

2J-04 デッドロック回復方式の違いによるルータの性能比較

中村さゆり 堀田 真貴 林 匡哉 吉永 努 大津 金光 馬場 敬信
宇都宮大学工学部

1 はじめに

適応ルータはデッドロック解消の観点から大きく2つにわかれる。1つはデッドロックの発生自体を防止するものであり、もう1つはデッドロックが発生した場合、それから回復する機構を装備するものである。本論文では、並列デッドロック回復機能を持つルータについて、アルゴリズムの相違による性能比較を行う。ルータ回路の論理合成とシミュレーション結果により、我々の提案する Recover-x の有効性を示す。また、近年の集積回路技術の発展により、多くのバーチャルチャネル (VC) が実装可能となりつつある。そこで、VC 数の増加による性能向上の評価も行う。

2 ルーティングアルゴリズム

2.1 ルーティングアルゴリズム

本論文で使用したルーティングアルゴリズムと、各ポートの VC 数を表 1 に示す。ネットワークは 2 次元トーラスネットワークであり、各ノードは 4 つのネットワークポートを持つ。

表 1: ルーティングアルゴリズムと VC 数

Router	Network ports	PE	DB
Dimension-order	n	2	-
Recover-x	n	2	-
DISHA-concurrent	n	1	2

n = 3, 4, 5

これらのルータは、各ネットワークポートに VC が最低 3 本あれば全て実装できる。今回は最少数の 3 本と、最少数 + 1 本である 4 本、最少数 + 2 本である 5 本の 3 種類を用意した。

2.2 デッドロック回復法の違いについて

前項で述べたルーティングアルゴリズムのうち、Recover-x^[2] と DISHA-concurrent^[1] (以降、DISHA-con と呼ぶ) は適応ルータであり、共に並列デッドロック回復方式を採用している。図 1 に、それぞれのデッドロック回復方式の違いを示す。

Recover-x は、X 次元の各ポートに装備されている非適応 VC を使用してデッドロック回復を行う。デッドロックサイクルを形成しているメッセージの中で、退避候補とするものを条件によって限定する。退避候補となったメッセージは、目的ノードのアドレスによって、X 次元のどちらかのポートの非適応 VC へと切替える。そのため、各非適応 VC への退避候補が比較的少なく、オーバーヘッドが小さい。

DISHA-con はデッドロックバッファ (DB) を複数ポートで共有し、ブロックされたメッセージを DB に退避することでデッドロックから回復する。すべてのメッセー

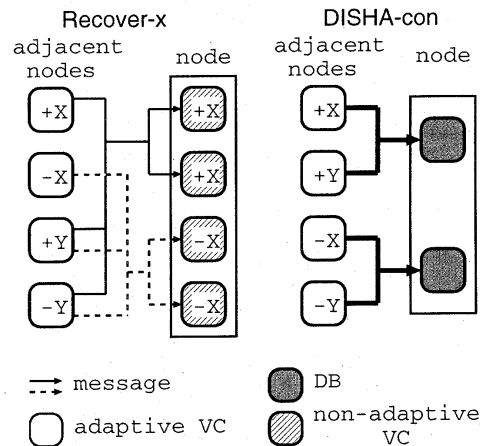


図 1: デッドロック回復法

ジが退避候補となり、進行方向によって 2 つの DB を使い分ける。

3 論理合成

3.1 論理合成条件

論理合成はクロック制約条件を除き、すべて次のような同一条件で行った。

シンセサイザ:

Synopsys HDL Compiler version 1999.05

回路の動作条件: 民生用最悪条件

配線負荷: セル面積による自動選択

マッピング最適化: Medium effort

ターゲットライブラリ:

LSI Logic 0.6 μ m Array-Based Gate Array

3.2 論理合成結果

表 2 に各ルータの論理合成結果を示す。

最高動作速度は、VC 切替えの組合せ数とマルチプレックス数の影響を受ける。そのため、ハードウェア構成の簡単な Dimension-order が最も高速となる。また適応ルータ間では、DISHA-con と比較して VC 切替えの少ない Recover-x がより高速となる。同一アルゴリズムで VC 数を変化させた場合、3 本から 4 本よりも、4 本から 5 本にしたときの方が最高動作速度が低下する。これはマルチプレックス数が増加し、調停回路がより複雑になるためである。

総面積は主にバッファ容量、ポート間結線などの影響を受ける。非適応ルータである Dimension-order はポート間結線が少なく、デッドロック検出/回復のための回路も持っていないので最小となる。また、DISHA-con は、他の 2 つのルータよりもバッファ数が 1 つ多いので、面積が最大となる。VC を増やしていく場合の面積の増加は、VC とそれに付随する配線の追加によるものである。

表 2: 論理合成結果

ルータ	Dimension-order			Recover-x			DISHA-concurrent		
	3	4	5	3	4	5	3	4	5
VCs/port	3	4	5	3	4	5	3	4	5
最高動作速度 (MHz)	161.2	156.2	147.0	142.8	133.3	117.6	108.6	100.0	90.2
総面積 (K Gates)	110.9	141.8	171.9	118.7	152.7	194.2	144.2	191.8	225.6

4 シミュレーション

4.1 シミュレーション条件

シミュレーションは以下のような条件で行った。

シミュレータ: Cadence Verilog-XL

ネットワーク: 10×10 2次元トーラス

動作速度: 表1に示す各ルータの最高動作速度

メッセージサイズ: 64バイト

4.2 シミュレーション結果

図2の(a)にネットワーク負荷が均一なRandom通信、(b)に負荷が均一でないHot-spot通信のシミュレーション結果を示す。図は横軸がバンド幅、縦軸が平均レイテンシとなっている。バンド幅は、単位時間当たりのネットワーク全体のメッセージ転送量である。また、平均レイテンシは、PE内でメッセージが生成されてから、そのメッセージが目的ノードのPEに到着するまでにかかった時間の平均と定義している。

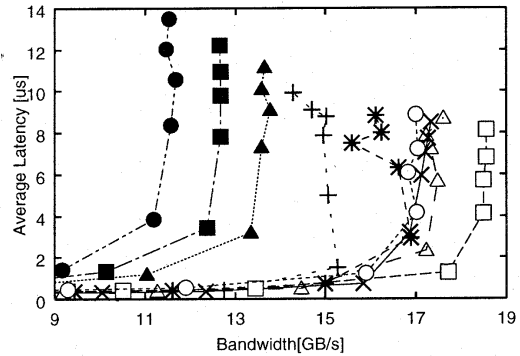
図2(a)のような偏りの少ない通信パターンでは、適応ルータは、空いている経路を探そうとしてもなかなか見つからず、その特性を十分に発揮することができない。しかし、Recover-xは動作速度の速いDimension-orderよりも高バンド幅低レイテンシを達成することができる。また、どのルータでもVC数を5本にすると良い結果が得られなかった。これはVCを増やしたことによる性能の向上よりも、動作速度の低下による影響がより大きいためである。

図2(b)のHot-spot通信では、ルーティングに柔軟性のある適応ルータと非適応ルータの差がRandom通信よりも大きくなる結果となった。ルーティングが柔軟であり、なおかつ高速動作可能なRecover-xが最も優れた性能を示した。また、Hot-spot通信における適応ルータは、VCを増やすほど性能が低下する傾向にある。これは、動作速度の低下をカバーできるほどの性能向上が見られなかったためである。

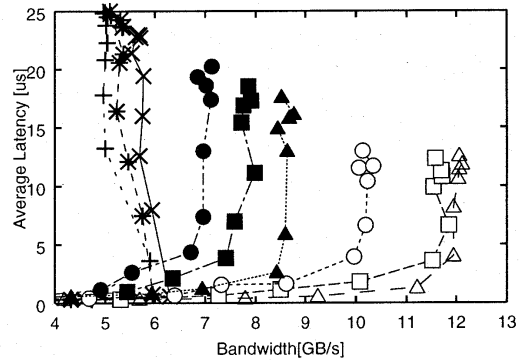
5 おわりに

本論文では、論理合成及びシミュレーション結果に基づき、デッドロック回復方式の異なるルータ間での性能比較を行った。その結果、高速動作可能でありルーティングに自由度のあるRecover-xは、同じ並列デッドロック回復方式のDISHA-conと比較しても、高バンド幅低レイテンシを達成することを示した。また動作速度とのトレードオフを考えると、VC数を増加することが性能の向上につながるとは限らない。それぞれのアルゴリズムに最適なVC数を探ることが良い結果をもたらす条件となる。

今後の課題としては、耐故障性を考えた迂回制御付きRecover-xの評価が挙げられる。また、更に大規模なネットワークでのシミュレーションを行う。



(a) Random 通信



(b) Hot-spot 通信

Dimension-order	Recover-x	DISHA-con
V3 - +	V3 - △	V3 - □
V4 - ×	V4 - □	V4 - ●
V5 - *	V5 - ○	V5 - ●

図 2: シミュレーション結果

謝辞

本研究の一部は東京大学大規模集積システム設計教育センターより提供して頂いたCADツールを使用しており、深く感謝する。本研究は、一部文部省科学研究費 基盤研究(B) 課題番号 010558039、奨励研究(A) 課題番号 11780190、及び、実吉奨学会の援助による。

参考文献

- [1] Anjan K.V., T.M Pinkston and J.Duato: "Generalized Theory for Deadlock-Free Adaptive Wormhole Routing and its Application to DishA Concurrent", *Proc.IPPS*, pp.815-821(1996).
- [2] 林匡哉, 堀田真貴, 吉永努, 大津金光, 馬場敬信: "適応ルータの効率的な並列デッドロックリカバリ方式の提案", 並列処理シンポジウム JSPP'99 論文集, pp.55-62(1999).