

## SCI を構成要素として用いた各種基本的相互結合網の評価

高橋 正人(†)、青山 和弘(‡)、宮田 裕行(†)、菅 隆志(†)

三菱電機株式会社 情報技術総合研究所(†) 三菱電機株式会社 鎌倉製作所(‡)

ユーザに対してプログラムし易い共有メモリプログラミングパラダイムを提供するという要件と、性能面で高水準の拡張性を備えるという要件を、一つの並列計算機において同時に実現することは、一般に難しい。しかし、この二要件を同時に満たす並列計算機を、従来を越える水準で実現することに対する潜在的需要は依然大きく、この二要件を実現するための実装技術が研究されている。この研究領域で、近年重視されているのが、IEEE が策定した、Scalable Coherent Interface(SCI)と呼ばれる要素技術規格である。SCI は、従来のバス構成を越える高い拡張性と、共有メモリプログラミングパラダイムの同時に実現という目的に徹したインタフェース規格である。しかし、SCI を用いて並列計算機を構築する場合に、最適な相互結合網は何かという設計選択に対する定量的なアプローチは、現在のところ見あたらない。そこで本稿では、SCI を用いて並列処理システムを構築しようとする並列計算機アーキテクトが設計初期段階に設計選択として直面せざるを得ない、どの型の相互結合網が最適かという問題に定量的指標を与えることのできる数理的手法を提案する。結果の一例として、ある条件下で 64 ノード程度の規模のシステムを構築するには、リング乗換コストと通常リンク伝達遅延コストの比が重要であり、それが 5 以下の場合に、双方向二次元トラスが優れた性能を発揮し、それを大きく上まわる際には双方向リングが優れた性能を発揮すること等がわかった。

## An Evaluation of Network Topologies using SCI interface

Masato TAKAHASHI(†) Kazuhiro AOYAMA(‡)

Hiroyuki MIYATA(†) Takashi KAN(†)

Mitsubishi Electric Corp. Information Technology R&D Center(†)

Mitsubishi Electric Corp. Kamakura Works(‡)

5-1-1 Ofuna, Kamakura, Kanagawa 247, Japan

It has been one of main goals for parallel computer architects to design and develop a parallel machine which simultaneously provides both good programability and high scalability. Recently, IEEE approved a standard Interface called "Scalable Coherent Interface", which can be an enabling technology to develop such a parallel machine in higher level than ever. In this paper, we propose a mathematical method useful for parallel computer architects to select the most suitable interconnect network topology using SCI as a component of the network. As a conclusion, we found double-linked 2 dimensional torus is a good candidate for 64 node systems when the ratio of SCI agent delay cost to SCI link delay cost can be kept less than 5.

## 1 背景と動機

### 共有バス構成マルチプロセッサの持つ限界と SCI の登場

共有バス / 共有メモリ構成並列マシンは、プログラマビリティに優れ、ユニプロセッサシステムの自然な拡張として実装も比較的容易であるため、これまで多くの商用マシンが登場している [1]。しかし、この構成は共有バス、共有メモリにおけるアクセス競合が性能隘路となり、拡張性に乏しいという限界を持っている。この限界を越えるために、IEEE の作業部会は検討を続け、Futurebus (IEEE 960)、Futurebus+ (IEEE986x) 等の高速・高機能バスを生み出した [1]。この検討は、その後もさらに P1596 ワーキンググループによって引きつがれたが、プログラマビリティと拡張性という相反する課題を、従来をより越える水準で実現するためには、バス構成で高機能化を目指す方向ではなく、共有分散メモリシステムを実現しうる、高速かつ高機能なリングベースの規格が必須との認識に至った。そこで策定された規格が、次世代標準となるべき通信インタフェース要素としての、Scalable Coherent Interface (SCI) [2] である。

### SCI の特徴と将来性

SCI の特徴とその将来性については次の三点が考えられている。第一に、SCI はリングベース規格であるため、ノード数の増加に際して、共有バスで見られるような競合の集中が生じないばかりか、共有バスでは不可能であったデータ並行転送ができる。これは、1 ノードあたりのバンド幅が、バス構成に比べての飛躍的に向上される結果を生む。1 Gbyte/sec という高速なデータ転送速度も、上記の特徴と相乗的に働き、SCI の将来性を高めている。第二に、SCI ではプログラマビリティの要請が視野に入っているため、ディレクトリベースのキャッシュ整合性制御方式が規定されている。これは、キャッシュ整合性制御のハードウェア実装への障壁を低くするものである。この制御をハードウェア実装し制御オーバーヘッドを縮小することで、システム性能劣化の主要因のひとつを回避できる。この特徴により SCI の実際の価値は極めて高いものとなった。第三に、SCI は IEEE 策定の標準規格であるため、ベンダー・サードパーティ各社の市場参入により、市場全体としての開発コストが低減され、さらにそれに伴って SCI 方式の採用の活性化という好循環が期待されている。図 1 に SCI のノードモデルを示す。

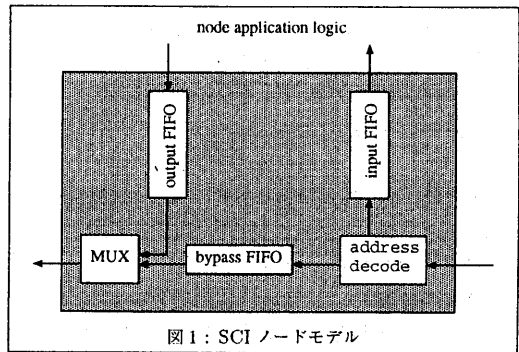


図 1: SCI ノードモデル

### SCI を構成要素とする相互結合網

SCI を構成要素として、ある相互結合網を構築する場合、SCI がリングベースの規格であるために、SCI での用語で SCI エージェントと呼ばれる、あるリングから別のリングへ乗り換える機能を持つノードが必要となることがある。この SCI エージェントをボードレベルで新規作製することができる場合には、SCI エージェント通過遅延時間と通常 SCI リンク通過遅延時間の比は、比較的小さく 5 以下であると考えられる。一方、既存の複数の計算機を、既成の SCI ボードを使用して接続しようとするとき、1 つの計算機に複数の SCI ボードを差して、SCI エージェントを構成する。このように構成された SCI エージェントをパケットが通過する際には、パケットが一旦計算機の内部バスを通過する。そのため、そのような SCI エージェント通過遅延時間は、SCI の通常のリンク通過遅延時間の 10 ~ 100 倍あるいはそれ以上であることが予想される。そこで、設計者が設計選択として、なんらかの相互結合網を選択する場合に、伝達遅延に関しての検討は、各相互結合網に関する、数学的に既知な距離の概念に基づく特性値 (最大距離・平均距離等) に基づく評価だけでは不十分で、その網の構成に必要な SCI エージェントでの遅延コストも含んだレベルでの、その相互結合網の特性値の評価が重要になってくる。本稿ではこの観点から、SCI を構成要素として相互結合網候補を構成するとした場合に現実的と思われる 6 種類の相互結合網候補について、それらの相互結合網の構築の際に必須の SCI エージェントの存在によって、それらの相互結合網がどのような伝達遅延特性になるかを数理的に検討する。

### 評価方法

相互結合網として、今回は、リング、二次元メッシュ、二次元トラスの各々につき、(一重リング構成) 単方向性構造と (二重リング構成) 双方向構造の、計 6 つの相互結合網を対象とする (図 2 参照)。まず、これらの相互結合網について、どこにどのような SCI エージェント機能が必要になるかを検討した。次いで、SCI エージェントを含む各相互結合網について、(1) 最大距離経路 (2) 平均距離経路 (3) 重みづけ平均距離経路、に相当する 3 経路での伝達遅延 ((2) および (3) では伝達遅延の期待値)、の理論式を導出し、わかりやすさのためにグラフ化する。最後に、この式から得られる評価の例として SCI を構成要素とする 16、ないし、64 ノード規模のシステムにおいて最適相互結合網の選択される例を示す。

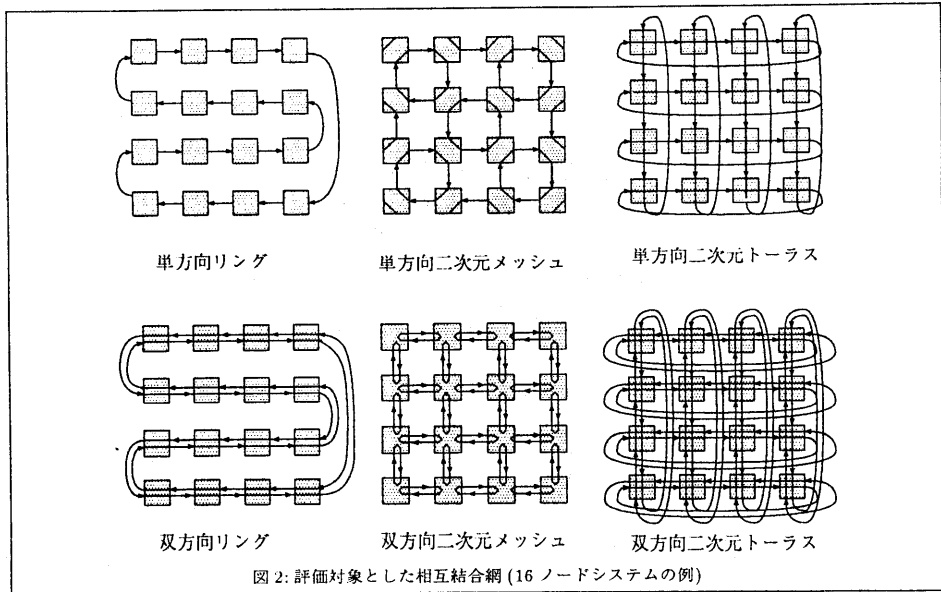


図2: 評価対象とした相互結合網 (16 ノードシステムの例)

## 2 評価方法

特定ノードへの伝達遅延時間 ノード0からノード*i*への伝達遅延時間  $T_{node_i}$  は、 $\Omega, \gamma, N, i, node_i, Nlink, Nagent, g, G$  を表1のように定義して、次のように与えられるとする。

$$T_{node_i}(\Omega, \gamma, N, g, G) = [g, G] \begin{bmatrix} Nlink(\Omega, \gamma, node_i, N) \\ Nagent(\Omega, \gamma, node_i, N) \end{bmatrix} \dots\dots (1)$$

変数 (関数) 名	定義	備考
$node_i$	ノード番号 $i$ ( $i=0,1,2,\dots,N-1$ ) のノード	
$N$	ノード総数	$N = n^2$ ( $n$ は自然数). $\Omega_{meshsingle}$ のみ $N = (4n)^2$
$T_{node_i}$	ノード0からノード <i>i</i> への伝達遅延時間	
$\Omega$	相互結合網のタイプ	図2の6網を対象
$\gamma$	ルーティングアルゴリズム	各 $\Omega$ に一対一に対応
$g$	SCIリング内で隣接するノード間でのデータ転送に必要な時間コスト	
$G$	SCI エージェントノードにおいて、SCI エージェント機能を使用する (SCIリングを乗り換える) 際の時間コスト	
$Nlink(\Omega, \gamma, node_i, N)$	SCIリングにより実現された相互結合網 $\Omega$ において、ルーティングアルゴリズム $\gamma$ に従うときに、ノード $node_0$ からノード $node_i$ へのアクセスを実現するために通過するSCIリンクの個数	算出方法は Appendix を参照
$Nagent(\Omega, \gamma, node_i, N)$	SCIリングにより実現された相互結合網 $\Omega$ において、ルーティングアルゴリズム $\gamma$ に従うときに、ノード $node_0$ からノード $node_i$ へのアクセスを実現するためにSCI エージェント機能を使用する (SCIリングを乗り換える) 回数	算出方法は Appendix を参照

表1: 式(1)における変数等の定義

### あるアクセスパターン下での伝達遅延時間の期待値

各ノードへのアクセスパターン (各ノードへのアクセス確率が一律、あるいは、距離に反比例する重みづけを持ったアクセス確率) での、1アクセス当りの伝達遅延の期待値を得るために、表2のように各変数を定義する。1アクセス当りの伝達遅延時間の期待

値は、次のようになる。

$$\begin{aligned}
 ET(\Omega, \gamma, pattern, N, g, G) &= \sum_{i=0}^{N-1} prob(pattern, node_i, N) T_{node_i}(\Omega, \gamma, N, g, G) \\
 &= \sum_{i=0}^{N-1} prob(pattern, node_i, N) [g, G] \begin{bmatrix} Nlink(\Omega, \gamma, node_i, N) \\ Nagent(\Omega, \gamma, node_i, N) \end{bmatrix} \\
 &= [g, G] \begin{bmatrix} \sum_{i=0}^{N-1} prob(pattern, node_i, N) Nlink(\Omega, \gamma, node_i, N) \\ \sum_{i=0}^{N-1} prob(pattern, node_i, N) Nagent(\Omega, \gamma, node_i, N) \end{bmatrix} \\
 ET(\Omega, \gamma, pattern, N, g, G) &= [g, G] \begin{bmatrix} ENlink(\Omega, \gamma, pattern, N) \\ ENagent(\Omega, \gamma, pattern, N) \end{bmatrix} \dots\dots (2)
 \end{aligned}$$

この右辺第二項の2行1列の行列は、「SCIで構成された各結合網Ωとそれを実現する最も現実的なルーティングアルゴリズムγにおいて、ノードへのアクセス様式がpatternによって反映されるとき、1アクセス当りの、リンク通過個数の期待値、および、SCIエージェント機能の使用回数の期待値」を示している。

この行列を、各Ω, γ, patternについて、本稿での評価対象である組合せ(それが表3に示されている)、について解いて、Nについての簡単な式として整理し、一覧表にまとめたものが、表4である<sup>1)</sup>。

変数(関数)名	定義	備考
$ET(\Omega, \gamma, pattern, N, g, G)$	SCIリングで実現された、総ノード数Nの相互結合網Ωが、ルーティングアルゴリズムγに従うときに、patternなるアクセス確率分布のアプリケーションが稼働する場合、一回のアクセスに関する伝達遅延時間の期待値	
pattern	ノードnode <sub>i</sub> アクセスパターンの指定	評価対象は、最大距離アクセス、一様アクセス確率、距離に反比例するアクセス確率の三通り
$prob(pattern, node_i, N)$	patternで指定されるアクセスパターン下での、各ノードnode <sub>i</sub> へのアクセス確率を定義する関数	式は Appendix を参照
$ENlink(\Omega, \gamma, pattern, N)$	SCIリングにより実現された、総ノード数Nからなる相互結合網Ωにおいて、ルーティングアルゴリズムγに従うときに、patternで指定されるアクセス確率分布を持つアプリケーションが稼働の場合に、1回のアクセス当りに通過するSCIリンクの個数の期待値	算出方法を Appendix に示す、また結果を表4に示す。
$ENagent(\Omega, \gamma, pattern, N)$	SCIリングにより実現された、総ノード数Nからなる相互結合網Ωにおいて、ルーティングアルゴリズムγに従うときに、patternで指定されるアクセス確率分布を持つアプリケーションが稼働の場合に、1回のアクセス当りにSCIエージェント機能を使用する(SCIリングを乗り換える)回数の期待値	算出方法を Appendix に示す、また結果を表4に示す。

表2：式(2)とその準備式における変数等の定義

Ω (相互結合網のタイプ)	γ (ルーティングアルゴリズム)	pattern (アクセスのパターン)	
単方向リング	ディメンジョンオーダールーティング ([5])	最大距離	一様 距離に反比例
双方向リング	ディメンジョンオーダールーティング	最大距離	一様 距離に反比例
単方向二次元メッシュ	特殊型ルーティング	最大距離	
双方向二次元メッシュ	ディメンジョンオーダールーティング	最大距離	
単方向二次元トラス	ディメンジョンオーダールーティング	最大距離	一様 距離に反比例
双方向二次元トラス	ディメンジョンオーダールーティング	最大距離	一様 距離に反比例

表3：評価対象としたΩ, γ, patternの組合せ

以降Ω, γ, patternに基づいて表4から行列を選択し、式(2)において、N, g, Gを代入することで、1アクセス当りの伝達遅延時間の期待値を比較的簡単に比較検討できる準備が整った。最適な網の選択を目的として、H/Wレベルの条件などに基づいてg, G, Nの変数に関するグラフを図3～図7について順次検討してゆくことにする。

<sup>1)</sup>なお、表4ではSCIを使用した相互結合網の設計選択という一覧の便宜のため、H/Wコスト見積りに重要となる項目、使用時のパーティショニングの利便性の実現の可否、ついで結論をさきどりして特定の(g, g, N)の組について(つまり我々が関心を持っているシステム規模・構成時に於ける)各網の総合評価の項目も並記した。

遅延時間期待値のグラフからの考察

図3、4、5においては、他のノードにあるデータへのアクセス頻度を変えた場合の各網における伝達遅延時間をノード数Nの関数として示した。図3では、最大距離にあるノードのみにアクセスするとした場合、図4では、自ノードを除いて他のノードに一樣にアクセスするとした場合、図5では、他ノードへのアクセスする確率を自ノードからの距離に反比例するとした場合に対応している。

最大距離へのアクセス まず、図3を検討する。これは最大距離に位置するノードへのアクセス確率が1でそれ以外が0であるケースを示す。結果を見ると網の形態によってははっきりと傾向がわかる。双方向二次元メッシュの性能が極端に悪い。ノードからノードへ転送される度に、必ずエージェントを使用するためにこのような結果となってくる。単方向二次元メッシュもNの増加につれて性能が悪くなる。エージェンを持たない二種類のリング構成はNにつれて直線的に遅延が増加している。単方向リングは、勾配が大きく、問題がある。双方向リングでは比較的良い。二種類の二次元トラスと、双方向リング、については、比較的接近した良い性能をどちらも維持していると言える。拡張性についても問題ないと考えられる。以上の検討と、H/Wコストつまり結線の複雑さからも、二次元メッシュについては有利は認められないことから、以降の検討では対象外とする。

SCIを採用した各網での、伝達遅延時間の見積り要素項目	クラスタにタスク投入時のタスク分配の様相を反映	分散共有メモリプログラミングパラダイムで作成されたタスクの稼働様相を反映	SCI キャッシュコヒーレンシトランザクションとデータ転送の混合したネットワークトラフィックを反映	ノードのH/Wレベルの作製コストを反映	
	最大距離における	一様アクセス確率下における	距離に反比例するアクセス確率下における	1ノード上のポート対の組数	1ノード上のリング乗り換え回路の対の組数
	1アクセス当りの、		SCIリンク通過個数の期待値	SCIエージェント使用回数の期待値	
単方向リング	$\begin{bmatrix} N-1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} \frac{N-1}{2} \\ 0 \end{bmatrix}$	$\begin{bmatrix} \sum_{i=1}^{N-1} \frac{1}{i} \\ 0 \end{bmatrix}$	1	0
双方向リング	$\begin{bmatrix} \frac{N}{2} \\ 0 \end{bmatrix}$	$\begin{bmatrix} \frac{N}{4} \\ 0 \end{bmatrix}$	$\begin{bmatrix} \frac{N-1}{N+2} \sum_{i=1}^{N/2-1} \frac{1}{i} \\ 0 \end{bmatrix}$	2	1
単方向二次元メッシュ	$\begin{bmatrix} 2\sqrt{N}-1 \\ \sqrt{N}-2 \end{bmatrix}$	<i>omitted</i>	<i>omitted</i>	2	1
双方向二次元メッシュ	$\begin{bmatrix} 2(\sqrt{N}-1) \\ (2\sqrt{N}-3) \end{bmatrix}$	<i>omitted</i>	<i>omitted</i>	4	4
単方向二次元トラス	$\begin{bmatrix} 2(\sqrt{N}-1) \\ 1 \end{bmatrix}$	$\begin{bmatrix} \sqrt{N}-1 \\ (1-\frac{1}{\sqrt{N}})^2 \end{bmatrix}$	$\begin{bmatrix} \sum_{i=0}^{\sqrt{N}-1} \sum_{j=0}^{\sqrt{N}-1} \frac{1}{i+j} \\ (i,j \text{ は同時に } 0 \text{ をとらない}) \\ (1 + \frac{2 \sum_{i=1}^{\sqrt{N}-1} \frac{1}{i}}{\sum_{i=1}^{\sqrt{N}-1} \sum_{j=1}^{\sqrt{N}-1} \frac{1}{i+j}})^{-1} \end{bmatrix}$	2	1
双方向二次元トラス	$\begin{bmatrix} \sqrt{N} \\ 1 \end{bmatrix}$	$\begin{bmatrix} \frac{\sqrt{N}}{2} \\ (1-\frac{1}{\sqrt{N}})^2 \end{bmatrix}$	$\begin{bmatrix} \frac{N}{\sqrt{N} + 4(f_1(N) + f_2(N) + f_3(N))} \\ (1 + \frac{4(\frac{1}{\sqrt{N}} + f_1(N))}{\sqrt{N} + 4(f_2(N) + f_3(N))})^{-1} \end{bmatrix}$	4	4

表4：SCIを用いた各網における評価項目 表中Nはノード総数を、また、 $f_1(N) = \sum_{i=1}^{\sqrt{N}/2-1} \frac{1}{i}$ 、 $f_2(N) = \sum_{i=1}^{\sqrt{N}/2-1} \frac{1}{i+\sqrt{N}/2}$ 、 $f_3(N) = \sum_{i=1}^{\sqrt{N}/2-1} \sum_{j=1}^{\sqrt{N}/2-1} \frac{1}{i+j}$ を表わすものとする。

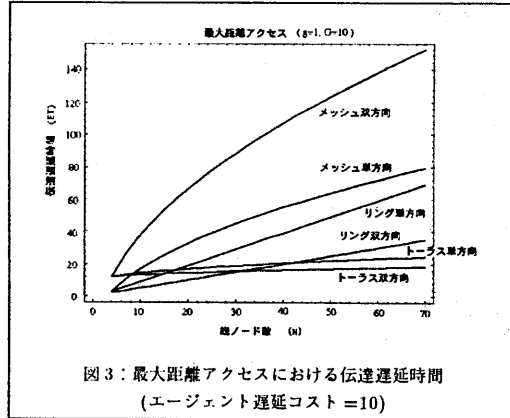


図3：最大距離アクセスにおける伝達遅延時間  
(エージェント遅延コスト=10)

一様アクセス確率 次いで、図4を検討する。他のノードに一様アクセス確率を持つ場合である。ここでの確率分布は、ユーザが分散共有メモリプログラミングパラダイムで作成したタスクの稼働様相を反映するものと考えられる。

先の評価で性能面でも悪く、H/Wコスト面でも実現性の乏しい、二次元メッシュの二種類は除いてある。

突出している単方向リングを除いて、二次元メッシュ(双方向・単方向)とリング双方向は同程度の伝達遅延時間が期待できると見られる。

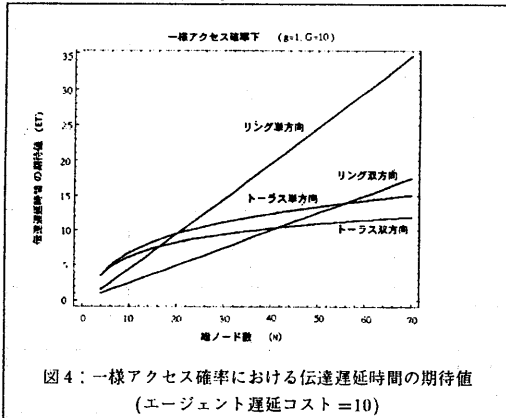


図4：一様アクセス確率における伝達遅延時間の期待値  
(エージェント遅延コスト=10)

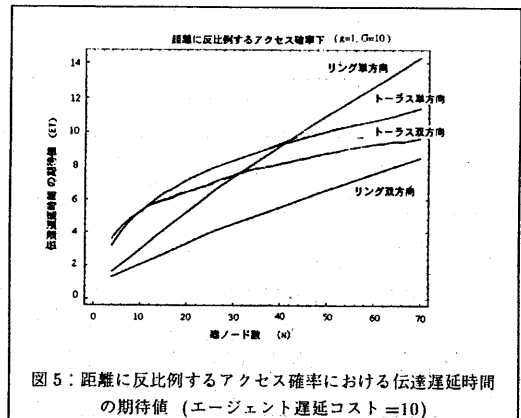


図5：距離に反比例するアクセス確率における伝達遅延時間の期待値  
(エージェント遅延コスト=10)

距離に反比例するアクセス確率 次いで、図5を検討する。距離に反比例するアクセス確率を持つ場合である。このようなアクセス確率の現実的裏付けとしては、近傍との比較的頻繁なデータ転送と、全ノードに拡散して薄く広く生起するキャッシュコヒーレンシと、いう両者のミックス負荷としての意味づけがあると考えられる<sup>2</sup>。

この状況下では、図4に比べて全体に遅延時間が減少していることを考慮にいと、網の間の差が少なくなっているとも言えるが、二次元トラス(双方向・単方向)より、双方向リングが良い性能を示していることが興味深い。

エージェント遅延コストからの検討 次いで、図6、図7を検討する。これは、N=16、64のとき、SCIエージェントの通過遅延コストの大小と、各網における性能の良否の交互作用の検出を主な関心としてグラフ化を行なったものである。距離に反比例するアクセス確率である。図6から、16ノード程度では、エージェント遅延のコストを10前後とすれば双方向リングで十分である。図7から、64ノード程度になると、エージェント遅延コストが小さな値では双方向二次元トラスか、双方向リングのどちらが優勢となるかがわかる。

<sup>2</sup>  $y = 1/x$  のグラフの形状が、原点に左下隅を持つ縦長の長方形と、同じく原点に左下隅を持つ横長の長方形の組合せでイメージされることを考える。

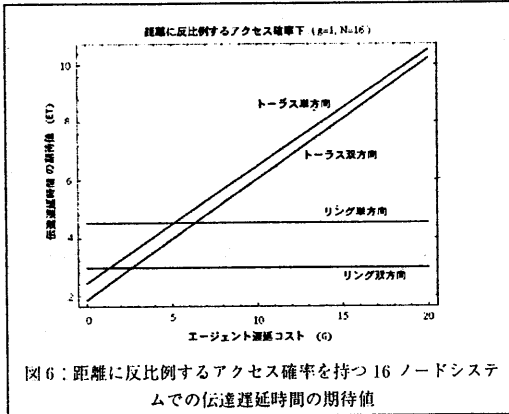


図6：距離に反比例するアクセス確率を持つ16ノードシステムでの伝達遅延時間の期待値

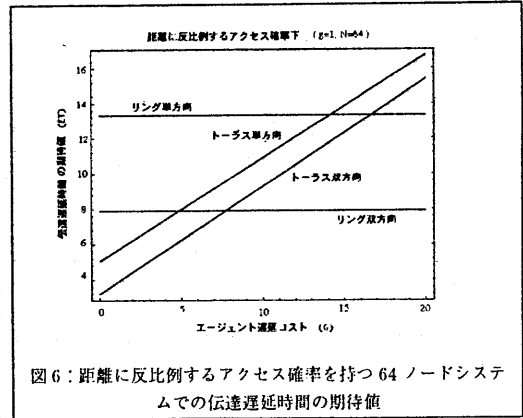


図6：距離に反比例するアクセス確率を持つ64ノードシステムでの伝達遅延時間の期待値

### 3 まとめ

#### 結論

リングベース規格であるSCIを構成要素としてなんらかの相互結合網を構築する際には、「ノード数(システム規模)の大小」、および、「SCIエージェント通過遅延コストとSCI通常リンク通過遅延コストの比の大小」により、最適網構成が異なってくる。このことを背景に、本稿では、上記二つの変数から、SCIが特定の相互結合網を構成した場合の、伝達遅延時間の期待値を得るための数理的手法を提案した。この手法を用いることにより、SCIによって構成される各種の相互結合網の特性を定量的に比較検討することができる。特に並列計算機設計者が設計初期段階で網を設計選択するとき上記の検討例1、検討例2、のような定量的指標を得られる点で有用である。

#### 検討例1：16ノードシステム

ノード数が16程度のシステムにおいては次のことがあきらかとなった。エージェント機能の遅延がリンクの伝達遅延の約5倍程度以上の場合には、SCIエージェントを持たない双方向二重リング構成が、SCIリンクの高速性を最も生かすことが可能となる。二次元メッシュは単方向、双方向にかかわらず、SCIエージェント機能の遅延のため非常に性能が劣化してしまう傾向にある。二次元トラスは、エージェント通過遅延がSCIの通常リンク通過遅延の2～3倍程度であるときに、双方向リング構成よりも性能面での優位性を発揮する。しかし、両者の性能面での程度は比較的小さく、性能面ではリングと同程度であることが予想されるため、実装のH/Wコストまでも含めると、双方向リングの方が有利であろう。

#### 検討例2：64ノードシステム

一方、ノード数が64台程度になると、エージェント通過遅延とSCI通常リンク通過遅延の比が7～8程度のところに臨界点が見いだされている。両者の比が5以下程度のものであれば、SCIをエージェントを持たないリング構成よりも、SCIエージェントを使用することで経路を有効に短縮しうる二次元トラスの方に伝達遅延の面で優位性が発生してくることが判明した。一方先述の比が、15程度を上まわるときには、16ノードの時と同様に単純な双方向リングの優位性が顕著となってくることが明らかとなった。

#### 今後の課題

本稿で行なった検討の範囲では、双方向リング構成の評価が比較的高かった。このことから、高速なSCIリングの利点を活かすことができる双方向リングを少数のエージェントで接続した構成が、良好な性能を発揮することが推察された。この意味での今後の課題として、複数の二重双方向リングの特殊な結合形態や、ツイステッドトラス構成などを評価対象に加えて行くことを計画している。また、SCIを使用したボードを開発予定であるため、それを用いた実機評価による精度検証を予定している。

#### 謝辞

本研究に関して、貴重な助言と示唆を数多く頂きました。川田圭一 三菱電機情報技術総合研究所 前次世代方式技術開発部長、下間芳樹 情報処理基盤部長、中島克人 知識処理チームリーダー、古市昌一氏、石塚裕一氏、山崎高日子氏、大谷治之氏の各氏に感謝いたします。

参考文献

- [1] "Advanced Computer Architecture with Parallel Programming (Preliminary Edition)", Kai Hwang, McGraw Hill, 1993
- [2] "IEEE Standard for Scalable Coherent Interface (SCI)", IEEE Computer Society, IEEE Std 1596-1992
- [3] "Proceedings of the First International Workshop on SCI-Based High-Performance Low-Cost Computing", Aug.17-18, 1994, at Snta Clara Univ., SCIzzL
- [4] "Tutorial Notes of The First Proceedings of the First International Workshop on SCI-Based High-Performance Low-Cost Computing", Aug.17-18, 1994, at Snta Clara Univ., SCIzzL
- [5] Planar-Adaptive Routing (PAR): Low-cost Adaptive Networks for Multiprocessors, Jae Hoon Kim, Master Thesis, Urbama, Illinois Univ. 1993

Appendix

prob(pattern, node<sub>i</sub>, N) の評価式

$$\begin{aligned}
 \text{prob}(\text{pattern}_{\text{maximumdistance}}, \text{node}_i, N) &= \begin{cases} 1 & \text{if } N\text{link}(\Omega, \gamma, N, \text{node}_i) = \max_{i=1, \dots, N-1} (N\text{link}(\Omega, \gamma, N, \text{node}_i)) \\ 0 & \text{otherwise} \end{cases} \\
 \text{prob}(\text{pattern}_{\text{uniform}}, \text{node}_i, N) &= \begin{cases} 0 & \text{if } i = 0 \\ \frac{1}{N-1} & \text{otherwise} \end{cases} \\
 \text{prob}(\text{pattern}_{\text{neighbourweighted}}, \text{node}_i, N) &= \begin{cases} 0 & \text{if } N\text{link}(\Omega, \gamma, N, \text{node}_i) = 0 \\ \frac{N\text{link}(\Omega, \gamma, N, \text{node}_i)}{\sum_{k=1}^{N-1} N\text{link}(\Omega, \gamma, N, \text{node}_k)} & \text{otherwise} \end{cases}
 \end{aligned}$$

Nagent(Ω, γ, node<sub>i</sub>) の評価式

SCI リングレットを基本要素として採用した網Ω上において、ノード node<sub>i</sub> に対し、ルーティングアルゴリズム γ に従ってアクセスする際の通過 SCI エージェントの個数。(floor(x) は x 以下の最大の整数(切り下げ)を与える関数、ceiling(x) は x 以上で最小の整数(切り上げ)を与える関数とする。)

$$\begin{aligned}
 N\text{agent}(\Omega_{\text{ringsingle}}, \gamma_{\text{anyalgorithm}}, i) &= 0 \\
 N\text{agent}(\Omega_{\text{ringdouble}}, \gamma_{\text{anyalgorithm}}, i) &= 0 \\
 N\text{agent}(\Omega_{\text{meshsingle}}, \gamma_{\text{specificalgorithm}}, (i, j)) &= \begin{cases} 2 \times \text{floor}(\frac{j}{2}) & \text{if } i = 0 \\ \max_{i,j}(i, j) - 1 & \text{otherwise} \end{cases} \\
 &= 2 \times \text{floor}(\frac{j}{2}) \times (-1) \times (\text{ceiling}(\frac{i}{j+1}) - 1) \\
 &\quad + (\max_{i,j}(i, j) - 1) \times \text{ceiling}(\frac{i}{j+1}) \\
 N\text{agent}(\Omega_{\text{meshdouble}}, \gamma_{\text{dimensionorder}}, (i, j)) &= \begin{cases} 0 & \text{if } i = j = 0 \\ i + j - 1 & \text{otherwise} \end{cases} \\
 N\text{agent}(\Omega_{\text{torussingle}}, \gamma_{\text{dimension_order}}, \text{node}_i) &= \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ 1 & \text{otherwise} \end{cases} \\
 N\text{agent}(\Omega_{\text{torusdouble}}, \gamma_{\text{dimension_order}}, \text{node}_i) &= \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ 1 & \text{otherwise} \end{cases} \\
 &= \text{ceiling}(\frac{i}{j+1}) \times \text{ceiling}(\frac{j}{j+1})
 \end{aligned}$$

Nlink(Ω, γ, node<sub>i</sub>) の評価式

網Ωのノード node<sub>i</sub> に対し、ルーティングアルゴリズム γ に従ってアクセスする際の通過リンク個数。網Ω毎の式は類番を参照。γ が γ<sub>dimensionorder</sub> であるものは基本的に Nlink(Ω, γ<sub>dimensionorder</sub>, node<sub>i</sub>) = |node<sub>i</sub>| である。  
 Ω<sub>meshsingle</sub> のときだけは、構造特異的なルーティングアルゴリズム γ<sub>meshsingle</sub> を要求する。その際の、Nlink(Ω<sub>meshsingle</sub>, γ<sub>meshsingle</sub>, (i, j)) を以下に示す。(関数 ceiling(x) は x 以上で最小の整数をあたえる関数とする。つまり小数点以下の切上げを示す。) Ω<sub>meshsingle</sub>, Ω<sub>meshdouble</sub> では、起点ノードを(0,0)は左上隅ないし右下隅としているが、本稿内の言及では、これらの関数は最大距離についてのみ使用されているので、この仮定で実際の問題はない。

$$\begin{aligned}
 N\text{link}(\Omega_{\text{ringsingle}}, \gamma_{\text{dimensionorder}}, N, i) &= i \\
 N\text{link}(\Omega_{\text{ringdouble}}, \gamma_{\text{dimensionorder}}, N, i) &= \min(i, N-i) \\
 &= \begin{cases} i & \text{if } i < N/2 \\ N-i & \text{otherwise} \end{cases} \\
 N\text{link}(\Omega_{\text{meshsingle}}, \gamma_{\text{meshsingle}}, N, (i, j)) &= i + j + 2 \text{mod}(\frac{i+j}{2}) - 2 \times \text{ceiling}(\frac{i}{j+1}) \times (\text{ceiling}(\frac{i}{j+1}) - 1) \\
 &= \begin{cases} i + j + 2 \text{mod}(\frac{i+j}{2}) + 2 & \text{if } i = 0 \text{ and } j \neq 0 \\ i + j + 2 \text{mod}(\frac{i+j}{2}) & \text{otherwise} \end{cases} \\
 N\text{link}(\Omega_{\text{meshdouble}}, \gamma_{\text{dimensionorder}}, N, (i, j)) &= i + j \\
 N\text{link}(\Omega_{\text{torussingle}}, \gamma_{\text{dimensionorder}}, N, (i, j)) &= i + j \\
 N\text{link}(\Omega_{\text{torusdouble}}, \gamma_{\text{dimensionorder}}, N, (i, j)) &= \min(i, N-i) + \min(j, N-j)
 \end{aligned}$$

以上