ送信回数と遅延時間

5 おわりに

A-NET マルチコンピュータの通信性能について、メッセージ送受信性能とトポロジ可変の有効性を評価し、以下のような結果を得た。

- メッセージ衝突がない場合の往復転送時間 T_d を得た。
- ランダム転送においては、送信間隔が $300\mu s$ 以下ではメッセージ衝突による遅延は認められたが、PEのマシンサイクル時間と比較し小さく抑えられていることを確認した。
- 通信パターンに応じてネットワークトポロジを選択することにより、メッセージ通信性能を向上させることができることを確認した。

今後の課題として、適応型ルーティングによるフォールトトレランス性の評価などが挙げられる。

謝辞

本研究は、一部文部省科学研究費、基盤(C)08680346および電気通信普及財団の援助による。

参考文献

- [1] 馬場敬信, 吉永努, “並列オブジェクト指向トータルアーキテクチャA-NETにおける言語とアーキテクチャの統合”, 信学会論文誌 Vol.J75-D-I No.8 pp.563-574(1992).
- [2] 吉永努, 馬場敬信, “並列オブジェクト指向トータルアーキテクチャA-NETのノードプロセッサ”, 信学論, J79-D-I, 2, pp.60-68(1996).
- [3] 廣田守, 吉永努, 馬場敬信, “A-NET マルチコンピュータにおける適応型ルータの通信性能”, 情処学計算機アーキテクチャ研報96-ARC-119, vol.96, no.80, pp.31-36(1996).