

NoC のための ロングエッジファーストルーティングの提案と評価

笹河 良介^{1,a)} 吉瀬 謙二^{1,b)}

概要：2次元メッシュ状のメニーコアプロセッサに複数の並列アプリケーションのタスク配置をおこなう場合、全ての並列アプリケーションが長辺や短辺の無い正方形のような配置になるとは限らない。さらに、Network on Chip において、タスク配置によって X 軸方向が長辺となる場合は XY 次元順ルーティング、Y 軸方向が長辺となるような場合は YX 次元順ルーティングで、パケットの転送をおこなう方が性能は高くなる傾向にある。以上のことから、本稿では、タスク配置によって長辺、短辺が生じた場合に、より好ましいルーティングになるようなロングエッジファーストルーティングを提案し、評価する。

1. はじめに

プロセス技術の進展により、1チップに複数のコアが集積されるようになった。設計プロセスは依然として微細化傾向にあり、数百ものコアが搭載されるメニーコアプロセッサが実現可能になると考えられる。また、メニーコアプロセッサにおいて、コア間の通信はネットワークオンチップ (Network on Chip, NoC) [1] と呼ばれる通信路を用いておこなわれる。このコア間の通信を担う Network on Chip に対しての高性能化の要求は高まると考えられる。

2次元メッシュ状のメニーコアプロセッサ上に、複数の並列アプリケーションのタスク配置を同時におこなう場合、全ての並列アプリケーションが長辺や短辺の無い $n \times n$ といった正方形のような配置になるとは限らない。

さらに、Network on Chip において、タスク配置によって X 軸方向が長辺となる場合は XY 次元順ルーティング、Y 軸方向が長辺となるような場合は YX 次元順ルーティングで、パケットの転送をおこなう方が性能は高くなる傾向にある。

本稿では、以上のことを踏まえ、タスク配置によって長辺、短辺が生じた場合に、より好ましいルートを選択するロングエッジファーストルーティングを提案し、評価する。

本稿の構成を以下に示す。2章では、予備評価として、タスク配置によって長辺、短辺が生じた場合の XY 次元順ルーティングと YX 次元順ルーティングの性能を評価する。3章では予備評価を踏まえて、ロングエッジファース

トルーティングを提案し、4章でその提案手法を評価する。5章で関連研究を述べ、最後に6章でまとめる。

2. 予備評価

2次元メッシュ状のメニーコアプロセッサ上に複数の並列アプリケーションのタスク配置を同時におこなう場合を考える。10×10ノードの2次元メッシュ状のメニーコアプロセッサ上に、タスクが36個あるアプリケーションA、タスクが40個あるアプリケーションB、タスクが24個あるアプリケーションCを同時に配置する。この時、アプリAは6×6の正方形にタスクが配置され、アプリBは10×4のX軸方向を長辺とする長方形にタスクが配置され、アプリCは4×6のY軸方向を長辺とする長方形にタスクが配置されたとする。以上のタスクの配置例を図1に示す。

図1にあるように、複数の並列アプリケーションのタスク配置をおこなう場合、全ての並列アプリケーションが長辺や短辺の無い正方形のような配置になるとは限らない。また、メニーコアプロセッサのノード構成自体、 $n \times n$ といった正方形のようなノード構成であるとも限らない。し

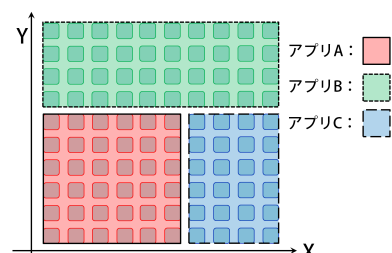


図1 複数並列アプリケーションのタスク配置

¹ 東京工業大学 大学院情報理工学研究科

^{a)} sasakawa@arch.cs.titech.ac.jp

^{b)} kise@cs.titech.ac.jp

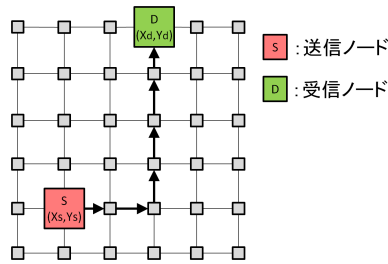


図 2 XY 次元順ルーティングの例

たがって、長辺、短辺が存在するような形にタスクが配置される可能性は十分に考えられる。

2 次元メッシュの NoC において、もっとも代表的な固定型ルーティングとして、次元順ルーティングの XY 次元順ルーティングおよび YX 次元順ルーティングがある [1]。XY 次元順ルーティングの例を 図 2 に示す。XY 次元順ルーティングとは、送信ノード (X_s, Y_s) から受信ノード (X_d, Y_d) へパケットを転送する場合、最初に (X_s, Y_s) から X 軸方向にパケットを移動して (X_d, Y_s) に転送させたのち、次に Y 軸方向にパケットを移動して (X_d, Y_d) まで転送させるルーティングである。逆に、YX 次元順ルーティングとは、最初に Y 軸方向にパケットを移動させてから、次に X 軸方向にパケットを移動させるルーティングとなる。

では、図 1 のように長辺、短辺が存在するようにタスク配置がなされた場合に、XY 次元順ルーティングおよび YX 次元順ルーティングの両ルーティング間で、長辺となる次元軸の違いによって、どの程度の通信性能差が存在するのかを評価する。

ルータの構成および評価環境を表 1 に示す。独自開発したサイクルアキュレイトなソフトウェアシミュレータを用いて、 16×8 ノードという X 軸が長辺となった場合での、XY 次元順ルーティングおよび YX 次元順ルーティングのパケット転送の平均レイテンシを評価する。各ノードにはそれぞれの位置により $(0,0)$ から $(15,7)$ までのノード ID を割り振る。 $(0,0)$ のノードが一番左下、 $(15,0)$ のノードが一番右下、 $(0,7)$ のノードが一番左上、 $(15,7)$ のノードが一番右上のノードとなる。以降も同様なノード ID の割り振り規則を用いる。トラフィックパターンには Uniform traffic と Hotspot traffic [2] を採用した。Uniform traffic はランダム通信である。Hotspot traffic は、ノード ID が $(7,3)$ と $(7,4)$ と $(8,3)$ と $(8,4)$ であるノードが hotspot と

表 1 評価環境

# Directions	5-input/output (North, East, West, South, Local)
# Virtual Channels	4
Pipeline	3-Stage Pipeline NRC/VA/SA: Routing Computation, Virtual Channel Allocation, Speculative Switch Allocation, ST: Switch Traversal, LT: Link Traversal
Packet Length	16
Input Buffer Size	4
Flow Control, Transfer	Credit-Base Flow Control, Warmhole
# Simulation Cycles	Warm-up: 10,000 cycles, Simulation: 100,000 cycles

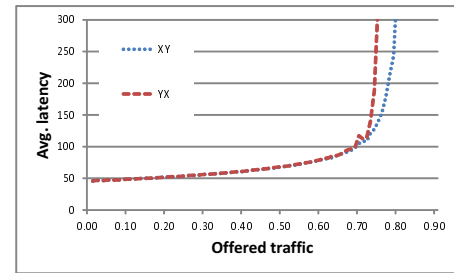


図 3 16×8 ノード構成 Uniform traffic

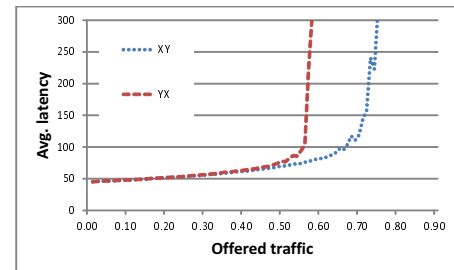


図 4 16×8 ノード構成 Hotspot traffic

なるようにし、hotspot となるノードには、他のノードに比べて 4 倍のパケットが転送されるようにする。

各トラフィックパターンでの平均レイテンシの結果は図 3 および図 4 である。Uniform traffic および Hotspot traffic の両トラフィックにおいて、YX 次元順ルーティングの方が、XY 次元順ルーティングよりも小さい Offered traffic の値でネットワークが飽和し、レイテンシが急激に増大している。したがって、図 3 および図 4 から、X 軸が長辺となるノード構成では、YX 次元順ルーティングよりも XY 次元順ルーティングの方が性能が良いことが分かる。特に図 4 ではその違いが顕著に表れている。

8×16 ノードというように長辺が Y 軸となる場合は、 16×8 ノード構成を 90 度回転させた構成と同じである。したがって、 8×16 ノード構成での XY 次元順ルーティングは、 16×8 ノード構成での YX 次元順ルーティングと同じであり、さらに、 8×16 ノード構成での YX 次元順ルーティングは、 16×8 ノード構成での XY 次元順ルーティングと同じである。つまり、 8×16 ノード構成では、先ほどとは逆に YX 次元順ルーティングの方が XY 次元順ルーティングよりも性能は良くなる。

3. 提案手法

3.1 ロングエッジファースト (LEF) ルーティング

前章で述べたとおり、タスク配置において長辺および短辺が存在する場合において XY 次元順ルーティングと YX 次元順ルーティングで性能に差が生じてしまう。よって本章では、タスク配置によって長辺、短辺が生じた場合に、より好ましいルーティングとなるルーティングアルゴリズムを提案する。

前章の評価から、X 軸が長辺となる場合は XY 次元順

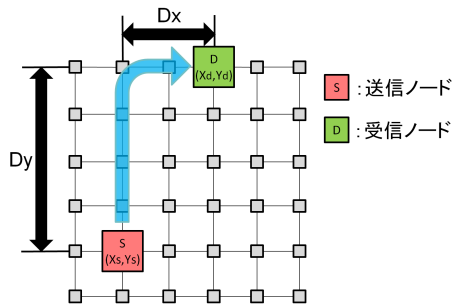


図 5 D_X と D_Y

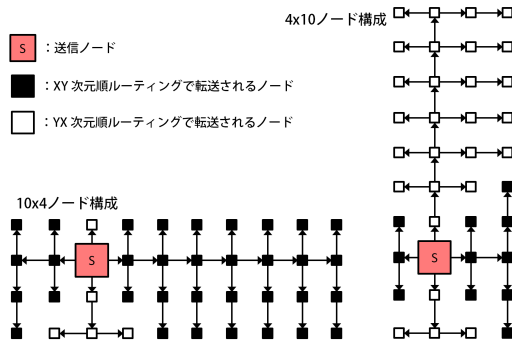


図 6 ある送信ノードから各受信ノードへのルーティングの分け方

ルーティング、Y 軸が長辺となる場合は YX 次元順ルーティングが好ましい。しかし、タスク配置された各アプリケーションに対し、長辺を判断して、それぞれ好ましい方の次元順ルーティングを採用する、という手法を採る場合は、どのノードがどのアプリケーションのタスクを担当しているのか、などといった様々な情報が必要となってしまう。また、全てのタスク配置が、長方形や正方形といった矩形になるとは限らない。矩形でないときにどちらの次元順ルーティングを採用するのかといった問題もある。

そこで、我々はロングエッジファースト (Long Edge First (LEF)) ルーティングを提案する。LEF ルーティングとは以下のようなルーティングである。アルゴリズムを Algorithm 1 として以下に示す。

Algorithm 1 LEF ルーティング

```

1:  $SrcID(X_s, Y_s), DstID(X_d, Y_d)$ 
2:  $D_X = |X_s - X_d|, D_Y = |Y_s - Y_d|$ 
3: if  $D_X \geq D_Y$  then
4:   XY Dimension order routing
5: else  $\{D_X < D_Y\}$ 
6:   YX Dimension order routing
7: end if

```

まず、図 5 の様に送信ノード (X_s, Y_s) と受信ノード (X_d, Y_d) のノード ID の X 成分と Y 成分の差分である D_X, D_Y を計算する。そして $D_X \geq D_Y$ の場合には XY 次元順ルーティングでパケットを転送し、 $D_X < D_Y$ の場合には YX 次元順ルーティングでパケットを転送する、というものである。

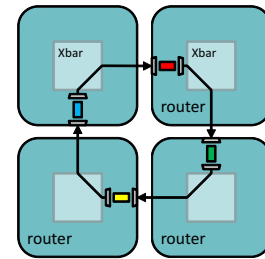


図 7 デッドロックの 1 例

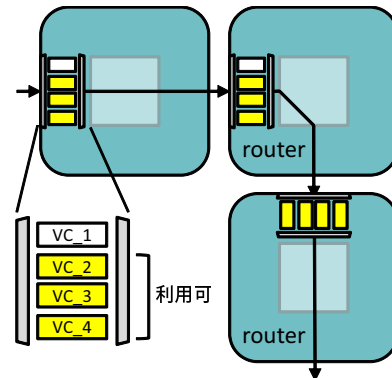


図 8 仮想チャネル 4 本での XY ルーティング

XY 次元順ルーティングを用いるか、YX 次元順ルーティングを用いるかは、パケット生成時またはネットワークに注入する時に、XY なのか YX なのかという情報をパケットのヘッダに 1 ビットだけ情報を付加する。また、パケット通信の送信ノード同士と受信ノード同士がそれぞれ同じなら、それらのパケットは毎回同じ次元順ルーティングとルートで送信される。

LEF ルーティングを利用して、ある送信ノードから各ノードへパケットを転送した場合に、どちらの次元順ルーティングで転送されるかの例を図 6 に示す。図 6 に示されているように、ある送信ノード S から各ノードへパケットを送信する場合、X 軸が長辺となる 10×4 ノード構成のような場合には多くのノードに対して XY 次元順ルーティングで送信され、Y 軸が長辺となる 4×10 ノード構成のような場合には多くのノードに対して YX 次元順ルーティングで送信されるようになる。

したがって、「タスク配置によって X 軸が長辺となる場合、YX 次元順よりも XY 次元順ルーティングの方が高い通信性能を示し、Y 軸が長辺となる場合は、XY 次元順よりも YX 次元順ルーティングの方が高い通信性能を示す」といった利点を、用いるようなルーティングとなっている。

3.2 デッドロック回避

LEF ルーティングでは、XY 次元順ルーティングと YX 次元順ルーティングの両方を利用しているので、そのままではデッドロックが起きてしまう。デッドロックが起きている例を図 7 に示す。図の右上のルータに格納されてい

る赤色のパケットは XY 次元順ルーティングにより、移動を X 軸方向から Y 軸方向にターンしようとしており、右下のルータに格納されている緑色のパケットが移動するのを待っている。右下のルータに格納されている緑色のパケットは YX 次元順ルーティングにより、左下のルータに格納されている黄色のパケットが移動するのを待っている。以上を繰り返すことで、互いが移動するのを待っている状況になってしまい、パケットの転送がおこなわれなくなってしまう。

そこで、LEF ルーティングではデッドロック回避のために、複数の仮想チャネル [3] が必要である。

LEF ルーティングでは、以下のようにしてデッドロックを回避する。1 本の物理チャネルにつき仮想チャネルが n 本 (VC_1 から VC_N まで) あったとする。XY 次元順ルーティングで転送されるパケットは、最初に X 軸方向に移動する時には VC_1 を除いた VC_2 から VC_N までの仮想チャネルを利用でき、次に Y 軸方向に移動する時は VC_1 から VC_N までの全ての仮想チャネルを利用できるとする。仮想チャネル 4 本の場合の例を図 8 に示す。逆に、YX 次元順ルーティングで転送されるパケットは、最初に Y 軸方向に移動する時には VC_1 を除いた VC_2 から VC_N までの仮想チャネルを利用でき、次に X 軸方向に移動する時は VC_1 から VC_N までの全ての仮想チャネルを利用できるとする。したがって、Y 軸方向の仮想チャネル VC_1 は XY 次元順ルーティング専用、X 軸方向の仮想チャネル VC_1 は YX 次元順ルーティング専用の仮想チャネルとなる。

以上のように仮想チャネルを利用すると、両次元順ルーティングにおいて、最初の次元方向を移動する際には、仮想チャネルを $n-1$ 本、次の次元方向を移動する際には、仮想チャネルを n 本利用できることになる。

図 7 のようにデッドロックが起こりそうな状況になっても、XY 次元順ルーティングで転送されているパケットは Y 軸方向の VC_1 に、YX 次元順ルーティングで転送されているパケットは X 軸方向の VC_1 に転送されることでデッドロックを回避することができる。以上のようにして、デッドロックを回避する。

しかし、 VC_1 が有限時間内で資源として解放されていて、利用できることが保証できなければいけない。したがって、 VC_1 が有限時間内で資源として解放され、利用できることを以下に証明する。

補題： VC_1 は資源として解放され、有限時間内で利用可能である。

証明：

$n \times m$ 構成と仮定し、まず、 $-y$ 方向のインプットチャネル (上のノードから来たパケットが利用するチャネル) の VC_1 を考える。

i) ノード ID (0,0) について

ノード ID (0,0) の $-y$ 方向の VC_1 を考える。たとえ、この VC_1 がパケットに占有されていたとしても、そのパケットは XY 次元順ルーティングでルーティングされているパケットである。したがって、これ以上 $-y$ 方向に進むことはできないので、そのパケットの受信ノードは (0,0) であり、ローカルに転送される。よって、ノード ID (0,0) の $-y$ 方向の VC_1 を占有していたパケットはローカルに転送され、ノード ID (0,0) の $-y$ 方向の VC_1 は有限時間内に解放される。

ii) ノード ID (0,1) について

次にノード ID (0,1) の $-y$ 方向の VC_1 を考える。たとえ、パケットに占有されていたとしても、そのパケットは XY 次元順ルーティングでルーティングされているパケットである。したがって、そのパケットの受信ノードは (0,0) か (0,1) である。

VC_1 を占有しているパケットのヘッドフリットが (0,1) の VC_1 に存在する場合を考える。受信ノードが (0,1) の場合は、先ほどと同様に、ローカルに転送される。受信ノードが (0,0) の場合は、i) の議論によってノード ID (0,0) の $-y$ 方向の VC_1 は有限時間内に解放されるので、ノード ID (0,0) の $-y$ 方向の VC_1 を利用して (0,0) に転送することが可能になる。したがって、ノード ID (0,1) の $-y$ 方向の VC_1 を占有していたパケットを次に転送することができる。

また、ノード ID (0,1) の $-y$ 方向の VC_1 を占有しているパケットのヘッドフリットが、既に先のインプットバッファに転送されている場合を考える。ノード ID (0,1) のローカルにヘッドフリットがある場合は、同様にローカルに転送される。ノード ID (0,0) にヘッドフリットがある場合でも、同様にローカルに転送される。したがって、ノード ID (0,1) の $-y$ 方向の VC_1 を占有していたパケットを次に転送することができる。

よって、ノード ID (0,1) の $-y$ 方向の VC_1 を占有していたパケットは次に転送され、ノード ID (0,1) の $-y$ 方向の VC_1 は有限時間内に解放される。

iii) ノード ID (0,2) について

さらにノード ID (0,2) の $-y$ 方向の VC_1 を考える。たとえ、パケットに占有されていたとしても、そのパケットは XY 次元順ルーティングでルーティングされているパケットである。したがって、そのパケットの受信ノードは (0,0) から (0,2) のいずれかである。

占有しているパケットのヘッドフリットが (0,2) の VC_1 に存在する場合を考える。受信ノードが (0,2) の場合は、ローカルに転送される。受信ノードが (0,1) 以降の場合は、ii) の議論によってノード ID (0,1) 以降の $-y$ 方向の VC_1 は有限時間内に解放されるので、ノード ID (0,1) の $-y$ 方向の VC_1 を利用して転送することが可能になる。したがって、ノード ID (0,2) の $-y$ 方向の VC_1 を占有して

いたパケットを次に転送することができる。

また、ノード ID (0,2) の $-y$ 方向の VC_1 を占有しているパケットのヘッドフリットが、既に先のインプットバッファに転送されている場合を考える。ノード ID (0,2) のローカルにヘッドフリットがある場合は、同様にローカルに転送される。ノード ID (0,1) 以降にヘッドフリットがある場合でも、ii) の議論によってノード ID (0,1) 以降の $-y$ 方向の VC_1 は有限時間内に解放されるので、ノード ID (0,1) 以降の $-y$ 方向の VC_1 を利用して転送することが可能になる。したがって、ノード ID (0,2) の $-y$ 方向の VC_1 を占有していたパケットを次に転送することができる。

よって、ノード ID (0,2) の $-y$ 方向の VC_1 を占有していたパケットはローカルに転送され、ノード ID (0,2) の $-y$ 方向の VC_1 は有限時間内に解放される。

iv) ノード ID (0,3) 以降について

同様の議論を用いることで、ノード ID (0,0) から (0, $m-1$) までの $-y$ 方向の VC_1 は有限時間内に解放されることが示せる。(0, m) については、上方向にノードは無いので、関係はない。

さらに同様に、ノード ID (1,0) から (1, $m-1$) までの $-y$ 方向の VC_1 は有限時間内に解放されることが示せ、最終的に、全ノードでの $-y$ 方向の VC_1 は有限時間内に解放されることが示せる。

v) iv) までの議論を受けて

そして、これまでの議論を $+x$ 方向, $-x$ 方向, $+y$ 方向でおこなうことで、全ノードで、全方向の VC_1 は有限時間内に解放されることが示せる。よって、 VC_1 は有限時間で資源として解放され、利用可能であることが示せた。

□

以上の補題を用いることで、デッドロックが起こるようなパケットの依存関係のループができないことが示せるので、以下の定理も証明できることになる。

定理：LEF ルーティングはデッドロックフリーである。

4. 評価

4.1 長辺, 短辺が存在する構成

長辺および短辺が存在する場合において、LEF ルーティングの性能を評価する。

評価環境は予備評価の時と同様の表 1 で評価をおこなう。LEF ルーティングでは、 $D_X = D_Y$ の場合は XY 次元順ルーティングを採用するので、 16×8 ノードという X 軸が長辺となった場合だけでなく 8×16 ノードという Y 軸が長辺となった場合の評価もおこなう。評価する項目は平均レイテンシとスループットである。トラフィックパターンには Uniform traffic と Hotspot traffic を採用した。 16×8 ノードの場合は、ノード ID を (0,0) から (15,7) まで振り分け, hotspot となるノードを (7,3) と (7,4) と (8,3)

と (8,4), 8×16 ノードの場合は、ノード ID を (0,0) から (7,15) まで振り分け, hotspot となるノードを (3,7) と (4,8) と (3,7) と (4,8) となるようにする。

また、XY 次元順ルーティングと YX 次元順ルーティングの両方を利用していることによる性能変化も評価するために、XY 次元順ルーティングと YX 次元順ルーティングをランダムで選択するルーティングも評価する。仮想チャネルの使い方は LEF ルーティングと同様とする。

評価結果は図 9 から図 16 に示す。XY 次元順ルーティングを XY, YX 次元順ルーティングを YX, LEF ルーティングを LEF, XY 次元順ルーティングと YX 次元順ルーティングをランダムで選択するルーティングを RANDOM として表記している。

それぞれのトラフィックパターンにおける各構成の平均レイテンシを示している図 9 および図 11, 図 13, 図 15 の全てにおいて、ネットワークが飽和し、レイテンシが急激に増加する時の Offered traffic の LEF の値は、XY 次元順ルーティングと YX 次元順ルーティングの間の値となっている。RANDOM の性能については LEF と同様に、XY と YX の間に位置しているが、特に Hotspot traffic において LEF の方がよりよい性能を示している。

Uniform traffic ではそれぞれのスループットに特に大きな差はみられない。しかし、Hotspot traffic において LEF は、 16×8 ノード構成では、よりスループットの高い XY 次元順ルーティングと同様のスループットを達成し、 8×16 ノード構成では、よりスループットの高い YX 次元順ルーティングと同様のスループットを達成していることが図 12 および図 16 で示されている。RANDOM の性能については全てにおいて XY と YX の間に位置している。しかし、Hotspot traffic においては、LEF ほどの高いスループットを達成できていない。

以上の結果より、タスク配置において長辺および短辺が存在する場合において、LEF ルーティングは、性能のより高い次元順ルーティングと同様のスループットが得られる、より好ましいルーティングであることが分かる。さらに、RANDOM の性能評価により、XY 次元順ルーティングと YX 次元順ルーティングの双方を利用することのみでも、性能向上させることは可能であると言える。しかし、XY 次元順ルーティングと YX 次元順ルーティングをただランダムに選択して平等に利用するよりも、LEF ルーティングのように長辺, 短辺を意識した利用の方が、より高いスループットを達成できることが示された。

4.2 $n \times n$ 構成

次に 8×8 ノード構成での性能評価をおこなう。評価環境は同様に表 1 で評価をおこない、評価する項目は平均レイテンシとスループットである。トラフィックパターンには同様に Uniform traffic と Hotspot traffic を採用し、

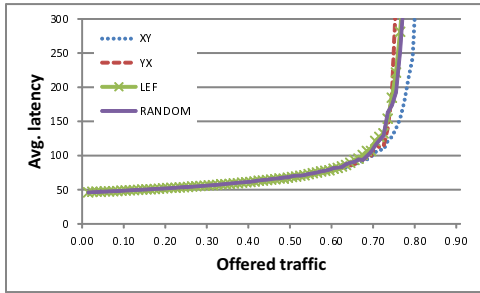


図 9 16x8 ノード構成 平均レイテンシ (Uniform traffic)

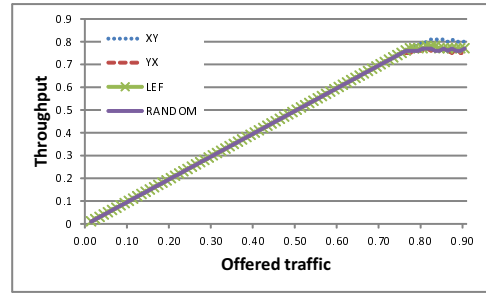


図 10 16x8 ノード構成 スループット (Uniform traffic)

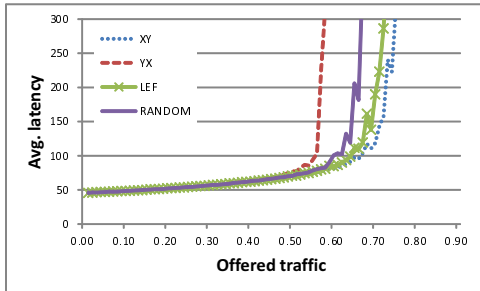


図 11 16x8 ノード構成 平均レイテンシ (Hotspot traffic)

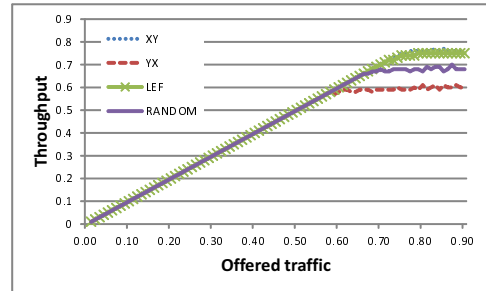


図 12 16x8 ノード構成 スループット (Hotspot traffic)

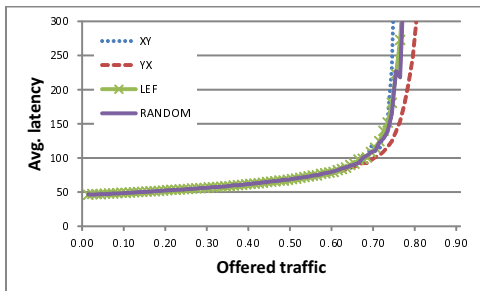


図 13 8x16 ノード構成 平均レイテンシ (Uniform traffic)

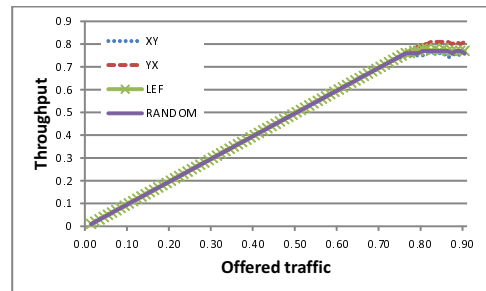


図 14 8x16 ノード構成 スループット (Uniform traffic)

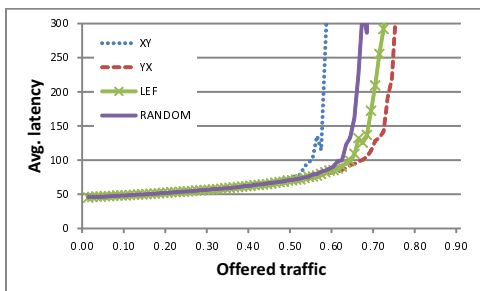


図 15 8x16 ノード構成 平均レイテンシ (Hotspot traffic)

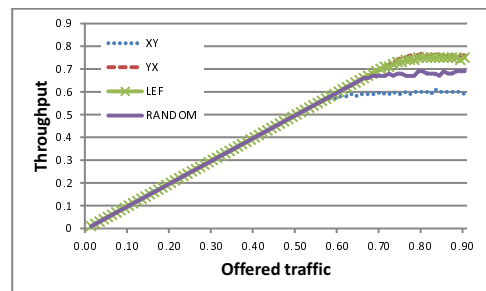


図 16 8x16 ノード構成 スループット (Hotspot traffic)

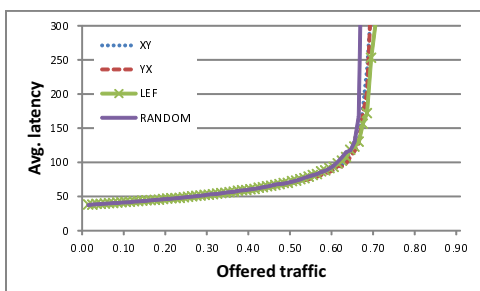


図 17 8x8 ノード構成 平均レイテンシ (Uniform traffic)

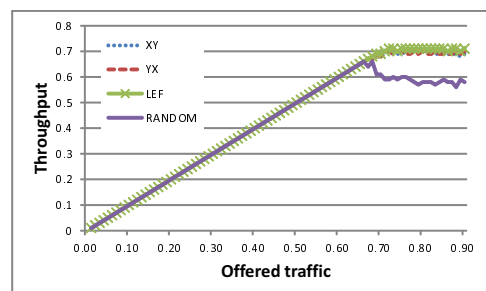


図 18 8x8 ノード構成 スループット (Uniform traffic)

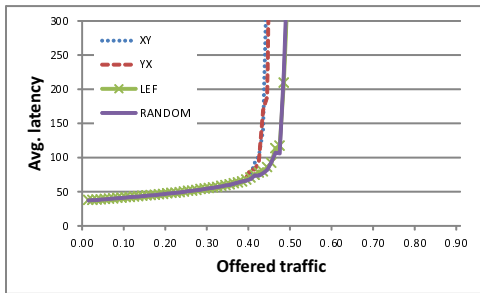


図 19 8x8 ノード構成 平均レイテンシ (Hotspot traffic)

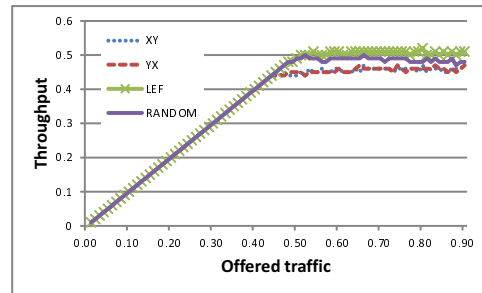


図 20 8x8 ノード構成 スループット (Hotspot traffic)

ノード ID を (0,0) から (7,7) まで振り分け, hotspot となるノードを (3,3) と (3,4) と (4,3) と (4,4), となるようにする.

評価結果は図 17 から図 20 に示す. Uniform traffic では大きな違いはみられないが, 先ほどとは異なり, 両トラフィックで LEF が XY と YX の両ルーティングよりも, 平均レイテンシとスループットの両方で, よりよい性能を示している. RANDOM は Hotspot traffic では LEF 同様に高い性能を示したが, Uniform traffic のスループットでは, 高い Offered traffic に対し, スループットの落ち込みが発生してしまっていることが図 18 から分かる.

以上のことから, LEF ルーティングは $n \times n$ ノードといった構成においても, XY 次元順ルーティングや YX 次元順ルーティングのみというルーティングより高スループットなルーティングであると言える. また, RANDOM と比較してもより高スループットなルーティングであることが示され, ランダム選択よりも送信受信位置を考慮することの優位性があったと考えられる.

5. 関連研究

5.1 Blue Gene/L および Blue Gene/Q

Blue Gene/L および Blue Gene/Q では, point-to-point 通信において, ノード構成において長い次元方向のリンクから順にルーティングする手法を採用している [4], [5], [6].

Blue Gene/L は 3 次元トラス, Blue Gene/Q は 5 次元トラスのネットワークとなっている. ノード構成によっては, Blue Gene/L では $64 \times 32 \times 32$ 構成, Blue Gene/Q は $16 \times 16 \times 16 \times 12 \times 2$ 構成といったように, 非対称なトラス構成となる. また, ノード構成によっては最大長となる次元も異なってくる.

このような非対称なトラス構成において, point-to-point 通信を利用した All-to-All 通信をおこなう場合, ノード構成において長い次元方向のリンクから順にルーティングする手法によってスループットが向上することが確認されている. 特に [5] では, $16 \times 8 \times 8 \times 8$ トラスの All-to-All 通信において, ピーク性能の 66% 程度のパフォーマンスが 93% にまで向上したという記述がある.

[5] や [6] においては, zone routing や longest-to-shortest

routing などといった名称が用いられており, 同じ長さの次元が複数ある場合は, それらの次元方向を移動が終わるまで, 最小ホップでの適応型ルーティングで転送する.

5.2 O1TURN

LEF ルーティングと同様に XY 次元順ルーティングと YX 次元順ルーティングの両方を利用したルーティングとして, O1TURN [7] がある. O1TURN はワーストケースでのスループットを近最適化するための 2 次元メッシュ状ネットワーク用のルーティングである.

ワーストケースの 1 例を図 21 に示す. 図 21 に示してあるように, 横一列に並んだ送信ノード群が XY 次元順ルーティングを用いて受信ノード群にパケットを転送しようとする, 全てのパケットはチャンネル C を利用することとなり, スループットは低いものになってしまう. また, YX 次元順ルーティングでは受信ノード群が横一列に並んだ場合に, 同様に同じチャンネルをパケットが利用することになり, スループットは低くなってしまふ.

こういったワーストケースでのスループットを近最適化するために, O1TURN は, XY 次元順ルーティングと YX 次元順ルーティングをランダムに使い分けるルーティングとなっている. パケットによって, XY 次元順ルーティングと YX 次元順ルーティングをランダムに使い分けることによって, 先ほど示したワーストケースに直面しても, 全てのパケットが同じチャンネルを利用することが無くなり, スループットを向上させることが可能となっている.

LEF ルーティングでは送信ノードと受信ノードが同じならば, 選択するルーティングおよびルートは同一のものとなっている. したがって, こういったワーストケースにおいては, 単純な次元順ルーティングよりもスループットは高くなる可能性はあるが, O1TURN ほど高スループットにはならないと思われる.

また, O1TURN ではデッドロック回避方法として, 仮想チャンネルの半分を XY 次元順ルーティング用に, もう半分を YX 次元順ルーティング用に割り当てるようにしている. この割り当てにより, パケットは仮想チャンネルを半分しか利用できなくなってしまう. しかし, O1TURN はワーストケースでのスループットに焦点をおいたルーティ

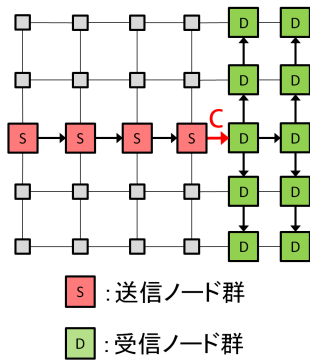


図 21 ワーストケースの 1 例

ングなので、どの物理チャネルでも両ルーティングが確実に使える仮想チャネルを用意するために、XY 次元順ルーティング用と YX 次元順ルーティング用に分けたと考えられる。

6. まとめ

2 次元メッシュ状のメニーコアプロセッサに複数の並列アプリケーションのタスク配置をおこなう場合、全ての並列アプリケーションが長辺や短辺の無い正方形のような配置になるとは限らない。またそのような形が最適とは限らない。タスク配置によって X 軸方向が長辺となる場合は XY 次元順ルーティング、Y 軸方向が長辺となるような場合は YX 次元順ルーティングで、パケットの転送をおこなう方が性能は高くなる傾向にある。

本稿では、タスク配置によって長辺および短辺が生じた場合に、XY 次元順ルーティングと YX 次元順ルーティングとで、性能の高い方と同様の性能を常に得るために、LEF ルーティングを提案した。評価によって、LEF ルーティングの有効性が示された。

謝辞

本研究の一部は、科学技術振興機構・戦略的創造研究推進事業 (CREST) の「ディペンダブルネットワークオンチッププラットフォームの構築」の支援による。

参考文献

- [1] Dally, W. J. et al.: Principles and Practices of Interconnection Networks (2004).
- [2] Fulgham, M. L. et al.: Performance of Chaos and Oblivious Routers Under Non-uniform Traffic, Technical report (1993).
- [3] Dally, W. J.: Virtual-Channel Flow Control, *IEEE Trans. Parallel Distrib. Syst.* (1992).
- [4] Kumar, S., Sabharwal, Y., Garg, R. and Heidelberg, P.: Optimization of All-to-All Communication on the Blue Gene/L Supercomputer, *Proceedings of the 2008 37th International Conference on Parallel Processing, ICPP '08*, pp. 320–329 (2008).
- [5] Chen, D., Easley, N. A., Heidelberg, P., Senger, R. M., Sugawara, Y., Kumar, S., Salapura, V., Satterfield, D. L.,

Steinmacher-Burow, B. and Parker, J. J.: The IBM Blue Gene/Q interconnection network and message unit, *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, SC '11*, pp. 26:1–26:10 (2011).

- [6] Chen, D., Easley, N., Heidelberg, P., Kumar, S., Mamidala, A., Petrini, F., Senger, R., Sugawara, Y., Walkup, R., Steinmacher-Burow, B., Choudhury, A., Sabharwal, Y., Singhal, S. and Parker, J. J.: Looking under the hood of the IBM blue gene/Q network, *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '12*, pp. 69:1–69:12 (2012).
- [7] Seo, D., Ali, A., Lim, W.-T., Rafique, N. and Thottethodi, M.: Near-Optimal Worst-Case Throughput Routing for Two-Dimensional Mesh Networks, *Proceedings of the 32nd annual international symposium on Computer Architecture, ISCA '05*, pp. 432–443 (2005).