

ルーティングアルゴリズム開発のための 環境構築とその初期評価

絵面 聡 西谷 雅史 横田 隆史 大津 金光 馬場 敬信 †
宇都宮大学工学部情報工学科 ‡

1 はじめに

相互結合網の転送方式は並列計算機の性能に大きな影響を与える．そのため、転送バンド幅の拡大と並行して、転送効率を向上させるための転送方式の改善検討を進める必要がある．そこで我々は、転送効率を向上させるためにルーティングアルゴリズムなどを容易に追加・改変でき、かつ、相互結合網内部の挙動を多角的に解析評価できるシミュレータ Chimera を作成した．本稿ではこのシミュレータを説明し、その初期評価を報告する．

2 ルーティングアルゴリズムの開発環境

近年、集積度の向上によりハードウェア資源が豊富に使用できるようになっている．そこで資源を有効に活用しながら、より効率的なパケット転送を行うための新たなルーティングアルゴリズムが求められている．このようなアルゴリズムとして、動的な情報を利用することで転送の効率化を図る SSR (Speculative Selection Routing)^[1] を用いた方式や我々が提案している Cross-Line^[2] がある．Cross-Line 開発の際、容易にルータを実装できないので、何種類ものルーティングアルゴリズムを取り入れ可能で、かつ、様々な角度からこれらを解析評価できるネットワークシミュレータを開発する必要がある．

3 相互結合網シミュレータ Chimera

3.1 設計思想

Chimera は多くのパラメータを設定可能にすることで柔軟なシミュレーションを可能にする．これにより、パラメータ条件を合わせることで Chimera に実装されていないルーティングアルゴリズムとの比較が可能となる．また Chimera に Cross-Line が実装されたとき、スムーズに解析評価を行えるようにする．

3.2 パラメータ

Chimera で設定可能なパラメータは以下の通りである．表 1 に実装予定パラメータと共にまとめる．

- トポロジー
ネットワーク構造を決定する．選択肢は基本的に実用的な 2 次元メッシュ構造か 2 次元トラス構造の 2 種類とした．
- フロー制御
パケット送信方式を決定する．選択肢はストアアンドフォワード方式、ワームホール方式、バーチャルカットスルーの 3 種類とした．

- ルーティングアルゴリズム
非適応ルーティングの Dimension-order と Cross-Line の 2 種類とした．
- 通信パターン
通信パターンとしてランダム通信、ホットスポット通信、ホットライン通信が選択可能である．また、ホットスポット通信の場合は何処のノードを、ホットライン通信の場合は何処のノード列をどれだけ混ぜるかという選択もする．
- ネットワークサイズ
シミュレーションを行う際のノード数を決定する．任意の 2 以上の整数を与えることで、その数の 2 乗個のノード数を持つ正方形の相互結合網を形成する．
- 最大パケット生成インターバル
シミュレータ Chimera はパケット送信後ランダムな間隔を経て次のパケットを生成する．その間隔の最大値が最大パケット生成インターバルである．
- 最終パケット生成クロック
このクロック以降のパケット生成を行わない．シミュレータは全パケットが到着したとき終了する．
- フリットサイズ
フリットサイズを決める．全てのパケットはここで定めた値によって固定長で生成される．

表 1: 実装済と実装予定パラメータ (* 印付は実装予定)

トポロジー	2 次元メッシュ構造 2 次元トラス構造
フロー制御	ストアアンドフォワード方式 ワームホール方式 * バーチャルカットスルー方式
ルーティングアルゴリズム	Dimension-order * Cross-Line
通信パターン	ランダム通信 ホットスポット通信 ホットライン通信
ネットワークサイズ	$n \times n$ (ノード数) ($n \geq 2$)
最終パケット生成クロック	n (clocks) ($n \geq 1$)
最大パケット生成インターバル	n (clocks) ($n \geq 0$)
フリットサイズ	n (flits) ($n \geq 1$)

3.3 処理の流れ

最初に 3.2 で示したパラメータの設定後、設定されたサイズのノードやリンクなどが生成される．その後イベントドリブン方式で処理が行われる．パケットの生成・送信・受信のイベントをイベントリストに登録しながら、リストに書かれた処理を順次処理していくことによりシミュレーションを進めて行く．最終パケット生成クロック以降はパケットの生成をやめる．リストの中にイベントがなくなったらシミュレーションを終了し、総送信パケット数と最終クロックと平均レイテンシを出力する．

* An environment for routing algorithm development and its preliminary evaluation

† Satoshi Ezura, Masashi Nishitani, Takashi Yokota, Kanemitsu Ootsu, and Takanobu Baba

‡ Department of Information Science, Faculty of Engineering, Utsunomiya University

3.4 機能

Chimeraは定常状態検出のための支援機能を備えている。シミュレーション開始直後のパケット数は少なく、ネットワークも混雑していないが、ある一定クロック以降はレイテンシが安定した定常状態になる。過渡状態を除外し定常状態でのレイテンシやスループットを測定するためにこの支援機能を備えた。計測方法は一定のクロック毎に到着したパケットの平均レイテンシを計測した。

さらに、パケットの動きを把握可能にするため、ネットワーク内の状態を濃淡によるグラフィック表示により可視化する機能を備えた。ある時間における各ノード内のバッファのブロック時間やリンクの使用率などを測定することで可視化を行った。

4 Chimeraの適用例

Chimeraによりシミュレーションを行った例を示す。以下の4.1と4.2では表2に示すパラメータで行ったものである。

表 2: シミュレーション条件

トポロジー	2次元メッシュ構造
フロー制御	ストアアンドフォワード方式
ルーティングアルゴリズム	Dimension-order
通信パターン	ランダム通信
ネットワークサイズ	32 × 32(1024 ノード)
最終パケット生成クロック	50000(clocks)
最大パケット生成インターバル	可変数(clocks)
フリットサイズ	16(fbits)

4.1 定常状態

定常状態検出支援機能により、図1のグラフを作成し、ここから評価する範囲を定める。この図はインターバルを100とした場合で20000クロック～50000クロックの間はレイテンシがほぼ一定の振幅になっているので、この間を定常状態とみなす。

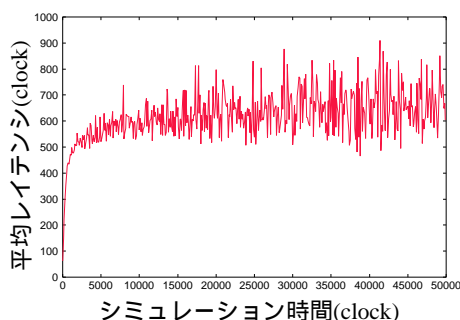


図 1: 定常状態検出補助機能によるグラフ

4.2 レイテンシとネットワークスループット

ネットワークスループットは式1の通りに定義する。式1は分母が計測時間内の最大処理能力を表し、分子がシミュレーション上の総処理量を表している。

$$\text{スループット} = \frac{\text{到着パケット数}}{\text{ノード数} \times \frac{\text{計測時間}}{\text{パケット転送時間}}} \quad \dots (1)$$

4.1で求めた範囲でインターバルを変化させたときの、レイテンシとネットワークスループットのグラフを図2に示す。図はパケット生成インターバルを小さくするとスループットが増加していることを示している。また、インターバルを小さくしすぎると輻輳し、スループットとレイテンシが悪くなっていることもこの図から分かる。

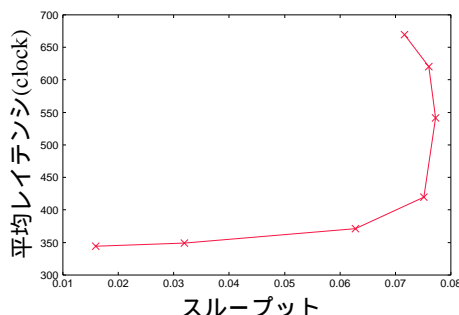


図 2: レイテンシとスループット

4.3 ネットワーク状態の可視化

ネットワークの状態を可視化した例を図3、図4に示す。図中の各ドットがノードに対応しており、色が薄いほどそのノードでのブロック時間が長く(図3)、リンクの使用回数が多い(図4)ことを表す。シミュレーション条件は表2の内容から通信パターンだけをホットスポット通信に変更した。ホットスポットは中央で全パケットの約5%を集中させた。Dimension-orderを用いて横軸優先で送信を行っているため、図3よりホットスポットの縦軸に沿った薄い線で著しく渋滞していることが分かる。ホットスポットに近い部分ではパケットが活発に転送されることも分かる。そして図4より縦方向に比べ横方向が活発に使用されていることが分かる。

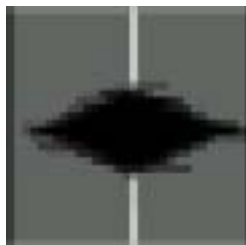


図 3: ブロック時間

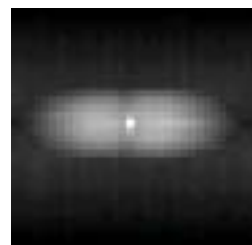


図 4: リンク使用率

5 おわりに

本稿では、ネットワークシミュレータChimeraについて説明した。さらにChimeraを用いることで定常状態時のレイテンシとネットワークスループットが評価可能であることを示し、ネットワーク状態のグラフィック表示を行った。今後はChimeraにCross-Lineやバーチャルカットスルー方式を実装する予定である。

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(B)14380135, 同(C)14580362, 若手研究14780186)の援助による。

参考文献

- [1] Tran Cong So, Shigeru Oyanagi, Katsuhiro Yamazaki: "Speculative Selection in Adaptive Routing on Interconnection Networks," IPSJ Symposium Series Vol.2003, No.8, pp.29-36(2003).
- [2] 西谷 雅史, 絵面 聡, 横田 隆史, 古川 文人, 大津 金光, 馬場 敬信: "動的な情報を用いたルーティングアルゴリズム Cross-Lineの検討," HOKKE2004(発表予定).