

流量制限による相互結合網の実効性能改善の検討

横田 隆史^{1,a)} 大川 猛¹ 大津 金光¹ 馬場 敬信¹

概要：大規模な相互結合網において転送性能を向上させるためには、輻輳の発生や成長を効果的に抑制する必要がある。この目的に則して、本稿では、ルーティングアルゴリズムと直交する手法を流量制御の観点から検討する。効果的な流量制御のためには、隣接ルータ間の局所的な情報だけでは不足であり広域・準広域的に輻輳に関連した情報を収集する必要がある。さらに膨大な数のルータが使われることに鑑み、集中制御ではなく個々のルータの自律的な動作の集積としての分散制御により実現することを検討する。そこで本稿では、まず我々の先行研究の成果である仮想制御チャネルをもとにして、分散制御による広域・準広域情報の収集機構を導入する。その上で、実際の流量制御手法として、広域情報を扱うエントロピースロットリングを導入するとともに、準広域情報を有効に使用するゲーティング手法を提案する。ゲーティング手法は準広域情報とその評価関数をもとにして本線に合流する経路のケット流を制御することで本線経路の輻輳を抑制するものである。各々の手法には様々なパラメータが考えられるため、本稿ではすべてのパラメータの組み合わせについて評価を行う。一斉同期通信、定常通信の通信スタイルそれぞれについて、転置をはじめとする様々な通信パターンを設定した包括的な評価を行う。その結果、特にゲーティングにおいて多くの場合に性能向上が得られることを示す。

1. はじめに

多数の計算ノードを擁する大規模並列計算システムでは、通信機能を司る相互結合網が性能上・コスト上の大きなファクターとなる。このために古くから多くの研究が行われ、様々な方式が提案・開発されてきている [4], [5]。

相互結合網に対する本質的な要求はシンプルである。すなわち計算ノードの数に応じたスケール性を性能やコストの面において確保できれば良い。しかし実際には、多数のスイッチ（ルータ）の組み合わせにより構成せざるを得ず、このために多くの課題がなお山積している。2つ以上のケットが同時にひとつのルータを通過しようとする衝突する可能性がある。衝突した場合は、いずれか1つのケットのみが通過し他は転送をブロックされる。結合網上のケット密度が低い場合、このブロックは一時的なもので済むが、頻繁に転送が行われておりケット密度が高い状況では、転送のブロックが後続のケットの転送を抑制してしまい、その結果として輻輳を引き起こす。

相互結合網中で輻輳が発生している領域は、輻輳のない領域に比べると、ケットの移動度や密度の点において顕著に劣る。この差は際立っており、輻輳の発生を相転移として捉えることも可能である [9], [15], [16], [17]。輻輳は上

述のようにケット間の衝突に伴う転送のブロックによって生じる。転送のブロックが上流に向けて連鎖する速度、すなわち輻輳が成長する速度は、ケットが転送の抑制なしに移動するときの速度と同等である。一方、輻輳の解消には、ケットが転送をブロックされている状況がすべて解決される状態になるまで待たねばならない。このように輻輳の成長速度と解消速度との間にも桁違いの差がある。そして、この影響は、相互結合網の規模が大きいくほど顕著になる [15], [17]。

こうした背景のもと、並列計算機の相互結合網においては、これまで以下の観点から検討が進められてきた。

- バンド幅を拡大する、
- スイッチ（ルータ）の経路段数を減らす、
- ケットの衝突を回避するように経路などを調整する。

バンド幅を増すと、ケットが通過するために必要な時間が短くなるため、結果的に他のケットと干渉し合う時間が短くなる。すなわちケット間で衝突が発生する可能性が低くなる。また、ケットの衝突は上述のようにルータ上で生じるから、ルータの経路段数を減らすことは、衝突の機会を減らすことになる。また適応ルーティングの導入により、ケットの経路が2通り以上あるとき、ケット間での衝突が発生する経路を避けることも検討されている。ケットペーシング [11], [12] もケットの衝突を回避する手法とみなすこともできる。このように、上記3つ

¹ 宇都宮大学大学院工学研究科
宇都宮市陽東 7-1-2

^{a)} yokota@is.utsunomiya-u.ac.jp

の観点はいずれも輻輳制御の観点からも効果のあるものと言える。

大規模な並列計算機システムの構築を考えると、すなわち多数の計算ノードを集積するシステムの構築を考えると、これら計算ノード相互を接続する相互結合網の重要性が位置付けられるが、その一方で現実性・実現性を考えると大きな制約が課せられる。

バンド幅は、物理デバイスの技術および回路技術、すなわち物理層（レイヤ）を担う技術の進展とともに今後も向上することが期待できる。一方で、スイッチ（ルータ）の経路段数を減らすことは個々のスイッチの入出力数（トポロジ的に言えば次数）が過大になり、実現性に乏しくなる。このため、比較的低次数のスイッチを多数組み合わせることで目的の相互結合網を構築しなければならない。これはすなわち、スイッチ（ルータ）の経路段数が増すことを意味する。経路段数が増すことは衝突の機会が増すことにほかならない。

さらに、多数のスイッチ要素を使用することから、集中制御による効果的な運用が現実的ではなくなる。このために、各スイッチ（ルータ）に制御機能を持たせ、個々のルータの自律的な制御動作の集積として相互結合網全体を効果的に制御する方法を求める必要がある。

以上のような状況を受け入れた上で、なお相互結合網の転送性能を向上するには、輻輳の発生を極力抑える技術、また、発生した輻輳を拡大させず解消に向かわせる技術を求める必要がある。

本稿では、それぞれのルータが個別に相互結合網内の状態を感知し、自律的な適応動作を行うことにより相互結合網全体として効果的な制御を実現する方法を検討する。本稿における検討の中心概念は、パケットの流量制御である。個々のパケットはルーティングアルゴリズムに従ってそれぞれ相互結合網中を移動し目的ノードに達する。ルータ間を結ぶパケットの経路を流路とみなせば、合流離脱をしながらパケットが流路を流れているものと考えられる。流路が持つ転送能力を超える流れがあった場合、輻輳が発生する。このため輻輳の発生や拡大を抑制するには、輻輳領域へのパケットの供給、すなわち流量を制限することが最も効果がある。本稿での流量制限の考えはこうした発想に基づくものである。

効果的な流量制御を行うには、たとえば隣接ルータ間での局所的な情報だけでは明かに不足である。本稿では、流量制御の初期検討として、我々が相互結合網に関してこれまでに提案してきた様々なメカニズムや手法を流量制御のために使うことを検討する。このため、まず次節で既提案の機構・手法のうち本稿での検討に用いるものを整理する。それらを受けて第3節で、流量制御の考え方を示すとともに、広域情報・準広域情報を利用した具体的な流量制御手法を検討する。そして、検討した手法についてシミュレー

タにより評価した結果を第4節に示す。最後に、第5節で本稿での議論並びに結果をまとめる。

2. 準備

動的な流量制御により輻輳の効果的な抑制を行うためには、相互結合網の広域に渡る状態検知が欠かせない。一方で、大規模システムでは多数の独立したルータの組み合わせにより相互結合網を実現しなければならないとの現実的な制約がある。さらに、検討の過程ではいくつもの方法について評価を行い相互に比較する必要がある。

本節では、我々がこれまでに提案してきた種々の手法の中から、本稿での議論にあたり有用なものを整理する。これらは、次節以降においてツールとして使用する。

2.1 仮想制御チャネル

ノード間のデータ転送には物理リンクが使われる。これはノード間を結ぶものとして並列計算機の構成上必要不可欠なものであるが非常にコストが高い。仮想チャネルは、物理的な結合網トポロジの上に論理的な転送路を複数重畳させ物理リンクを共有させることにより、転送性能の改善やデッドロックフリー性の保証に利用するものである。

実際に並列計算機で相互結合網が使用される状況では、パケット密度が低く仮想チャネルで送るべきパケットがそもそも存在しない場合や、衝突によりブロックされているため転送可能な状況にない場合があり、すべての物理リンクが100%使われ続けているわけではない。

一方で現実の並列計算機では、並列プログラムの実行に伴いやり取りされるメッセージ通信のためのパケット以外に、同期制御など制御目的での通信を効率良く行うことが求められることも多い。仮想制御チャネル (virtual control channel, VCC, [13]) は、もともとの仮想チャネルの概念を制御用の通信にまで拡大したものである。その基本的な考えは、制御のための論理的な通信経路を仮定しメッセージ通信用の（通常の）仮想チャネルと共存することにある。このため仮想制御チャネルでは仮想チャネルとともに物理リンクを共有する。

仮想制御チャネルの増設は、新規に物理リンクを増設することに比べれば格段に低コストで実現可能であり、また、制御性能の面でも専用網を用いた場合と同等であることが期待される。

2.2 エントロピースロットリング

非輻輳領域ではパケットの存在密度は比較的低いが衝突によりブロックされる機会が少ないために移動度が高い。一方、輻輳領域ではパケットが相互に転送を阻む状況になっており、存在密度が高い一方で移動度は極めて小さくなる。こうした状況は、パケットが相互結合網中を移動する状況を、分子が空間中を飛翔する状況に対比させ、熱エネ

ルギー的なエントロピーを定義することにより定量的に表すことができる。輻輳の発生が相転移と言えるものであることはエントロピーの劇的な変化により示されている [16]。

系内の輻輳の状況は、エントロピーにより定量的に表現することができる。系内に輻輳が発生するとエントロピー値が大きく変わる。こうした観点から、我々は、系のエントロピーを常時求めておき、輻輳の検出とともに計算ノードからのパケット送出を抑制するエントロピースロットリングを提案している [18]。文献 [18] ではエントロピーの定義そのものではなく、仮想チャネルのバッファの使用状況をもとにした近似手法を用いている。この近似エントロピーの値により系内に輻輳が発生したことを検知すると、計算ノードからの新たなパケット送出を抑制し系内のパケット密度を下げる。これにより輻輳の状態が緩和し、あるいは輻輳そのものが解消される。こうした変化は近似エントロピーの変化として検出可能であるから、輻輳を抑制し相互結合網を効果的に使う手法として有用である。

この手法では、近似エントロピーを求める際に 4 種の仮想制御チャネルを用いている。図 1 に 8×8 メッシュ/トラス網での例を示す。まず、各ルータで求めた値の総和を行単位で求め、同一行内にブロードキャストする。次に、個々の列の単位で系全体の総和を求め、同一列内にブロードキャストする。図は 8×8 に並べられたルータでの仮想制御チャネルを表している。ルータ内にある四角形が仮想制御チャネル用のレジスタを表す。そのうち、ハッチがついているものは、各ルータでのチャネル使用状況を保持しているレジスタである。また図では、総和を求める演算 (リダクション) の経路を赤色で、ブロードキャスト経路を青色で表している。ブロードキャストされた近似エントロピーの値をもとに、各ルータに対応する計算ノードでスロットリング制御を行う。

2.3 準広域情報 (VCinfo)

輻輳を効果的に制御するには、隣接するルータだけでなく物理的に離れた場所にあるルータの状況を知る必要が生じる。理想的には、各ルータそれぞれが自己を中心とした広い範囲のルータの状況を詳しく知ることが望ましいのだが、現実的には情報収集のための時間的・資源的なコストのため制限される。さらに、この情報収集機能は各ルータの自律的な機能動作により実現できなくてはならない。

我々は Cross-Line ([14]) と呼ぶルーティングアルゴリズムの提案にあたり、パケットの進行方向にあるノードでの仮想チャネルの状況、すなわち仮想チャネルを割り当てられたバッファに空きがあるか否かの情報を収集する機構を提案した。これは、情報の収集範囲をパケットの進行方向のみに限定し、さらに、仮想チャネルのバッファの状態を空きの有無に対応した 1 ビットで表現し、パケットの進行方向にある一連のルータのバッファ状況をビット列に重畳

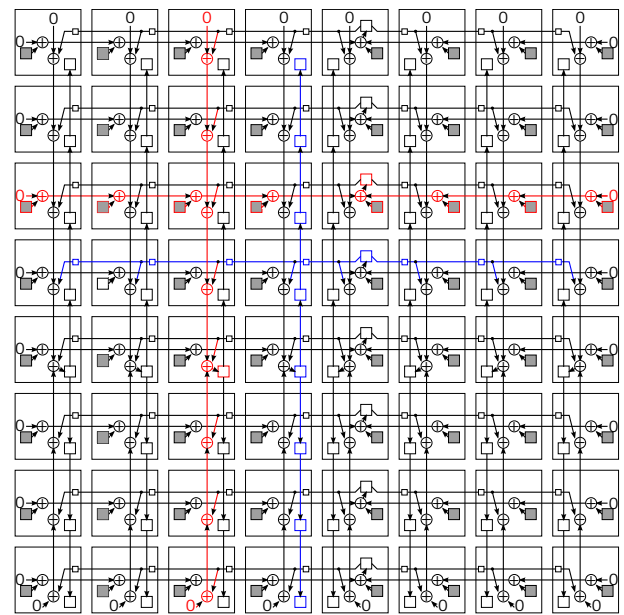


図 1 エントロピースロットリングのための仮想制御チャネル

して表したものである。

上流ルータからの情報を 1 ビット左シフトしたうえで自ノードの情報を LSB にセットし、この情報を下流ルータに向けて送出する。パケット進行方向にあるルータが上記の操作を行うと、LSB から MSB に向かう各ビットが、それぞれ、隣接ルータおよびその先の各ルータでのバッファ使用状況を表すことになる。我々はこうした情報を準広域情報 (VCinfo) と呼んでいる。

VCinfo の伝播に仮想制御チャネルを用いることにより低コストで実現することができる。

2.4 定量的評価手法

相互結合網の方式の評価は、特定の通信パターンを仮定したうえで、各計算ノードからのパケット生成の頻度をパラメータとしながら連続的な転送を行い、スループットと転送遅延 (レイテンシ) を性能指標として計測する手法が多く用いられている。パケット生成頻度は相互結合網に対する転送負荷に他ならないから、様々なパケット生成頻度の条件のもとで行ったシミュレーション評価の結果を集積することにより、転送負荷に対する性能指標の関係を求めることができる。

性能指標にはスループットとレイテンシの 2 面が必須であるから、結局、評価の結果は、転送負荷対スループットおよび転送負荷対レイテンシの 2 つのグラフで表されることになる。このため、方式間の比較を行うには、2 つのグラフに描かれた内容を見比べなければならない。グラフの形状をもとに 2 者を比較し、相対的な優劣を議論することはできるが、優劣の度合いを定量的に議論することができなかった。

こうした問題意識のもと、我々は、臨界転送負荷と網性

能指標 (network performance measure, NPM) の2つの定量的指標を提案した [21]。前者は、輻輳により転送性能の飽和が始まる転送負荷を示す。後者は、スループットを平均レイテンシで割った値を特徴量とし、この特徴量を転送負荷の区間で積分して得られる値として定義される。転送負荷の定義域での特徴量の平均を表すものと考えてもよい。この値は、スループットが高いほど、またレイテンシが低いほど高くなる。

臨界転送負荷の値により、比較しようとしている結合網方式がどこまでの負荷に耐えられるのかを定量的に比較することができる。また、網性能指標 (NPM) により、スループット・レイテンシの両方を考慮した性能を定量的に比較できる。

3. 流量制限による輻輳制御

3.1 検討の基本方針

3.1.1 前提条件

基本的に、並列計算機システム上での実行されるプログラムに関する一切の事前知識を使わない。各ルータは通信パターン等に関して無知識の状態を制御を行うものとする。このために結合網内の動的な状態を検知し、動的に対応することを基本とし、静的な制御は本稿での検討範囲から除外する。

本稿では基本的に流量制限にのみ着目し、流量制限のための手法と、その効果について議論を行うことを目的とする。このため、本稿で検討する流量制限の手法は、従来からあるフロー制御手法 (store-and-forward, wormhole, virtual cut-through) ならびに仮想チャネル (virtual channel, vc) と独立であるものとする。

また、本稿で検討する流量制限の考え方は、パケットの経路を変えるものではない。このため、本稿ではルーティングアルゴリズムとは独立に議論する。パケットの転送経路を変えることにより輻輳領域へのパケット流入を抑制することが、即ち輻輳領域への流量制限であると考えられることも可能ではあるが、上述のように流量制限による正味の効果を評価するために、ルーティングアルゴリズムにまで踏み込んだ議論は行わない。

本稿の検討は上述のようにルーティングアルゴリズムとは独立するが、パケットの流れを一時的に制限することから、ベースとするルーティングアルゴリズムのデッドロックフリー性をそのまま継承することができない。このため、新たにデッドロックを生じないことを確認する必要がある。これについては、4.4節において言及する。

さらに、本稿では議論の簡単化のため2次元トラス網を前提とする。

3.1.2 広域情報の収集

輻輳の発生や拡大を効果的に抑制するためには、結合網内の広域的な状況を正しく把握する必要がある。大規模並

列システムでは、その構成要素の多さからシステム内の詳細な情報を集中的に管理するのは現実的ではない。相互結合網を構成する多数のルータそれぞれの自律的な動作の集積により機能することが求められる。このために、2.1節に示した仮想制御チャネルを有効に使う。

次に検討せねばならないのは、どのような情報を収集するかである。本稿では相互結合網の輻輳制御が中心テーマであるから、収集対象の情報は必然的にルータの状態となる [8], [10]。第一に検討すべきは収集範囲と収集した情報の表現方法、それから流量制限への適用方法である。

本稿では、システム全体に渡る情報と収集範囲を制限した準広域情報の2通りについて検討する。前者には、エントロピースロットリング (2.2節) で用いている近似エントロピーを用いる。後者には2.3節で述べた準広域情報 (VCinfo) を用いる。

前後いずれの場合についても、各々の方式に固有のパラメータが複数設定される。本稿の評価では、網羅的な評価が行える範囲でパラメータの粗い組み合わせを設定し、パラメータ・スタディを行う。

3.1.3 その他の検討事項

相互結合網の方式の評価においては、ランダムな通信パターンによる定常的な (steady) 通信を行う場合の性能 (スループットと平均レイテンシ) を問題にすることが多い。実際の運用では主に集合通信 (collective communication) が行われるであろうから、本稿での評価は集合通信を中心とする。ランダム通信を含め典型的と考えられる通信パターンを設定し、その通信パターンに従って各ノードがパケットを所定の量だけ送出し、それらが全て目的ノードに受信されるまで同期する一斉同期通信を扱う。また、設定された通信パターンおよび送出頻度に従って継続的にパケットを送出する定常通信についても評価を行い、一斉同期通信での結果と比較する。

3.2 広域情報に基づくパケット投入の抑制

相互結合網におけるスロットリングとは、一般に、計算ノードでのパケット生成 (あるいはパケット送出) を一時的に抑制することをいう。これは、計算ノードから結合網に流入するパケットの量を絞る (throttle) ものであるから、本稿ではスロットリングを流量制御の一手法と位置づけ、2.2節で示したエントロピースロットリングを対象に効果を検討する。

このスロットリング手法では、具体的には以下のようにパケット投入を制御している。

- (1) 各ルータ内で、フリットを1片以上保持しているバッファの数 (n_{ob})、そのうちブロックされずに転送できたものの数 (n_{ab}) を求める。
- (2) 相互結合網にある全ルータについて、 n_{ob} , n_{ab} それぞれの総和を求め、 N_{ob} , N_{ab} とする。これらの値を求め

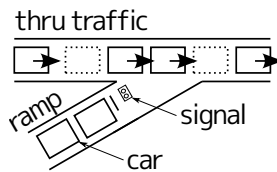


図 2 ランプメータリング

るために 2.2 節で示した仮想制御チャネルを用いる。

- (3) N_{ob} , N_{ab} からパケット移動度 $R_m = \frac{N_{ab}}{N_{ob}}$ を求める。
- (4) 系内のバッファの総数を N_b として, $P_n < \frac{N_{ob}}{N_b}$, ($0 < P_n \leq 1$) の場合にはスロットリングしない。ここで P_n は強制解除パラメータと呼ばれる。
- (5) パケット移動度 R_m が閾値 R_{th} に満たないとき, すなわち $R_m < R_{th}$, ($0 \leq R_{th} \leq 1$) のとき, スロットリングを行う。

文献 [18] では, ランダム通信および 2%ホットスポット通信*1の 2 種類のための通信パターンでの定常通信の性能評価にとどまっていたほか, ルーティングアルゴリズムも単純なものだった。本稿では, 文献 [18] での評価対象を大幅に拡大し, 広域情報を元にしたスロットリングの効果について検討する。

なお, 本稿でのスロットリングでは, パケットの生成そのものを抑制する。したがって, スロットリングが発動されたとき既に生成されているパケットは抑制の対象としていない。

3.3 準広域情報による流量制限—ゲーティング

3.3.1 ランプメータリング

パケットが相互結合網上を通過するときの性質は, パケットの転送過程において物理的な資源 (すなわちバッファ) を必要とする点, 衝突や輻輳が発生した状況下においても廃棄されないことなどの点において, 移動粒子問題, 特に交通流と共通する特徴を持つ。粗い議論を恐れずに言えば, 交通流における渋滞は, 相互結合網における輻輳と等価である。交通流の制御において効果のある方法は, 相互結合網においても同様に効果があるものと期待できる。

本稿では, 交通流の制御手法としてランプメータリング (ramp metering) に着目する。これは, 高速道路などにおいて, 本線での渋滞を抑制するために合流する量を抑制するものである。図 2 に模式図を示す。本線での流量に対し合流路から過大な流入があると渋滞を引き起こす。このために, 本線への合流路に信号を設置し, 本線で渋滞を起こさない範囲で合流路の信号を操作する*2。

*1 全体の 2%のパケットが特定ノードに向けて送出される通信パターン。他のパケットは送信先をランダムに選ばれる。

*2 ほんの僅かな時間だけ青信号にし, 1 台ずつ合流させる制御を行うのが基本のようである。合流路からの流量は青信号が現れる間隔で制御される。現実には様々な本線・支線が相互に合流・分離しているため, 最適な解を求めるには関係する本支線の流れをモデル化し, DP などの方法を用いる必要がある。

3.3.2 相互結合網への適用

ランプメータリングは現実の道路網で適用された実績があり, 系全体のパフォーマンスを著しく向上させる効果があることが報告されている [1], [2], [3]。ここで用いられている考え方を相互結合網の制御に導入する。

ルータ内をパケットが通過する際の流路の中から図 2 の本線に相当するものを選び, この本線と同一の出力ポートを要求する流路を合流路とみなすことができる。図 3 に簡単な例を示す。ここでは, 2 次元メッシュ/トラス網において, 東方向の流路を本線とし, 北上してきたパケットが同一ルータの中で東方向に向きを変える経路を合流路とみなしている。図に示されていないが, 南下してきたパケットが東方向に向きを変える場合も同様に合流路になり, また, ルータに対応した計算ノードからパケットが送出される際の経路もまた同様に合流路とみなすことができる。

このように, 隣接ルータへの出力ポートを共有する流路がある場合, 一つを本線とし他を合流路として考えることができる。本稿では, 2 次元トラス網を前提としていることから x 軸および y 軸方向に直進する流路を本線とし, y 軸方向から x 軸方向 (ないしはその逆の方向) に向きを変える流路と, 計算ノードから新たに投入される流路を合流路として考える。

交通流におけるランプメータリングをそのまま拡張すると, すべての合流路に対して流量制限を課すことになる。しかし, 相互結合網においては, デッドロックが起きないことを保証する観点から, 流量制限を課するのは一部の合流路にとどめることも考えねばならない。

さらに, 流量制限の有無を判断するもとなる転送状況の取得方法についてもここで議論しなければならない。2.3 節で説明した仮想チャネル情報は, パケットの直進進行方向にあるルータでのバッファの使用状況をバッファひとつにつき 1 ビットの情報で表したものであり, この情報はパケットの進行とは逆の方向に伝播される。つまり, 準広域情報 (VCinfo) の機構により, パケット進行方向での輻輳状況が検知できる。交通流でのランプメータリングでは, 本線の流量が判断基準になることから, この点において本稿での手法と異なる。

準広域情報 VCinfo が, 進行方向の輻輳状況をビットマップにより表している。そこで次に, この情報をもとに流量を制御する方法を検討する。このために, 本稿では文献 [20] で検討した評価関数を導入する。この文献では, VCinfo がパケット進行方向にあるルータ (バッファ) の輻輳状況をまとめたものであることを考慮し, そのビット列をもとに「その方向にパケットを転送した場合のコスト」を見積もる方法を論じている。本稿では, 上記文献中に示されている評価関数のうち, 以下の 4 つを試行する。

group-1: 自ノードから m ホップ以内に 1 つ以上の busy バッファがあれば 1, なければ 0 とする。なお, 探索



図 3 ランプメータリングの導入

範囲内の busy バッファの個数は問わない．

$$f_n(VCI, m) = \sum_{i=0}^m VCI_i \quad (1)$$

$$f_1(VCI, m) = \begin{cases} 1 & \text{if } f_n(VCI, m) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

group-2: VCinfo 中の各ビットに自ノードからの距離に比例（減少）する重みをかける．busy バッファのビットの重みの総和を求め評価値とする．

$$f_2(VCI, a) = \sum_{i=0}^m VCI_i (1.0 - a \cdot i), \quad m = 16, \text{ただし } (1.0 - a \cdot i) \geq 0 \quad (3)$$

group-3: group-2 と同様だが，自ノードからの距離に対し指数関数的に重みを減らす．group-2 に比べ，自ノードから近距離にある busy ビットがより大きな比重を持つ．

$$f_3(VCI, a) = \sum_{i=0}^m VCI_i \cdot a^i, \quad m = 16 \quad (4)$$

group-4: 自ノードから m ホップ以内にある busy ビットの個数を求める．

$$f_4(VCI, m) = f_n(VCI, m) = \sum_{i=0}^m VCI_i \quad (5)$$

上において， VCI は準広域情報 VCinfo の値を示し， VCI_i は VCinfo の第 i ビット目の値を意味する．

2次元トラス網においては，上述のように x 軸 y 軸方向に直進する流路をランプメータリングにおける本線とみなせるが，本来は高速道路網とは異なり本支線概念は存在しない．また，「メータリング (metering)」の語には測定の意味合いが強く，本稿が主題としている積極的に流量を制御するとの意味が伝わりにくい．本稿の狙いは，上述のように，入力ポートに到着したパケットの各々に対して，その進行方向により水門（ゲート）の開閉を行うことで，その進行方向先にある輻輳の抑制を図るものである．このため，本稿ではゲーティングと呼ぶことにする．

本稿では，2.3 節に示した準広域情報 VCinfo を用い，

- ゲーティング対象とする流路の選択
- 評価関数の選択

の組み合わせについて評価する．

3.3.3 バッファ占有準位の導入

本稿ではさらに準広域情報 VCinfo そのもの定義に立ち返り検討し直す作業を行う．2.3 節に示した VCinfo は，単

に仮想チャネルのバッファが busy（満杯）であるか否かを表していた．バッファが busy のとき，そのルータの該当仮想チャネルに向けてのパケットは転送できないから，バッファに空きができるまで転送がブロックされる．

実際にパケットの転送がブロックされ輻輳が生じている状況を検出するのであれば，この方法が妥当と言える．しかし本稿で検討したいのは，輻輳が発生した後の制御よりもむしろ輻輳の発生を抑制することにある．輻輳の成長はパケットの自由行程での速度と同等の速度で成長するが，いったん発生し成長した輻輳の解消には成長に比べ桁違いの時間を要するためである．

このため，バッファが既に満杯になった状態ではなく，満杯になりそうな状態にあることを検出したい．本稿では，仮想チャネルにあるバッファがある程度の容量を持つことを仮定したうえで，バッファの残り容量の水準を $1/0$ の 2 ビットで表すことを検討する．たとえば，パケット長を 8 フリットとし，バッファの容量を 16 フリット（パケット 2 個分）とする．ここでフロー制御はパーチャルカットスルーを仮定する．

バッファの残り容量が，予め定めておいた閾値を下回った場合「ほとんど満杯」になったものとして VCinfo を構成すれば良い．パケットがルータを通過する際，他のパケットとの衝突がなければ，そのまま即座に次の経由ルータに向けて転送される．パケット密度が低くてもパケット間の衝突は発生するが，その頻度は十分に低いのでパケットの一部が一時的に保持されるに留まるはずである．逆に，結合網が輻輳傾向にあるときは，パケットの次ルータへの転送が滞りがちになるため，バッファに滞留するパケット（フリット）が多くなる．したがって，バッファの使用率により閾値を設けることは，輻輳傾向にあるか否かを検出する有望な手段となる．さらに，このメカニズムの実現のために必要となるコストは大きくはなく，実用上許容範囲と考えられる．

本稿ではこの閾値をバッファ占有準位 (buffer occupancy level) と呼ぶ．以下の検討では，3.3.2 節で述べた検討事項の組み合わせに加え，さらにこのバッファ占有準位を加味する．

3.4 専用通信路による制御情報の伝播

これまでの検討では，仮想制御チャネルを使用することを前提として議論してきた．仮想制御チャネルは，パケットによるデータ通信と物理リンクを共用することに起因する制約がある．データ通信性能への影響を避けるには，データ通信が物理リンクを占有していない時間を用いてサイクルスチール式に制御情報を転送するしかない．このため，データ転送の状況によっては制御情報の転送が著しく遅れる可能性がある．こうした転送遅れが許容できない場合は，データ通信と競合することになる．言うまでもなく

データ通信の性能に干渉する。

仮想制御チャネルを用いる場合は、上記のようなトレードオフを勘案する必要がある。本稿の検討では、仮想制御チャネルによる制御情報の伝送が本来のデータ通信の性能に一切の影響を与えないよう、サイクルスチール式に転送されるものとする。無論シミュレーション評価もその前提の下に行う。

こうした評価は、結合網の制御のために多大なコストを払えないとの現実的な制約から来たものである。本稿で検討する方式の効果について、本質的な議論を行うには、制御チャネルを仮想化したことによるオーバーヘッドを除いた状態で評価したい。このために、比較の対象として、専用通信路により制御情報を伝播するケースについてもシミュレーション評価する。

4. 評価

4.1 評価環境ならびに評価条件

4.1.1 シミュレータ

本稿での評価には、文献 [19] で用いた相互結合網シミュレータを用いる。

1 パケットは 8 フリットから構成されており、最初のフリットに送信先ノード等の制御情報が格納されているものと仮定している。

各ルータは、入力ポートにおいて仮想チャネルごとに 1 個のバッファを備えている。各々の容量は 16 フリット分ある。ただし、リングバッファにより構成することを前提にしているため、保持できるフリットの量は最大で 15 である。

本稿で評価対象とする相互結合網のトポロジおよび規模は、 32×32 の 2 次元トーラスである。

4.1.2 通信パターン

転送パターン設定の都合上、 $N \times N = 2^n \times 2^n$ のトーラス網とする。ノードの 2 次元座標 (X, Y) をそれぞれ n ビットの 2 進数 $X = x_{n-1}x_{n-2} \cdots x_1x_0$, $Y = y_{n-1}y_{n-2} \cdots y_1y_0$ で表す。また転送パターンによって、両者を 1 次元にまとめた $2n$ ビットのアドレス $W = w_{2n-1}w_{2n-2} \cdots w_1w_0 = y_{n-1}y_{n-2} \cdots y_1y_0x_{n-1}x_{n-2} \cdots x_1x_0$ も用いる。

転置 (transpose, trns)

$$(X, Y) \rightarrow (Y, X)$$

シャッフル (shuffle, shfl)

$$w_{2n-1}w_{2n-2} \cdots w_1w_0 \rightarrow w_{2n-2} \cdots w_1w_0w_{2n-1}$$

ビット反転 (bit complement, bcmp)

$$w_{2n-1}w_{2n-2} \cdots w_1w_0 \rightarrow \overline{w_{2n-1}} \overline{w_{2n-2}} \cdots \overline{w_1} \overline{w_0}$$

ビット逆順 (bit reverse, brev)

$$w_{2n-1}w_{2n-2} \cdots w_1w_0 \rightarrow w_0w_1 \cdots w_{2n-2}w_{2n-1}$$

ビット回転 (bit rotation, brot)

$$w_{2n-1}w_{2n-2} \cdots w_1w_0 \rightarrow w_0w_{2n-1}w_{2n-2} \cdots w_1$$

トルネード (tornado, torn)

$$W \rightarrow \text{mod}(W + N/2, N^2)$$

全ノードが同一の距離だけ離れたノードに送出する。文献 [5], [6], [7] 等では、距離をパラメータとして様々なケースについて評価しているが、本稿では、ノードの座標を W で表したとき $N/2$ だけ離れたノードに向けて通信するものとしている。

ランダム (random pair, rpar)

システム中にある全ノードに対し送受信相手となるノード 1 つをランダムに選び、2 つのノードの組で通信する。

全対全 (all-to-all, atoa)

各ノードが全ノードに対してパケットを送出する。送出は $(0,0)$ ノードから順に機械的に行う。他のパターンでは一対一での通信が行われるのに対し、このパターンでは全対全であることが異なる。

4.1.3 ルーティングアルゴリズム

本稿の評価では、以下の 5 種類のルーティングアルゴリズムを使う。

次元順 (dimension order, do)

まず x 軸方向に進んだ後、 y 軸方向に進む非適応ルーティング。

単純適応ルーティング (simple adaptive, sa)

下記の Cross-Line のアルゴリズムと同様だが準広域情報を一切使わずに、隣接ルータの情報のみでパケット転送方向を決めるものである。

Cross-Line (cl)

文献 [14] で提案した最短経路での適応ルーティングである。このルーティングアルゴリズムでは、パケットの転送方向を決めるために 2.3 節で述べた準広域情報を用いている。準広域情報を流量制御のためにも使用する場合は、同じ情報を 2 つの目的で使用することになる。

Duato's protocol (最短経路) (dp)

最短経路での適応ルーティングである。適応ルーティング用の仮想チャネル 1 本と、非適応ルーティング (次元順ルーティング) 用の仮想チャネルを持つ。パケットは適応仮想チャネルに投入される。パケットを最短経路で転送するとき、転送先候補となる隣接ルータの状態を検知し、ブロックされない方向に転送する。転送候補のいずれもブロックされる場合には、非適応仮想チャネルに移行し、以降次元順ルーティングで転送される。

Duato's protocol (非最短経路) (df)

上記 Duato's protocol を非最短経路に拡張したものである。

4.1.4 一斉同期通信

各ノードが通信パターンにより規定される送り先に向

け、8フリット長の packets を 10 個連続して送出する*3。
 $t = 0$ で全ノードが一斉に送出を始め、所定の packets がすべて目的地ノードに到着し終えるまでにかかる時間 (duration) を計測する。本稿で用いたシミュレータでは、packets の出力ポートが競合した際、当該ルータに到着した時間 (age) をもとに全ての入力ポート×仮想チャネルについてラウンドロビン式に割り当てる方法をとっている。このため、packet 制御は以前のルーティング動作の履歴に影響される。この理由により、一斉同期通信を 100 サイクルの間隔をおいて 10 回行い、それらの平均値をとった。なお、評価の過程では 4.2 節に示すようにスループットや平均レイテンシなどの性能指標をウィンドウ時間を設けて時系列に取得している。ウィンドウ時間は事前の試行により定める。一対一通信では通信時間が数 100 から多くても 2500 [cycles] 程度であったため、ウィンドウ時間を 10 [cycles] とする。ただし、全対全通信 (atoa) ではこれに比して 50 倍程度長い場合があるためウィンドウ時間を 100 [cycles] とする。またこれに伴い、一斉同期通信 1 回毎の間隔 (ギャップ時間) を 1000 [cycles] とする。

4.1.5 定常通信

傾斜負荷法 (Ramp Load Method, RLM [21]) により評価を行う。転送負荷率をシミュレーション時刻 (クロック数) と共に徐々に増していき、時間あたりの受信 packet 数 (スループット) と平均レイテンシを求め、2.4 節に示した方法により、定量的評価のための 2 指標である臨界転送負荷と網性能指標を求める。

4.1.6 エントロピースロットリングのパラメータ

3.2 節で述べたように、エントロピースロットリングは

- packet 移動度の閾値 R_{th} ,
- 強制解除パラメータ P_n

の 2 つのパラメータで制御される。文献 [18] では主にランダム通信 (定常通信) について検討を行っていたが、本稿では一斉同期通信ならびに種々の転送パターンに拡大して評価を行うことに鑑み、パラメータ空間を拡大して評価する。具体的には、 $R_{th} = \{1.0, 0.8, 0.6, 0.4\}$, $P_n = \{0.6, 0.4, 0.2, 0.1\}$ の全組み合わせである。

4.1.7 ゲーティング方法

本稿では、2次元トーラスの性質から、 x 軸方向、 y 軸方向に直進する経路を本線とみなす。このため、

- 軸を変える場合 (turn-gating)
- 投入する場合 (injection-gating)
- 両者の組み合わせ (combinatorial gating)

の経路が合流路に相当する。なお、本稿でのゲーティング手法では、packets の転送をフリットの単位で抑制する。このため、上記の packet 投入の抑制 (injection-gating)

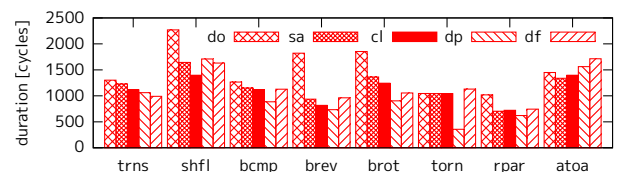


図 4 一斉同期通信性能 (各ルーティングアルゴリズムでの基本性能)

は、3.2 節に示したスロットリングとは異なり、packets の生成自体は抑制せず、生成された packets の転送をフリット単位で抑制する。

本稿では、以下の 4.1.8 節に示す評価関数と、上記の合流路の全組み合わせについて評価を行う。

4.1.8 準広域情報の評価関数

文献 [20] で用いた評価式パラメータのうち、

- $f_1(VCI, 16) > 0$,
- $f_2(VCI, 0.2) > 1.1$,
- $f_3(VCI, 0.8) > 1.1$,
- $f_4(VCI, 16) > 4$

の 4 通りを用いた。

4.1.9 バッファ占有準位

本稿での評価において、バッファ占有準位は 0, 4, 8 [flits] とした。たとえば、バッファ占有準位 4 のとき、バッファの空き容量が 4 以下になると、当該バッファに相当する準広域情報のビットが busy となる。

4.2 評価結果：一斉同期通信

それぞれのルーティングアルゴリズムのみを用いた場合の、一斉同期通信の性能 (通信時間) を図 4 に示す。なお全対全通信 (atoa) の場合のみ、通信時間が長いことから一律に 1/50 を乗じている。本稿では、これらの値をベース性能 (BA) と呼ぶ。

各ルーティングアルゴリズムに対して、エントロピースロットリング (ET) ないしゲーティング (GT) のいずれかを適用する。エントロピースロットリングの場合は、4.1.6 節に示したパラメータ (packet 移動度の閾値 R_{th} , 強制解除パラメータ P_n) の全組み合わせで評価を行う。ゲーティングの場合は、評価関数 (4.1.8 節)、合流路 (4.1.7 節)、バッファ占有準位 (4.1.9 節) の全ての組み合わせで評価を行う。

通信パターンごとの一斉同期通信性能を図 5 に示す。この図では、エントロピースロットリングについてはパラメータの全組み合わせの中で、ゲーティングについてはバッファ占有準位ごとに、最も通信時間の短いもののみを示している。BA がベース性能、TE がエントロピースロットリング、GT(n) がゲーティング (バッファ占有準位 n) を表している。

図 5 から、非適応ルーティングアルゴリズム (do) でも、流量制御を導入することにより性能改善できるケースがあ

*3 本評価でのスロットリングは packets の単位で行った。packet 数 1 の一斉同期通信ではスロットリングの効果を得られないことから、長い packet 1 個ではなく 10 個に分割している。

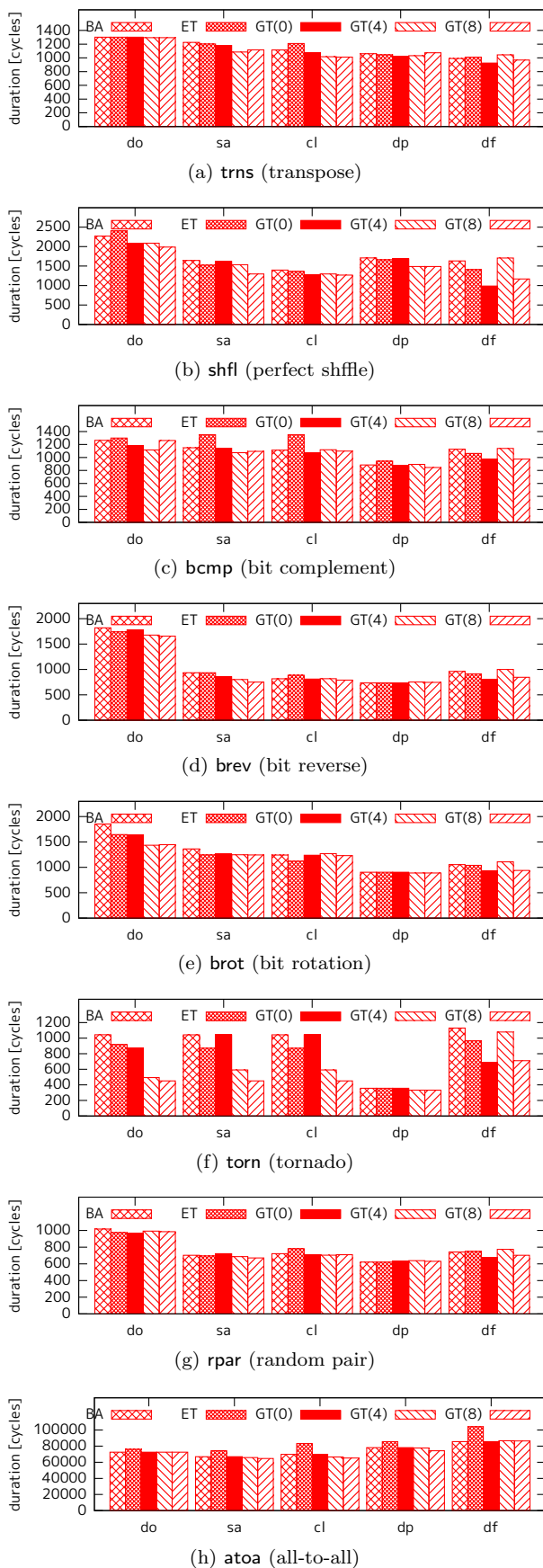


図 5 通信パターンごとの一斉同期通信性能

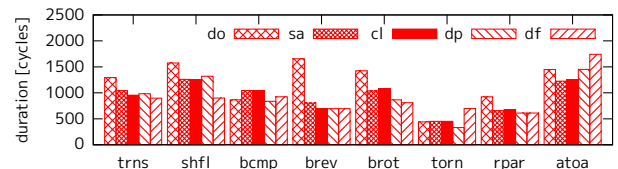


図 6 一斉同期通信性能 (本稿での評価の最良値)

ることがわかる．効果の有無やその程度は，通信パターンによって大きな差がある．たとえば trns や atoa では大きな改善は得られていないが，その一方で torn では通信時間が半分近くまで改善している．

また同図から，適応ルーティングについても流量制御が有効であることがわかる．

バッファ占有率の導入により，輻輳が実際に発生する前に局所的に輻輳傾向にあることを検出し，輻輳の発生や成長を抑える効果が期待された (3.3.3 節)．図 5 では，多くの場合について，バッファ占有率が大いほど一斉同期通信時間が短くなる傾向が見られる．これは同率の導入が輻輳制御に有効であることを示している．

各通信パターンについて，ベースのルーティングアルゴリズムごとに一斉同期通信時間の最良のものを表したものが図 6 である．この図においても，全対全通信 (atoa) の場合のみ通信時間に一律に 1/50 を乗じている．

次に仮想制御チャンネルが通信性能に与える影響を調べた．流量制御のための制御情報を流通させるための専用の通信路 (制御網) を仮定したときの通信性能を上記と同様のシミュレーションにより求め，比較する．その結果を図 7 に示す．制御網以外の条件を揃えたうえで，専用の制御網を用いた場合の通信性能 (一斉同期通信時間)，仮想制御チャンネルを用いた場合の通信性能を求め，それぞれ図中の x 軸， y 軸として表している．直線は，両者の性能が同じ ($y = x$) 場合を示す補助線であり，点がこの直線よりも上側にある場合は，専用制御網のほうが性能が高いことを表す．

この図から，専用の制御網を用いたほうが高い性能を得られる傾向が強いが，わずかな例外を除き $y = x$ の近辺に分布している，すなわち，本来のデータ通信に干渉しない仮想制御チャンネルでも十分な制御性能が得られることがわかる．

本稿で用いた流量制御手法は，一斉同期通信を行う際にどのような効果をもたらしているか，10 サイクルごとの送受信パケット数の経緯を図 8 に示す．各グラフの上段が送出パケット数であり，ノードから結合網に新たに送出された時点でカウントしている．バッファが満杯の場合は，空きができて実際に投入できた時点でカウントする．下段が受信パケット数である．一斉同期通信が終了した時刻を x 軸上の丸印で示している．

図 8(a) は trns の場合である．この通信パターンでは，

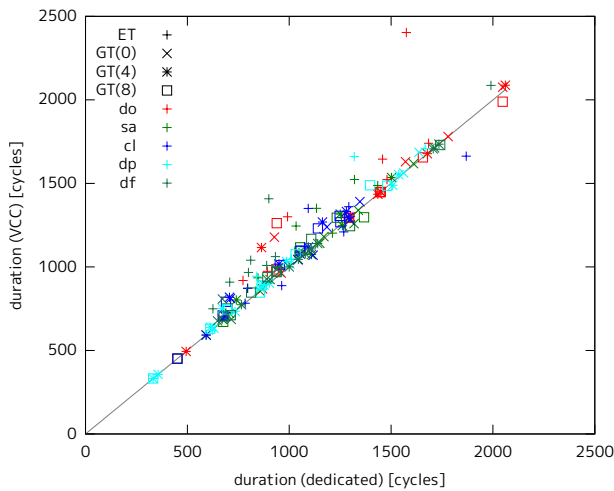


図 7 仮想制御チャネルの影響評価

ベース性能とゲーティングが全く同じになっている。すなわち、この図のゲーティングパラメータでは、効果が全く得られていない。このように、通信パターンによっては期待した効果が得られない場合も多い。

エントロピースロットリングにより流量制御を行う場合、パケットの送出を過剰に抑制してしまう傾向が見られる。図 8(b) に示す shfl の例が典型的である。スロットリングにより送出パケット数が低く抑えられる状態が続いた後、スロットリングが解除されるが、そこで一気にパケットが流入するため輻輳を招き、再度スロットリングが発動している様子が見られる。送信パケット数は時間により極端な変動を示すが、受信パケット数はそれに比べ平準化している。スループットは受信パケット数により計られるため、受信パケットのみに着目しがちであるが、それだけではこうした極端な時間挙動を捉えにくい。文献 [18] でも議論されたように、スロットリング手法では、パケットの送出を抑えすぎると最大のスループットまで達せず、一方で輻輳が起きてしまうとそのためにスループットが落ちる問題がつきまとう。本稿の評価においてもエントロピースロットリングにより性能を向上させるのが難しいのは、過剰な動的挙動が大きなオーバーヘッドを生じたためと考えるのが妥当である。

1 回の通信量（各ノードが送出するパケット数）が多い場合、エントロピースロットリングでは、上述のような過剰な抑制と、抑制を解除した後の輻輳の発生とが繰り返し起きる状態になる。全対全通信の場合（図 8(c)）では、送受信パケット数が大きく振動していることがわかる。

4.3 評価結果：定常通信

定常通信の評価結果は、定量化性能指標である臨界転送負荷と網性能指標を用い、両者の値の 2 次元空間上の点として表す。

図 9 に結果を示す。図 9(a) が次元順ルーティング (do)

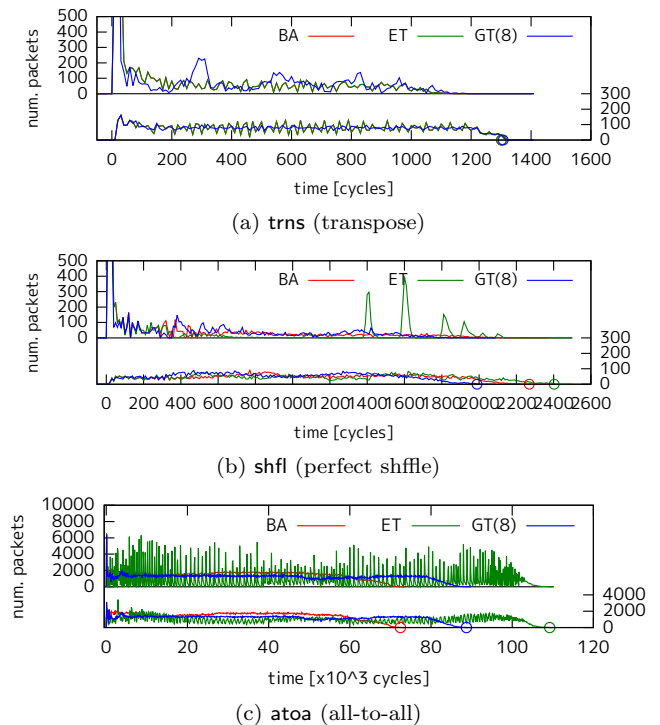


図 8 一斉同期通信の時系列性能

の場合、同図 (b) が最短経路 Duato's protocol (dp) の場合である。ベース性能、エントロピースロットリング、ゲーティングを点の形状で、通信パターンを色で区別して表現している。図では、ベース性能を表す点から x 軸に向けて線を引いている。同じ通信パターンでの性能点がこの垂直線の右側にある場合は、ベース性能より臨界転送負荷において上回っていることを示す。この結果を一斉同期通信の結果（図 5）と比較すると、性能の改善の度合いや順位が必ずしも符合しないことがわかる。

図 9(a),(b) とともに、ベース性能に比べ臨界転送負荷や網性能指標が際立って悪いケースが見られる。これらのほとんどはエントロピースロットリングの場合である。エントロピースロットリングでは、設定パラメータによって、本来の臨界転送負荷に達する前であっても、局所的に生じたわずかなパケットブロック状況を輻輳として検出してしまい、その結果スロットリングの発動と解除が繰り返され、一斉同期通信において図 8(c) に見られたような振動が同様に生じるものと考えられる。

4.4 デッドロックフリー性の保証

本稿での評価の過程で、適応ルーティングにゲーティング手法を適用した場合においてデッドロックが発生することが確認された。

デッドロックが発生するのは、パケットが進行方向の軸を変えるときにゲーティングする場合 (turn-gating) と、投入抑制との組み合わせの場合 (combinatorial gating) である。ゲーティングの機構では、輻輳の検出によりターン

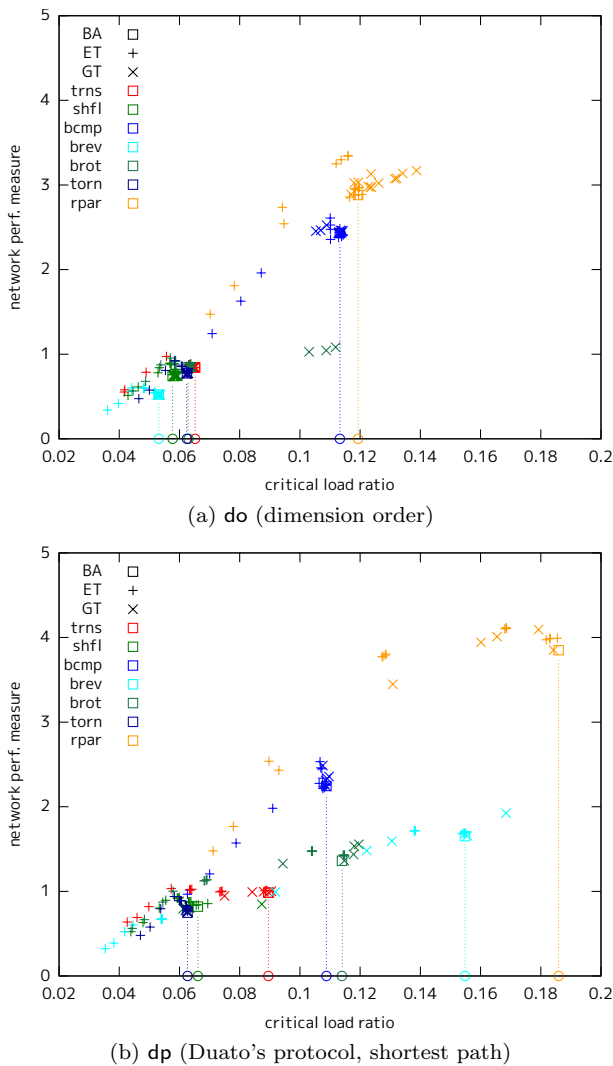


図 9 定常通信性能

が禁止される．ターンを禁止されたパケットはバッファ中で待たねばならず，ターン元のパケット転送が抑制され転送がブロックされる．この転送ブロックは準広域情報 VCinfo により輻輳として周辺ルータに検知されるから，別の箇所でゲーティングが発動されそこでターンが禁止される．こうした転送のブロックによるターンの禁止が連鎖していき，その依存グラフが閉路を構成するためにデッドロックが生じるのである．このため，本稿の評価においては，適応ルーティングでは，実質上，パケットの投入抑制 (injection-gating) のみにとどまった．

なお，次元順のような非適応ルーティングでは，そもそも可能なターンが限られているため，上述のようなデッドロックは生じない．しかし適応ルーティングにゲーティング手法を適用しようとすると，デッドロックフリー性を保証するため合流路の制御 (turn-gating) を制限する必要が生じ輻輳制御能力を引き出せない問題がある．この点は課題として残る．

5. おわりに

本稿では，相互結合網を構成する個々のルータの自律的な動作により内部状態を動的に観測する機能を導入し，パケットの流量制御を行うことで輻輳の発生や成長を抑制し，結合網全体の性能を向上させる手法を検討した．このために本稿では，広域的な輻輳の指標 (エントロピー) を基にしたエントロピースロットリングと，個々のルータ上で準広域情報により輻輳が検出された方向へのパケットの転送を抑制するゲーティング手法を導入し，種々の通信パターンについて一斉同期通信および定常通信での転送性能を評価・比較した．

その結果，特に準広域情報を用いるゲーティング手法において改善効果が認められた．本稿で検討をした手法は，基本的にルーティングアルゴリズムと直交するものであり，任意の組み合わせが可能である．

性能向上の有無やその度合いは，通信パターンに強く依存しており，大きな改善効果が認められなかったものもあった．たとえば本稿で考案したゲーティング手法では，ランプメータリングにおける本線に相当する通信経路の存在が前提になる．本稿では x 軸ないし y 軸を直進する経路を本線とみなし議論を進めたが，通信パターンによっては本線を明確に設定できないケースがある．期待した効果が得られなかったのは，このようなケースである．たとえば，転置 (trns) では，直進路ばかりでなく，進行方向の軸を変えるターンを本線として優先させた方が良い箇所もある．本稿で検討した手法ではこのようなケースに対応できず，今後の課題とする．

また，本稿ではパラメータの膨大な組み合わせについて評価を行う都合から，比較的小規模 (32×32) の相互結合網を扱った．[19], [22] で行ったように，規模の大小と効果の度合いとの関連を確認しておく必要がある．

謝辞 本研究は，一部日本学術振興会科学研究費補助金 (基盤研究 (C)24500054, 24500055) の援助による．

参考文献

- [1] L. S. Yuan, J. B. Kreer, "Adjustment of Freeway Ramp Metering Rates to Balance Entrance Ramp Queues," *Transportation Research*, Vol.5, No.2, pp.127-133, 1971.
- [2] M. Papageorgiou, A. Kotsialos, "Freeway ramp metering: an overview," *Proc. 2000 IEEE Intelligent Transportation Systems*, pp.228-239, 2000.
- [3] 赤松隆, "交通ネットワーク流の動的制御モデル," *土木計画学研究・講演集* 35, 311 (CD-ROM), 2007.
- [4] J. Duato, S. Yalamanchili, L. Ni, "Interconnection Networks: An Engineering Approach," Morgan Kaufmann Pub., 2003.
- [5] W. J. Dally, B. Towles, "Principles and Practices of Interconnection Networks," Morgan Kaufmann Pub., 2004.
- [6] B. Towles, W. J. Dally, "Worst-case Traffic for Oblivious Routing Functions," *IEEE Computer Architecture*

- Letters, Vol.1, No.1, 2002.
- [7] A. Singh, et al., "Globally Adaptive Load-Balanced Routing on Tori," IEEE Computer Architecture Letters, Vol.3, No.1, 2004.
 - [8] T.C. So, S. Oyanagi, K. Yamazaki, "Speculative Selection in Adaptive Routing on Interconnection Networks," IPSJ Trans. Advanced Computing Systems, Vol.44, No.SIG 11(ACS-3), pp.147-156, 2003.
 - [9] T. Ohira, R. Sawatari, "Phase Transition in a Computer Network Traffic Model," Physical Review E, Vol.58, No.1, pp.193-195, 1998.
 - [10] M. Thottethodi, A. R. Lebeck, S. S. Mukherjee, "Exploiting Global Knowledge to Achieve Self-Tuned Congestion Control for k -Ary n -Cube Networks," IEEE Trans. Parallel and Distributed Systems, Vol.15, No.3, pp.257-272, 2004.
 - [11] 柴村英智ほか, "パケットベーシングによる全対全通信の最適化とシミュレーション評価," 情報処理学会論文誌: コンピューティングシステム (ACS), Vol.4, No.3, pp.56-65, 2011.
 - [12] 柴村英智ほか, "パケットベーシングを用いた集団通信に対するロード/ネットワークインバランスの影響," 情処研報, Vol.2012-HPC-135, No.2, pp.1-8, 2012.
 - [13] 横田隆史ほか, "相互結合網のトポロジを活かしたシステム支援機能とその評価," 情報処理学会論文誌, Vol.38, No.4, pp.873-882, 1997.
 - [14] 横田隆史ほか, "大域的な情報を用いる相互結合網方式 Cross-Line," 情報処理学会論文誌: コンピューティングシステム, Vol.46, No.SIG 16 (ACS-12), pp.28-42, 2005.
 - [15] 横田隆史ほか, "セルオートマトンによる相互結合網の輻輳の解析," 情報処理学会論文誌: コンピューティングシステム, Vol.47, No.SIG 7 (ACS-14), pp.21-42, 2006.
 - [16] T. Yokota, et al., "Phase Transition Phenomena in Interconnection Networks of Massively Parallel Computers," Journal of Physical Society of Japan, Vol.75, No.7, p.078401 (7 pages), 2006.
 - [17] 横田隆史ほか, "セルオートマトンによる相互結合網の間欠的な輻輳の解析," 情処研報, Vol.2006, No.88, pp.91-96, 2006.
 - [18] 横田隆史ほか, "エントロピー・スロットリング: 相互結合網のパケット移動度に着目した輻輳制御手法," 情報処理学会論文誌: コンピューティングシステム, Vol.47, No.SIG 12 (ACS-15), pp.1-11, 2006.
 - [19] 横田隆史, 大津金光, 馬場敬信, "相互結合網のスケール効果に関する初期評価," 情処研報, Vol.2008, No.75, pp.91-96, 2008.
 - [20] 横田隆史, 大津金光, 馬場敬信, "準広域情報を用いるルーティングアルゴリズムの改善検討," 情処研報, Vol.2009-ARC-184, No.30, pp.1-8, 2009.
 - [21] T. Yokota, K. Ootsu, T. Baba, "A Quantitative Evaluation Methodology of Interconnection Networks," IPSJ Trans. Advanced Computing Systems, Vol.2, No.3, pp.58-70, 2009.
 - [22] 横田隆史, 大津金光, 馬場敬信, "相互結合網の転送パターンと定常・非定常性能について," 情処研報, Vol.2010-ARC-190, No.18, pp.1-8, 2010.