ハイパクロスバ・ネットワークにおけるVirtual Channelの動的選択による適応ルーティング

曽根 猛[†] 朴 泰祐[††]
中村 宏[††] 中澤 喜三郎[†††]

本論文では，並列計算機用ネットワークにおけるルーティング手法が，virtual channelの使用方法と経路決定のアルゴリズムにより，一般的に4種類に分類できることを示す．そして，それら4種類のルーティング手法を超並列計算機向きのプロセッサ間ネットワークのひとつであるハイパクロスバ・ネットワークに適用する手法を提案し，その転送性能を計算機シミュレーションにより評価する．ハイパクロスバ・ネットワークにおいて，これまでに提案されている適応ルーティングはvirtual channelの使用方法が静的に決定されており，それらが必ずしも有効に使われていなかった．本論文において，同ネットワークへのvirtual channelの使用方法を動的に決定する適応ルーティングの導入を提案する．評価は，転送先がランダムに決定される場合と転送先に偏りが存在する場合を対象とした．その結果，virtual channelの使用方法を動的に決定する適応ルーティングは，単純な固定ルーティングに比べて，メッセージの転送先がランダムに決定される転送では約46%，メッセージの転送先に偏りが存在する場合では，約2倍の性能向上が確認された．

Adaptive Routing by Dynamic Selection of Virtual Channels on Hyper-Crossbar Network

TAKESHI SONE,[†] TAISUKE BOKU,[††] HIROSHI NAKAMURA[††]
and KISABURO NAKAZAWA[†††]

In this paper, we show that routing schemes on inter-processor communicating network can be classified into four types depending on the usage of virtual channel and routing algorithm. We evaluate the performance of these routing schemes on Hyper-Crossbar Network by computer simulation, which is an inter-processor communicating network for massively parallel processing systems. In the adaptive routing scheme on Hyper-Crossbar Network proposed so far, virtual channels were selected deterministicly, then the usage of virtual channels was not so effective. In this paper, we propose dynamic virtual channel selection schemes on that network. We evaluated various routing schemes on both uniform and nonuniform data transfer patterns. On uniform data transfer pattern, it is confirmed that adaptive routing with dynamic virtual channel selection improves the network performance about 46% compared with the simple routing scheme. On nonuniform data transfer pattern, it is also confirmed that adaptive routing with dynamic virtual channel selection achieved twice of data transfer throughput compared with the simple routing scheme.

1. はじめに

高性能なRISCプロセッサをProcessing Unit（以下PUと省略）として用い，それらを相互結合ネットワークにより数千台以上結合し，高性能な分散メモリ型超並列計算機を実現しようという試みがさかんに行われている．このようなアーキテクチャでは，ネットワークの転送能力はシステムの性能を左右する大きな要因のひとつであり，これまでに多くのネットワークが提案されている．しかし，基本特性や拡張性を考えた場合，数千台規模の超並列計算機に採用できるものはかなり限られている．その中でも，ハイパクロスバ・ネットワーク[1]（以下HXBと省略）は，他の典型的な超並列計算機向きのネットワークに比べてPU間距離が比較的小さく，また，クロスバスイッチを多段

† 株式会社日立製作所汎用コンピュータ事業部
General Purpose Computer Division, Hitachi Ltd.
†† 筑波大学電子・情報工学系
Institute of Information Sciences and Electronics, University of Tsukuba
††† 電気通信大学情報工学科
Department of Computer Science, University of Electro-Communications

に組み合わせていることから通信チャネル数も多いため，大規模システムにおける各種の複雑な転送パターンにおいて高い転送性能を保つことができる．特に，送信相手がランダムに決定されるような転送では，他のネットワークに比べて高い転送性能を実現できることが知られている．さらに，各次元方向のサイズや次元数を任意に決定することができる（全次元方向のサイズがすべて等しいものは base-m n-cube として知られている[2]）ため拡張性にも優れている[3)~5)]．

HXB に対する過去の研究では，store & forward 方式または wormhole 方式による無衝突転送時の評価とランダム転送時の評価が行われてきた[5]．メッセージの転送方式やルーティング・アルゴリズムは，ネットワークの形状と同様に転送性能を決定する重要な要因であり，これらを改良することによりネットワークの転送性能を向上できることが知られている．最近の研究において，HXB の転送方式として，virtual cut-through 方式を導入したり，経路決定のアルゴリズムとして virtual channel を用いた適応ルーティングを導入することにより，同ネットワークの性能をさらに向上できることが分かってきた[6),7)]．

しかしながら，HXB において，これまでに提案された適応ルーティングは virtual channel の使用方法が静的に決定されていたため，その利用率に偏りが生じており，それらが必ずしも有効に使われていなかった．そこで本論文では，HXB の転送性能をさらに向上させる手法として，virtual channel の使用方法を動的に決定する適応ルーティングの導入を考える．以下では，2 章で HXB の構成について説明し，3 章で超並列計算機用ネットワークにおけるルーティング手法が，virtual channel の使用方法と経路決定のアルゴリズムにより一般的に 4 つに分類できることを示す．そして，4 章でそれら 4 種類の手法を HXB に適用する手法を示し，5 章においてその転送性能を計算機シミュレーションにより評価する．

2. ハイパクロスバ・ネットワーク

HXB の構成について述べる．**図 1** に 3 次元の HXB (4 × 4 × N) の構成を示す．

n 次元の HXB は，n 次元の直交座標上に PU を配置し，1 次元方向に並んだ各 PU をクロスバスイッチ（以下 XB と省略）により結合する．これにより，送受信 PU のアドレスが 1 次元だけ異なる場合は，XB を 1 回通過するだけで相手 PU にメッセージを転送することができる．送受信 PU のアドレスが 2 次元以上で異なる場合は，XB を 2 回以上通過しなければならないが，このときの各次元方向の XB の乗り換えは，エクスチェンジャ（以下 EX と省略）と呼ばれるルータ・スイッチによって行われる．また，EX は各 PU に対応して 1 つずつ用意され，各次元方向の XB の乗り換えと同時に，PU と XB 間の転送も行う．3 次元 HXB における転送の例として，X 次元と Y 次元の 2 次元の転送（送受信 PU の Z 次元のアドレスが等しい場合）を考える．送信 PU アドレスが (x_s, y_s, z)，受信 PU アドレスが (x_d, y_d, z) の場合，メッセージは PU (x_s, y_s, z) → EX (x_s, y_s, z) → X 次元 XB → EX (x_d, y_s, z) → Y 次元 XB → EX (x_d, y_d, z) → PU (x_d, y_d, z) というようにして XB を 2 回通過することにより転送される．このとき，Z 次元の PU アドレスは等しいので，その次元における転送は必要ない．このように，n 次元の HXB において PU から送出されたメッセージは最大 n 個の XB を通過することにより転送が終了する．このとき，EX は小規模なクロスバスイッチによって構成されるので，メッセージが通過するクロスバスイッチの最大数は $2n+1$ となり，PU 間の最大距離は $2 \times (n+1)$ となる．

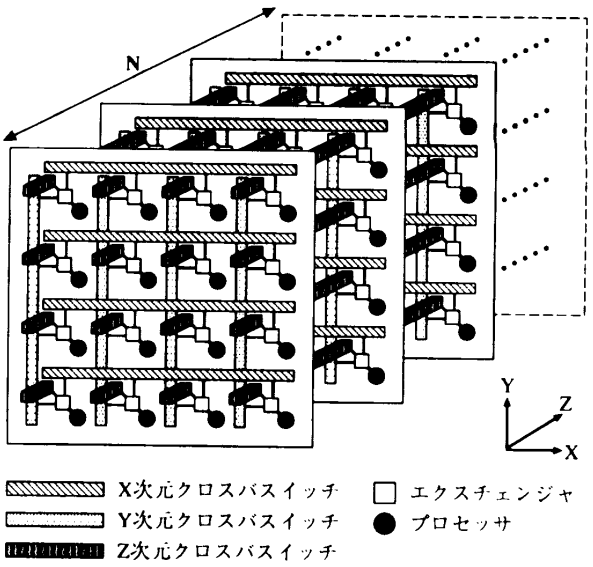


図 1 3 次元ハイパクロスバ・ネットワーク (4 × 4 × N)
Fig. 1 3 dim. Hyper-Crossbar Network (4 × 4 × N).

3. ルーティング手法の分類

これまでに，多くのネットワークにおいて，デッドロックを回避するためにメッセージの送信 PU と受信 PU の位置関係のみにより経路が決定される固定ルーティングが用いられている．しかし，固定ルーティングでは，衝突したメッセージは他の空いている経路を利用せずその場でブロックされてしまう．このとき wormhole 方式では，ブロックされたメッセージは経

表1 ルーティング手法の分類
Table 1 Classification of routing schemes.

		virtual channelの使用方法	
		静的に決定	動的に決定
経路決定	固定	(1) fix-static	(2) fix-dynamic
	適応	(3) adp-static	(4) adp-dynamic

路上の各ノードのバッファを占有したままなので，メッセージの後続部分がさらに他のメッセージをブロックしてしまう．これにより，メッセージ転送のレイテンシが増加し，ネットワークの転送性能が低下する．そこで，いくつかのネットワークに対してメッセージが転送中に動的に空きチャネルを見つけ，そちらに転送するような適応（動的）ルーティングが提案されている．その際，virtual channelを用いてチャネルを多重化し，デッドロック・フリーを保証するのが一般的である[8]．これまでに提案されている各種ルーティング手法を系統的に整理すると，virtual channelの使用方法と経路決定のアルゴリズムの組合せにより，それらは一般的に**表1**に示す4通りに分類できることが分かる．

以下では，これまで提案さてれきたルーティング手法とこの分類との対応を示す．

(1) virtual channelの使用方法を静的に決定する固定ルーティング

以下では，このルーティング手法をfix-staticと省略する．この代表例としては，次元オーダによる固定ルーティング（*e*-cubeルーティング，X → Yルーティング等）があげられる．次元オーダの固定ルーティングを行うことでデッドロック・フリーが保証されるネットワーク（ハイパキューブ，メッシュ等）は，virtual channelの本数が1本の場合として，このルーティング手法に分類できる．

さらに，virtual channelが複数本ある場合として，トーラス・ネットワークを考える．同ネットワークにおいては，一般的にvirtual channelを2本用いてchannel dependency graphにサイクルが生じないように，その使用方法を制限することでデッドロックを回避している[8]．さらに，AP1000[9]に用いられているトーラス・ネットワークは，ネットワークのサイズに応じて複数本のvirtual channelを用意する必要があり，それらの使用方法は各次元方向ごとのホップ数により静的に決定されている．

(2) virtual channelの使用方法を動的に決定する固定ルーティング

以下では，このルーティング手法をfix-dynamicと省略する．このルーティング手法は，(1)の場合に対して必要最低限のvirtual channelに，さらに余分なvirtual channelを付加することにより実現できる．このとき，(1)と同様に固定ルーティングを行うため，デッドロック・フリーが保証される．たとえば，トーラス・ネットワークにおいて，複数のvirtual channelを2組に分け，それらを(1)の場合と同様にして使用することによりデッドロックが回避できる．これにより，wormhole方式においてメッセージがブロックされた場合，後続のメッセージは他のvirtual channelを用いることにより移動することが可能となるので，レイテンシを低減することできるようになる．

また，トーラス・ネットワークにおいてvirtual channelを2本だけしか用いない場合でも，それらを動的に使用する手法が提案されている[10]．

(3) virtual channelの使用方法を静的に決定する適応ルーティング

以下では，このルーティング手法をadp-staticと省略する．このルーティング手法は，メッセージが選択可能な複数の経路において，使用できるvirtual channelを制限することにより，チャネルの接続にサイクルを生じないようにする．

たとえば，*k*-ary *n*-cubeでは，メッセージがどのような経路を通過してもデッドロック・フリーが保証されるトポロジーを持つvirtual networkを複数構成し，メッセージが使用できるvirtual networkを送受信PUの位置関係により制限することで，システム全体のデッドロック・フリーを保証できる[11]．

(4) virtual channelの使用方法を動的に決定する適応ルーティング

以下では，このルーティング手法をadp-dynamicと省略する．このルーティング手法として，文献12)で提案されている手法がある．その概略を以下に示す．

virtual channelを2組に分け，そのうちの1組により全PU間で相互にメッセージ転送が可能でデッドロック・フリーなvirtual networkを構成し，残りのvirtual channelはその使用方法を動的に決定する．以下では，前者のvirtual channelをcoreのvirtual channel，後者のvirtual channelをfreeなvirtual channelと呼ぶことにする．メッセージは移動の際，各ノードにおいて使用するvirtual channelをcoreとfreeの中から任意に選ぶことができる．メッセージが次のノードに移動する際，つねにcoreのvirtual channelが転送経路の選択肢の中に含まれているため，システム全体のデッドロック・フリーが保証される．また，バッファが空いていれば，ネットワーク中のメッセージはfreeなvirtual channelを任意に使用することができるので，

システム全体として適応ルーティングが実現される．

文献12)においては，ハイパキューブ・ネットワークにこの手法を適用することにより，同ネットワークの性能を向上できることが報告されている．

4. ハイパクロスバ・ネットワークへの適用

ここでは，3章に示した4種類のルーティング手法のHXBへの適用，および過去の研究との関連について述べる．また，5章でこれらのルーティング手法における有効性を計算機シミュレーションにより確認する．

(1) HXBにおけるfix-static

HXBは，ネットワークの形状により次元オーダ（たとえば，X → Y → Zの順に従う）による固定ルーティングを行うことでデッドロック・フリーが保証される．よって，このルーティング手法を適用した場合，virtual channelの本数が1本として考えることができる．HXBにおいて，このルーティング手法の研究はすでになされており，基本的な転送特性が示されている[2),4),5)]．

(2) HXBにおけるfix-dynamic

(1)で示したように，HXBでは次元オーダによる固定ルーティングを行えば，デッドロック・フリーが保証されるため，複数本のvirtual channelを持つHXBにおいて，メッセージは各次元方向を移動する際，任意のvirtual channelを使用することができる．

2本のvirtual channelを持つ3次元HXBにおいて，メッセージ（3次元とも通過する場合）が使用できる経路を**図2**に示す（1回のXBの通過を1ステップとする）．図中のX，Y，Zの添字はvirtual channelの番号を表す．

このルーティング手法は(1)の場合に比べて，どのvirtual channelを使用するかという制御を付加することにより実現される．

(3) HXBにおけるadp-static

このルーティング手法のHXBへの適用は文献7)で提案されており，その有効性が確認されている．その概要は以下のとおりである．

n次元のHXBにおいて，virtual channelをn本用意する．このとき，メッセージが使用できるvirtual channelをそのメッセージの転送ステップ数により決定する．n次元HXBでは，最大nステップ（1回のXB通過が1ステップ）でメッセージの転送が終了するため，n本のvirtual channelを各ステップごとに使い分ければ，経路上にサイクルを生じないので，デッドロック・フリーな適応ルーティングが可能となる．

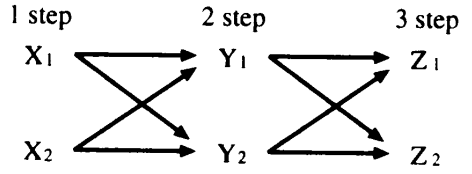


図2 3次元HXBにおける転送経路：(2) fix-dynamicの場合
Fig. 2 Routing on 3 dim. HXB: (2) fix-dynamic.

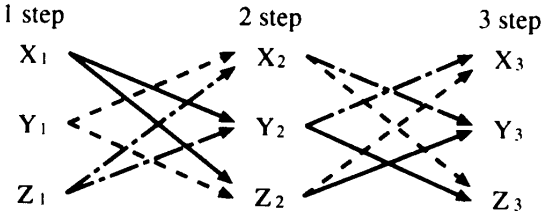


図3 3次元HXBにおける転送経路：(3) adp-staticの場合
Fig. 3 Routing on 3 dim. HXB: (3) adp-static.

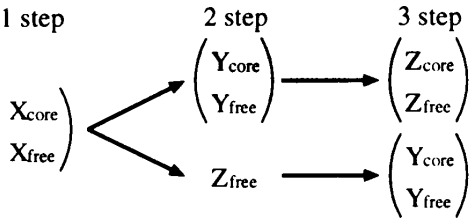
この手法を3次元HXBに適用した場合（virtual channelは3本必要となる）のメッセージの転送経路（3次元とも通過する場合）を**図3**に示す．

ここで，HXBにおいてEXでの出力先を決定すると，その先に接続しているXBの出力先も一意に決定される．したがって，ネットワーク中のメッセージが各EXにおいて次のステップで使用するvirtual channelを選ぶ際に，接続している各XBを通じて，その先に接続しているEXの入力バッファの情報を取得する必要がある．メッセージは，XBの入力バッファとその先のEXの入力バッファがともに利用可能であれば，そのvirtual channelを用いて転送を行う．また，実際のメッセージ転送ではメッセージのヘッダがEXの入力に到達するまでに遅延が生じるので，それまでの間はバッファを予約するような制御が必要となる．

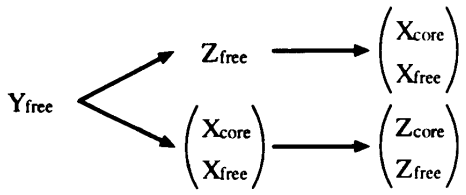
(4) HXBにおけるadp-dynamic

このルーティング手法はDuatoにより最近提案された概念であり，本論文においてHXBへの適用を提案する．この手法を導入することにより，(3)の手法で問題となるvirtual channelの利用率の偏りを解消し，HXBの転送性能のさらなる向上が期待できる．

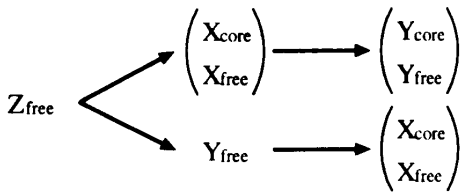
このルーティング手法をHXBに適用したときのメッセージの転送経路を以下に示す．例として，3次元HXBにおいて2本のvirtual channelを持つ場合を考え，coreとfreeのvirtual channelをそれぞれ1本とする．coreのvirtual channelでは，次元オーダ（X → Y → Z）の固定ルーティングを行うものとする．3次元転送を行うメッセージの転送経路を**図4**に示す．図において，X_{core}，Y_{core}，Z_{core}はcoreの



(a) 第1ステップでX次元方向に移動する場合



(b) 第1ステップでY次元方向に移動する場合



(c) 第1ステップでZ次元方向に移動する場合

図4 3次元HXBにおける転送経路：(4) adp-dynamicの場合
Fig. 4 Routing on 3 dim. HXB: (4) adp-dynamic.

virtual channel, X_{free}, Y_{free}, Z_{free} はfreeなvirtual channelをそれぞれ表す.

これらのvirtual channelを使用するときの原則は,以下のとおりである.各ステップにおいて使用できるvirtual channelは,アドレスが異なる全次元のfreeなvirtual channelと,アドレスが異なる次元のうち次元番号が最小の次元(ここではX < Y < Z)のcoreのvirtual channelである.図4において,第1ステップではX, Y, Zのどの次元方向にも移動できるが,coreのvirtual channelでは次元オーダの固定ルーティングを行うため,次元番号が最小であるX次元でのみ,freeなvirtual channelに加えてcoreのvirtual channelも使用できる(図4).図4(a)において,第1ステップ終了後はYZ平面の転送となる.この場合も,Y < Zなので第2ステップのメッセージは,Y次元方向に移動するときはfreeとcoreのどちらのvirtual channelも使用できるが,Z次元方向に移動するときはfreeなvirtual channelしか使用できない.同様なことが,図4(b), (c)の第2ステップについてもいえる.

図4から分かるように,各ステップにおいてメッセージが経路を決定する際に必ずcoreのvirtual channelが選択肢の中に含まれている.たとえば,第1ステップにおいて選ぶことができるvirtual channelはX_{core}, X_{free}, Y_{free}, Z_{free}であり,coreのvirtual channelとしてX_{core}が含まれている.これにより,coreのvirtual channelの選択がつねに確保されるので,システム全体としてのデッドロック・フリーが保証される.また,(3)の適応ルーティングでは,各ステップ数において使用できるvirtual channelが制限されるため,その利用率に偏りが生じてしまうが,この手法では同じvirtual channelを各ステップで使用できる(たとえば,X_{free}は第1, 第2, 第3の各ステップで使用できる)ため,virtual channelの利用率の偏りが解消され,virtual channelを効率的に利用できるようになる.

このルーティング手法においても,(3)の場合と同様なXBの状態の先読み制御が必要となる.

5. シミュレーションによる性能評価

ここでは,4章で述べた4種類のルーティング手法について評価を行う.評価は一様なランダム転送の場合と,メッセージの転送先に偏りが存在する場合を対象として,計算機シミュレーションによって行った.

システムの規模は4096 PUとし,3次元(16×16×16)構成のHXBを対象とした.XB, EXおよびPUの各チャネルのバンド幅を正規化し,1 flit/clockでメッセージが転送されるものとする.したがって,ここではメッセージ長をflit単位で表す.また,メッセージの転送方式としてはwormhole方式を用いる.

4章において示した4種類の手法において,(2) fix-dynamic, (3) adp-static, (4) adp-dynamicの3種類については,全通信チャネル(EX ↔ XB間)に各々3本のvirtual channelを設けることにした.また,各手法において,PUから送出されたメッセージがX, Y, Zの各次元方向に移動できることを考慮して,PUとEX間のスループットをネットワーク中の3倍とした(物理的にチャネルの本数を3倍とした).また,適応ルーティングにおいて,あるEXから複数の次元方向へのXBのチャネルが空いている場合は,X, Y, Zの優先順位で転送先を決定することとした.

各ルーティング手法における仮定を以下に示す.

- fix-staticでは,次元オーダ(X → Y → Z)による固定ルーティングを採用し,ネットワーク中のvirtual channelは1本とする.
- fix-dynamicでは,次元オーダによる固定ルーティングを採用する.ネットワーク中を移動しているメッセージは各次元において3本のvirtual channelのうち空いている任意のvirtual channelを使

用するものとする．

- adp-static では，4 章に示した手法を用いて，3 本の virtual channel をメッセージのステップ数（= これまでに通過した XB の数）に従って使用する．
- adp-dynamic では，1 本の virtual channel を core，残り 2 本の virtual channel を free として用いる．core の virtual channel では，次元オーダの固定ルーティングを行う．また，core と free の virtual channel がどちらも使用可能の場合は，core の virtual channel を優先して使用する．

すべてのルーティング方法について，PU に到着したメッセージは割り込み処理により適宜処理されるものとし，明示的な受信処理は行わない．通常，メッセージの送受信は DMA コントローラによって行われるので，ここでは送信処理と受信処理が同時にできるものとした．

5.1 ランダム転送時の性能評価

システム中の各 PU の動作を以下のように仮定した．システム中の全 PU は完全に独立に動作し，シミュレーション時間内の各クロックにおいて，ある一定の確率によりメッセージの送信を行う．その際，各 PU はランダムに選んだ相手 PU に対して一定長（ここでは，10 flit または 100 flit）のメッセージを転送する．以上のような動作を全 PU が繰り返す．

以上のような仮定の下で，ネットワークの転送性能を評価する．評価としては，ネットワーク・スループットとメッセージ・レイテンシを用いる．スループットは理想的な無衝突転送の場合を 1 として正規化し，1 PU あたりの実際に受信した総メッセージ受信量の比率で表す．ここで理想的な無衝突転送とは，システム中の全 PU が毎クロック，1 flit 分のデータを受信した場合のことである．ここでは，10000 クロックのシミュレーションを行い，各 PU が受信した平均総メッセージ受信量（flit）を 10000 で割った値をスループットとした．また，レイテンシはメッセージが生成されてから，受信 PU にメッセージの先頭が到着するまでの時間である．評価結果を**図 5** と**図 6** に示す．グラフの横軸はネットワーク・スループットで，縦軸がそのときのメッセージ・レイテンシである．

fix-static とそれ以外の手法では virtual channel の本数が等しくなく，コスト・パフォーマンスを考慮すると単純には比較できないが，図 5 と図 6 から adp-static, fix-dynamic, adp-dynamic は従来の単純な手法である (1) fix-static に比べ転送性能が向上していることが分かる．adp-static, fix-dynamic, adp-dynamic の最大スループットは fix-static に対して，それぞれ約 34%，約 54%，約 46%向上している（メッセージ長が 10flit のとき）．

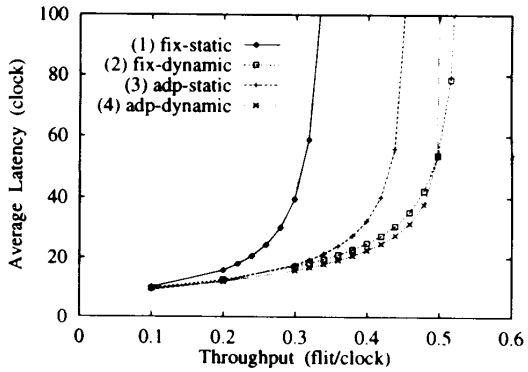


図 5 ランダム転送時の転送性能（メッセージ長 = 10 flit）
Fig. 5 Network Performance on uniform transfer pattern (messagelength = 10 flits).

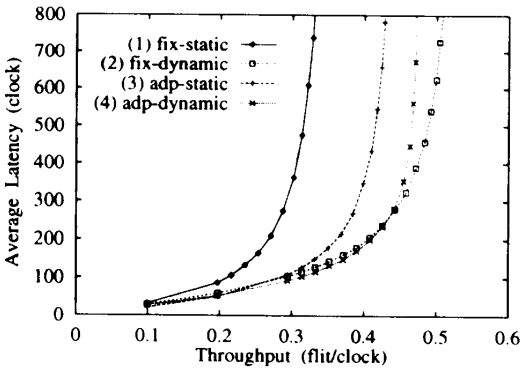


図 6 ランダム転送時の転送性能（メッセージ長 = 100 flit）
Fig. 6 Network Performance on uniform transfer pattern (message length = 100 flits).

適応ルーティング（adp-static, adp-dynamic）の性能向上に関しては，以下の 2 点が大きく影響していると考えられる．

- 今回対象とした規模の HXB においてランダムな転送が起こった場合，固定ルーティングにおいては，PU から送出されたメッセージの約 94%（= 15/16：X 次元のアドレスが異なる場合）が最初に X 次元方向に移動する．これに対し，適応ルーティングを導入することにより，PU から送出されたメッセージは X, Y, Z の各次元方向に移動できるので，PU は複数のメッセージを同時に送出でき，PU と EX 間の 3 倍のスループットが有効に使われている．
- メッセージがネットワーク中を移動する際，空きチャネルがあればそちらにメッセージを転送することにより他のメッセージとの衝突を回避するこ

とができる．

fix-dynamicの性能向上に関しては，virtual channelによる効果が大きい．$n \times n$の単一のクロスバ・スイッチを考えた場合，メッセージが他のメッセージと衝突する確率は$1-(1-1/n)^{n-1}$と表すことができる[7]（簡単化のため，クロスバ・スイッチの各入力に必ずメッセージが存在する場合を仮定している）．ここで，スイッチのサイズが16のとき，衝突確率は約0.6となる．HXBの各次元のXBにおいても，同じような確率で衝突が起こっていると考えられる．fix-staticにおいては，これにより後続のメッセージもブロックされてしまうが，virtual channelを用いることによりfix-dynamicにおいては，後続のメッセージがブロックされず，移動可能なことがあるため転送性能が向上していると考えられる．

次に，virtual channelを動的に使用した場合と静的に使用した場合を比較する．fix-dynamicとadp-dynamicがadp-staticよりも高いスループットを示しているが，これは以下に示すように，ネットワーク中を移動しているメッセージが使用できるvirtual channelの本数が影響していると考えられる．各次元方向でのvirtual channelをX_1，X_2，X_3のようにして表すと，3次元すべてを通過するメッセージの場合，

- adp-staticでは第1ステップでX_1，Y_1，Z_1のいずれかを使用できる．第1ステップでX次元方向に移動したメッセージは第2ステップではY_2，Z_2のどちらかを使用でき，さらに第2ステップでY次元方向に移動したら第3ステップではZ_3しか使用できなくなる．
- fix-dynamicでは第1ステップでも$X_{1\sim3}$，第2ステップでも$Y_{1\sim3}$，第3ステップでも$Z_{1\sim3}$をそれぞれ使用でき，転送が進んでも，使用できるvirtual channelの本数が少なくなることがない．
- adp-dynamicも図4に示すように，転送のステップ数が進んでいっても，使用できるvirtual channelの本数が少なくなることがない．

このように，virtual channelの使用方法が動的に決定される手法ではvirtual channelが有効に利用されるため，高い転送性能が実現されている．

fix-dynamicとadp-dynamicを比べると，ネットワークが過負荷な状態となったときの最大スループットはfix-dynamicの方がadp-dynamicより高くなる．これは，ランダム転送を行った場合，$16 \times 16 \times 16$の構成の3次元HXBにおいては約82%（$= \frac{15\times15\times15}{4095}$）のメッセージが3次元すべてを通過することになるため，

- fix-dynamicでは，ほとんどのメッセージがX → Y → Zの順で移動するため，EXにおける衝突がそれほど生じない．
- adp-dynamicでは適応ルーティングを行うため，EXにおいても衝突が起こる．

以上のことが影響してネットワークが過負荷な状態においてはfix-dynamicの方がadp-dynamicよりも最大スループットが高くなっているものと考えられる．逆に，ネットワークの負荷が比較的小さい範囲（スループットが0.5より低い範囲）では，adp-dynamicの方がfix-dynamicに比べてメッセージ・レイテンシが小さくなっている．先に述べたように，固定ルーティングではPUから送出されたメッセージの9割以上が最初にX次元方向に移動するのに対して，適応ルーティングでは，これらのメッセージは各次元方向に移動することができるため，ネットワーク中の空いているチャネルを有効に使用することでメッセージ間の衝突による待ち時間が減少しているものと考えられる．

最後に図5と図6から，メッセージ長による影響を考察する．各ルーティング手法間での全体的な傾向はメッセージ長が変わってもほぼ同じだが，固定ルーティング（fix-static，fix-dynamic）の場合に比べて，適応ルーティング（adp-static，adp-dynamic）ではメッセージ長が長くなることにより最大スループットが減少している．これは，1回の経路決定に対してどのくらいのデータ量が転送されるかが影響している．メッセージ長が100 flitの場合，いったん経路を決定すると100 flit分のデータ転送が終了するまで経路を変更することができないのに対して，メッセージ長が10 flitの場合，100 flit分のデータを転送する際，その間に9回の経路変更が可能となる．このことから考えて，メッセージ長が短い方が適応ルーティングの効果を十分に発揮できることが確認できる．

以上のことから，その使用方法が動的に決定されるvirtual channelの本数を多くすることにより，転送性能の大幅な向上が達成できることが分かる．実際にネットワークが使用される状況を考えた場合，過負荷な状態で使用される場合は比較的稀であることから，virtual channelの使用方法を動的に決定する適応ルーティングは，メッセージ・レイテンシを低くおさえることができるので有効な手法であるといえる．

5.2 転送に偏りが存在する時の性能評価

システム中のメッセージの転送先に何らかの偏りがある場合を想定して，以下のようなモデルを考えた．システム中の各PUは転送する相手PUをランダムに決定するが，その際，各次元方向についてある一定の

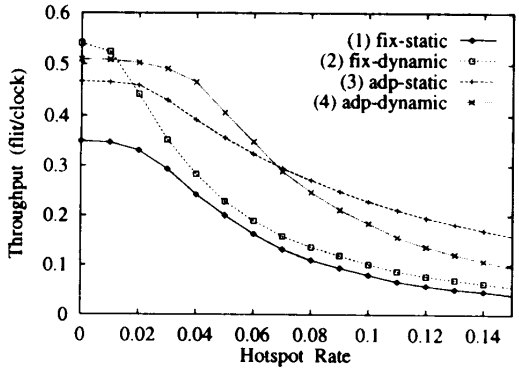


図7 転送に偏りがある時の転送性能（メッセージ長 = 10 flit）
Fig. 7 Network Throughput on nonuniform transfer pattern (message lentgh = 10 flits).

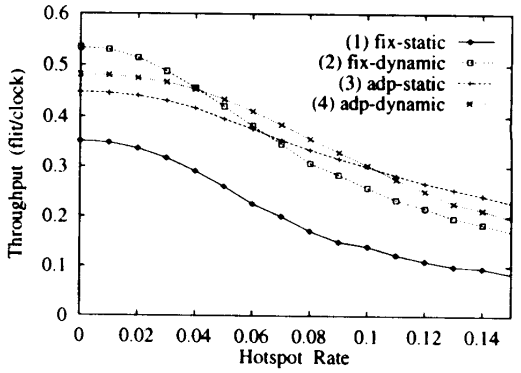


図8 転送に偏りがある時の転送性能（メッセージ長 = 100 flit）
Fig. 8 Network Throughput on nonuniform transfer pattern (messagelength = 100 flits).

割合で1つのアドレス（たとえば0番）にメッセージが集中するようにする．この割合をHotspot率と定義する．各PUはメッセージを転送する際，Hotspot率に従い，ある一定の確率あらかじめ決められたアドレスのPUにメッセージを転送し，そうでないときは，ランダムに決定されたアドレスのPUにメッセージを転送する．これを各次元ごとに独立に行う．したがって，全次元方向ともあらかじめ決められたアドレスにメッセージを転送する場合は，1つのPUにメッセージが集中することになる．今回は，簡単化のため各次元方向でのHotspot率を等しくおき，それを変化させてメッセージの転送先に偏りがある場合のランダム転送と見なして，シミュレーションを行った．ホットスポットが存在するような転送では，ネットワークには比較的高い負荷がかかるので，ここでは以下のような状況で評価を行った．システム中の各PUは，あるメッセージの送出が完了したら，すぐに次のメッセージの転送を開始することにより，ネットワークに比較的高い負荷を与える．各PUが転送するメッセージの長さは10 flitと100 flitで，ネットワーク・スループットを評価する．評価結果を**図7**と**図8**に示す．グラフの横軸はHotspot率で，縦軸がそのときのネットワーク・スループットである．

図7から分かるように，各手法ともHotspot率が大きくなるに従って，スループットが低下しているが，その中でもfix-dynamicの性能低下が著しい．固定ルーティングではメッセージの転送経路が一意に決まってしまうため，メッセージの転送先に少しでも偏りが生じると転送性能が著しく低下してしまう．それに対して，adp-staticやadp-dynamicの適応ルーティングでは空いているチャネルがあればそちらにメッセージを転送するので，メッセージの転送先にある程度の偏りが生じてもそれほど性能が低下しないことが確認できる．Hotspot率が5%の付近では，fix-staticに比べfix-dynamicは約14%の性能向上であるのに対して，adp-staticは約78%の性能向上，adp-dynamicは約2倍の転送性能を示している．

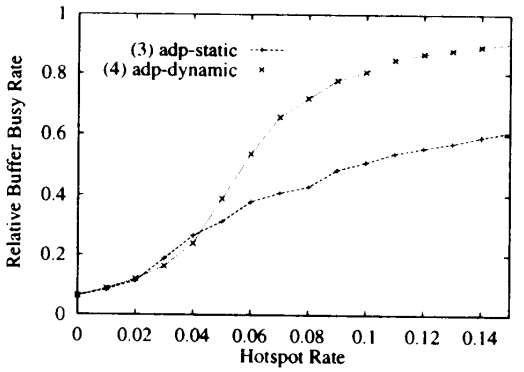


図9 EXのY次元方向における入力バッファの相対利用率（メッセージ長 = 10 flit）
Fig. 9 Relative buffer busy rate of Y dim. input buffer on EX (messagelength = 10 flits).

また，Hotspot率が7%の付近でadp-staticとadp-dynamicの性能の逆転が起こっている．これは，adp-dynamicではvirtual channelの使用が動的に決定されるため，特定のPUに向かうメッセージにより，混雑している付近のvirtual channelがすべて占有されてしまうことが影響している．このことを確認するため，EXにおけるY次元方向の入力バッファの利用率を調べてみた（メッセージ長は10 flit）．結果を**図9**に示す．横軸はHotspot率，縦軸は全EXのY次元方向の入力バッファ利用率のうち，転送先のY次元アドレスが0のメッセージが占めている割合を表す．Hotspot率が0のとき，メッセージの転送先のY次元アドレスは0～15の範囲で均等に分散されるため，EXのY次元方向の入力バッファ利用率のうち，転

送先のY次元アドレスが0のメッセージが占める割合は0.0625 (= 1/16) となる．図9から分かるように，adp-dynamicではadp-staticに比べ，Hotspot率が高くなるにつれてホットスポットに向かうメッセージの占める割合が高くなり，その傾向は図7におけるスループットの減少傾向とほぼ一致している．すなわち，ホットスポットに向かうメッセージの量は，どちらのルーティング手法でも等しいにもかかわらず，バッファ中におけるそれらのメッセージのflitの滞在時間はadp-dynamicの方が相対的に長いことが分かる．そして，wormhole方式を対象としてるため，それらのメッセージにより後続のメッセージがさらにブロックされ，結果的にadp-dynamicのスループットがより小さくなっていると考えられる．

図7と図8からメッセージ長の変化による影響を考察する．図7と図8を比べて分かるように，固定ルーティング (fix-static, fix-dynamic) ではメッセージ長が短くなると，転送性能が著しく低下している．それに対して，適応ルーティングでは (adp-static, adp-dynamic) は固定ルーティングに場合に比べてメッセージ長の変化による影響を受けにくいことが分かる．

以上のことから，メッセージの転送先に偏りが存在するような状況において適応ルーティングは有効な手法であり，メッセージ長の変化によって転送性能がそれほど影響を受けないことが分かる．

6. おわりに

本論文では，virtual channelの使用方法と経路決定のアルゴリズムに従い，一般的にルーティング手法を4種類に分類できることを示し，それらをハイパクロスバ・ネットワークに適用した場合の転送性能を評価した．その結果，その使用方法が動的に決定されるvirtual channelの本数を増やすことにより固定ルーティング・適応ルーティングの双方において，同ネットワークの転送性能を大幅に向上できることが分かった．また，メッセージの転送先に偏りがある場合では，固定ルーティングの場合はわずかな偏りで転送性能が著しく低下するのに対し，適応ルーティングでは空きチャネルがあればそちらにメッセージを転送するため，メッセージの転送先にある程度の偏りが存在してもそれほど性能が低下しないことも確認された．以上のことから，virtual channelの使用方法を動的に決定する適応ルーティングは非常に有効な手法であり，単純な固定ルーティングに比べて，ランダム転送では約46%，メッセージの転送先に偏りが存在する場合では，Hotspot率が5%程度の時に約2倍の性能向上が確認された．さらに，メッセージ長の変化による影響を考えた場合，なるべくメッセージ長が短い方が，適応ルーティングの有効性を発揮できることが分かった．

今回はメッセージの転送方式としてwormhole方式を対象としたが，ネットワーク中のリソースの活用を考えた場合，各ノードにメッセージをストアするためのバッファをより有効に利用することが考えられる．今後はバッファの利用法も考慮に入れた評価も行う予定である．また，今回はランダム転送とメッセージの転送先に偏りが存在する場合の転送性能のみを評価したが，今回提案した各ルーティング手法における制御の複雑さの評価，および各種転送パターンにおける転送性能の評価も行っていく予定である．さらに，今回行った4種類のルーティング手法の分類は，他のネットワークにも適用可能なので，他のネットワークにおける4種類のルーティング手法の適用の検討と，その時の転送性能の評価も行っていく予定である．

謝辞 本研究を進めるにあたり貴重なご意見をいただいた筑波大学西川博昭助教授ならびにアーキテクチャ研究室諸氏に深く感謝します．なお，本研究の一部は創成的基礎研究 (07NP0401) の補助によるものである．

参考文献

1) 田中輝雄ほか：識別子を用いたデータ転送方式を基本とするMIMD型並列計算機アーキテクチャ，並列処理シンポジウムJSPP '89論文集，pp.115-122 (1989).
2) 鈴木 節ほか：並列AIマシンProdigyの相互結合網の評価，電子情報通信学会論文誌D，Vol.J71-D, No.8, pp.1496-1501 (1988).
3) 斉藤哲也ほか：超並列計算機のネットワークの実現可能性と性能評価，情処研報，92-ARC-95 (1992).
4) 保田淑子ほか：ハイパクロスバネットワークの通信性能評価，信学技報，CPSY93-25 (1993).
5) 朴 泰祐ほか：ハイパクロスバ・ネットワークの性能評価，信学技報，CPSY93-40 (1993).
6) 朴 泰祐ほか：ハイパクロスバ・ネットワークにおける転送性能向上のための手法とその評価，情報処理学会論文誌，Vol.36, No.7, pp.1610-1618 (1995).
7) 朴 泰祐ほか：ハイパクロスバ網における適応ルーチングの導入とその評価，電子情報通信学会論文誌D-I，Vol.J78-D-I, No.2, pp.108-117 (1995).
8) Dally, W.J., et al.: Deadlock-Free Message Routing in Multiprocesser Interconneciton Networks. *IEEE Trans. Comput.*, Vol.C-36, No.5, pp.547-553 (1987).
9) Ishihata, H., et al.: An Architecture of Highly

Parallel Computer AP1000, *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp.13–16 (1991).
10) Duato, J.: A Necesssary and Sufficient Condition for Deadlock-Free Adaptive Routing in Wormhole Networks, *International Conference on Parallel Processing* (1994).
11) Linder, D.H., et al.: An Adaptive and Fault Tolerant Wormhole Routing Strategy for k-ary n-cubes, *IEEE Trans. Comput.*, Vol.40, No.1, pp.2–12 (1991).
12) Duato, J.: A New Theory of Deadlock-Free Adaptive Routing in Wormhole Networks, *IEEE Transactions on Parallel and Distributed System*, Vol.4, No.12, pp.1320–1331 (1993).

(平成7年 9 月11日受付)
(平成7年11月 2 日採録)

曽根 猛（正会員）

昭和46年生．平成6年筑波大学第三学群情報学類卒業．平成8年同大学院工学研究科修士課程修了．同年（株）日立製作所入社，現在に至る．超並列計算機アーキテクチャ，並列処理等に興味を持つ．

朴 泰祐（正会員）

昭和59年慶應義塾大学工学部電気工学科卒業．平成2年同大学院理工学研究科電気工学専攻後期博士課程修了．工学博士．昭和63年慶應義塾大学理工学部物理学科助手．平成4年筑波大学電子・情報工学系講師，平成7年同助教授，現在に至る．超並列計算機アーキテクチャおよび並列処理言語・アルゴリズムの研究に従事．電子情報通信学会会員．

中村 宏（正会員）

昭和38年生．昭和60年東京大学工学部電子工学科卒業．平成2年同大学院工学系研究科電気工学専攻博士課程修了．工学博士．同年筑波大学電子・情報工学系助手．平成3年同講師，平成7年同助教授，現在に至る．計算機アーキテクチャ，並列処理，計算機の上位レベル設計支援に関する研究に従事．本学会平成5年度論文賞，本学会平成6年度山下記念研究賞各受賞．電子情報通信学会，IEEE，ACM各会員．

中澤喜三郎（正会員）

昭和30年東京大学工学部応用物理卒業．昭和35年同大学院数物系博士課程応用物理修了．同年日立製作所入社．TAC，HITAC 5020，E/F，8800/8700，M-200H/280H，680H，S-810等，超大型コンピュータ・スーパーコンピュータの開発に従事．平成元年より筑波大学電子・情報工学系教授，計算物理学センター向きの超並列処理システムCP-PACSの研究開発に従事．平成8年より電気通信大学情報工学科教授，現在に至る．工学博士．電子情報通信学会，IEEE，ACM各会員．平成5年度本学会論文賞受賞．