

パイプラインステージ統合による 省電力・可変パイプラインルータに関する研究

枚田 優人^{†1} 松谷 宏紀^{†1}
鯉 淵 道 紘^{†2} 天 野 英 晴^{†1,†2}

チップ内ネットワークはパケットをパイプライン処理するルータを多数用いることで、高スケーラビリティ、高スループットを実現している。本論文では、トラフィック負荷に応じて動作周波数とパイプライン構成を切り替えることが可能な省電力オンチップルータを提案する。各動作周波数におけるパイプラインステージあたりの処理量を大きくするために、低い動作周波数ではパイプライン段数を浅くする。これを、単純なカスケードによるパイプラインステージの統合ではなく、各ステージ処理の並列性を最大限に利用する投機実行によって実現する。評価結果より、性能、電力面で最適なパイプライン構成は、1 段から 3 段とアプリケーションごとに大きく異なることが分かった。さらに、パイプライン構成を変更しない既存のルータに DVFS (Dynamic Voltage and Frequency Scaling) を用いた場合に比べ、提案ルータを用いることで消費電力の増加を最大 46%、抑えることができた。

A Study of Low-power Variable-pipeline Router Using Pipeline-stage Integration

YUTO HIRATA,^{†1} HIROKI MATSUTANI,^{†1}
MICHIIHIRO KOIBUCHI^{†2} and HIDEHARU AMANO^{†1,†2}

Network-on-chip achieves both high scalability and high throughput, by using a large number of pipelined routers for packet processing. In this paper, we propose a low-power on-chip router which can dynamically adjust the depth of its pipeline structure in response to the operating frequency and traffic loads. To increase the workload per pipeline stage at each frequency, it reduces the pipeline depth when the frequency is low by speculatively executing multiple pipeline stages in parallel unlike a simple cascading that drastically decreases the frequency. Evaluation results show that the best pipeline structure optimized to the performance or power is varied from one stage to three stages depending on applications. The proposed router reduces up to 46% of the power

consumption compared with that of router using the traditional fixed-pipeline structure with DVFS (Dynamic Voltage and Frequency Scaling).

1. はじめに

半導体技術の進歩によって単一チップ上にプロセッサやメモリ、I/O など複数の設計モジュールをタイル状に実装できるようになり、このようなタイルどうしの接続にチップ内ネットワーク (Network-on-Chip: NoC) が用いられるようになった¹⁾。

オンチップルータは、高い動作周波数、高スループットを実現するためにパケット処理を複数に細分化するパイプライン方式を採用している。そして、パケットはルータ内においてルーティング計算、出力ポートの設定、アービトレーション、クロスバ転送などの複数のステージを経て入力ポートから出力ポートへ転送される。

一般的に、パイプラインサイクル数 (以後、段数と呼ぶ) を含むパイプラインの構成は、チップの動作周波数、ルータのスループット、遅延などの処理能力面、ひいては消費電力に大きな影響を及ぼす。パイプライン段数が少ない場合、1 サイクルあたりの処理量が大きくなるためにパイプライン段数の多いルータに比べて最大動作周波数が低下する。しかし、1 クロックサイクルあたりのスループットは、特にパケット長が短い場合は段数の少ないルータの方が大きくなる²⁾。クロックサイクルあたりのスループットの性能差にはトラフィックパターンが大きく影響するため、最大のスループットを提供するルータのパイプライン構成はアプリケーションに大きく依存する。同様に、一定のスループットを達成するために必要となるルータの消費電力の最適化も動作周波数と電圧に大きく影響を受けるため、段数を含むパイプラインの構成が鍵を握ることになる。

マルチ/メニーコアプロセッサでは様々なアプリケーションが実行されるため、トラフィックパターンが設計時に一意に定まらない。しかし、性能、電力消費量の面で最適なルータのパイプライン構成は、トラフィックパターンに大きく影響を受ける。

そこで本研究では、動作周波数と段数を含むパイプライン構成をトラフィック負荷、パターンに応じて切り替えることが可能な可変パイプラインルータを提案する。最近では DVFS

^{†1} 慶應義塾大学理工学部

Faculty of Science and Technology, Keio University

^{†2} 国立情報学研究所

National Institute of Informatics

(Dynamic Voltage and Frequency Scaling) を用いることにより、チップへの供給電圧と動作周波数を動的に変化させることが可能となりつつあるため、可変パイプラインルータにおいても DVFS あるいは VFS の利用を前提とする。各動作周波数におけるパイプラインサイクルあたりの処理量を大きくするために、可変パイプラインルータでは低い動作周波数において単純なカスケードによるパイプラインステージの統合ではなく、各ステージ処理の並列性を最大限に利用する投機実行を行う。

可変パイプラインルータを運用するためには、トラフィック負荷の測定あるいは予測方法や、負荷に応じてパイプライン構成を切り替えるタイミングなどを検討する必要がある。ただし、これらは既存の DVFS を用いたマイクロプロセッサにおいても生じる課題であり、前者は、アプリケーションの設計段階におけるシステムレベルシミュレーションや、pilot execution などからノード間の通信パターンを解析することが可能である。

本論文では可変パイプラインルータに特化した最も重要な課題である、(1) ネットワークの性能を最大限引き出すためのルータパイプライン構成、(2) 一定のスループットを提供するうえで消費電力の面において最適なパイプライン構成、の 2 つに焦点をあて、可変パイプラインルータの有効性とそのオーバーヘッドを明らかにする。

本論文の構成は次のとおりである。2 章で関連研究を述べ、3 章では可変パイプラインルータのアーキテクチャと設計について述べる。4 章では可変パイプラインルータの面積、動作周波数、スループット、消費電力についての評価を行う。最後に、5 章でこれらをまとめる。

2. 関連研究

CPU アーキテクチャでは、そのパイプライン構造を負荷に応じて動的に変更する研究がさかんである。プロセッサの負荷が低い場合、動作周波数を低下させて複数の演算を 1 サイクルで実行させることで、大きな性能低下をとまわずに、消費電力を低下させるカスケードあるいはパイプライン統合³⁾が提案されている。

一方でオンチップルータに関しても、トラフィック状況に応じてレイテンシや消費電力を削減することを目的とした、ルータの機能を動的に変更して適応させる手法が提案されている。適応型ルーティングを採用することでトラフィックを分散させ、性能低下することなく動作周波数を抑える⁴⁾、ルータにチャンネル単位のパワーゲーティングを適用し、混雑状況に応じて動的にチャンネル数を変化させる^{5)–7)}、ルータ内のデータパスを変更させることで耐故障性を向上させる Default Backup Path⁸⁾、仮想的に非隣接ルータ間でバイパス経路を構成することにより中継ルータにおける所要パイプライン段数を削減する Express VC⁹⁾ な

どがこの例である。

しかし、これらの方法は、ルータの機能やバッファ数を切り替えたり、バイパス経路を利用するかどうかを切り替えたりするもので、パイプラインの構造自体を動的に変化させる CPU のカスケードに近い手法の適用例は報告されていない。

3. 可変パイプラインルータ

本章では、可変パイプラインルータのアーキテクチャを提案する。

3.1 パイプライン構造

一般的にパイプラインルータにおけるパケット処理は 4 つのプリミティブに分けることができる。ここでは、説明のために図 1 に示した典型的な 4 サイクルパイプライン処理を採用しているルータ¹⁰⁾ について簡単に説明する。パケットが入力ポートに到着すると、(1) 入力ポートにおいてパケットのヘッダから出力ポート情報を解読、あるいはルータの制御ユニットから出力ポート情報を獲得し (Routing Computation: RC), (2) 出力仮想チャネルを割り当て (Virtual-Channel Allocation: VA), (3) 出力ポートへのクロスバの設定を行い (Switch Allocation: SA), (4) パケットを転送 (Switch Transfer: ST) することで、パケットは出力ポートへ転送される。この処理を各ルータが行うため、パケットヘッダが 3 ホップ先のルータに到着するためには、12 サイクル必要となる。4 サイクルルータにおける一連の流れを図 2(a) に示す。横軸は経過時間を表し、3 フリットで構成されたパケット

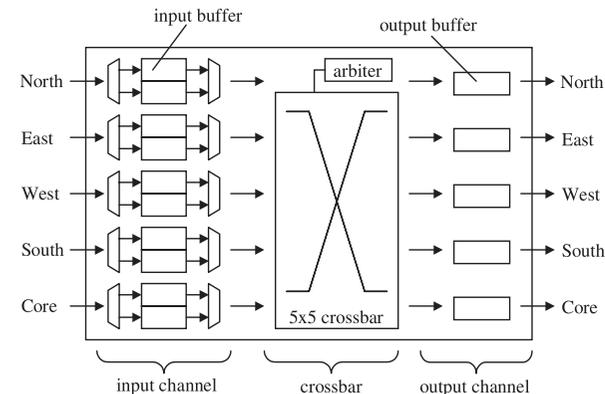


図 1 典型的な 4 サイクルルータ

Fig. 1 Typical 4-cycle router.

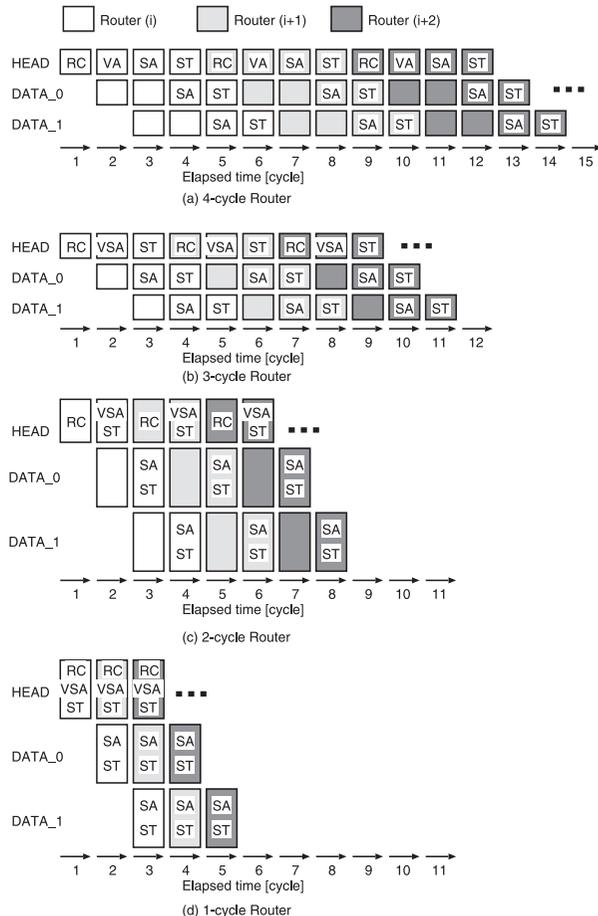


図2 ルータの packets 処理の流れ
Fig.2 Packet processing flow.

が転送されていく場合を表している。

可変パイプラインルータは、任意のルータアーキテクチャに適用可能であるが、ここでは、典型的なルータの例としてこの4サイクルルータを基とする。

次に、可変パイプラインルータにおいてパイプライン段数を変更することを考える。4サ

イクルパイプラインから3サイクルパイプライン構成へ変更するために、VAとSAを投機 (Speculative) 実行²⁾ することで1つのサイクルに統合 (VSAと以後呼ぶ) する。これは、SAはVAの処理の完了を待たなくても、VAが成功すると仮定して実行することが可能である点を利用している。VSAは、VAとSAを統合しているが、どちらか一方、あるいは両方が失敗した場合にはいったんVAとSAの要求をとりやめ、1クロックの間をおいてから再試行する。

さらに、VSAとSTを投機的に処理することで2サイクルルータを実現することができ、究極的には、すべての動作を同じサイクルの中で行うことで1サイクルルータを実現することができる。なお、3サイクルルータのVSAとSTはRCにおいてパケットが転送される出力ポートが定まった後のみ実行可能である。そのため、2サイクルルータのパイプラインサイクルはRCとVSA+STという構成とした。

図2に各サイクルにおけるパイプライン処理の例を示す。このように、サイクルあたりのステージ処理量を変化させることでパケットの転送に必要なサイクル数は大きく異なる。

3.2 可変パイプラインルータのアーキテクチャ

前節で述べたとおり、3サイクルルータのVSAは4サイクルルータにおけるVAとSAを統合している。そしてVAとSAを投機的に実行した場合、両方の結果を処理する新しいロジックが必要になる。ただし、軽量のオンチップルータではVAとSAを統合した場合でも最大動作周波数が低下しないことが分かっている⁷⁾。同一周波数におけるスルーputは、ルータのパイプライン段数が小さいほど性能が高くなる。

これにより、ルータのパイプライン段数を深くした場合には最大動作周波数が向上しなければ、性能、消費電力の面で効果が得られない。そのため、以後、本論文で扱う可変パイプラインルータは4サイクルパイプライン以上の長さのパイプラインは利用せず、1~3サイクルパイプラインの議論に焦点をあてる。

3サイクルパイプラインから、2サイクルパイプラインに変更する場合、出力ポートのチャネルバッファをバイパスし、同時に、統合するVSA、STのステージ処理を並列に投機実行させる。ここから、投機処理の成功の可否によるフリットの管理が必要となる。

さらに、1サイクルパイプラインに変更する場合は、入力ポートのチャネルバッファをもバイパスする。ただし、RC、VSA+STのステージ処理には依存関係があるため、これらは直列に実行する。

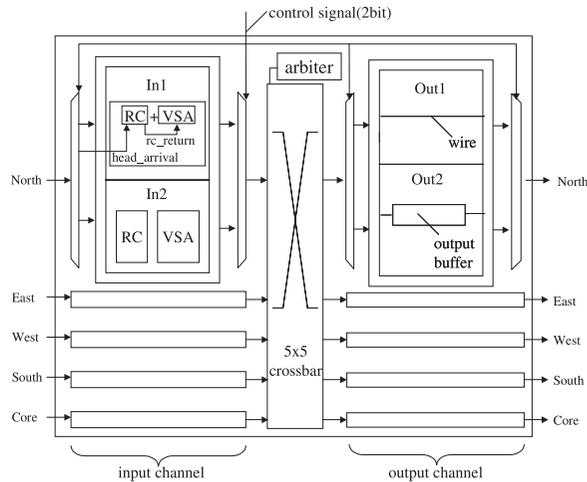


図 3 可変パイプラインルータの概要
Fig. 3 Diagram of the variable-pipeline router.

3.3 パイプライン段数の切替え

可変パイプラインルータは、単純なカスケードリングと異なり、パイプライン段数を減らす場合、複数のステージ処理を可能な限り並列に投機実行させることで、最大動作周波数の低下を抑えつつ 1 つのサイクルに統合する点が特徴である。この可変パイプラインルータを実現するため、我々は図 1 に示したオリジナルルータに対して、図 3 のように、モード切替え信号と呼ばれる 1~3 サイクルパイプラインを識別するための 2 ビットの入力を持たせることで外部からのパイプライン段数の制御を可能にし、1 サイクルから 3 サイクルルータ間の動的な切替えを行う。モード切替え信号は、プロセッサコアもしくはネットワークインタフェースから設定されることを想定している。

可変パイプラインルータで、入力チャンネルと出力チャンネルはすべてのモードで使われる共通回路を持つ。さらに、入力チャンネルは 1 サイクル専用回路 In1 および 2, 3 サイクル専用回路 In2 を、出力チャンネルは 1, 2 サイクル専用回路 Out1 および 3 サイクル専用回路 Out2 を持つ。そして、可変パイプラインルータはモード切替え信号によって指定されたモードに応じて使用する回路を切り替えて動作する。これにより、たとえばモード切替え信号が 3 サイクルモードを示しているときに入力チャンネルでは共通回路と 2, 3 サイクル専用回路 In2 が使われ、出力チャンネルでは共通回路と 3 サイクル専用回路 Out2 が使われる。

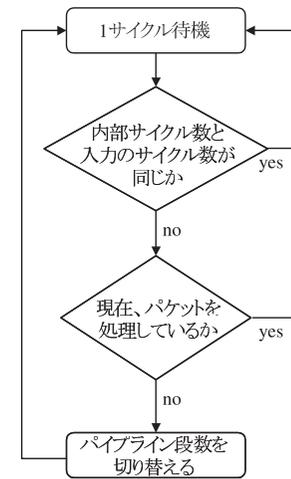


図 4 可変パイプラインルータのサイクル数切替えフローチャート
Fig. 4 Switching of the cycle number.

専用回路については 3.4 節で解説する。

可変パイプラインルータは、パイプライン段数の変更にも備えモード切替え信号の値をつねに監視している。動作サイクル数切替えの処理の流れを示すフローチャートを図 4 に示す。ルータがモード切替え信号の変化を検知した瞬間に段数を切り替えると、誤作動を招く原因となる。たとえばパケット処理中に 2 サイクルから 1 サイクルモードに切り替えると、フリットを格納する入力ポートにあるチャンネルバッファに溜まっているフリットが処理されずにデータの欠損が生じる。これを回避するため、ルータはモード切替え信号の変化を検知してパケットを処理していない場合にのみ段数を切り替える。可変パイプラインルータは、入出力の仕様がサイクルモードによらない。よって、チップ内の複数のルータのサイクルモードが各々異なっても、パケットの転送は正しく行われる。

ルータの動作サイクル数の切替えは、フリットが通るルータの内部データパスを切り替える処理であるため、そのルータがパケットを処理していない場合には 1 サイクルで切替えが完了する。ルータがパケットを処理していた場合は、そのルータがパケット中の最後のフリットの処理を終えた次のサイクルにデータパスの切替えが行われる。

3.4 回路構成

図 3 に示したとおり、パイプライン構成を更新する場合、入力チャンネルと出力チャンネルの

2つのモジュールで回路の構成が変化する。

3.4.1 入力チャンネルにおける専用回路

入力チャンネルは、1 サイクル専用回路 In1、2、3 サイクル専用回路 In2 および共通回路を持つ。

In1 は、VSA ステージに RC の機構を組み込む。このため、In1 では、VSA ステージに RC の結果が返ってきたことを示す信号線 rc_return を追加する。また、In2 ではパケットのヘッダフリットの到着を受けて状態レジスタを書き換えることにより、VSA ステージに遷移する。しかし、In1 はヘッダフリットの到着と同時に VSA を開始しなければならないため、状態レジスタを使用することができない。そこで、ヘッダフリットの到着を示す信号線 head_arrival を追加する。

最終的に、In1 では、rc_return、head_arrival と VA、SA の要求線との論理積を入力チャンネルから出力する VA、SA 要求線とする。また、テールフリット処理後に VSA ステージに戻る処理を組み込む。In2 には、RC のみ行うステージと、テールフリット処理後に RC ステージに戻る処理を組み込む。

3.4.2 出力チャンネルにおける専用回路

出力チャンネルは、1、2 サイクル専用回路 Out1、3 サイクル専用回路 Out2 および共通回路を持つ。

Out2 では出力チャンネルにおいてバッファリングを行う。バッファリングにより、フリットの転送にかかるサイクル数に 1 サイクルが加算されるが、2つのルータの入力チャンネルどうしを結び長いクリティカルパスを短縮することができる。

Out1 において出力チャンネルは出力バッファを持たず、クロスバを通過してきたデータは次のサイクルを待たずに直接次のルータの入力チャンネルへと転送される。これによりクリティカルパスが伸び、最大動作周波数の低下を招くが、フリットの転送にかかるサイクルを 1 サイクル節約することができる。

Out1 を用いる 1、2 サイクルモードでは VSA と ST を同時に行うため、出力バッファを使用しない。このため、Out1 は、Out2 に存在する (1) 出力先の仮想チャンネルを格納するレジスタ、(2) 入力チャンネルにデータの出力を指示する値を持つレジスタ、(3) 出力データを格納するレジスタ、と同じ役割を持つ信号線を追加する。

4. 評価

この章ではまず、Verilog を用いて実装を行った可変パイプラインルータの面積、動作周

波数の結果を示す。次に、フリットレベルシミュレーションによりスループットを測定し、その結果から消費電力を算出する。

本評価において、可変パイプラインルータは軽量のオンチップ 3 サイクルルータ⁷⁾を前章の設計により拡張することで実現した。本ルータはスイッチング技術としてワームホールを採用し、フリットはデータ部が 64 ビット幅であり、ヘッダ/ボディ/テールを識別する 2 ビットの制御部を加えた計 66 ビットで構成される。入力チャンネルが 5 フリット分のバッファ、出力チャンネルは 1 フリット分のバッファを持つ。ただし、パイプライン構成によってはこれらはバイパスされる。ルーティング実装は、ソースルーティング法を仮定し、ネットワークサイズに大きさが依存するルーティングテーブルは持たない。また、仮想チャンネル数は 4 本とした。これにより、パワーゲーティングによって使用可能なリンク数が増える省電力チップ内ネットワークにおいても、任意のトポロジにおいて適用できる様々な最短期デッドロックフリールーティング¹¹⁾の実装が可能となる。

4.1 面積、動作周波数と電圧

本ルータは Synopsys 社 Design Compiler 2007.12-SP3 を使って合成を行った。合成には供給電圧 1.2 V、温度 25 度の 65 nm CMOS プロセスライブラリを使用した。

1 サイクルルータ、2 サイクルルータ、3 サイクルルータ、可変パイプラインルータの面積と、モジュールごとの内訳を図 5 に示す。

可変パイプラインルータと異なり、1、2、3 サイクルルータはパイプライン構成の変更を行うことがないため、図 2 に示した各々のパイプライン構成に特化した設計となっている。図 5 より、いずれのルータも入力チャンネルが面積の多くを占めていることが分かる。ルー

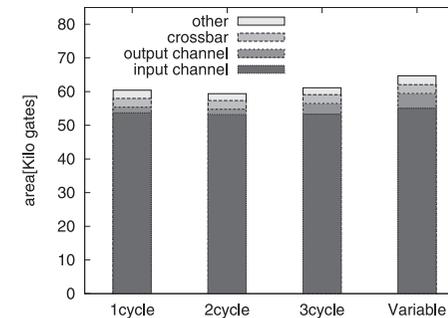


図 5 各ルータの面積と内訳
Fig. 5 Hardware amount and breakdown.

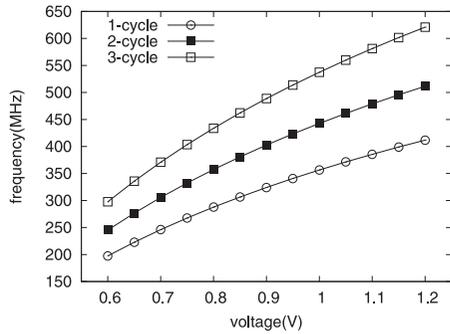


図 6 供給電圧に対する最大動作周波数 (MHz)
Fig. 6 Maximum operating frequency (MHz).

タ間のレジスタ利用の有無などにより、細かな面積の違いはあるが、オリジナルの 3 サイクルルータと改造後の可変パイプラインルータの面積の差は 5.8%程度であった。

次に、1 サイクルモード、2 サイクルモード、3 サイクルモードの可変パイプラインルータに対して、供給電圧を 1.20 V から 0.60 V まで変化させたときの最大動作周波数を図 6 に示す。また、電圧変更時の最大動作周波数の導出には以下の式を用いた。

$$T_{\text{delay}} \propto \frac{CV}{(V - V_{\text{th}})^\alpha} \quad (1)$$

ただし、 T_{delay} はゲート遅延、 C はスイッチング容量、 V は供給電圧、 V_{th} は閾値電圧、 α はキャリアの速度飽和パラメータ¹²⁾である。文献 13) より $\alpha = 1.3$ とした。また、Predictive Technology Model が提供している 65 nm のモデル¹⁴⁾ より $V_{\text{th}} = 0.4$ とした。

動作周波数に関して、1.2 V の供給電圧を与えたときの最大動作周波数は、1 サイクルモードは 411 MHz、2 サイクルモードは 511 MHz、3 サイクルモードは 614 MHz であった。

4.2 スループット

本節では、C++ で記述されたフリットレベルネットワークシミュレータ irr_sim⁷⁾ を用いて 16、64 コアのチップ内ネットワークのスループットと遅延の評価を行う。

図 7、図 8 に、1 サイクルモードから 4 サイクルモードでの遅延と accepted traffic [flit/cycle/core] を、各図のキャプション末尾の大括弧内にそのアプリケーションにおける平均ホップ数を示す。

遅延とは、出発地のコアがパケットを生成してから目的地のコアが受信するまでのサイ

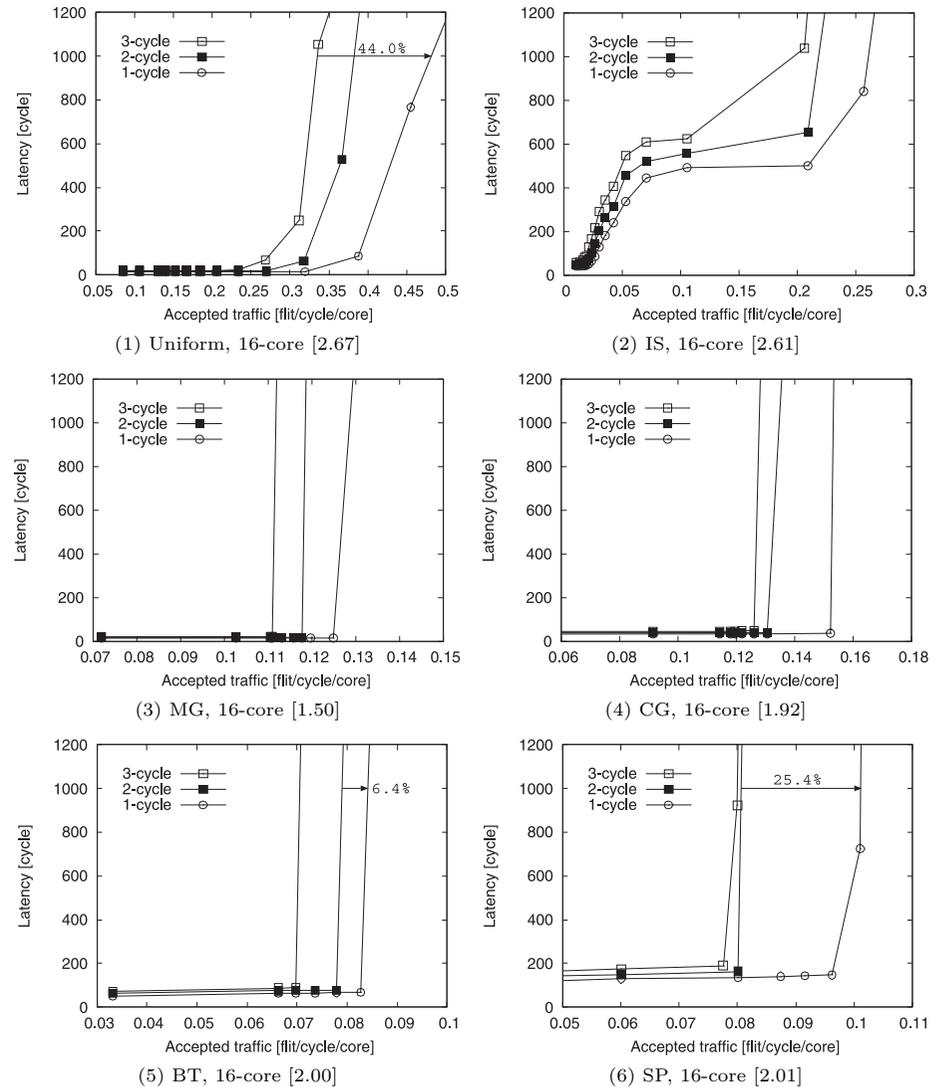


図 7 スループット (A)
Fig. 7 Throughput (A).

77 パイプラインステージ統合による省電力・可変パイプラインルータに関する研究

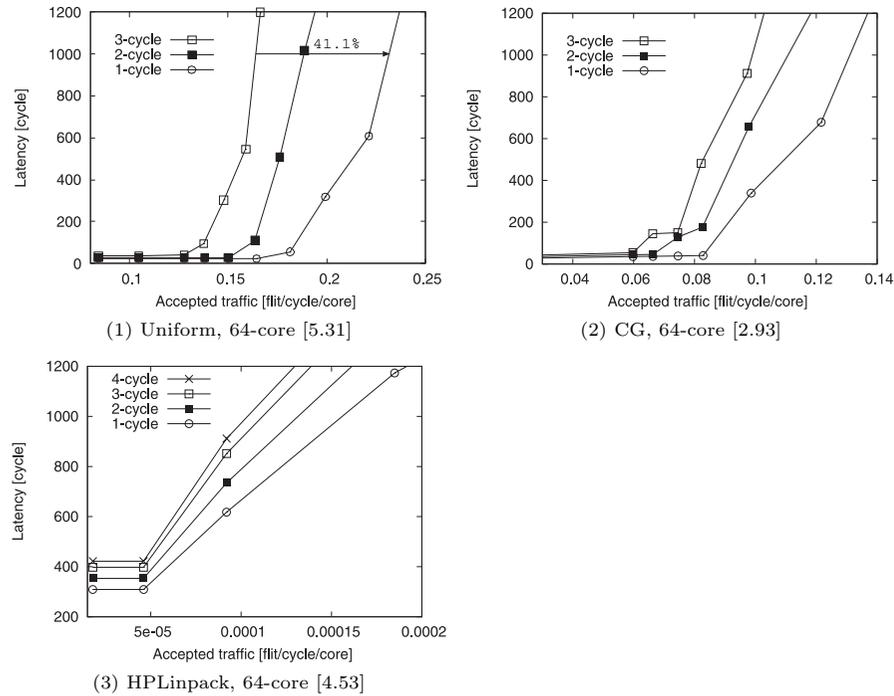


図 8 スループット (B)
Fig. 8 Throughput (B).

クル数であり、ネットワーク滞在時間だけでなく、ネットワークへの注入時間も含まれる。また、accepted traffic とは、各コアが 1 サイクルあたりに受信する平均フリット数であり、通信遅延が 1,000 サイクル未満の場合における accepted traffic の最大値をネットワークのスループットとした。これはシミュレーション結果を解析した結果、遅延が 1,000 サイクル以上の場合に最大の accepted traffic が生じたが、この場合にはネットワークが飽和していることを確認したため、これらを除くために 1,000 サイクルという上限を設けた。

ルータはパケット転送方式にウォームホールを用いるなど、Verilog で記述したルータとほぼ同じアーキテクチャを採用した。また、トポロジは 4×4 あるいは 8×8 の 2 次元メッシュである。トラフィックパターンはパケット長を 7 フリットとした Uniform traffic (16 コア, 64 コア) と HPLinpack (64 コア)(以下 HPL と呼ぶ), NAS Parallel Benchmarks

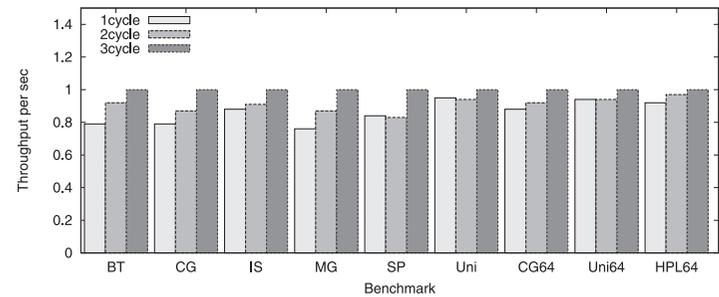


図 9 供給電圧 1.2 V における時間あたりの正規化スループット
Fig. 9 Normalized throughput per second (1.2 V power supply).

Class W (以下 NPB と呼ぶ) より IS, MG, CG, BT, SP (いずれも 16 コア, CG のみ 16 コアと 64 コア), の計 9 つのトレースを用いた。NPB のトレースは文献 7) と同様のものを用い、HPL については同志社大学の PC クラスタにおいて MPICH1.2.7p1, 数値演算ライブラリ GotoBLAS1.22 を用いて実行したトレースであり、パラメータは文献 15) にまとめられている。HPL は ClearSpeed CSX600 などの大規模チップ内ネットワークを採用した実システムで実装されているが、チップ内ネットワークで実行した際のトレースが入手できないため PC クラスタを用いたトレースで代替した。

IS はバースト的に全対全通信があるため、アプリケーションの性能が通信のレイテンシおよびスループット性能によって相対的に大きな影響を受ける。一方, CG, BT, SP は全対全通信ではなく隣接間通信が多く、通信のピークは低い。CG では定常的な通信が起こる。また, MG も IS にくらべると通信のピークが低い。一方 HPL トレースではアプリケーション開始時と終了時にデータ分配, 収集のための大規模なデータ転送が生じている。

サイクルあたりのスループットでは、コア数によらず、1 サイクルルータを用いたチップ内ネットワークが最高のスループットを達成している。これは、フリットをルータあたり 1 サイクルで転送できるため、通信遅延が削減され、より多くのパケットをネットワークに注入することができるようになるためである。

次に、これらのサイクルあたりのスループットの結果から、時間あたりの正規化スループットを算出した結果を図 9 に示す。動作周波数は 4.1 節で示した供給電圧 1.2 V における最大動作周波数を使用した。

図 9 より、1 サイクルモードの時間あたりのスループットが総じて低いことが分かる。1 サイクルモードは 1 サイクル中に処理できるパケットの量は多いが、最大動作周波数が低い。

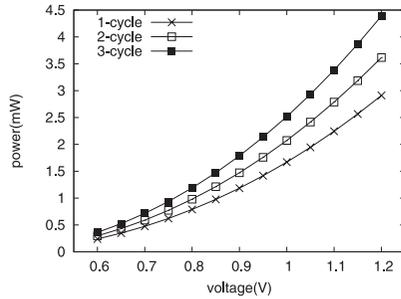


図 10 供給電圧に対する消費電力

Fig. 10 Power consumption versus supply voltage.

このため、NPB, HPL においては時間あたりのスループットが最も低い。反面, Uniform traffic では 1 サイクル, 2 サイクル, 3 サイクルでの時間あたりのスループット差が非常に小さい。このように, アプリケーションごとに特性は異なるが, おおむねサイクル数が多いモードほど時間あたりのスループット性能が高いことが分かる。

4.3 消費電力

図 10 に, 1 サイクルモードから 3 サイクルモードでの供給電圧ごとの消費電力を示す。本研究においてはリーク電力については考慮せず, ダイナミック電力に焦点をあてる。ダイナミック電力の導出には式 (2) を用いた。

$$P_{sw} = \alpha \times C \times f \times V^2 \quad (2)$$

ただし, α は回路の活性化率, C はキャパシタンス, f は動作周波数である。供給電圧が 1.2V のときの P_{sw} は回路シミュレーションにより取得し, C と α は一定であることを利用して, V と各 V における最大動作周波数 f を式 (2) に代入することにより消費電力を算出した。

次に, 各時間あたりのスループットを達成するために必要となる電力をアプリケーションごとに算出するために, (1) 図 7, 図 8 に示したサイクルあたりのスループットと, (2) 図 6 に示した動作周波数, 電圧を用いた。時間あたりのスループットは, サイクルあたりにスループットにそのサイクルのルータの動作周波数を掛けあわせて算出した。1 サイクル, 2 サイクル, 3 サイクルの各モードにおいてあるスループットを満たすために必要な消費電力を図 11, 図 12 に示す。

これらの結果から, 各モードにおいてスループットが減少するにつれて, 消費電力が減少

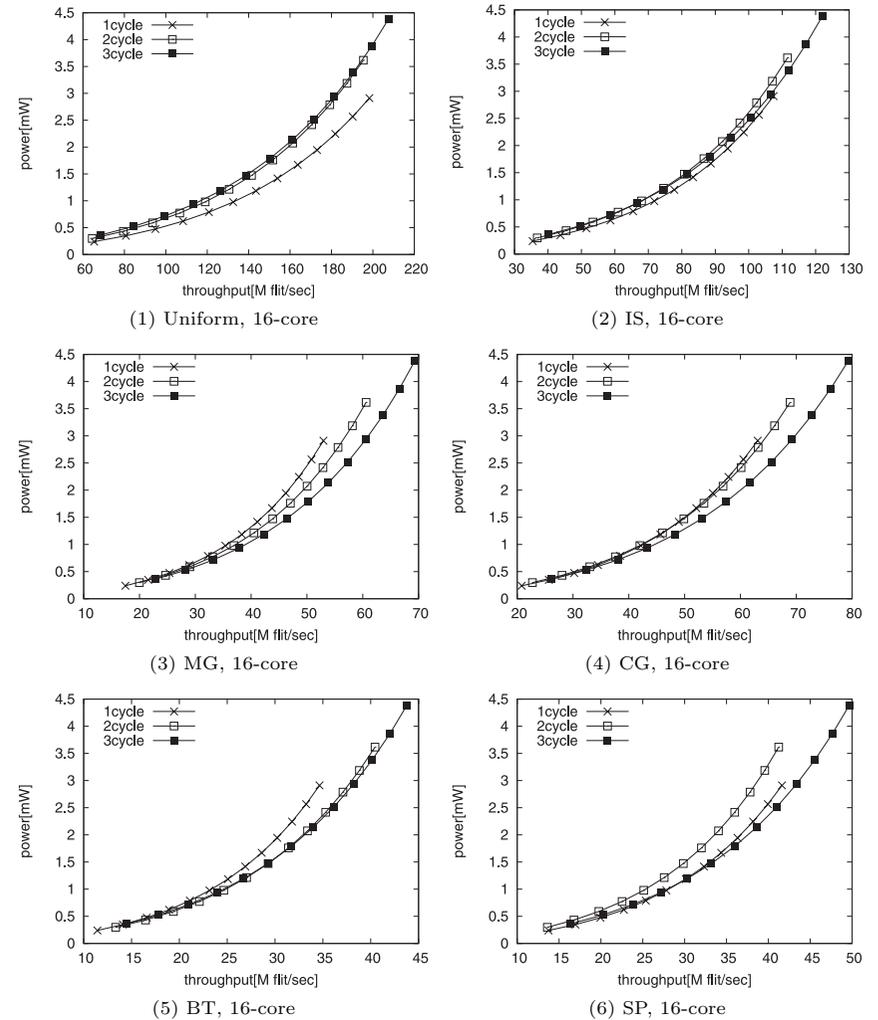


図 11 消費電力 (A)

Fig. 11 Power consumption (A).

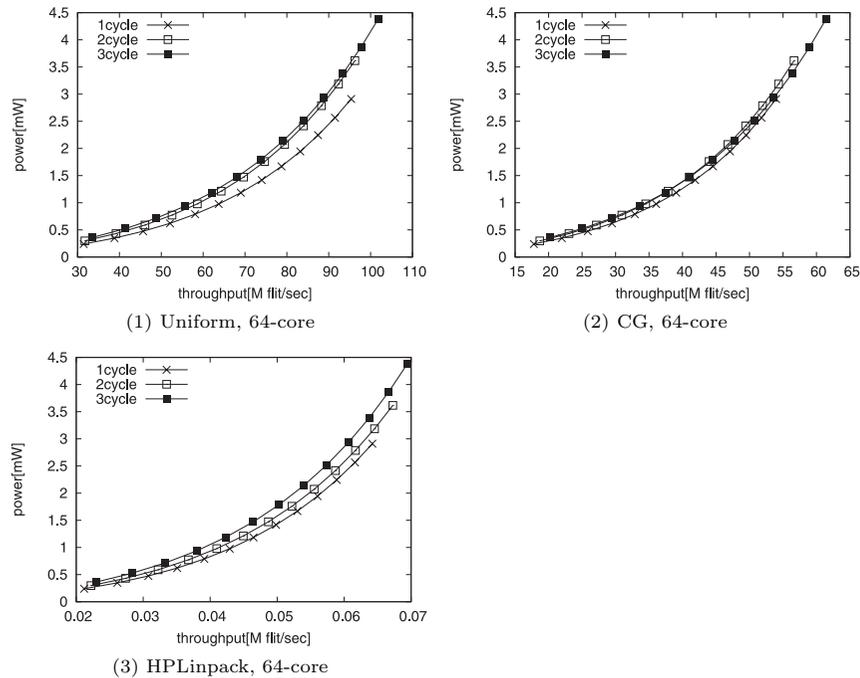


図 12 消費電力 (B)
Fig. 12 Power consumption (B).

していくことが分かる．この主な原因は動作周波数と電圧の低下による．この消費電力の減少は，可変パイプラインでないルータに対して DVFS を使った場合と同じ結果であり，スループットを犠牲にしたときの消費電力の低下を読み取ることができる．またモード間の差に着目すると，たとえば 16 コアの MG トラフィックパターンにおいて必要となるスループットが 20 [M flit/sec] から 50 [M flit/sec] に変化した場合，1 サイクルモードの場合，消費電力は 698% の増加となる．一方，可変パイプラインルータでは，同様の場合，3 サイクルモードに変更することにより消費電力の増加を 477% に抑えることができる．つまり，単純にチップ内ネットワークに DVFS を適用する場合に比べ，可変パイプラインルータを用いることで消費電力の増加を最大 46% ，抑えることができる．

また，Uniform traffic では 1 サイクルモードと 3 サイクルモードのサイクルあたりスルー

プットの差が図 7 (1) ，図 8 (1) に示したとおり 16 コアでは 44.0% ，64 コアでは 41.1% と大きく，最大動作周波数を考慮して時間あたりスループットに換算した際に，1 サイクルモードの電力効率が最も良いという結果が得られた．

一方，NAS Parallel Benchmark では 3 サイクルモードが最大のスループットを達成し，さらに多くのスループットにおいて最小の消費電力を達成している．BT と SP において同じスループットのときの 1 サイクルモードと 2 サイクルモードの消費電力の大小関係が逆転しているが，これは図 7 (5) ，(6) において BT の 1 サイクルモードと 2 サイクルモードのサイクルあたりスループットの差が 6.4% であるのに対し，SP ではこの差が 25.4% と大きく開いているために，最大動作周波数を用いて図 11 (5) ，(6) の時間あたりスループットに換算した際に消費電力の大小関係の入れ替わりが起こったといえる．

4.4 無負荷時レイテンシ

最後に，可変パイプラインルータにおけるパイプライン段数と無負荷時レイテンシについて考察する．無負荷時レイテンシとは，パケットの衝突が起きないときの通信レイテンシのことである．したがって，ウォームホールスイッチングの場合，無負荷時レイテンシ L_0 [sec] は次式で計算できる¹⁰⁾ ．

$$L_0 = (HT + l)D \quad (3)$$

ただし， H をホップ数， T をルータの転送遅延 (パイプライン段数) ， l をパケット長 (ボディフリット数) ， D を 1 サイクルの時間 [sec] とする．

各トラフィックパターンにおける H は図 7 ，図 8 に示したとおりである．また，1 サイクル，2 サイクル，3 サイクルルータの動作周波数を 411 MHz ，511 MHz ，614 MHz とすると， D はそれぞれ 2.43 nsec ，1.96 nsec ，1.63 nsec となる．本章のシミュレーションではパケット長は 7 フリット，うちボディ部分は 6 フリットとしたため $l = 6$ である．

上記のパラメータを用いて各トラフィックパターンにおける各ルータの無負荷時レイテンシを計算した．結果を図 13 に示す．1 サイクルモードと 3 サイクルモードを比べると，1 サイクルモードのほうが D が大きいものの T は $1/3$ である．そのため，図 13 に示すとおり，1 サイクルモードの無負荷時レイテンシは他のモードと同等かそれ以下である．ここで注目すべきは各パイプラインモードの動作周波数であり，1 サイクルモードは他と同等かそれ以下の無負荷時レイテンシを 411 MHz で実現している．式 (2) で示したとおり，消費電力は動作周波数に比例して増加するため，1 サイクルモードは 2 サイクルモードの 80.4% ，3 サイクルモードの 66.9% の消費電力でこれを実現できたことになる．

4.2 節および 4.3 節では，主にスループットに着目して可変パイプラインルータの有用性

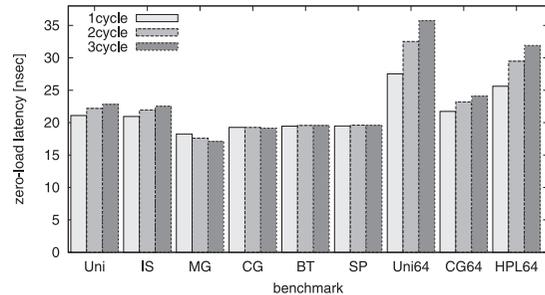


図 13 各ルータモードにおける無負荷時レイテンシ
Fig. 13 Zero-load latency of each router mode.

について議論してきたが、近年では無負荷時レイテンシを削減する研究も注目されている⁹⁾。このように無負荷時レイテンシの削減が重要となる場合は、1 サイクルモードを用いたほうが消費電力の点で有利となる。

4.5 既存の手法との比較

前節の評価結果より、可変パイプラインルータはパイプライン段数が固定されている通常のルータに比べて面積が 5.8% 増加したものの、既存のルータに DVFS を用いた場合に比べて消費電力の増加を最大 46% 抑えることができた。

2 章で述べたとおり、文献 4) では DVFS と適応型ルーティングを組み合わせることで消費電力を削減している。本論文では NoC 向けのシンプルな固定型ルーティングを想定しているが、文献 4) のように適応型ルーティングを用いると、出力チャネルの選択機構、パケットのリオーダ回路などハードウェアのオーバーヘッドが大きくなる。

DVFS 以外の省電力技術として、オンチップルータへの電力供給を部分的に止める方法が報告されている^{5),6)}。文献 5), 6) ではハードウェア量のオーバーヘッドについて十分に評価していないが、実際にはパワーゲーティングのためのパワースイッチやパワースイッチをドライブするためのロジックが必要となり、面積オーバーヘッドは無視できない¹⁶⁾。しかも、パワーゲーティング技術は主にルータのリーク電力削減を主眼としている。一方、DVFS における可変パイプラインルータでは動的電力および静的電力（リーク電力）の双方の削減を目的としている。実際、提案手法は、ネットワークの動作時において、トータルの消費電力の増加を最大 46% 抑えることができる点で、電力削減効果がきわめて大きいといえる。

ほかに、ネットワーク、もしくは、ルータ内部のデータパスを動的に変更することでレ

イテンシを削減する技術がいくつか報告されている^{8),9)}。文献 8) では、通信遅延が小さい軽量のネットワークを新たに追加することで通信遅延を削減しているため、面積の増加量は無視できない。文献 9) においては、非隣接ノード間に仮想的なバイパス経路を構築するため、追加のハードウェア量は小さいと考えられるが、非隣接ノード間にフロー制御が新たに必要となる。

5. まとめ

チップ内ネットワークは、パケットをパイプライン処理するルータを多数用いることで高スケーラビリティ、高スループットを実現している。本論文では、性能、あるいは消費電力の面で最適なルータを実現するために、動作周波数とパイプライン構成とをトラフィックパターンに応じて切り替えることが可能な可変パイプラインルータを提案した。各動作周波数におけるパイプラインサイクルあたりの処理量を大きくするために、低い動作周波数では、単純なカスケードによるパイプラインステージの統合ではなく、各ステージ処理の並列性を最大限に利用する投機実行を行う。

合成トラフィックである Uniform traffic や並列アプリケーションの NPB, HPL のトレースを用いた評価結果より、多くの場合最大スループットが得られたのは 3 サイクルパイプラインを使用した場合であった。一方、性能、電力面で最適なパイプライン構成は 1 段から 3 段とアプリケーションごとに異なることが分かった。

可変パイプラインルータはパイプラインを切り替えるオーバーヘッドが最小で 1 サイクルであり、通常のパイプラインが固定されているルータに比べて面積が 5.8% 増加する。しかしパイプライン構成を変更しない既存のルータに DVFS を用いた場合に比べ、提案ルータを用いることで消費電力の増加を最大 46% も抑えることができることが分かった。よって、これらのトレードオフから様々なトラフィックパターンを扱うメニーコア/マルチコアシステムのチップ内ネットワークにおいて、可変パイプラインルータは性能、電力面できわめて有効であるといえる。

また、可変パイプラインルータはスイッチのポートを Power Gating により On/Off する手法⁵⁾⁻⁷⁾ などの既存の省電力手法と共存でき、省電力チップ内ネットワークを実現するための 1 つの要素技術となるものである。

今後はルータごとに動作周波数、パイプライン構造を変更することで、チップ内ネットワークのさらなる性能、電力面の最適化を行う予定である。さらに、パイプライン構成を切り替えるタイミングやそのためのトラフィック負荷の測定法を開発し、可変パイプライン

ルータをさらに実用的なものとする予定である。

謝辞 本研究は東京大学大規模集積システム設計教育研究センターを通じ、株式会社半導体理工学研究センター、(株)イー・シャトル、富士通株式会社の協力で行われた。

参 考 文 献

- 1) Dally, W.J. and Towles, B.: Route Packets, Not Wires: On-Chip Interconnection Networks, *Proc. Design Automation Conference (DAC'01)*, pp.684–689 (2001).
- 2) Peh, L.-S. and Dally, W.J.: A Delay Model and Speculative Architecture for Pipelined Routers, *Proc. International Symposium on High-Performance Computer Architecture (HPCA'01)*, pp.255–266 (2001).
- 3) 嶋田 創, 安藤秀樹, 島田俊夫: パイプラインステージ統合と DVS の併用による消費電力の削減, *情報処理学会論文誌*, Vol.48, No.3, pp.75–87 (2007).
- 4) Stine, J.M. and Carter, N.P.: Comparing Adaptive Routing and Dynamic Voltage Scaling for Link Power Reduction, *IEEE Computer Architecture Letters*, Vol.3, No.1, pp.14–17 (2004).
- 5) Soteriou, V. and Peh, L.-S.: Exploring the Design Space of Self-Regulating Power-Aware On/Off Interconnection Networks, *IEEE Trans. Parallel and Distributed Systems*, Vol.18, No.3, pp.393–408 (2007).
- 6) Matsutani, H., Koibuchi, M., Wang, D. and Amano, H.: Run-Time Power Gating of On-Chip Routers Using Look-Ahead Routing, *Proc. Asia and South Pacific Design Automation Conference (ASP-DAC'08)*, pp.55–60 (2008).
- 7) Matsutani, H., Koibuchi, M., Wang, D. and Amano, H.: Adding Slow-Silent Virtual Channels for Low-Power On-Chip Networks, *Proc. International Symposium on Networks-on-Chip (NOCS'08)*, pp.23–32 (2008).
- 8) Koibuchi, M., Matsutani, H., Amano, H. and Pinkston, T.M.: A Lightweight Fault-tolerant Mechanism for Network-on-Chip, *Proc. International Symposium on Networks-on-Chip (NOCS'08)*, pp.13–22 (2008).
- 9) Kumar, A., Peh, L.-S., Kundu, P. and Jha, N.K.: Express Virtual Channels: Towards the Ideal Interconnection Fabric, *Proc. International Symposium on Computer Architecture (ISCA'07)*, pp.150–161 (2007).
- 10) Dally, W.J. and Towles, B.: *Principles and Practices of Interconnection Networks*, Morgan Kaufmann (2004).
- 11) Jouraku, A., Koibuchi, M. and Amano, H.: An Effective Design of Deadlock-Free Routing Algorithms Based on 2-D Turn Model for Irregular Networks, *IEEE Trans. Parallel and Distributed Systems*, Vol.18, No.3, pp.320–333 (2007).
- 12) Sakurai, T. and Newton, A.R.: Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas, *IEEE Journal of Solid-State Circuits*, Vol.25, No.2, pp.584–594 (1990).
- 13) Khellah, M.M. and Elmasry, M.I.: Power Minimization of High-Performance Sub-micron CMOS Circuits Using a Dual-Vd Dual-Vth (DVDV) Approach, *Proc. International Symposium on Low Power Electronics and Design*, pp.106–108 (1999).
- 14) Predictive Technology Model: 65 nm BSIM4 model card for bulk CMOS V1.0. <http://www.eas.asu.edu/~ptm/>
- 15) 廣安知之, 渡辺崇文, 中尾昌広, 大塚智宏, 鯉淵道紘: PC クラスタにおける VLAN イーサネットのトポロジの評価, 先進的計算基盤システムシンポジウム (SACIS'09) 論文集, pp.283–291 (2009).
- 16) Seki, N., Zhao, L., Kei, J., Ikebuchi, D., Kojima, Y., Hasegawa, Y., Amano, H., Kashima, T., Takeda, S., Shirai, T., Nakata, M., Usami, K., Sunata, T., Kanai, J., Namiki, M., Kondo, M. and Nakamura, H.: A Fine-Grain Dynamic Sleep Control Scheme in MIPS R3000, *Proc. International Conference on Computer Design*, pp.612–617 (2008).

(平成 21 年 1 月 27 日受付)

(平成 21 年 4 月 27 日採録)



枚田 優人 (学生会員)

平成 21 年慶應義塾大学理工学部情報工学科卒業。現在、同大学大学院理工学研究科開放環境科学専攻修士課程。オンチップネットワークにおけるルーティングの研究に従事。



松谷 宏紀 (正会員)

平成 16 年慶應義塾大学環境情報学部卒業。平成 20 年同大学大学院理工学研究科開放環境科学専攻博士課程修了。博士 (工学)。現在、東京大学大学院情報理工学系研究科特別研究員。平成 21 年度より日本学術振興会特別研究員 SPD。オンチップネットワークの研究に従事。



鯉淵 道紘（正会員）

平成 12 年慶應義塾大学工学部情報工学科卒業．平成 15 年同大学大学院理工学研究科開放環境科学専攻博士課程修了．博士（工学）．平成 14 年度より 16 年度まで日本学術振興会特別研究員．現在，国立情報学研究所助教，総合研究大学院大学複合科学研究科情報学専攻助教（兼任）．ハイパフォーマンスコンピューティングとインターコネクトに関する研究に従事．IEEE Computer Society Japan Chapter Young Author Award 2007，平成 19 年度情報処理学会論文賞受賞．IEEE，電子情報通信学会各会員．



天野 英晴（正会員）

昭和 56 年慶應義塾大学工学部電気工学科卒業．昭和 61 年同大学大学院理工学研究科電気工学専攻博士課程修了．工学博士．現在，慶應義塾大学工学部情報工学科教授．計算機アーキテクチャの研究に従事．