

Title	超並列計算機向き相互結合網SRTにおける適応型ルーティング
Author(s)	川井,雅之; 井口,寧; 堀口,進
Citation	情報処理学会論文誌, 41(7): 2010-2017
Issue Date	2000-07
Type	Journal Article
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/3319">http://hdl.handle.net/10119/3319</a>
Rights	<p>社団法人 情報処理学会, 川井雅之 / 井口寧 / 堀口進, 情報処理学会論文誌, 41(7), 2000, 2010-2017. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は(社)情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	

# 超並列計算機向き相互結合網 SRT における 適応型ルーティング

川井 雅之<sup>†</sup> 井口 寧<sup>††</sup> 堀口 進<sup>†</sup>

超並列システムに適合する結合網には、科学技術計算によく用いられる 2 次元格子結合を含み、ノードあたりのリンク数が少数であるなどの実装性、耐故障性などの要件が求められている。SRT (Shifted Recursive Torus) はグリッドの大きさが異なるトーラス結合を再帰的にシフトして構成された、超並列計算機に適した結合網である。SRT は、トーラス結合網に遠距離ノード間通信のためのバイパスリンクを付加しノードあたりのリンク数を固定した階層構造を有する結合網であり、従来の相互結合網に比べて遜色ない次数や直径を有している。SRT におけるルーティング (再帰ルーティング) は直径や平均距離などの点で十分に高い性能を有しているが、転送経路が固定であるため混雑や故障に対応できない。そこで、本論文では、SRT のデッドロックフリーな適応型ルーティング手法を提案する。提案する適応型ルーティングは仮想チャネルを増設する必要がない。また、シミュレーションにより適応型ルーティングの性能評価を行い従来手法と比較検討を行った。その結果、デッドロックフリーな再帰ルーティングに比べ非常に高い転送能力を有していることを示す。

## An Adaptive Routing of Shifted Recursive Torus Networks

MASAYUKI KAWAI,<sup>†</sup> YASUSHI INOBUCHI<sup>††</sup> and SUSUMU HORIGUCHI<sup>†</sup>

Massively parallel computers require interconnection networks with excellent features of a small diameter, a small number of links, expendability and fault-tolerance. Shifted Recursive Torus (SRT) consists of torus networks which are shifted recursively. SRT has the advantage that the number of links a node is fixed and the diameter is relatively small. We have proposed a deadlock-free routing of SRT and proved the recursive routing is a near-optimal static routing. However, the proposed recursive routing does not have an adaptability and a fault-tolerance. This paper addresses a deadlock-free algorithm for adaptive routing of SRT without additional virtual channels. This algorithm allows a detour routing on the same dimension. The adaptive routing algorithm has been proved as a deadlock-free adaptive routing and performances are evaluated by computer simulation. It's seen that the proposed adaptive routing achieves much better dynamic communication performance than a statistic recursive routing.

### 1. はじめに

自然科学におけるシミュレーションや VLSI 設計など、先端科学技術分野における大規模科学技術計算の需要は増大しており、多数のマイクロプロセッサ (MPU) を用いた超並列計算機による高速化が求められている。超並列計算機では、プロセッサ要素 (PE) 間の通信性能が並列処理の効率に大きな影響を与える

ために、様々な視点から数多くの相互結合網が提案されている。

科学技術計算の多くは 2 次元または 3 次元構造のデータを対象とするため、格子型の結合網と適合しやすい。しかし、格子結合網は、システムの規模が大きくなると通信性能が急速に低下してしまう。

Inoguchi ら<sup>1),2)</sup>により提案された Shifted Recursive Torus (SRT) は、トーラス結合網にノード間距離の異なるバイパスリンクを再帰的に付加した結合網である。SRT は少ない次数で RDT<sup>3)</sup>や PEC<sup>4)</sup>と同程度の直径を実現できる。SRT の再帰ルーティングは、再帰ネットワーク構造を十分に考慮した手法であり、高い通信性能<sup>5)</sup>を有している。しかし、再帰ルーティングは固定型であるため混雑や故障を回避できず、超並列計算機において重要な適応性や耐故障性を有

<sup>†</sup> 北陸先端科学技術大学院大学情報科学研究科  
School of Information Science, Japan Advanced Institute  
of Science and Technology, Hokuriku  
現在、株式会社日本総合研究所

Presently with The Japan Research Institute, Limited

<sup>††</sup> 北陸先端科学技術大学院大学情報科学センター  
Center for Information Science, Japan Advanced Institute  
of Science and Technology, Hokuriku

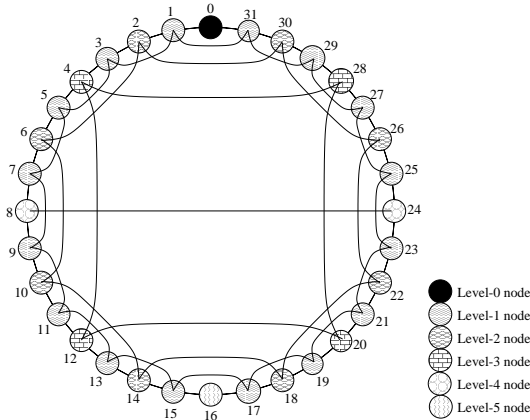


図 1 32 ノードから成る基本型 1D-SRT

Fig.1 Standard 1D-SRT consists of 32 nodes.

していない．相互結合網における適応手法としては Duato の手法<sup>6),7)</sup>や Turn Model<sup>8)</sup>などが代表的であるが，いずれも SRT に適用するにはいくつかの問題がある．そこで本論文では SRT の構造的特徴を考慮した同次元迂回ルーティングによる適応型ルーティングを提案<sup>9)</sup>する．また，適応型ルーティングがデッドロックフリーであることを証明する．さらに，シミュレーションにより動的通信性能について性能評価を行い，適応型ルーティングが高い通信性能を有することを示す．

## 2. SRT の構成

### 2.1 1次元 SRT の構成

1次元 SRT (1D-SRT) は，ノード数  $N$  ( $N = 2^n$ ) から成る環状網を基に構成される．環状網のあるノードを番号 0 とし，ノードを昇順に番号付ける．

定義 1 (1D-SRT( $n, T$ ))  $n$  ノードから成る環状網から，

$$(x_l - 2^{l-1}) \bmod \min(2^l, 2^T) = 0 \quad (1)$$

を満たすノード  $x_l$  を取り出し，これをレベル  $l$  のノードと呼ぶ．ただし，ノード 0 は，どの  $l$  でも式 (1) を満たさないため，ノード 0 のレベルを 0 と定める．各ノード  $x_l$  は，レベル 0 のリンクにより，隣接する左右のノード  $(x_l \pm 1) \bmod N$  と結合され，さらにレベル  $l$  のリンクによって  $2^l$  だけ離れたノード  $(x_l \pm 2^l) \bmod N$  と結合される．自己回帰する結合は省略し，結合が重複する場合は 1 本のリンクで結合する．

$T$  は，1D-SRT の構成を決定するパラメータであり，ノード  $0, \frac{1}{4}N, \frac{1}{2}N, \frac{3}{4}N$  の結合が決められる<sup>2)</sup>．

```

FindMedLevel( $x_s, x_d$ ) {
     $sp = x_d - x_s$ ;
    if(  $sp \leq 2$  ) return( 0 );
     $l = \text{int}(\log_2(sp)) + 1$ ;
     $x_1 = 2^{l-1}$ ;
     $x_2 = x_1 * 2$ ; /*  $x_1 \leq sp < x_2$  */
    if(  $x_2 - sp \leq sp - x_1$  )  $l++$ ;
     $c = \lfloor (\sqrt{8 * l + 1} - 1) / 2 \rfloor$ ;
    return(  $l - c$  );
}
    
```

図 2 再帰ルーティングに用いるレベルの決定  
Fig.2 Level for recursive routing.

1D-SRT( $n, n$ ) を基本型 SRT と呼ぶ．これに対し，1D-SRT( $n, n - 2$ ) を LongSpan 型と呼び，ノード  $0, \frac{1}{2}N$  間に新たなリンクを設ける．1D-SRT( $n, n - 3$ ) を ShortSpan 型と呼び，ノード  $\frac{1}{4}N, \frac{3}{4}N$  間の結合を除くかわりにノード  $0, \frac{1}{4}N, \frac{1}{2}N, \frac{3}{4}N$  を結ぶリンクを設ける．

図 1 に 32 ノードから成る 1D-SRT(5,5) のリンク結合の様子を示す．

### 2.2 2次元 SRT の構成

本節では，1D-SRT を 2次元に拡張した 2D-SRT について述べる．2D-SRT は，1D-SRT を幅  $s_x$  だけずらしながら並べ，四方へのリンクを設けることにより構成される．

定義 2 (2D-SRT( $n, T, s_x$ ))  $N \times N$  ( $N = 2^n$ ) ノードから成るトーラス網から，

$$(x_l - 2^{l-1} + s_x \cdot y_l) \bmod \min(2^l, 2^T) = 0 \quad (2)$$

を満たすノード  $(x_l, y_l)$  を取り出し，これをレベル  $l$  のノードと呼ぶ．ただし，式 (2) を満たす  $l$  を持たないノード  $((-s_x \cdot y_l) \bmod N, y_l)$  のレベルは 0 と定める．

各ノード  $(x_l, y_l)$  は，レベル 0 およびレベル  $l$  のリンクにより，隣接ノードおよび  $2^l$  離れたノード  $\{((x_l \pm 1) \bmod N, (y_l \pm 1) \bmod N), ((x_l \pm 2^l) \bmod N, (y_l \pm 2^l) \bmod N)\}$  と結合される．1D-SRT の場合と同様に，自己回帰リンクはこれを省略し，重複する場合は 1 つのリンクで結合する．

定数  $s_x$  は，2D-SRT の構成を決定するためのパラメータであり，このとり方により 2D-SRT の通信性能が変化する<sup>2)</sup>．

### 2.3 再帰ルーティング

SRT におけるノード  $x_s$  から  $x_d$  へのルーティングは，まずルーティングに使用するリンクのレベルを

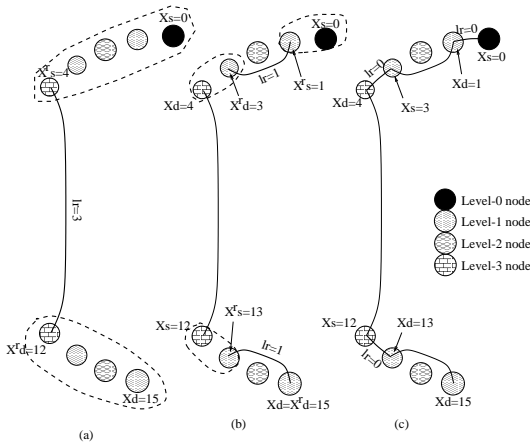


図3 1次元SRTのルーティング例  
Fig. 3 Example of 1D-SRT routing.

図2に示す手続きによって求め、環状網上の距離として始点ノード  $x_s$  に最も近い同じレベルのノードを探す。次に、得られたノードと始点ノードの間に存在する最大レベルを求める。この手続きを始点ノードと接続するノードが見つかるまで再帰的に繰り返す。同様な手続きを目的ノード  $x_d$  側でも行い、経路を算出する。この方針に基づくルーティングを再帰ルーティングと呼ぶ。

32ノードから成る1D-SRTにおける、ノード0 ( $x_s = 0$ ) からノード15 ( $x_d = 15$ ) までの例を、図3に示す。SRTでは再帰ルーティングを行うことにより、最適ではないものの実用上十分な性能を得ることができる。再帰ルーティングはデッドロックフリーを保証していないが、SRTの基本結合がトラス結合であるため、代表的なデッドロック回避の手法であるdimension-order routingを適用することによって、デッドロックフリーとすることができる<sup>5)</sup>。しかしながら、再帰ルーティングは始点ノードと終点ノードの位置によって経路が決定される固定型ルーティングであるため、適応性や耐故障性を有していない。

### 3. 1D-SRTの適応型ルーティング

1D-SRTのデッドロックフリー・ルーティングは、monotonic order routingを適用することで簡単に得ることができた<sup>5)</sup>。しかし、1D-SRTにおける適応型ルーティングを考える場合、従来の代表的な適応化の手法であるDuatoの手法<sup>6),7)</sup>やTurn Model<sup>8)</sup>を適用するにはいくつかの問題がある。

Duatoの手法<sup>6),7)</sup>は、一次元ネットワークに対して、チャンネル選択の自由度は上がるが、経路選択の自由度の適応性が得られない。Turn Model<sup>8)</sup>は、メッセージ

のターンに対する制限を最小限に抑えることで、迂回が可能な適応ルーティングが実現できる。しかし、迂回ルーティングが可能となる条件が明確ではないためTurn Modelを用いると目的ノードまでのパスが保証されないという問題がある。

方向転換が180度の方向のみの1D-SRTに対し、Turn Modelを適用することは不可能である。そこで本論文ではTurn Modelを用いずに、1D-SRTでターンを可能にする手法を提案する。

#### 3.1 諸定義

本論文で用いる用語、諸定義について述べる。

定義3 任意の2つのノード  $n_1, n_2$  に対し次のように大小関係を定義する。

$$n_1 < n_2 \iff n_1 \leq n_2 \wedge n_2 - n_1 < N/2 \vee \\ n_1 > n_2 \wedge n_1 - n_2 > N/2$$

また、ノード番号が  $n_1 < n_2$  ( $n_1 > n_2$ ) なる方向を正(負)方向と呼ぶ。

定義4 パケットがある次元で1方向のみ(monotonic order<sup>10)</sup>)に転送されるルーティングをmonotonic order routingと呼ぶ。monotonic order routingを次元順で用いたルーティングをdimension order routingと定義する。

定義5 任意のチャンネルに割り当てられた番号(チャンネル番号)を  $n$  次元ベクトル  $C = (c_{n-1}, c_{n-2}, \dots, c_1, c_0)$  で表したとき、任意の2つのチャンネル番号  $C_1 = (c_{1n-1}, c_{1n-2}, \dots, c_{11}, c_{10})$ ,  $C_2 = (c_{2n-1}, c_{2n-2}, \dots, c_{21}, c_{20})$  に対し、次のように大小関係を定義する。

$$C_1 > C_2 \iff \exists i (c_{1i} > c_{2i} \wedge \forall j > i (c_{1j} = c_{2j}))$$

定義6 パケットの存在するノードを  $n_{cur}$ 、パケットの次の転送先のノードを  $n_{next}$  としたとき、次の条件を満たす2つのノード間のリンクをラウンドトリップループと呼ぶ。

$$(条件) (n_{cur} < n_{next} \wedge n_{cur} > n_{next}) \vee (n_{cur} > n_{next} \wedge n_{cur} < n_{next})$$

#### 3.2 チャンネル割当

Turn Modelに限らず、従来のルーティングアルゴリズムは、アルゴリズムの提案が先に行われ、デッドロックフリーは、そのあとでチャンネルに適当な番号を与え、循環が生じないことを示すことで保証してきた。本論文では、Turn Modelを用いずにチャンネル番号が昇順となる領域(ノードの組合せ)を導き出し、180度のターンを行ってもデッドロックフリーが保証される条件を明確にする。そして、条件が満たされるときは180度のターンと従来からの再帰ルーティングとを選択的に使用するデッドロックフリーな適応型ルーティングを提案する。したがって、提案手法では先に

チャンネル番号を与える必要がある．

1D-SRT の任意のチャンネルに対し次のように番号を割り当てる．

定義 7 任意のノード  $n$  の出力チャンネルに対し，次のようなチャンネル番号  $(v, m, l)$  を割り当てる．

$v$  : 仮想チャンネル番号

$m$  :  $- n < n_{next}$  の場合 : ノード番号  $n$

$- n > n_{next}$  の場合 : サイズに対するノード番号の補数  $(N - 1 - n)$

ここで， $n_{next}$  は出力先のノード番号である．

$l$  : 出力チャンネルが結合されるリンクのレベル  $l$  □

### 3.3 適応型ルーティングの導出

本節では，定義 7 に基づきチャンネル番号が割り当てられた SRT においてデッドロックフリーが保証されるルーティング手法とその条件を明確にする．まず，monotonic order routing が可能であることを示し，次に 180 度のターンが可能となる条件を導出する．なお，ラウンドトリップループの使用法については，以後断りのない限り，次の方針に従うものとする．

(i)  $n_{cur} < n_{dst}$  のとき

$$v_{next} = \begin{cases} v_{cur} & \text{if } n_{cur} < n_{next} \vee v_{cur} = N(v) \\ v_{cur} + 1 & \text{if } n_{cur} > n_{next} \wedge v_{cur} < N(v) \end{cases}$$

(ii)  $n_{cur} > n_{dst}$  のとき

$$v_{next} = \begin{cases} v_{cur} & \text{if } n_{cur} > n_{next} \vee v_{cur} = N(v) \\ v_{cur} + 1 & \text{if } n_{cur} < n_{next} \wedge v_{cur} < N(v) \end{cases}$$

ここで， $n_{dst}$  は目的ノード， $N(v)$  は物理リンク 1 つあたりの仮想チャンネル数を表し， $N(v) \geq 2$  である．

定理 1 定義 7 に従いチャンネル番号を割り当てた 1D-SRT では，デッドロックフリーな monotonic order routing が可能である．

証明 パケットのヘッダが存在するカレントノードを  $n_{cur}$ ，転送先のノードを  $n_{next}$ ，目的ノードを  $n_{dst}$  とする．また， $n_{cur}$ ， $n_{next}$  の各出力チャンネルを  $(v_{cur}, m_{cur}, l_{cur})$ ， $(v_{next}, m_{next}, l_{next})$  とする．

(i)  $n_{cur} < n_{dst}$  のとき

monotonic order routing ではメッセージは一方方向のみ進むことが許される．そのため，チャンネルは正の方向のみ使用され，チャンネル番号は

$$n_{cur} < n_{next} \implies$$

$$(v_{cur}, m_{cur}, l_{cur}) = (v_{cur}, n_{cur}, l_{cur})$$

$$\begin{aligned} &< (v_{cur}, n_{cur} + 2^{l_{cur}}, l_{next}) = \\ (v_{next}, n_{next}, l_{next}) &= (v_{next}, m_{next}, l_{next}) \\ n_{cur} > n_{next} &\implies \end{aligned}$$

$$(v_{cur}, m_{cur}, l_{cur}) = (v_{cur}, n_{cur}, l_{cur})$$

$$\begin{aligned} &< (v_{cur} + 1, (n_{cur} + 2^{l_{cur}}) \bmod N, l_{next}) = \\ (v_{next}, n_{next}, l_{next}) &= (v_{next}, m_{next}, l_{next}) \end{aligned}$$

となり，一樣に昇順である．ここで  $l_{cur}$  は現在のパケットが選択しているチャンネルのレベルを， $l_{next}$  は次にパケットが選択するチャンネルのレベルを示す．

たとえば，32 ノードから成る 1D-SRT で， $n_{cur} = 4$ ， $n_{dst} = 14$  だと仮定すると，以下のようになる．

$$(v_{cur} = 1, m_{cur} = 4, l_{cur} = 3)$$

$$< (1, 4 + 2^3, l_{next} = 1)$$

$$= (v_{next} = 1, m_{next} = 12, l_{next} = 1).$$

また， $n_{cur} = 28$ ， $n_{dst} = 6$  の場合，

$$(v_{cur} = 1, m_{cur} = 28, l_{cur} = 3)$$

$$< (1 + 1, (28 + 2^3) \bmod 16, l_{next} = 1)$$

$$= (v_{next} = 2, m_{next} = 4, l_{next} = 1).$$

(ii)  $n_{cur} > n_{dst}$  のとき

この場合も，(i) の場合と同様な議論ができ，つねに，チャンネル番号は昇順となる．

(i)，(ii) よりつねにチャンネル番号は昇順となる．つまりデッドロックフリーな monotonic order routing が可能である． □

以上より，再帰ルーティングに monotonic order routing を施したルーティングはデッドロックフリーであることがいえる．

次に，180 度のターンが可能な領域を示す．この領域では，パケットは目的ノードをバイパスリンクを使用して飛び越え，その後で目的ノードへ引き返すことが可能である．つまり，この領域では行き過ぎと逆方向へのルーティングが可能となる．

定義 8  $n_0$  から  $n_k$  へのルーティング  $n_0, n_1, \dots, n_k$  において，少なくとも 1 つのノード  $n_i$  ( $0 \leq i < k - 1$ ) について  $n_i < n_k < n_{i+1}$  または  $n_i > n_k > n_{i+1}$  であるならば，このルーティングを同次元迂回ルーティングといい， $n_i$  を飛越ノードという．

補題 任意の飛越ノード  $n_{cur}$  が次の条件を満たすとき，転送先のノード  $n_{next}$  は飛越ノードではない．

(条件)

$$\begin{aligned} &\left( n_{cur} < n_{dst} \wedge n_{cur} < \frac{N-1}{2} - \frac{2^{l_{cur}}}{2} \right) \vee \\ &\left( n_{cur} > n_{dst} \wedge n_{cur} > \frac{N-1}{2} + \frac{2^{l_{cur}}}{2} \right) \quad (3) \end{aligned}$$

ここで、 $N$  はノード数を表す。

証明 転送先のノード  $n_{next}$  も飛越ノードであると仮定する。

- (1)  $n_{cur} < n_{dst}$  のとき  
このとき、定義 8 より、 $n_{next}$  と目的ノード  $n_{dst}$  に関して、 $n_{next} > n_{dst}$  が成り立ち

$$n_{next} > \frac{N-1}{2} + 2^{l_{cur}} \quad (4)$$

が満たされる。

一方で、飛越ノード  $n_{cur}$  と転送先ノード  $n_{next}$  には  $n_{next} = n_{cur} + 2^{l_{cur}}$  なる関係が成り立つ。したがって、 $n_{next}$  が存在する領域は

$$n_{next} < \frac{N-1}{2} + 2^{l_{cur}-1}$$

となり、式 (4) に矛盾する。

- (2)  $n_{cur} > n_{dst}$  のとき  
この場合も、(i) の場合と同様な議論ができ、仮定が矛盾していることが示される。

以上より飛越ノード  $n_{cur}$  が条件を満たすとき、転送先のノード  $n_{next}$  は飛越ノードにはなりえない。□  
定理 2 定義 7 に従いチャンネル番号を割り当てた 1D-SRT において、すべての飛越ノード  $n_{cur}$  が条件式 (3) を満たせば、同次元迂回ルーティングはデッドロックフリーである。

証明 (i)  $n_{cur} < n_{dst}$  のとき

仮定より条件式 (3) は、 $n_{next} = n_{cur} + 2^{l_{cur}}$  と書けるので、

$$\begin{aligned} n_{cur} &< \frac{N-1}{2} - 2^{l_{cur}-1}, \\ 2n_{cur} &< N-1 - 2^{l_{cur}} \\ n_{cur} &< N-1 - n_{cur} - 2^{l_{cur}}, \\ n_{cur} &< N-1 - n_{next}, \end{aligned} \quad (5)$$

はつねに成り立つ。また、補題より、 $n_{next}$  が飛越ノードとなることはなく必ずレベル 0 のリンクが使用されるため、 $n_{next}$  でつねに  $(v, N-1-n_{next}, 0)$  なるチャンネルが使用される。したがって、

$$(v, n_{cur}, l_{cur}) < (v, N-1-n_{next}, 0), \quad (6)$$

がつねに成り立ち、負方向のチャンネル  $(v, N-1-n_{next}, 0)$  への転送はつねにデッドロックが保証される。

(ii)  $n_{cur} > n_{dst}$  のとき

(i) 同様、条件式 (3) は、 $n_{next} = n_{cur} - 2^{l_{cur}}$  と書けるので、

$$\begin{aligned} n_{cur} &> \frac{N-1}{2} + \frac{2^{l_{cur}}}{2}, \\ 2n_{cur} &> N-1 + 2^{l_{cur}}, \\ n_{cur} &> N-1 - n_{cur} + 2^{l_{cur}}, \\ N-1 - n_{cur} &< n_{cur} - 2^{l_{cur}}, \\ N-1 - n_{cur} &< n_{next}, \end{aligned} \quad (7)$$

となる。また、補題より、 $n_{next}$  が飛越ノードとなることはなく、必ずレベル 0 のリンクを使用されるため、 $n_{next}$  でつねに  $(v, n_{next}, 0)$  なるチャンネルが使用される。したがって、

$$(v, n_{cur}, l_{cur}) < (v, N-1-n_{next}, 0), \quad (8)$$

がつねに成り立ち、デッドロックフリーが保証される。

以上より条件を満たす領域では、デッドロックフリーな同次元迂回ルーティングが可能である。□

### 3.4 1D-SRT の適応型ルーティング

定理 1, 2 を用いると、monotonic order routing による通常の再帰ルーティングと同次元迂回ルーティングとを選択的に使用する適応ルーティングが可能である。1D-SRT における適応型ルーティング (Adaptive Routing) は、再帰ルーティング (Recursive Routing) を用いると次のようなアルゴリズムとなる。

$$\begin{aligned} & AdaptiveRouting(n_{cur}, n_{dst}) \{ \\ & \quad \text{if}(n_{cur} < n_{dst}) \{ dir = +1 \} \\ & \quad \quad \text{else} \{ dir = -1 \} \\ & \quad n_{next} = RecursiveRouting(n_{cur}, n_{dst}) \\ & \quad \text{if}(IsBusy(n_{next}) \\ & \quad \quad \&\& Leapable(n_{cur}, n_{dst}) \\ & \quad \quad \&\& |n_{dst} - n_{cur}| > \frac{2^{l_{cur}}}{2}) \{ \\ & \quad \quad \quad n_{next} = n_{cur} + dir \times 2^{l_{cur}} \\ & \quad \quad \} \\ & \} \end{aligned}$$

ここで  $IsBusy(n_{next})$  はノード  $n_{next}$  の状態を返す関数であり、 $Leapable(n_{cur}, n_{dst})$  は  $n_{cur}, n_{dst}$  が条件式 (3) を満たすか否かを判別する関数である。また、 $|n_{dst} - n_{cur}| > \frac{2^{l_{cur}}}{2}$  は行き過ぎによって目的ノードまでの距離が増大してしまうを防ぐための条件である。 $n_{cur}$  が、 $n_{dst}$  から  $n_{next}$  に行くのに使うレベル  $l_{cur}$  のリンク長の半分の範囲に入った時点で再帰を止める。

次にチャンネルの使用法であるが、再帰ルーティングでは 2 つ以上ある仮想チャンネルの使用法はあらかじめ決められている。最初は番号が最小の仮想チャンネルを使用し、ラウンドトリップループを通過したとき、より番号の大きい方のチャンネルに切り替える。しかし、この方法では最初に使用される方が使用頻度が大きく

なる可能性がある。そこで本論文では、ラウンドトリップループを使用する場合は従来どおりにチャンネルを使用し、それ以外はパケットがネットワーク内に投入されるとき、空いている方を選択できる手法をとる。このチャンネルの使用方法はデッドロックフリーが保証される。つまり monotonic order routing はラウンドトリップループを通過しない限り、仮想チャンネルのクラスは固定である。ここでクラスとは論理的なチャンネル番号を示すものである。したがって、パケットの投入の際、2つ以上あるチャンネルのどれを選択してもチャンネル番号は一樣に昇順となり、デッドロックフリーは保証される。

### 3.5 2D-SRT の適応型ルーティング

定義 9 2D-SRT の任意のノードの出力チャンネルに対し、次のようにチャンネル番号  $(d, v, m_y, m_x, l)$  を割り当てる。

- $d$ : - チャンネルが  $x$  方向に属する場合: 0  
 - チャンネルが  $y$  方向に属する場合: 1  
 $v$ : 仮想チャンネル番号  
 $m_i$  ( $i \in (x, y)$ ):
- $n_i < n_{i_{next}}$  の場合: ノード番号  $n_i$
  - $n_i > n_{i_{next}}$  の場合: 各方向のサイズに対するノード番号の補数  $(N_i - 1 - n_i)$

ここで、 $n_{i_{next}}$  は各方向の出力先のノード番号である。また、 $N_i$  はそれぞれの方向のノード数である。

$l$ : チャンネルのレベル □

定理 3 2D-SRT について、各次元におけるルーティングを 1D-SRT の同次元迂回ルーティングとしたとき、dimension order routing はデッドロックフリーである。 □

証明 ルーティングは dimension order routing なので、パケットが  $x$  方向にあるときは  $y$  方向に関する番号を無視できる。つまり、チャンネル番号の体系は 1D-SRT のそれと等価である。また、 $y$  方向にあるときも同様である。したがって、同次元迂回ルーティングを次元オーダで用いれば 2D-SRT に対してデッドロックフリーは保証される。 □

## 4. 動的通信性能の評価

### 4.1 シミュレーション概要

同次元迂回ルーティングの性能を評価するために、C++によりシミュレータを作成し、性能評価を行った。シミュレータはフリットレベルでのシミュレーションが可能であり、トポロジ、ルーティングアルゴリズム、パケットサイズ、メッセージの発火確率などが変

表 1 シミュレーション概要  
Table 1 Parameter for simulation.

相互結合網トポロジ	1D-SRT(8, 5), 2D-SRT(5, 2, 1) MESH, ハイパーキューブ(HC)
サイズ	256(1D-SRT, MESH) 1024(2D-SRT, HC)
フロー制御方式	Wormhole
パケット長	16 flit
仮想チャンネル数	2, 3, 4(1D-SRT), 2(2D-SRT) 1(MESH) 1, 2(HC)
転送パターン	ランダム転送
シミュレーション時間	10000 clock
評価	パケット発生確率 - 平均レイテンシ

更可能である。

シミュレーションの概要を表 1 に示す。ここで、1D-SRT, 2D-SRT は Short-Span 型を対象としている。これは、基本型や Long-Span 型では最大レベルリンクの基本トラス上でのホップ数が  $N/2$  であり、ルーティングが 1 方向に制限されているので、大半がそのリンクを使用しないからである。

なお、SRT の基本構造である MESH を比較の対象とした。また、MESH およびハイパーキューブ網に関しては、デッドロック回避のために dimension order routing を適用した。

### 4.2 適応型ルーティングの動的性能

ここでは、各ルーティングアルゴリズムの一樣なランダム転送を対象とした動的通信性能評価を行う。シミュレーションは、デッドロックフリーな再帰ルーティング(DLF)とそれに同次元迂回ルーティングを付加した適応型ルーティング(ADP)に対して行った。

なお、シミュレーションはある発生確率でランダムな宛先にメッセージを発生させ、その平均通信時間を測定した。シミュレーション時間は各発生確率において 10000 クロック行った。ここで、発生確率はフリットが毎クロック投入される確率を 1.0 とする。

図 4 に 1D-SRT に同次元迂回ルーティングを適用した際のメッセージ発生確率と平均メッセージ遅延時間の関係を示す。使用した仮想チャンネル数は、MESH の場合 1, SRT の場合は 2 である。提案した同次元迂回ルーティングによる適応型ルーティング(ADP)は、通常のデッドロックフリーな再帰ルーティング(DLF)に比べ高い性能が得られた。ネットワーク内の通信量が最大となり平均通信時間が飽和する付近では、再帰ルーティングが約 0.05 (flit/clock) であるのに対し適応型ルーティングでは約 0.065 (flit/clock) であり、性能向上比は約 1.3 倍あった。これは、256 PEs の 1D-SRT では同次元迂回ルーティングを適用できる範囲が

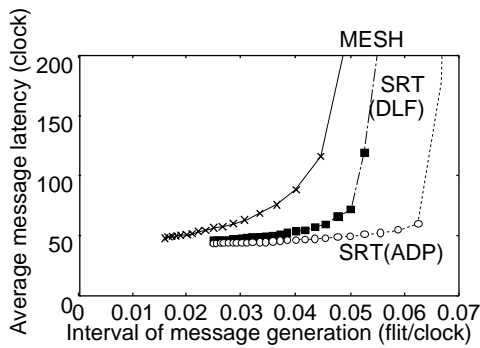


図4 1D-SRTとメッシュ網の平均メッセージ遅延時間

Fig. 4 Average message latencies of 1D-SRT and mesh as a function of interval of message generation.

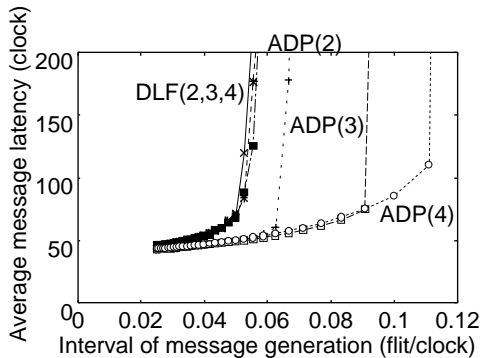


図5 仮想チャネル数による平均メッセージ遅延時間

Fig. 5 Average message latency for the number of virtual channels.

十分に大きく、迂回により混雑が回避できたためであると考えられる。また、メッシュ網 (MESH) と比較しても、平均通信時間が飽和する付近では、メッシュ網が約 0.035 ( flit/clock ) であるのに対し 1D-SRT では約 0.065 ( flit/clock ) と約 1.8 倍の性能向上が見られた。

図 5 にさらに仮想チャネルを付加した場合の動的通信性能を示す。図中、DLF ( $L$ ) はデッドロックフリーな再帰ルーティングを示し、ADP ( $L$ ) は同次元迂回ルーティングを示す。ここで、括弧内  $L$  は、仮想チャネル数を表す。図 5 より、適応型ルーティングでは仮想チャネルを増設することにより、さらに性能を上げることができた。特に、仮想チャネル数を 4 にしたときの性能向上が著しく、再帰ルーティング (DLF(4)) が約 0.045 ( flit/clock ) で平均通信時間が飽和してしまうにもかかわらず、適応型ルーティング (ADP(4)) の飽和点は約 0.105 ( flit/clock ) である。

図 6 に 2D-SRT に同次元迂回ルーティング ( 仮想

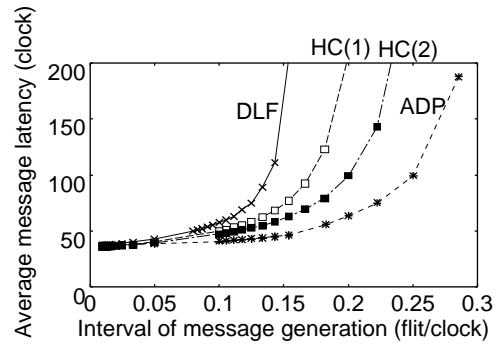


図6 2D-SRTとHC網の平均メッセージ遅延時間

Fig. 6 Average message latencies of 2D-SRT and HC as a function of interval of message generation (2D-SRT, HC).

チャネル数 2) を適用した際のメッセージ発生確率と平均メッセージ遅延時間の関係を示す。提案する適応化手法は通常のデッドロックフリーな再帰ルーティング (DLF) に比べ高い性能が得られた。また、HCと比較しても、再帰ルーティングでは性能が劣るものの、同次元迂回ルーティングでは HC よりも高い性能を得ることができた。ネットワークのノード数が 1024 PEs の場合、HC のリンク次数が 10 であるのに対し 2D-SRT のリンク次数は 8 と少なく、動的通信特性に優れており、SRT は並列計算機向きの相互結合網である。

## 5. まとめ

SRT の適応型ルーティングの提案を行い、デッドロックフリーを証明した。この適応型ルーティングは、既存の再帰ルーティングとの併用が可能で、仮想チャネルを新たに付加する必要がない。また、シミュレーションによる動的通信性能評価を行った。その結果、1D-SRT、2D-SRT とともに従来の再帰ルーティングに比べ高い性能が得られることを示した。特に、2D-SRT は 1024 PEs のネットワークサイズで、ハイパーキューブよりもリンク次数が少ないにもかかわらず高い動的通信性能を得ることができた。

今後の課題はリンクやノードの故障に対する通信性能の解析、および仮想チャネルを付加した場合の性能評価である。

謝辞 本論文の一部は、文部省科学研究助成金、ならびにセコム科学技術振興財団研究奨励金を用いて行われた。関係各位に感謝する。

## 参考文献

- 1) Inoguchi, Y. and Horiguchi, S.: Shifted Recursive Torus Network for Mesh-Oriented In-



terconnections, *Proc. 31st Conference on Information Sciences and Systems*, Baltimore (Mar. 1997).

- 2) Inoguchi, Y. and Horiguchi, S.: Shifted Recursive Torus Interconnection for High Performance Computing, *IEEE High Performance Computing in Asia Conference*, Seoul, pp.61–66 (Apr. 1997).
- 3) 楊 愚魯, 天野英晴, 柴村英智, 末吉敏則: 超並列計算機に向き結合網: RDT, 信学論 (D-I), Vol.J78-D-I, No.2, pp.118–128 (1995).
- 4) Kirkman, W.W. and Quammen, D.: Packed Exponential Connections – A Hierarchy of 2D-Meshes, *Proc. 5th International Parallel Processing Symposium*, pp.464–470 (Apr. 1991).
- 5) 川井雅之, 井口 寧, 堀口 進: 超並列計算機向き相互結合網 SRT のデッドロックフリー・ルーティング, 情報処理学会論文誌, Vol.40, No.5, pp.1977–1984 (1999).
- 6) Duato, J.: A New Theory of Deadlock-Free Adaptive Routing in Wormhole Networks, *IEEE Trans. Parallel and Distributed System*, Vol.4, No.12 (Dec. 1993).
- 7) Duato, J.: A Necessary and Sufficient Condition for Deadlock-Free Adaptive Routing in Wormhole Networks, *IEEE Trans. Parallel and Distributed System*, Vol.6, No.10 (Dec. 1995).
- 8) Glass, C.J. and Ni, L.M.: Maximally Fully Adaptive Routing in 2D Meshes, *Proc. ISCA92*, pp.278–287 (1992).
- 9) 川井雅之, 井口 寧, 堀口 進: 超並列計算機向き相互結合網 SRT における適応型ルーティング, *HOKKE-99*, pp.13–18 (Mar. 1999).
- 10) Ni, L.M. and McKinley, L.P.: A Survey of Wormhole Routing Technique in Direct Networks, *IEEE Trans. Comput.* (1993).

(平成 11 年 6 月 16 日受付)

(平成 12 年 5 月 11 日採録)



川井 雅之

昭和 49 年生。平成 9 年法政大学工学部システム制御工学科卒業。平成 11 年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了。現在 (株) 日本総合研究所。



井口 寧 (正会員)

昭和 42 年生。平成 3 年東北大学工学部機械工学科卒業。平成 9 年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了。現在、同大学院情報科学センター助手。また、平成 6～9 年日本学術振興会特別研究員として研究に従事。この間並列システムに関する研究を行う。IEEE、電子情報通信学会各会員。



堀口 進 (正会員)

昭和 51 年東北大学工学部通信工学科卒業。昭和 56 年同大学院博士課程修了。昭和 57 年同大学情報工学科助手。昭和 64 年同助教授。現在、北陸先端科学技術大学院大学情報科学研究科教授。この間、並列処理、超並列システム、ウェーハ規模集積システム、並列アルゴリズム、マルチメディア統合システムに関する研究を行う。昭和 61 年 6 月～昭和 62 年 7 月米国 IBM ワトソン研究所・客員研究員として並列計算アルゴリズムの研究に従事。平成 6 年 7～8 月ルイジアナ州立大学 (ラファエット) 客員教授、平成 9 年 7～8 月テキサス A&M 大学客員教授。著書「UNIX と Pascal」(共著)、「UNIX と C」(共著) など。昭和 64 年～平成 7 年 IEEE 学会 WSI 国際会議組織委員会委員。平成 4～7 年同国際会議アジア地域議長。IEEE シニア会員、電子情報通信学会会員。