

平成 22 年 6 月 10 日現在

研究種目：若手研究 (B)
 研究期間：2008～2009
 課題番号：20700054
 研究課題名 (和文) 予測と部分二重化による超低遅延チップ内ネットワークに関する研究
 研究課題名 (英文) A Study of Ultra Low-latency Network-on-Chips using Prediction and Partially Duplications
 研究代表者 鯉淵 道紘 (KOIBUCHI MICHIIHIRO)
 国立情報学研究所・アーキテクチャ科学研究系・准教授
 研究者番号：40413926

研究成果の概要 (和文)：本研究では、マルチコア、メニーコアプロセッサなどの複雑化するシングルチップ計算機プラットフォームを実現するために急務である革新的な超低遅延インターコネクト技術の実現を目的とした。そのために、(1) 予測と部分二重化による低遅延ルーターアーキテクチャと(2) そのルーター間の協調動作による更なる遅延削減のためのトポロジ、ルーティングの部分再構成技術などを研究開発した。本研究成果は、従来のPCクラスター、並列計算機などのシステムレベルの相互結合網の研究領域の技術を活用して、マイクロシステムの大幅な低遅延化を達成するなどの成果を挙げることができた。

研究成果の概要 (英文)：The objective of this study is to develop an innovative ultra-low latency interconnect technology in order to achieve a complex single-chip computer-system platform, such as multi-core and many-core processors. We thus proposed and evaluated (1) low-latency router architectures using a prediction and partially duplication, and (2) partially reconfiguration techniques of topology and routing for further reducing the latency with inter-router co-operations. The results of this study enable to reduce the latency of micro-systems by making the best use of the interconnection network techniques in system-level research regions, such as traditional PC clusters and massively parallel computers.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,100,000	330,000	1,430,000
2009年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総計	2,000,000	600,000	2,600,000

研究分野：相互結合網，並列分散システム

科研費の分科・細目：情報学・計算機システム・ネットワーク

キーワード：相互結合網，システムオンチップ，トポロジ，計算システム，ルーティング

1. 研究開始当初の背景

半導体技術の進歩によって、複数、あるいは多数の処理コアを持つマルチコア、メニーコアプロセッサが登場し、これらのプロセッサシステムは計算機プラットフォームとしての地位を確立し、今後も益々重要になっていくと考えられる。例えば、2007年、インテルは80コアのシングルプロセッサを開発し、同プロセッサはスーパーコンピュータに匹敵するTFlops(1秒当たり数兆回の演算速度)性能を目指している。従来のプロセッサ研究がスーパースカラ、アウトオブオーダー命令実行等のマイクロアーキテクチャに関連するものが多かったが、今後は、マルチ/メニーコア化するにあたり処理コアの相互結合方式、通信機構といった課題が極めて重要になっていくと考えられる。つまり、従来は、研究代表者の鯉淵の2005年ころの研究テーマであるPCクラスタ、並列計算機などの大規模システム・ネットワークの課題が、マイクロシステムの課題へと広がりを見せている。情報家電などの組み込み機器で使われるシステムオンチップ(SoC)の分野においても、実装可能なモジュール数が飛躍的に増加しているため、同様の課題が生じている。

今後は、より多くのモジュールをチップ内に実装可能となるため、チップ内パケットネットワークは通信オーバーヘッドが大きいルータを多数接続する構成を取ることになる。その結果、これまではさほど問題にならなかったネットワークの遅延が、オンチップメモリにアクセスする場合などには特に、深刻な性能低下を引き起こすことになる。つまり、ネットワーク遅延を削減することができれば、アプリケーションの並列性を最大限に引き出

すことが可能となり、システム設計者、プログラマは多数のコアを効果的に使うことが容易になる。

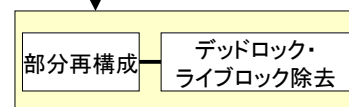
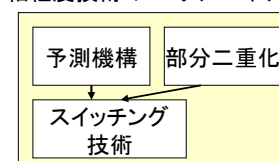
2. 研究の目的

本研究では、マルチコア、メニーコアプロセッサなどの複雑化するシングルチップ計算機プラットフォームを実現するために急務である革新的な超低遅延インターコネクト技術の実現を目的とした。そのために、(1) 予測と部分二重化による低遅延ルータアーキテクチャと(2) そのルータ間の経路最適化による更なる遅延削減のためのトポロジ、ルーティングの部分再構成技術などを提案、評価する。

3. 研究の方法

チップ内ネットワークにおける超低遅延技術を効率良く順序立てて実現するために、図1に示した通り、ルータ内アーキテクチャの細粒度技術、ルータ間を扱うネットワーク全体の設計という粗粒度技術の2つに分類する。

細粒度技術: ルータアーキテクチャ



粗粒度技術: チップ内ネットワーク設計

図1: 研究方法の概略

まず、細粒度技術として、ルータにおける本研究の2つの重要なアイデア“部分二重化”、“予測機構”を個々に検討し、本提案アーキテ

クチャが軽量なハードウェアで実現できることを明らかにする。次に、これら2つの要素技術に適した別途スイッチング、リンクレベルフロー制御を検討する。また、シミュレーションによる定量的な検証、評価を研究協力者である慶應義塾大学理工学研究科後期博士過程松谷、Wang と協力して行う。

なお、シミュレーションにおいて、万が一、計算機の計算能力が不足した場合、あるいは、開発が遅れた場合には、対象とするシステムのコア数を減らし、簡素化して評価を取る。

また、待ち行列理論、マルコフモデルを用いてルータとネットワークの性能を解析し、性能要因を確認し、かつ提案手法にフィードバックする。既存の多くの研究がマクロシステムの評価にシミュレーションのみを用いた総合評価を行っているが、本研究では理論的にも性能要因を特定できるように工夫を行った。

次に、粗粒度技術として、予測、二重化によるネットワーク経路の部分再構成技術、その経路群導入に伴うデッドロック、ライブロック除去技術を提案、評価する。

4. 研究成果

研究方法にしたがって、細粒度技術と粗粒度技術の2つの要素について順に成果を述べる。

(1) 細粒度技術

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換えることが可能な可変パイプラインルータを提案、評価した[2][4]。

最近では DVFS (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。

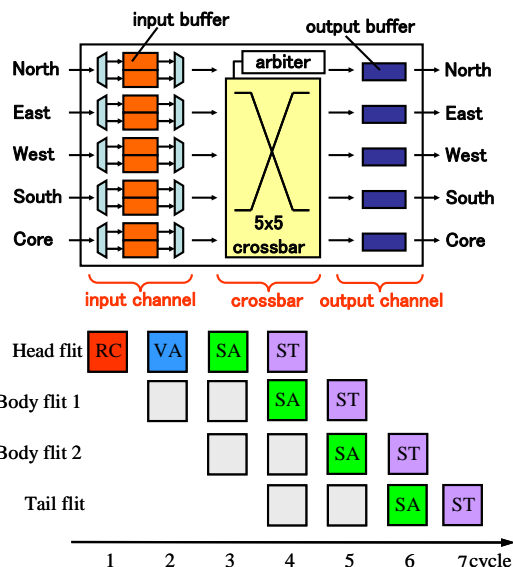


図2：オンチップ・ルータの概要

図2に典型的な既存のオンチップ・ルータを示す。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。そのため、図2に示した通り、4フリットで構成されるパケットを転送する場合、7サイクル必要となる。

この既存のオンチップ・ルータを低遅延化するためには、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最大限に大きくする最適化が不可欠である。そこで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。具体的には、図3に示した通り、可変パイプラインルータでは、チップの動作周波数が低くなるにつれて、パケット処理に必要なサイクル数を3、2、1サイクルと小さくする。

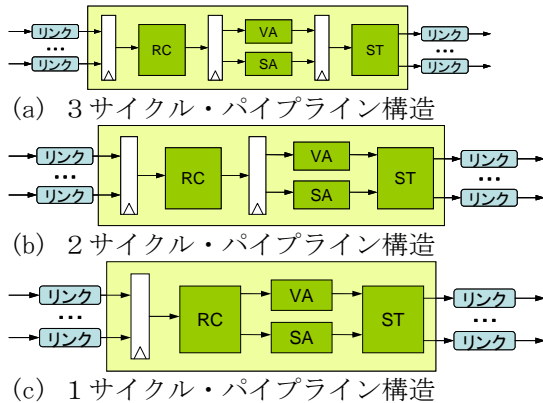


図3：パイプライン・ステージの統合

本研究ではまず、動作周波数の低下とパケット転送に必要なサイクル数のトレードオフについて調査するためにオンチップルータを設計、評価を行った。

可変パイプラインルータは、各サイクル転送に用いるデータパスを最大限共有するため、面積(ハードウェア)オーバーヘッドは6%未満であった。

次に、各パイプライン段数における最大動作周波数を図4に示す。

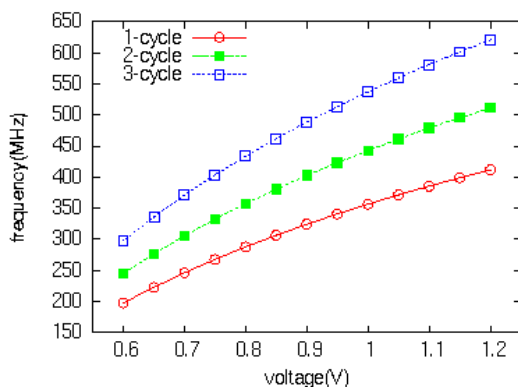


図4：各パイプライン構造の最大動作周波数

電圧が一定の場合、3サイクル・パイプライン構造が最も動作周波数を高くできていることが分かった。そして2サイクル・パイプライン構造、1サイクル・パイプライン構造となるにつれて動作周波数が緩やかに落ちていくことが分かった。

最後に、最大動作周波数を用いてNAS並列ベンチマークの実行トレース、HPL、合成トラフィックにおけるチップ内ネットワークの遅延削減の限界値である無負荷(zero-load)レイテンシを算出し、その結果を図5に示す。

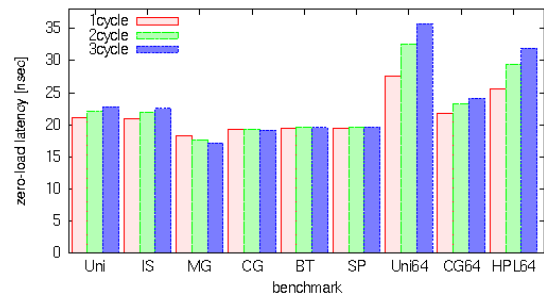


図5：NPB, HPLのトレース、および合成トラフィックにおける各パイプライン構造の無負荷レイテンシ

図5より、NAS並列ベンチマークでは最大動作周波数は落ちるが、各ルータが1サイクルでパケット転送を行う場合、もっとも遅延を削減できることが分かった。

さらに、可変パイプラインルータを運用するために必要となる、トラフィック負荷の測定あるいはトラフィックの予測方法と、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等も検討を行った。具体的には、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。

そして、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ために、消費電力の面で最適なパイプライン構成、の2つについて、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し、その上で低遅延化を実現することができた。また、可変パイプラインルータは複数のパイプライン構成を合わせることで、部分的に二重化することも実現できる。

(2) 粗粒度技術

ルータ間通信技術として、任意のトポロジに適用可能なデッドロックフリー適応型ルーティングである L-turn ルーティングをもとにした基盤技術を提案、評価した。現在、チップ内ネットワークでは、2次元メッシュなどの規則性の強いトポロジ、デッドロックフリールーティングを採用する人が多い。

そのため、ルータの一部の入出力ポート間の転送パターンに極めて強い偏在性が生じる。そこで、頻繁にトラフィックが転送される入出力間コネクションをルータ内のデータパスとは別に実装し(部分二重化)、ルータ内の複雑なパケット処理を迂回する。つまり、プロセッサのキャッシュのように、頻繁にアクセスされるルータ内入出力対データパスについてはその情報、構成を記憶(迂回路の利用)し、超低遅延化する。この場合、部分的にトポロジ構成が変更されるかデッドロックフリーを保障する経路群については若干の更新に留めることができた[6]。これは2次元グラフにルータ間コネクションをマッピングし、循環構造をできるだけ規則的に除去する。図5に、提案手法による2次元メッシュを元にしたトポロジにおける部分再

構成によるデッドロック除去例を示す。

図5において矢印はルータにおいてパケットの転送を禁止する入出力ポート対を示し、各出発地・目的地対の経路はこの禁止ターンを避けることによりデッドロックフリー転送が保証される。

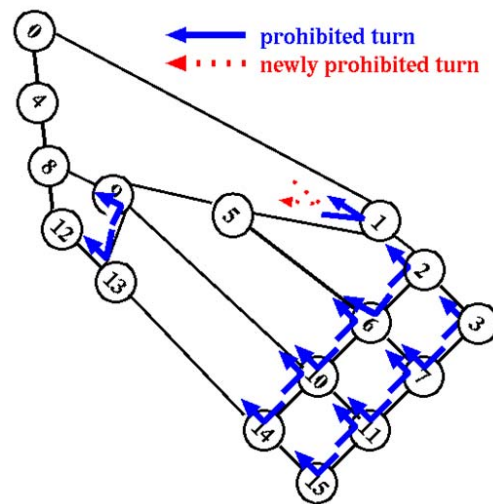


図5：部分再構成による16ルータ・チップ内ネットワークのデッドロック除去

また、FPGAでの実装を想定した基本となる各トポロジ、ルーティングについて総合的な評価も行った[1][3][5]。その結果、次数の大きなトポロジが遅延のみならず優れていることが分かった。

以上より、本研究成果は、従来のPCクラスタ、並列計算機などのシステムレベルの相互結合網の研究領域の技術を活用して、マイクロシステムの大幅な低遅延化を達成するなどの成果を挙げることができた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 6件)

(1) Sen In, Hiroki Matsutani, Michihiro Koibuchi, Daihan Wang, and Hideharu Amano, "Performance, Cost, and Power Evaluations of On-Chip Network Topologies in FPGAs", The IASTED

International Conference on Parallel and Distributed Computing and Networks (PDCN), pp.181-189, Feb 2010, 査読有

(2) 枚田 優人, 松谷 宏紀, 鯉渕 道紘, 天野英晴, ``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'', 情報処理学会論文誌: コンピューティングシステム (ACS), Vol.2, No.3, pp.71-82, Sep 2009, 査読有

(3) Jose Miguel Montanana, Michihiro Koibuchi, Hiroki Matsutani, Hideharu Amano, ``Balanced Dimension-Order Routing for k-ary n-cubes'', Proc. of the 4th International Symposium on Embedded Multicore Systems-on-Chip (MCSoc'09), Sep 2009, 査読有

(4) 枚田 優人, 松谷 宏紀, 鯉渕 道紘, 天野英晴, ``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'', 第7回先進的計算基盤システムシンポジウム (SACIS'09) 論文集, pp.19-26, May 2009, 査読有

(5) Sen In, Hiroki Matsutani, Daihan Wang, Michihiro Koibuchi, Hideharu Amano, ``Performance and Cost Evaluations of On-Chip Network Topologies in FPGAs'', IEICE Technical Reports, RECONF2009-3, Vol.109, No.26, pp.13-18, May 2009, 査読無

(6) Michihiro Koibuchi, ``A Partially Network Reconfiguration Mechanism on Two-dimensional Mesh and Torus with Faults'', The IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN), pp.91-96, Feb 2009, 査読有

〔学会発表〕 (計 0 件)

〔図書〕 (計 1 件)

Michihiro Koibuchi, Hiroki Matsutani, "Chapter 3: Networks-on-Chip Protocols" in "Networks-on-Chips: Theory and Practice", Edited by Fayez Gebali, Haytham Elmiligi, Mohamed Watheq El-Kharashi, pp.65-94, CRC Press, Mar 2009

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

6. 研究組織

(1) 研究代表者

鯉渕 道紘 (KOIBUCHI MICHHIRO)

国立情報学研究所・アーキテクチャ科学研究系・准教授

研究者番号: 40413926

(2) 研究分担者

該当なし

(3) 連携研究者

該当なし