

# オンチップルータにおける仮想チャネル単位の走行時パワーゲーティング

松谷 宏紀 (慶大), 鯉渕 道紘 (NII), 王 代涵 (慶大), 天野 英晴 (慶大)



## Network-on-Chip (NoC)

プロセッサコアはルータを介してネットワーク接続される

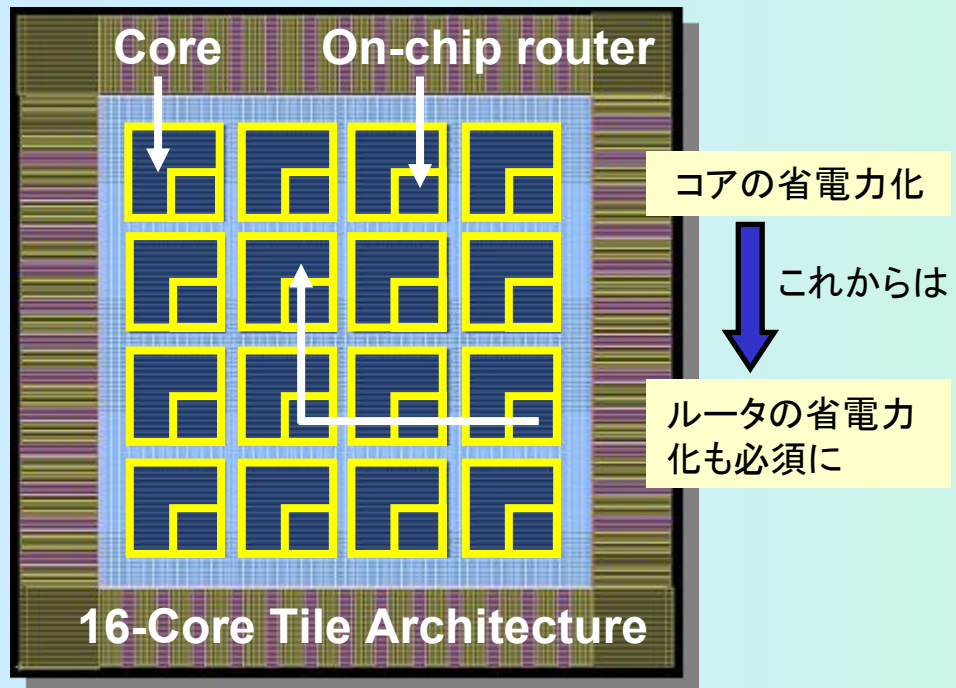
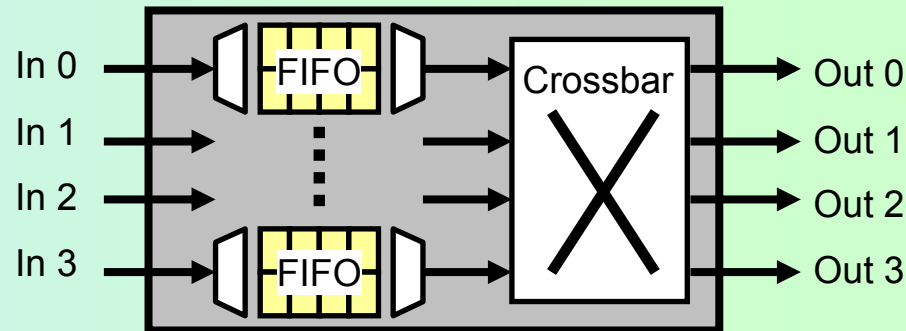


Table 1: タイルアーキテクチャの概要

Process	ASPLA 90nm (Low Vth) with supply voltage 1.0 V
Freq.	Max 500MHz
Core	Packet generator (将来は MIPS プロセッサを接続)
Router	Wormhole router (data width 64-bit, 3-stage pipeline)

## On-Chip Router

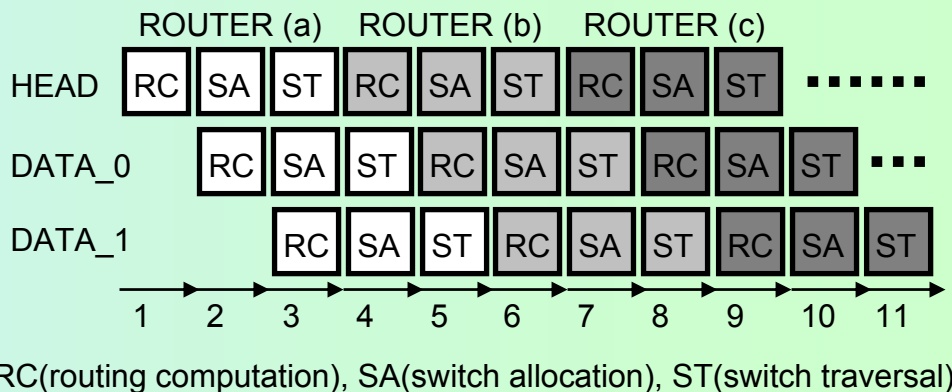
4入力4出力, データ幅 64-bit, ワームホール方式



入力ポートごとに2系統のバッファ (仮想チャネル) を持つ  
仮想チャネルごとに64-bit x 4 のFIFO バッファ (FFで実装)

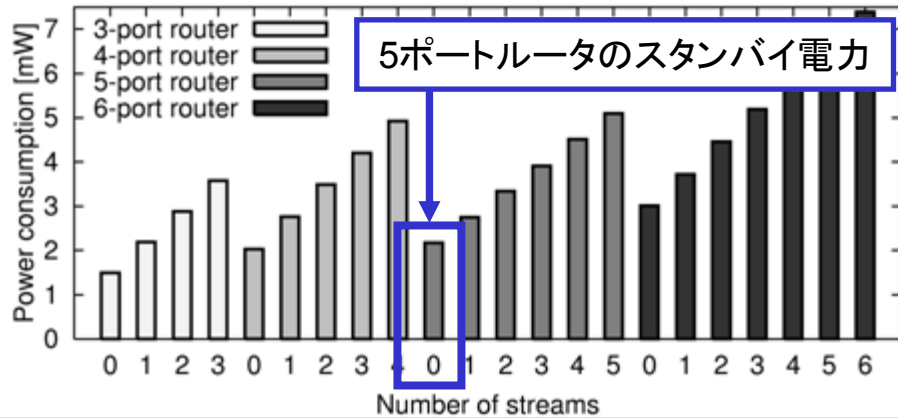
## Router Pipeline (3-stage)

ヘッドフリットが1ホップ移動するのに最低3サイクルかかる



## Active Power (200MHz 動作時)

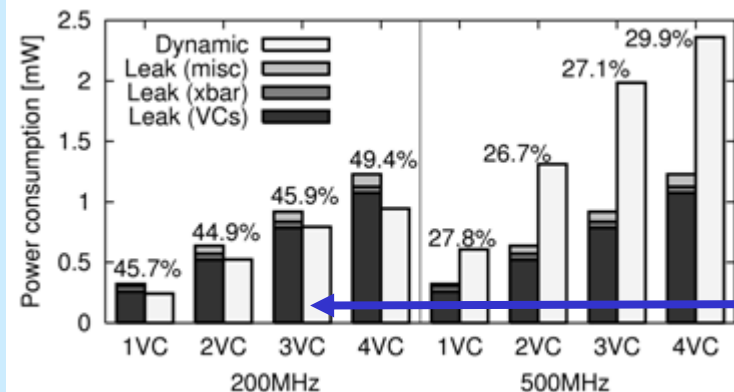
N-port ルータにおいて、n 本のポートでパケット転送が行われているときの消費電力 ( $0 \leq n \leq N$ )



→ スタンバイ電力 (負荷が0のとき) がかなり大きい

## Standby Power

仮想チャンネル数が1~4本のときのスタンバイ電力の内訳



→ 仮想チャンネルバッファのリークが全体の 49.4% (最大)

## Slow-Silent Virtual Channels の提案

仮想チャンネルの本数を増やすことで消費電力の削減を狙う

- 1) 仮想チャンネルの本数を増やす → スループットが向上
- 2) 個々の仮想チャンネルの動作周波数と電圧を下げる
- 3) 仮想チャンネル単位のパワーゲーティングでリーク削減

## 仮想チャンネルの本数と性能

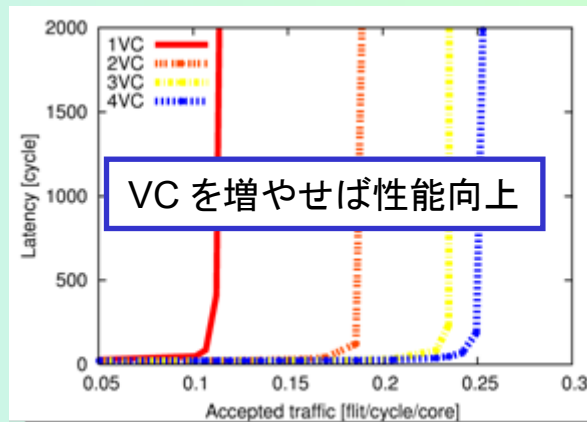


Table 2: 評価環境

# of cores	64
Traffic	Uniform
# of VCs	1, 2, 3, 4
Packet len.	5 flits

## Voltage and Frequency Scaling

性能が向上した分、動作周波数と電圧を低減 → 電力削減

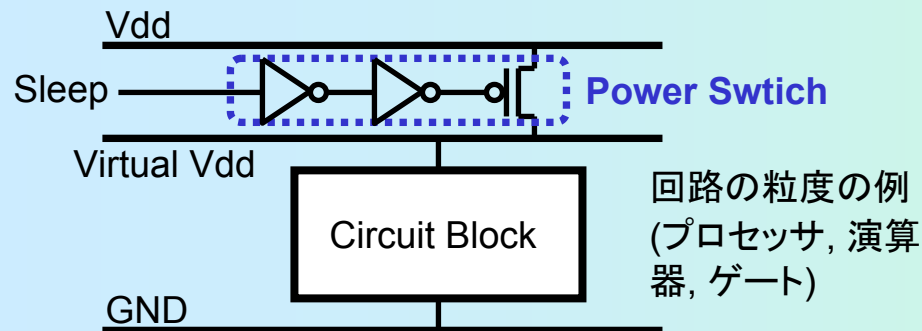
$$f \propto \frac{(V - V_{th})^\alpha}{CV} \quad P_{switching} = a \cdot C \cdot f \cdot V^2$$

しかし、仮想チャンネルの追加はリーク電力の増大を招く

→ 仮想チャンネル単位のパワーゲーティングでリークを削減 (本ポスタ発表では 3. について説明致します)

## Power Gating Technique

回路への電力供給を遮断することで、リーク電流を削減

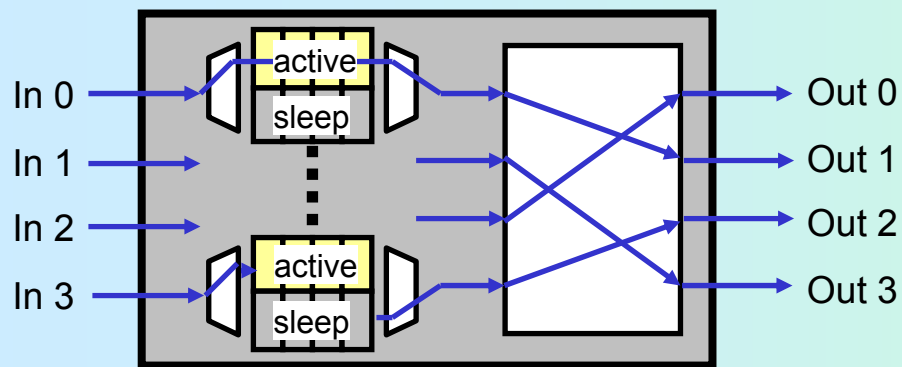


パワーゲーティングのコスト:

- 1) パワースイッチ挿入による面積の増加
- 2) ウェイクアップ遅延による性能の低下
- 3) パワースイッチ On/Off によるオーバヘッド電力

## Runtime Power Gating of VCs

パケットが来たらアクティブ, パケットが行ったらスリープ

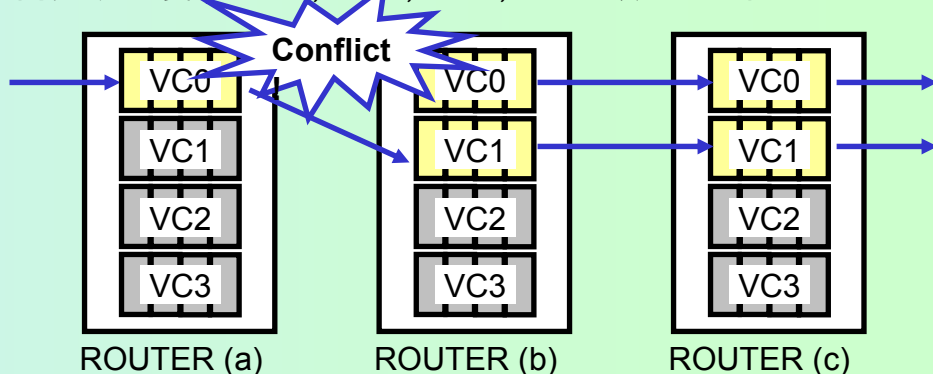


## Look-ahead Wakeup [松谷, ASPDAC'08]

パケットが2ホップ手前に来た時点でウェイクアップ開始

## Routing Strategy

パケット注入時は VC0 を使い, パケットが他のパケットと衝突する度に VC1, VC2, VC3, ... の順に切り替えていく



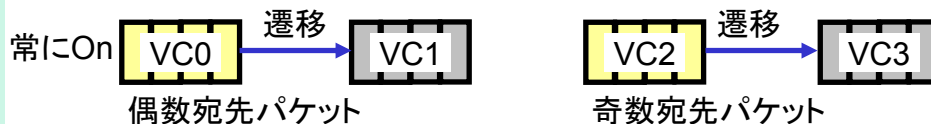
→ 負荷が高くなるに従い VC1, VC2, ... が活性化される

## Routing Strategy: 提案ルーティング(4VCx1)と比較

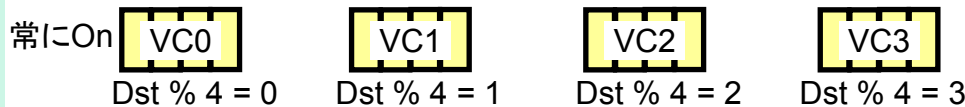
4VC x 1 (レーン数1): 先詰まりには強いが, ウェイクアップ遅延(大)



2VC x 2 (レーン数2):



1VC x 4 (レーン数4): 先詰まりには弱い, ウェイクアップ遅延(小)



## Evaluation #1 (leakage modeling)

どれだけスリープ出来ればオーバーヘッド電力を償却できるか？

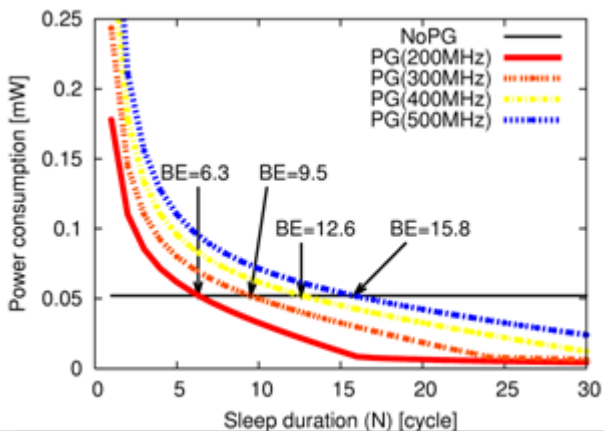


Table 3: VC の消費電力 (200MHz)

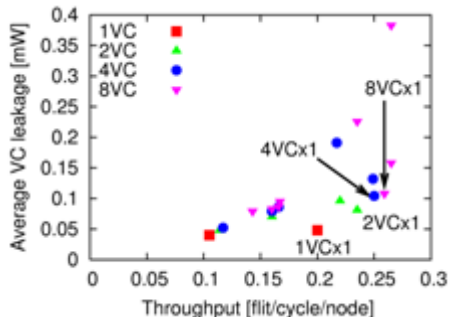
Voltage	1.0 V
SW activity	0.12
Leakage P.	52 uW
Switching P.	78 uW

[Hu, ISLPED'04] のパワーゲーティングモデルを用いて算出

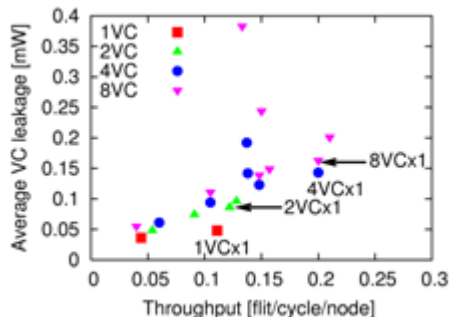
→ ネットワークシミュレータにリーク電力の計算機能を実装

## Evaluation #2 (routing strategy)

4VCx1, 2VCx2, 1VCx4 ... どのルーティング法が良いか？



16-coreの場合

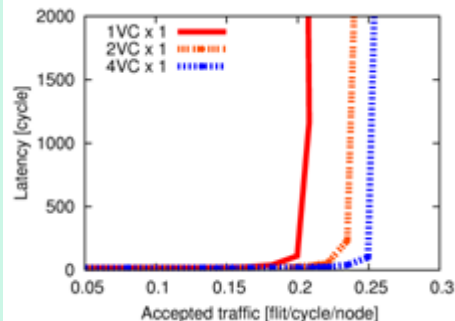


64-coreの場合

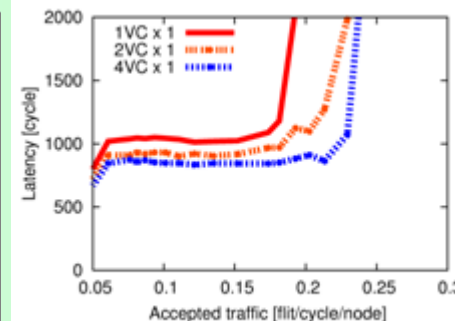
→ レーン当りの仮想チャネル数が多いほうが性能が高い  
→ リーク電力と性能の比で1VCx1, 2VCx1, 4VCx1が良い

## Evaluation #3 (throughput)

1VCx1, 2VCx1, 4VCx1 のスループットを評価



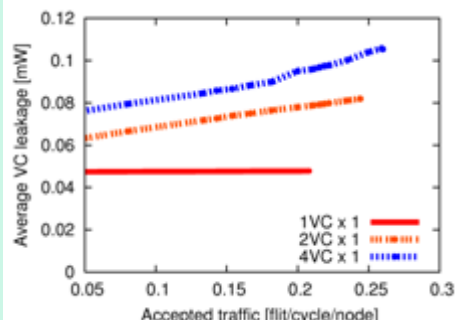
Uniform (16-core)



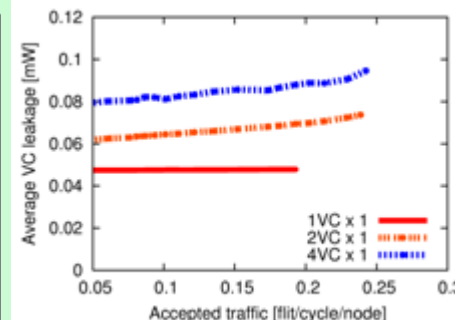
IS.W (16-core)

## Evaluation #4 (ave. leakage power)

1VCx1, 2VCx1, 4VCx1 のVC当りのリーク電力を評価



Uniform (16-core)



IS.W (16-core)

## まとめ - 仮想チャネル単位の PG -

- 1) 負荷に応じて VC を徐々に活性化するルーティングを提案
- 2) 1VCx1, 2VCx1, 4VCx1 ルーティングの電力性能比が高い
- 3) 4VCx1 の場合、ピーク時でも 36~53% のリークで動作
- 4) 今後は、VC 数に応じて周波数と電圧を制御して電力削減