

## NoC のための多電源可変パイプラインルータ

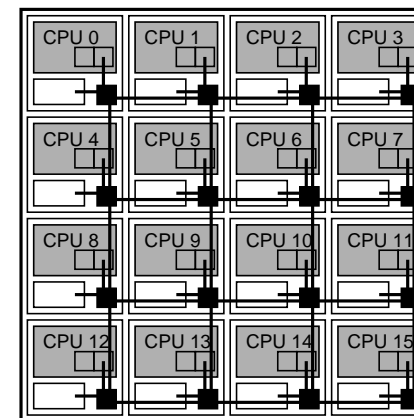
松谷 宏 紀<sup>†1,†2</sup> 枚田 優 人<sup>†3</sup> 鯉 淵 道 紘<sup>†4</sup>  
宇佐美 公 良<sup>†5</sup> 中 村 宏<sup>†1</sup> 天 野 英 晴<sup>†3</sup>

本論文では、メニーコアにおけるプロセッサ-キャッシュ間ネットワーク (Network-on-Chip, NoC) を低消費電力かつ高信頼化するために多電源可変パイプラインルータを提案する。多電源可変パイプラインルータでは、トラフィック負荷に応じて個々のルータのパイプライン段数 (最大遅延) と供給電圧を切り替える。Dynamic voltage and frequency scaling (DVFS) ルータとは異なり、すべてのルータは同じ動作周波数で動作するため、異なる周波数ドメイン間の同期を考慮する必要がないというメリットがある。本論文では、2 電源 2 モード切り替え可能な多電源可変パイプラインルータを 65nm プロセスを用いて設計し、さらに CMP シミュレータ上で評価した。その結果、アプリケーション性能が 1.0% ~ 2.1% 落ちたが、NoC のスタンバイ電力を 10.4% ~ 44.4% 削減できた。

## A Multi-Vdd Variable-Pipeline Router for NoCs

HIROKI MATSUTANI,<sup>†1,†2</sup> YUTO HIRATA,<sup>†3</sup>  
MICHIMIRO KOIBUCHI,<sup>†4</sup> KIMIYOSHI USAMI,<sup>†5</sup>  
HIROSHI NAKAMURA<sup>†1</sup> and HIDEHARU AMANO<sup>†3</sup>

We propose a Multi-Vdd variable pipeline (MVP) router in order to reduce the power consumption and improve the reliability of Network-on-Chip (NoC) designed for many-core processors. MVP router adjusts its pipeline depth (i.e., communication latency) and supply voltage level in response to the applied workload. Unlike Dynamic voltage and frequency scaling (DVFS) routers, all MVP routers share the same operating frequency, and thus there is no need to synchronize neighboring routers working at different frequencies. In this paper, we implemented a MVP router that can select two supply voltage levels (and pipeline modes) using a 65nm process, and we evaluated it using a full-system CMP simulator. Evaluation results show that although the application performance is degraded by 1.0% to 2.1%, the standby power of NoC is reduced by 10.4% to 44.4%.



■ L1 D/I cache □ L2 cache bank ■ On-chip router

図 1 16 コア CMP の例。チップは 16 個のタイルに分割され、各タイルはプロセッサ (CPU)、L1 キャッシュ、L2 キャッシュバンクから成る。L2 キャッシュはタイル間で共有する場合と共有しない場合が考えられる。

### 1. はじめに

半導体技術の微細化にともない 1 チップ上に複数のマイクロプロセッサを実装できるようになった。コンシューマ用途においても 2 コアや 4 コアのチップマルチプロセッサ (CMP) が広く普及しており、コアの数は今後も増え続けると予想される。図 1 に 16 コア CMP の例を示す。このような CMP では、プロセッサやキャッシュを Network-on-Chip (NoC)<sup>1)</sup> で接続し、データ転送はオンチップルータを介したパケット転送によって行う。

<sup>†1</sup> 東京大学大学院 情報理工学系研究科

Graduate School of Information Science and Technology, The University of Tokyo

<sup>†2</sup> 日本学術振興会特別研究員 (SPD)

Research Fellow of the Japan Society for the Promotion of Science (SPD)

<sup>†3</sup> 慶應義塾大学大学院 理工学研究所

Graduate School of Science and Technology, Keio University

<sup>†4</sup> 国立情報学研究所 / 総合研究大学院大学

National Institute of Informatics / The Graduate University for Advanced Studies

<sup>†5</sup> 芝浦工業大学 情報工学科

Department of Information Science and Engineering, Shibaura Institute of Technology

近年はほぼすべての用途において低消費電力化が求められており、NoC やそこで使われるオンチップルータも例外ではない。プロセッサの低消費電力化技術としては Dynamic voltage and frequency scaling (DVFS) が古くから使われており、オンチップルータにも一部応用されてきた<sup>2)</sup>。しかし、DVFS では動作周波数を変えるため、チップ内に複数の電源ドメインがある場合、隣接する 2 個のドメインの動作周波数を 1:k (k は整数) にしなければならず、選択可能な周波数 (電圧) のレンジに制限が生じる。さもなければ、電源ドメイン間で非同期通信が必要になったり、ドメイン間の周波数の違いを吸収するための FIFO バッファが必要になる。とりわけ、NoC では通信負荷に偏りが生じやすく、ルータ単位で独立に電圧および周波数を調整することが望ましいため、既存の DVFS 手法は向かない。

そこで、本研究では、個々のルータの動作周波数ではなく、個々のルータのパイプライン段数 (最大遅延) を切り替えることで DVFS と同様の効果を実現する。そのために、多電源可変パイプラインルータを提案する。多電源可変パイプラインルータでは、負荷が高いときはパイプライン段数を浅くすることで通信サイクル数を減らし高性能を狙う。パイプラインを浅くするとクリティカルパス遅延が延びるため、高い電圧で駆動する必要がある。一方、負荷が低いときはパイプラインを深くして低い電圧で駆動する。このように動的にパイプライン構成と電圧を変えるが、動作周波数は一定とする点が、これまでの可変パイプラインルータ<sup>3)</sup> との違いである。

さらに、多電源可変パイプラインルータは、チップ製造時もしくは温度変化によるタイミングばらつきにも対処できる。具体的には、プロセッサの発熱等で高温になった場合あるいはタイミング違反を検出した場合、ルータのパイプライン段数を深くすることでタイミングマージンを稼ぎタイミング違反を回避する。

本論文の構成は以下のとおりである。2 章で多電源可変パイプラインルータを設計、実装する。3 章でパイプラインおよび電圧の切り替えポリシーを 2 種類提案する。4 章で多電源可変パイプラインルータの回路レベルの評価、CMP シミュレータを用いたシステムレベルの評価を示す。最後に 5 章で本論文をまとめる。

## 2. 多電源可変パイプラインルータの設計

図 2 に多電源可変パイプラインルータの概要を示す。以降、多電源可変パイプラインルータの設計および実装について説明していく。

### 2.1 ベースラインルータ

ここでは、データ (フリット) 幅 64-bit のワームホールルータを用いる。物理チャンネル数

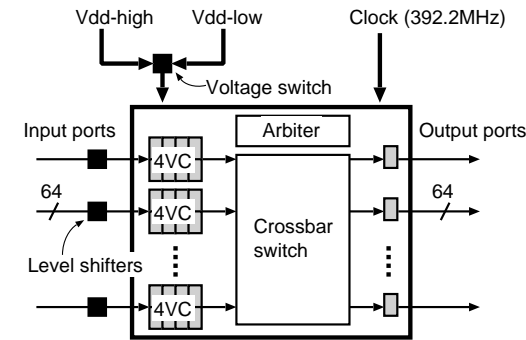


図 2 多電源可変パイプラインルータ。

表 1 切り替え可能なパイプラインモード。

Mode	Pipeline	Vdd	Freq
2-cycle	[RC/VA/SA] [ST,LT]	1.20V	392.2MHz
3-cycle	[RC/VA/SA] [ST] [LT]	0.83V	392.2MHz

5 本、仮想チャンネル数 4 本、仮想チャンネルごとに 4-flit 分の入力バッファを持つものとした。

ルータ内のパケット処理は、1) 経路計算を行う routing computation (RC)、2) 仮想チャンネルの割り当てを行う virtual-channel allocation (VA)、3) 出力ポートの割り当てを行う switch allocation (SA)、4) クロスバ上のフリット転送を行う switch traversal (ST) の 4 ステージに分割される。さらに、5) リンク上のフリットを転送を行う link traversal (LT) に 1 サイクル必要となる。LT を含め、以上の 5 ステージから成るルータをここでは 5-cycle ルータと呼ぶことにする。

### 2.2 可変パイプライン機構

可変パイプラインルータでは、表 1 に示す 2-cycle mode と 3-cycle mode を 1 サイクルで切り替えることができる。表 1 において、括弧は 1 サイクルの処理「/」は並列実行、「,」は逐次実行を表す。VA と SA の投機実行<sup>4)</sup>により両者は並列実行できる ([VA/SA])。また、先読みルーティング<sup>4)</sup>を行うことで RC と VA/SA をオーバーラップさせることができる ([RC/VA/SA])。2-cycle mode では ST を LT にマージして逐次実行しているため、[ST,LT] がクリティカルパスとなり、最大動作周波数が落ちる。

### 2.3 パイプラインモード

上述の 2 段可変パイプラインルータを Verilog-HDL で設計し、Fujitsu 65nm プロセス

を用いて Synopsys Design Compiler で合成, Synopsys IC Compiler で配置配線した。配置配線後の静的タイミング解析の結果, 電源電圧が 1.2V のとき 3-cycle mode の最大遅延は 1830.8 psec, 2-cycle mode の最大遅延は 2549.7 psec となった。

提案ルータでは, 動作周波数は変えずに, 負荷に応じてパイプライン段数と供給電圧を切り替える。2-cycle mode を 1.2V (Vdd-high) で 392.2MHz で動作させるとき, 3-cycle mode は 0.83V (Vdd-low) まで落としても 392.2MHz で動作可能である。切り替え可能なパイプラインモードを表 1 にまとめる。

パイプラインモード切り替えは以下の手順で行う。

- 3-cycle  $\rightarrow$  2-cycle : Vdd-low から Vdd-high に電圧が上がったのち, 2-cycle mode に切り替える。
- 2-cycle  $\rightarrow$  3-cycle : 3-cycle mode に切り替えたのち, Vdd-high から Vdd-low に電圧を下げる。

この手順に従わないと, 切り替え時にタイミング違反が生じ, 通信データの値化け等が生じる恐れがある。なお, 電圧レベルの遷移に要す時間については 4.1.2 節で評価する。

### 3. パイプライン切り替えポリシー

本章では, スタンバイ電力を減らすためのパイプライン切り替えポリシー (3.1 節), 温度変化などによるタイミング違反を防ぐパイプライン切り替えポリシー (3.2 節) を提案する。4 章の評価では, このうち前者のスタンバイ電力削減ポリシーについて CMP シミュレータを用いたシステムレベル評価を行う。

#### 3.1 スタンバイ電力削減ポリシー

DVFS では, 一般的に, 高負荷な状態では高い周波数, 高い電圧にしてスループットを高め, 低負荷な状態では低い周波数, 低い電圧にして消費電力を抑える。しかし, CMP 向け NoC では, 通信遅延が増えるとアプリケーションの実行時間が延びてしまううえに, NoC よりもプロセッサの消費電力のほうが圧倒的に大きいことが多い。そのため, アプリケーションの実行時間は延ばさないうえに, 可能な限り 3-cycle mode @ Vdd-low の状態を維持することが望ましい。

スタンバイ電力を減らすために以下のポリシーでパイプライン段数を切り替える。

- パケットを処理していない状態 (もしくは, 負荷が非常に低い状態) では, 可能な限り 3-cycle mode @ Vdd-low にしてスタンバイ電力を減らす。
- パケットを処理している状態では, 可能な限り 2-cycle mode @ Vdd-high にして通信

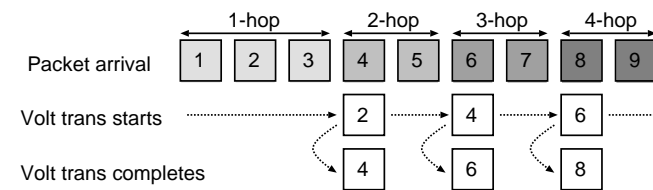


図 3 先読みルーティングによる昇圧スケジューリングの例。図中の数字はそのイベントが起きる時間 (サイクル) を示す。

遅延を減らす。

4.1 節の回路レベル評価で示すとおり, low-to-high の電圧切り替えに要す遅延は 392.2MHz 動作時で 2 サイクル程度である。つまり, パケットが到着する 2 サイクル前に電圧を low-to-high に切り替え, パケットが通過したら high-to-low に戻せばよい。このようなパケット到着の事前検出には先読みルーティングを応用できる<sup>5)</sup>。ただし, 多電源可変パイプラインルータの場合, low-to-high の昇圧が間に合わなかったとしてもパケットを (2-cycle mode ではなく) 3-cycle mode で転送するだけなので転送遅延のペナルティは高々 1 サイクルだけであり, パワーゲーティングルータ<sup>5)</sup> のように数サイクルの待ちが生じることはない。

文献 5) で指摘したように 2 ホップ目以降は先読みルーティングの応用で事前に昇圧できるが, 1 ホップ目に関しては事前予測ができない。そこで本論文では 1 ホップ目に関しては無理に昇圧せず, 3-cycle mode @ Vdd-low で転送し, 2 ホップ目以降は先読みルーティングを応用してパケット転送時のみ 2-cycle mode @ Vdd-high に切り替える。

図 3 を使って, 先読みルーティングによる昇圧スケジューリングを例示する。図中の数字はそのイベントが起きる時間 (サイクル) を示す。例えば, 1 サイクル目にパケットが 1-hop 目に到達している。同様に 4 サイクル目に 2-hop 目, 6 サイクル目に 3-hop 目, 8 サイクル目に 4-hop 目に到達している。1-hop 目のみ 3 サイクル転送で, それ以降は 2 サイクル転送となっている。先読みルーティングによって次ホップ (ルータ) が判明するので, 2 サイクル目に 2-hop 目の low-to-high が開始され, 実際にパケットが到着する 6 サイクル目までには昇圧および 2-cycle mode への切り替えが完了する。同様に, 4 サイクル目に 3-hop 目の low-to-high が開始され, 実際にパケットが到着する 8 サイクル目までには昇圧および 2-cycle mode への切り替えが完了する。

4 章の評価では, このスタンバイ電力削減ポリシーについて CMP シミュレータを用いたシステムレベル評価を行う。

### 3.2 タイミング違反回避ポリシ

半導体技術の微細化にともない、チップ製造時のタイミングばらつきが問題になっている。また、動作時のチップ温度によって遅延が変動することもあり、タイミング違反を起こす原因となっている。

多電源可変パイプラインルータの場合、電圧レベルは一定でパイプライン段数を 2-cycle mode から 3-cycle mode にすれば、それだけタイミングに余裕が生じるため、タイミング違反を回避できる。

タイミング違反を回避するには、以下に示すように、タイミング違反を未然に防ぐ方法、タイミング違反が生じてから対処する方法が考えられる。

- タイミング違反を事前に防ぐ方法： ルータ内に温度センサを実装し、温度が一定値を越えた場合、パイプライン段数を 3-cycle mode に切り替える。
- タイミング違反が生じてから対処する方法： タイミング違反により値化けが生じた場合、パイプライン段数を 3-cycle mode に切り替える。

前者のメリットはタイミング違反を未然に回避できる点であるが、温度と遅延変動の割合を正確に把握しておかないとエラーを未然に防げない、もしくは、必要以上にタイミングマージンを取ることになる。

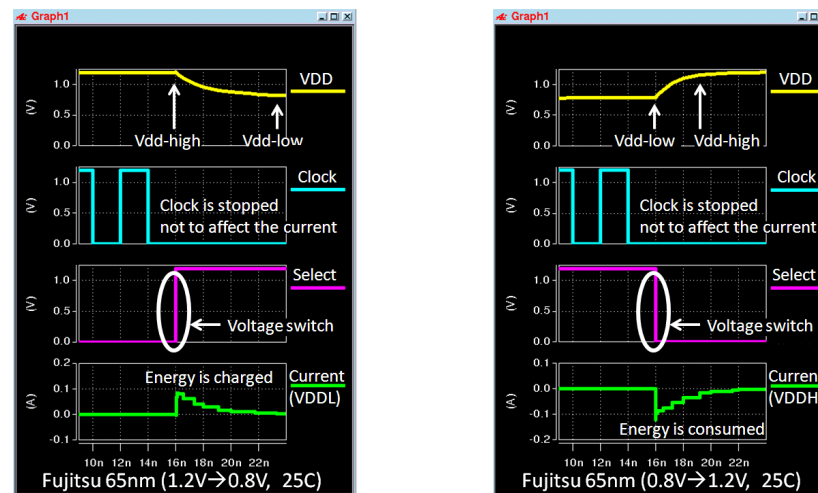
一方、後者は値化けを検出するため、リンクレベルでのビットエラーからの回復方法（エラー訂正符号もしくは検出・再送機構）が必要となる。つまり、フリット（もしくはパケット）ごとに CRC 符号などを付与し、ルータに到着する毎にチェックを行う。タイミング違反により値化けが生じた場合、パイプライン段数を 3-cycle mode に切り替える。実際にタイミング違反が生じるといふデメリットがあるものの、ビットエラーから回復さえできれば、効率的にタイミングエラーを回避できる。我々は、文献 6) で追加のバッファスペースが一切不要なリンクレベル再送方式を提案しており、現在、後者の方法と文献 6) の再送方式とあわせて用いることを検討している。

## 4. 評価

まず、多電源可変パイプラインルータの回路レベルの評価を行い、そこで得られた回路パラメータをもとに CMP のフルシステムシミュレータを用いたシステムレベルの評価を行う。

### 4.1 回路レベルの評価

まず、2章で実装した多電源可変パイプラインルータのレイアウト結果（GDS ファイル）から Cadence QRC Extraction を用いて SPICE ネットリストを抽出する。これを Synopsys



(a) Vdd-high から Vdd-low への遷移。

(b) Vdd-low から Vdd-high への遷移。

図 4 電圧切り替え時の波形（回路シミュレーションの結果）。

HSIM を用いて回路シミュレーションすることで、電圧スイッチを操作してからルータの電圧が切り替わるまでの遅延、電圧切り替えに要すオーバヘッドエネルギーを求める。

図 4 に電圧切り替え時の波形を示す。図 4(a) が Vdd-high (1.2V) から Vdd-low (0.8V) への遷移、図 4(b) が Vdd-low (0.8V) から Vdd-high (1.2V) への遷移である。それぞれのグラフにおいて 1 番上の波形はルータの電源電圧 (Vdd) を表す。2 番目と 3 番目の波形はそれぞれクロック信号 (Clock) と電圧切り替え信号 (Select) である。4 番目の波形は電流量を示す。Select 信号が 0 のとき電圧スイッチにより供給電圧は Vdd-high となり、Select 信号が 1 のときは Vdd-low となる。また、電圧切り替えに伴う電流量を正確に測定するために、クロックは途中で止めている。

電源電圧を切り替えてから、ルータの電位がターゲット電圧に到達するまでに一定の遅延がかかる。この遅延によってパイプラインモード切り替えが遅れるので、性能に影響があることがある。high-to-low の遷移と low-to-high の遷移を比べると、high-to-low の遷移のほうが変化が緩やかで遷移に時間がかかることが分かる（図 4）。

#### 4.1.1 ハードウェア量

多電源可変パイプラインルータでは、電圧を切り替えるための電圧スイッチセル、電源ド

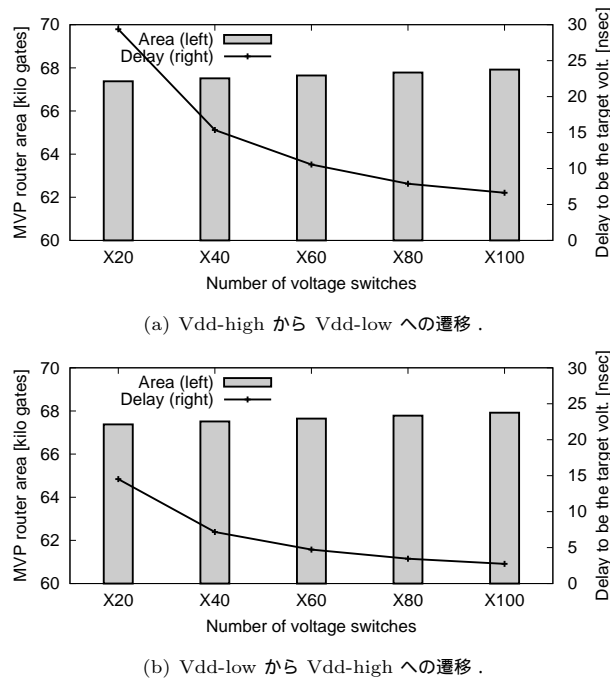


図 5 電圧切り替え時間とルータ面積 (電圧スイッチの個数を変えた評価) .

メイン間の電圧レベルの違いを吸収するためのレベルシフタセルが必要となる (図 2) . 当然 , 電圧スイッチの面積を増やせば , 電圧切り替えに要す遅延は小さくなる .

図 5 に電圧切り替え時間とルータ面積を示す . ここでは電圧切り替え信号を操作してから Vdd がターゲット電圧の  $\pm 0.05V$  に達するまでを電圧切り替え時間とする . グラフの X 軸は電圧スイッチの個数であり , 20 個から 100 個の範囲で変化させている . 棒グラフ (目盛は Y 軸左側) はルータのハードウェア量 [kilo gates] を示し , 折れ線グラフ (目盛は Y 軸右側) は電圧切り替え時間 [nsec] である . 電圧スイッチを増やせば電圧切り替えに要す遅延は小さくなる . それでも , 電圧スイッチのハードウェア量はルータ全体と比べると十分小さいため , 以降の評価ではルータごとに電圧スイッチを 80 個 (X80) 入れることにした .

既存の 3 サイクルルータと提案ルータの面積の内訳を表 2 に示す . ルータごとに電圧スイッチを 80 個 (X80) , レベルシフタをすべての入力ポートに挿入している . その結果 ,

表 2 ハードウェア量の内訳 (単位は kilo gate) .

	Router	Level shifter	Voltage switch	Total
Original	59.41	0.00	0.00	59.41
Proposed	63.13	4.11	0.54	67.78 (+14.1%)

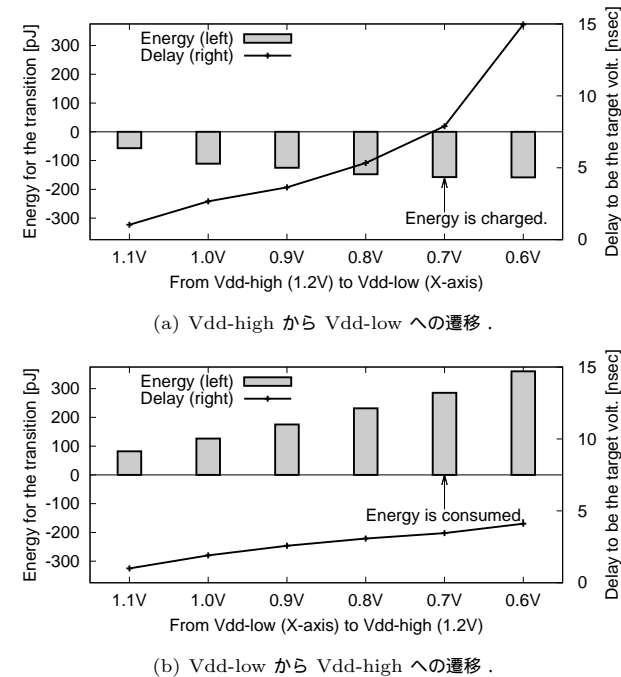


図 6 電圧切り替え時の遅延とオーバーヘッドエネルギー (Vdd-low を変えた評価) .

ハードウェア量が 14.1%増加した . 面積オーバーヘッドの大部分を 1) 通常のルータを可変パイプライン化するための面積 , 2) レベルシフタの面積が占めていることが分かる .

#### 4.1.2 切り替えレイテンシ

図 6 に high-to-low および low-to-high の電圧切り替え時間 (折れ線グラフ , 目盛は Y 軸右側) を示す . Vdd-high が 1.2V , Vdd-low が 0.8V のとき , high-to-low の遅延は 5.3nsec , low-to-high の遅延は 3.1nsec となった . つまり , 3-cycle から 2-cycle のモード切り替えでは , まず , 電源電圧を切り替えてから 3.1nsec (392.2MHz 駆動では 2 サイクル) 待ち , パ

イブラインを 2-cycle mode に切り替える必要がある。

#### 4.1.3 切り替えオーバーヘッドエネルギー

提案ルータの電圧切り替え時に、一定のエネルギーが消費される。図 6 に high-to-low 切り替えおよび low-to-high 切り替え時の消費エネルギー（棒グラフ、目盛は Y 軸左側）を示す。low-to-high 切り替えでは、切り替え直後に、Vdd-high からルータ回路へ電荷が流れ込み、オーバーヘッドエネルギーが消費される。一方、high-to-low 切り替えでは、ルータ回路内にたまった Vdd-high の電荷が、切り替え後に、Vdd-low 側の電源に流れ込む。この電荷は Vdd-low 側のキャパシタンスに蓄積され、電圧切り替え後に徐々に消費されていくため、電圧切り替え時のオーバーヘッドエネルギーはマイナス（つまりチャージされる）となる。Vdd-high が 1.2V、Vdd-low が 0.8V のとき、low-to-high の消費エネルギーは 231.4pJ、high-to-low の消費エネルギーは -147.6pJ となった。つまり、1 回の電圧切り替え（low-to-high + high-to-low）で消費されるエネルギーは 83.8pJ である。

#### 4.1.4 切り替えの損益分岐点

次に、この 83.8pJ のオーバーヘッドエネルギーを償却するには、どれだけの期間 3-cycle mode @ Vdd-low で駆動すれば良いか計算する。この期間を電圧切り替えの break-even time (BET) と呼ぶ。

配置配線後シミュレーションの結果、2-cycle mode @ Vdd-high の消費電力は、無負荷時において、2.78mW となった。消費電力は電圧の 2 乗に比例するため、3-cycle mode @ Vdd-low の消費電力は 1.33mW である。つまり、3-cycle mode にすることで、消費電力を 1sec で 1.45mJ セーブできる。したがって、83.8pJ を償却するには 57.8nsec がかかるため、BET は 23 サイクルとなる。これより短期間の電圧変更は消費電力の増加を招く。

### 4.2 システムレベルの評価

3章で提案したスタンバイ電力削減ポリシーを実際の CMP を想定したフルシステムシミュレータ上で評価する。

#### 4.2.1 シミュレーション環境

##### CMP アーキテクチャ

図 1 に示した 16 コア CMP で使われる NoC をシミュレーションする。ここでは以下の 2 種類の CMP アーキテクチャに多電源可変パイプラインルータを適用する。

- 共有 L2: 各タイルの L2 キャッシュバンクを全タイルで共有し、チップ全体として 1 つの大きな共有 L2 キャッシュを形成する。
- 非共有 L2: 各タイルの L2 キャッシュバンクを他のタイルと共有せず、同一タイル上

表 3 シミュレーション条件 (プロセッサ, メモリ関連)

Processor	UltraSPARC-III
L1 I-cache size	16 KB (line:64B)
L1 D-cache size	16 KB (line:64B)
# of processors	16
L1 cache response	1 cycle
L2 cache size	256 KB (assoc:4)
# of L2 cache banks	16
L2 cache response	6 cycle
Memory size	4 GB
Memory response	160 (± 2) cycle
# of memory ports	16

表 4 シミュレーション条件 (NoC, オンチップルータ関連)

Topology	4×4 mesh
Routing	dimension-order
# of VCs	4
Buffer size	4 flit
Flit size	64 bit
Control packet	1 flit
Data packet	9 flit

のプロセッサのプライベート L2 キャッシュとして振る舞う。

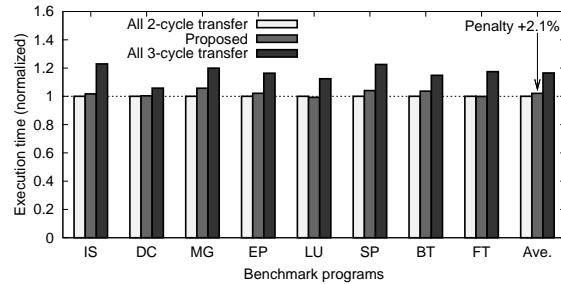
キャッシュのコヒーレンスを取るため、ここではディレクトリベースのコヒーレンスプロトコルを用いる。end-to-end のプロトコル (要求・応答) デッドロックを防ぐため、仮想チャネルを 3 本 (もしくは 4 本) 用いる。当然、共有 L2 ではネットワーク上のトラフィック量は多く、非共有 L2 ではトラフィック量は少ない。

##### シミュレーションモデル

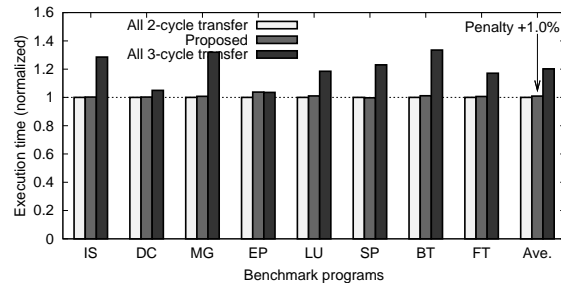
このような CMP のフルシステムシミュレータとして GEMS<sup>7)</sup> と Wind River Simics<sup>8)</sup> を組み合わせて使用する。NoC モデルとして GEMS に付属の Garnet<sup>9)</sup> を拡張することで、多電源可変パイプラインルータを cycle-accurate でシミュレーションできるようにした。表 3 にプロセッサとメモリ関連、表 4 に NoC とオンチップルータ関連のシミュレーション条件を示す。ここでは主記憶アクセスの影響を排除するため、各ルータにメモリポートを持たせている。

##### ベンチマークプログラム

多電源可変パイプラインルータの性能とスタンバイ電力を評価するため、OpenMP 版の



(a) 16 コアタイル (L2 共有) .



(b) 16 コアタイル (L2 非共有) .

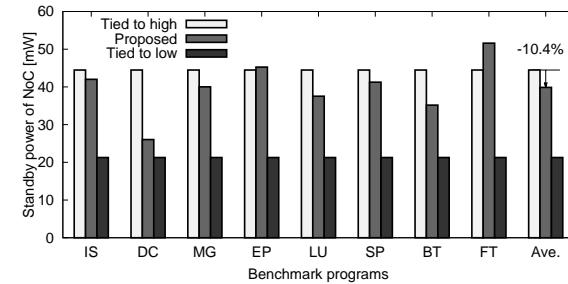
図 7 NPB アプリケーションの実行時間 (完全 2 サイクルルータの実行時間を 1.0 として正規化) .

NAS Parallel ベンチマーク (NPB)<sup>10</sup> から 8 種類のプログラム (IS, DC, MG, EP, LU, SP, BT, FT) を用いる . OS として Sun Solaris 9 を 16 コア CMP のうえで動作させる . NPB ベンチマークプログラムを Sun Studio 12 でコンパイルし , Solaris 9 上で動作させる . 問題サイズは S , スレッド数はすべて 16 とする .

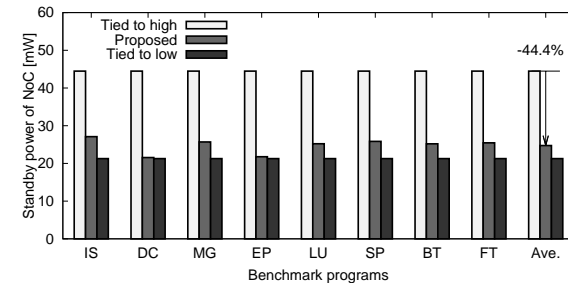
#### 4.2.2 アプリケーション実行時間

CMP 向け NoC に多電源可変パイプラインルータを適用する . そのうえで 8 種類の NPB プログラムを走らせ , その実行サイクル数を求める . ここでは以下の 3 つのルータモードを比較する .

- **All 2-cycle transfer** : 常に 2-cycle mode @ Vdd-high に固定
- **Proposed** : スタンバイ時は 3-cycle mode @ Vdd-low だが , パケットが来たらできるだけ 2-cycle mode @ Vdd-high で転送



(a) 16 コアタイル (L2 共有) .



(b) 16 コアタイル (L2 非共有) .

図 8 NPB アプリケーション実行時の NoC のスタンバイ電力 (オーバーヘッド電力を含む) .

- **All 3-cycle transfer** : 常に 3-cycle mode @ Vdd-low に固定

図 7 に NPB アプリケーションの実行時間を示す . 図 7(a) が L2 共有の場合 , 図 7(b) が L2 非共有の場合である . 実行時間 ( グラフの Y 軸 ) は All 2-cycle transfer の実行時間を 1.0 として正規化してある . グラフより , 提案手法は常に 2-cycle mode @ Vdd-high の場合と比べて 1.0% ~ 2.1% 性能が落ちている . これはパケット転送の 1 ホップ目は Vdd-low から Vdd-high への昇圧が間に合わず , 3-cycle 転送が生じたためである . 次節では , 常に 2-cycle mode @ Vdd-high の場合と比べて提案手法がどれだけスタンバイ電力を減らすことができたか評価する .

#### 4.2.3 スタンバイ電力の削減量

4.1 節で述べたとおり , 1 回の電圧切り替え ( low-to-high + high-to-low ) で消費されるエネルギーは 83.8pJ である . また , ルータ単体のスタンバイ電力は 2-cycle mode @ Vdd-high

時で 2.78mW, 3-cycle mode @ Vdd-low 時で 1.33mW である。ここでは前節で実行した NPB アプリケーションごとに、電圧切り替えが起こった時刻を記録しておき、電圧切り替えの回数と 1 回の電圧切り替えで消費されるエネルギーを掛け合わせて提案手法のオーバーヘッド込みのスタンバイ電力を見積もった。

図 8 に NPB アプリケーション実行時の NoC のスタンバイ電力を示す。前節の All 2-cycle transfer が Tied to Vdd-high に対応し、All 3-cycle transfer が Tied to Vdd-low に対応する。提案手法 (Proposed) のスタンバイ電力には電圧切り替えのオーバーヘッド電力が含まれている。

図 8(a) が L2 共有の場合、図 8(b) が L2 非共有の場合である。トラフィック量の少ない L2 非共有のほうが、電圧切り替え回数つまりオーバーヘッド電力が少なく、スタンバイ電力を平均 44.4%削減できている。一方、トラフィック量が多い L2 共有ではスタンバイ電力の削減量は平均 10.4% と効果があまり大きくない。実際、FT などとくにトラフィック量の多いアプリケーションでは BET 以下の短い間隔での電圧切り替えが頻発してしまっている。無駄なオーバーヘッド電力を減らすためにも BET を考慮した効率的な電圧切り替えが必要であり、これは今後の検討課題としたい。

## 5. まとめと今後の課題

多電源可変パイプラインルータを設計し、面積オーバーヘッド、電圧切り替え時の遅延およびエネルギーオーバーヘッドを測定した。また、このエネルギーオーバーヘッドを償却するために必要な BET を求めた。その結果、可変パイプライン化および多電源化 (電圧スイッチとレベルシフタ) の面積オーバーヘッドは 14.1%、切り替え遅延は 3.1nsec ~ 5.3nsec、1 回の電圧切り替えで消費されるエネルギーは 83.8pJ、BET は 23 サイクルとなった。

次に、パイプラインおよび電圧切り替えポリシーとして、スタンバイ電力削減ポリシーとタイミング違反回避ポリシーを提案した。今回は、前者のスタンバイ電力削減ポリシーを CMP のフルシステムシミュレータ上に実装し、16 スレッドの NPB アプリケーションを走らせて評価した。その結果、2-cycle mode @ Vdd-high に固定する場合と比べ、アプリケーション性能は 1.0% ~ 2.1% 落ちたが、スタンバイ電力を 10.4% ~ 44.4% 削減できた。ただし、トラフィック量が多い場合は電力削減量が著しく低下したため、今後は BET を考慮したより効率的な電圧切り替えポリシーを検討する必要がある。

謝辞 本研究は科学技術振興機構「JST」の戦略的創造研究推進事業「CREST」における研究領域「情報システムの超低消費電力化を目指した技術革新と統合化技術」の研究課題「革新的電源制御に

よる次世代超低電力高性能システム LSI の研究」による。本研究は日本学術振興会特別研究員奨励費の助成を受けている。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半導体設計ライブラリを利用させていただいた。最後に、本研究で使用した電圧スイッチおよびレベルシフタセルを提供くださった芝浦工業大学 馬橋雄祐氏に感謝致します。

## 参考文献

- 1) Dally, W.J. and Towles, B.: Route Packets, Not Wires: On-Chip Interconnection Networks, *Proceedings of the Design Automation Conference (DAC'01)*, pp.684-689 (2001).
- 2) Beigne, E., Clermidy, F., Lhermet, H., Miermont, S., Thonnart, Y., Tran, X.-T., Valentian, A., Varreau, D., Vivet, P., Popon, X. and Lebreton, H.: An Asynchronous Power Aware and Adaptive NoC Based Circuit, *IEEE Journal of Solid-State Circuits*, Vol.44, No.4, pp.1167-1177 (2009).
- 3) Hirata, Y., Matsutani, H., Koibuchi, M. and Amano, H.: Low-Power Variable-Pipeline Router using Pipeline-Stage Integration, *Proceedings of the International Symposium on Low-Power and High-Speed Chips (COOL Chips XII)*, p.155 (2009).
- 4) Dally, W.J. and Towles, B.: *Principles and Practices of Interconnection Networks*, Morgan Kaufmann (2004).
- 5) Matsutani, H., Koibuchi, M., Ikebuchi, D., Usami, K., Nakamura, H. and Amano, H.: Ultra Fine-Grained Run-Time Power Gating of On-Chip Routers for CMPs, *Proceedings of the International Symposium on Networks-on-Chip (NOCS'10)*, pp.61-68 (2010).
- 6) 小島 悠, 松谷宏紀, 鯉淵道紘, 天野英晴: エラー訂正・検出符号を用いた Network-on-Chip の低消費電力化, 第 8 回先進的計算基盤システムシンポジウム (SACSYS'10) 論文集, pp.3-10 (2010).
- 7) Martin, M. M.K., Sorin, D.J., Beckmann, B.M., Marty, M.R., Xu, M., Alameldeen, A.R., Moore, K.E., Hill, M.D. and Wood, D.A.: Multifacet General Execution-driven Multiprocessor Simulator (GEMS) Toolset, *ACM SIGARCH Computer Architecture News (CAN'05)*, Vol.33, No.4, pp.92-99 (2005).
- 8) Magnusson, P.S., Christensson, M., Eskilson, J., Forsgren, D., Hallberg, G., Hogberg, J., Larsson, F., Moestedt, A. and Werner, B.: Simics: A Full System Simulation Platform, *IEEE Computer*, Vol.35, No.2, pp.50-58 (2002).
- 9) Agarwal, N., Peh, L.-S. and Jha, N.: Garnet: A Detailed Interconnection Network Model inside a Full-system Simulation Framework, Technical Report CE-P08-001, Princeton University (2008).
- 10) Jin, H., Frumkin, M. and Yan, J.: The OpenMP Implementation of NAS Parallel Benchmarks and Its Performane, *NAS Technical Report NAS-99-011* (1999).