

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 24 日現在

機関番号：17102

研究種目：若手研究（A）

研究期間：2009～2012

課題番号：21680005

研究課題名（和文） オンチップ・スーパーコンピューティングを可能にするメニーコア・プロセッサの研究

研究課題名（英文） Manycore Processors for On-chip Supercomputing

研究代表者

井上 弘士（INOUE KOJI）

九州大学・大学院システム情報科学研究所・准教授

研究者番号：80341410

研究成果の概要（和文）：

本研究では、高性能かつ低消費電力な 3 次元積層メニーコア・プロセッサを見据えたアーキテクチャ技術を開発した。主な成果として、①3次元積層 LSI において大きな問題となる発熱に着目した性能評価、②複数のコアが強調して動作することで性能を改善する新しいメニーコア向け実行方式の確立、③メモリ性能を大幅に改善するキャッシュメモリ構成法の確立、ならびに、④将来のスーパーコンピュータの実現を見据えたインターコネクタ・シミュレータの開発、などの成果を挙げることができた。

研究成果の概要（英文）：

This research has developed several architectural techniques for three-dimensional integrated manycore processors. Main contributions of this research are: 1) we have evaluated the potential performance of 3D manycore processors by considering thermal impacts, 2) a novel manycore execution technique for high performance computing has been proposed, 3) new cache memory architectures for 3D manycore processors has been developed, and 4) an interconnect simulator for future supercomputing has been developed.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	6,500,000	1,950,000	8,450,000
2010年度	5,000,000	1,500,000	6,500,000
2011年度	5,300,000	1,590,000	6,890,000
2012年度	3,500,000	1,050,000	4,550,000
年度			
総計	20,300,000	6,090,000	26,390,000

研究分野：総合領域

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：メニーコア、3次元積層、高性能計算

1. 研究開始当初の背景

1 個の LSI に多くのプロセッサコアを搭載するメニーコアが大きな注目を集めている。

オンチップでの超並列処理を実現することで、消費電力を抑制しつつ、大幅な性能向上を期待できるためである。その一方、LSI デ

パイス技術としては、複数のダイを積層することで、高いメモリバンド幅の実現や、微細化に頼らない高集積化が可能となる3次元積層技術が実用レベルに近づいてきた。このように、プロセッサ・アーキテクチャ、ならびに、LSIデバイスの発展に伴い、今までにないプロセッサ・システムの実現が可能な時代へと突入した。

2. 研究の目的

本研究では、「メニーコアといったアーキテクチャ技術」ならびに「3次元積層といったデバイス技術」の融合を念頭に、次世代の高性能コンピューティングを支えるプロセッサ・アーキテクチャ技術の確立を目的とする。具体的には、3次元積層によって得られる利点を最大限活用しつつ、生じる欠点をアーキテクチャ技術でカバーする技術の開発、多数のコアを効率良く活用しより高い性能を実現する実行法の確立、さらには、将来のスーパーコンピューティングを見据えた評価環境の構築、を目指す。

3. 研究の方法

3次元積層デバイスの特性を考慮した各種詳細な評価を実施し、解決すべき課題を精査する。そして、従来の単純な並列処理に捕らわれない新しいプログラム実行法を考案し、その有効性をシミュレーションにより明らかにする。

4. 研究成果

以下、本研究で得た主な成果を示す。実際には、これらの他にも、近年注目を浴びているメニーコアの1種であるGPUでの電力性能特性解析など、多くの成果を挙げることができた。

(1) 3次元積層メニーコア・プロセッサの性能評価

複数コアの3次元積層を前提とし、温度制約を考慮したマルチコアの性能評価を行った。また、コア数の増加による並列性の向上と温度制約に起因する動作周波数の低下の間に存在するトレードオフを解析することで、今後の3次元積層マルチコア/メニーコアにおいて解決すべき課題を明らかにした。

本評価で対象とする3次元積層構造を図1に示す。また、前提とするプロセッサ・ダイの積層法は図2の通りである(Flip法とNon-flip法)。動作周波数が0.6GHzにおける3次元積層8コア(4層構造)の温度分布を図3に示す。これらの結果より、プロセッサ・ダイ数の増加に伴いプロセッサの温度が上昇していることが分かる。多くのダイを積層する場合にはより低い動作周波数に設定しな

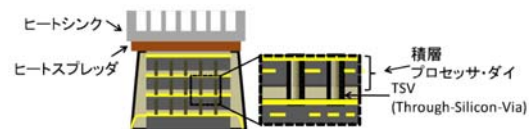


図1: 想定する3次元積層構造

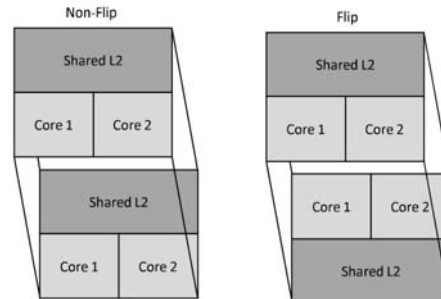


図2: 想定するプロセッサダイの積層法

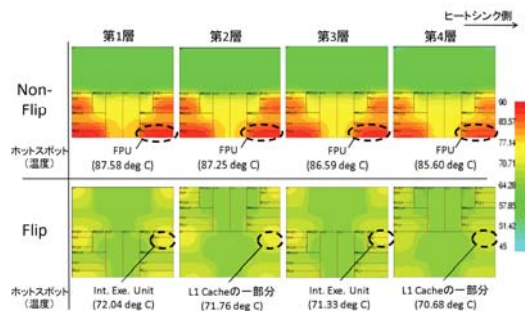


図3: 3次元積層プロセッサの温度分布

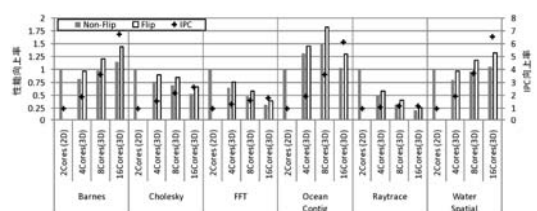


図4: 3次元積層プロセッサの性能

なければならない。一方、同じ層数においてコアが重複しないFlip方式では、Non-Flip方式と比較して低い温度となる。これは、コア内部のホットスポットである機能ブロック(浮動小数点演算ユニットなど)が垂直方向に重複せず、熱分散が達成されているためである。Flip方式ならびにNon-Flip方式のいずれの場合においても、ヒートシンクから一番離れたダイにおいて最も温度が高い場所が存在する。しかしながら、Non-Flip方式においてはコア内の浮動小数点演算ユニットを中心としたホットスポットが発生しているのに対し、Flip方式では熱が比較的拡散している様子が分かる。

3次元積層マルチコアは、既存の2次元実

装プロセッサと比較し多数のコアを有しているため、並列度を活かした処理の高速化が狙える。その反面、温度制約のため低い動作周波数となり、性能が低下する可能性がある。すなわち、コア数の増加による並列性の向上と温度制約に起因する動作周波数の低下の間にトレードオフが存在する。評価結果を図4に示す。各ベンチマークにおいて、2次元実装デュアルコア(2Cores-2D)の結果で正規化している。図中左の縦軸はGIPSに基づく性能向上率であり、各プロセッサ・モデルにおける2本の棒グラフ(Non-Flip方式とFlip方式)に対応する。図中右の縦軸はIPC向上率であり、メーカーで表している。これは、3次元積層化に伴う動作周波数の低下を加味しない値であり、3次元積層マルチコアにおいて2次元実装デュアルコアと同じ動作周波数を想定した場合に達成できる性能向上率となる。これらの結果より、十分な並列性を確保できない場合には、発熱に起因する動作周波数の低下の影響が顕著に現れ、3次元積層化による恩恵を得ることができないことが分かる。

(2) マルチ/メニーコア向けコア間協調実行技術

低ハードウェアコストでマルチコア/メニーコア・プロセッサのメモリ性能を改善することを目的として、プログラムの特徴に応じてプロセッサコアの動作を適切に変更する実行方式を提案し、その有効性を明らかにした。従来の並列処理では、全てのコアで並列プログラムを実行する。これに対して本提案方式では、実行対象プログラムがメモリ性能ボトルネックであるか否かを見極め、必要に応じて一部のコアをメモリ性能向上に用いる。これにより、ハードウェアコストを抑えつつ、演算時間とメモリ参照時間の和を小さくし、性能向上を達成する。

非共有オンチップメモリ搭載型マルチコアを対象としたオンチップメモリ貸与法を開発した。本手法では一部のコアをメモリ性能向上用に活用する。当該コアでは、プログラム実行を強制的に禁止し、オンチップメモリをプログラム実行用のコアに貸し出す。メモリ性能向上用コアを増やすことでメモリ性能が改善する一方、並列プログラム実行用のコア数が減るため演算性能が低下する。そのため、高性能化を達成できるようメモリ性能向上用コアの数を適切に決定する必要がある。また、メモリ性能向上用コアの追加による性能改善効果を高めるために、貸与メモリに有用なデータを読み込む。本手法は以下の流れで適用される。まず、対象プログラムを用いて事前実行することで並列性やメモリ参照トレースなどの統計情報を取出す。そし

て、データ毎の参照回数を元にメモリ性能向上用コアのデータ配置を求める。次に、各統計情報を入力として各々のコア配分における性能を予測し、メモリ性能向上用コア数を決定する。本手法では、主記憶アクセス回数削減による高性能化、ならびに、一部のコアの動作停止による低消費エネルギー化を実現できる。本手法をハードウェアの変更が不要なソースコード変換により実現し、実機において評価した結果、最大46%の実行時間の削減と32%の消費エネルギーの削減を達成した。

また、共有オンチップメモリ搭載型マルチコア・プロセッサを対象とするソフトウェア・プリフェッチ法を開発した。本手法では一部のコアをメモリ性能向上用に用いる。メモリ性能向上用のコアでは、ソフトウェア・プリフェッチを行うヘルパースレッドを実行する。ヘルパースレッド実行によるプリフェッチ効果によりメモリ性能が改善する一方、並列プログラム実行コア数が減るため、演算性能が低くなる。したがって、メモリ参照時間の減少が演算時間の増加を上回るよう、適切にヘルパースレッドを実行する必要がある。これを実現するために、ヘルパースレッドを占有実行するコア数を適切に決定する。また、コアがアイドルの時ヘルパースレッドを実行することで演算性能の低下を抑えつつ、メモリ性能を改善する。このヘルパースレッド実行法をOSのスレッドスケジューラの修正と3kB程度のFIFOバッファの導入により実装し、定量的に評価した結果、最大42%の性能向上を達成した。

(3) 3次元積層キャッシュ・アーキテクチャ

キャッシュメモリを大容量化するための手段の一つとして3次元積層の活用が注目されている。垂直方向にダイを積層し、TSV(Through-Silicon-Via)と呼ばれる層間金属柱などにより層間を直接接続することで、短い配線長を維持しつつ回路の大規模化が可能となる。しかしながら、キャッシュメモリを積層した場合には熱によるリーク消費電力の増加が問題となる。一般に、プロセッサコアなどのロジック部と比較して活性化頻度が低いL2キャッシュでは消費電力密度が低い。このため、平面実装時のL2キャッシュは比較的低温状態となる。これに対し、3次元積層キャッシュメモリでは、下層のプロセッサ・ダイの熱が伝導し、上層にあるキャッシュメモリが高温化する。リーク消費電力は温度に対して指数関数的に増加するため、特に3次元積層されたL2キャッシュではリーク消費電力の増加が深刻な問題となる。

そこで本研究では、3次元積層L2キャッシュのリーク消費電力削減を目的とし、温度分布の偏りを考慮した電源遮断方式を提案し

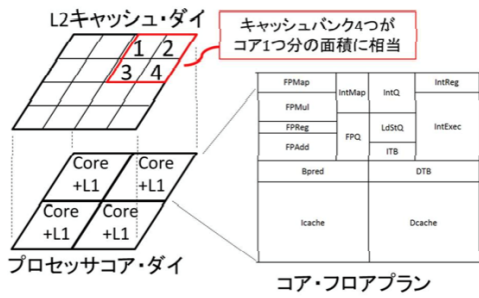


図 5：キャッシュメモリの積層

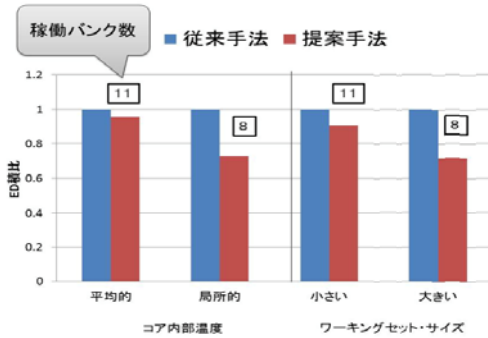


図 6：キャッシュ・リサイジングの効果

た。多数のコアを搭載したプロセッサを前提とし、特性の異なるプログラムを同時実行した場合の温度分布に基づき、キャッシュ・リサイジングを適用する。この際、下層プロセッサでのホットスポット発生状況を考慮して、性能低下を抑制しつつ、リーク消費電力を効率良く削減する。図 5 に示すように、プロセッサ・ダイに積層されるラストレベル・キャッシュはバンク構造をとっており、プロセッサ・ダイの発熱状況に応じて適宜キャッシュバンクの電源を遮断する。評価の結果、図 6 に示すように、エネルギー遅延積を 30% 程度削減できることが分かった。

(4) ライン共有キャッシュ・アーキテクチャ

現在、1 チップに複数のコアを搭載するマルチコア・プロセッサが主流となっている。また、更なる集積度の向上により、例えば 64 個以上のコアを搭載したメニーコアも現実味を増してきた。このようなマルチ/メニーコアでは、チップ内スレッドレベル並列処理により高い演算性能を達成できる。しかしながら、パッケージ制約による I/O ピンの制限等によりオフチップ・メモリバンド幅はコア数に対してスケールしない。また、主記憶として使用される DRAM の動作速度はプロセッサ速度と比較して極めて遅い。その結果、プロセッサとオフチップメモリの性能差拡大(いわゆる、メモリウォール問題)がより深刻化する。この問題を解決するために、多

くのマイクロプロセッサでは大容量なラストレベル・キャッシュ(LLC : Last Level Cache) を搭載している。例えば、Intel 社の Core i7 では 8MB、AMD 社の Barcelona においては 2MB の LLC を搭載しており、オフチップ・アクセス回数の削減を目的に年々大容量化する傾向にある。

一般に、LLC の実装においてはオフチップ・アクセス回数の削減が最重要課題となる。そのため、高い連想度かつ大容量なキャッシュ構成を採る場合が多い。特に、大容量化は容量性ミス削減するきわめて直接的かつ効果的なアプローチであり、高性能プロセッサの LLC 容量は増加傾向にある。たとえば、商用プロセッサである Intel 社の Core i7 では 4MB から 12MB もの大容量 LLC を搭載している。しかしながら、LLC の大容量化は多くのトランジスタ資源を必要とするため、チップ面積の増大やリーク消費電力の増加を招く。したがって、これらの問題をともなうことなく容量性ミス削減できる新しいキャッシュ構成法の考案が求められる。

通常、従来のキャッシュ構成法では、参照アドレス(タグ)とキャッシュラインが 1 対 1 に対応する。すなわち、異なるアドレスを有するデータは、その値そのものに関係なく、異なるキャッシュラインに格納される。その結果、キャッシュ内には、まったく同じデータ値を有する複数のラインが存在するといった状況が発生する。ベンチマーク・プログラムを用いた定量的解析を行った結果、一意なデータ値を有するラインは、1,000 回の書き込みデータに対して 10%以下の場合があることを確認した(すべてのベンチマークの平均では 26%)。これは今後 1,000 回の書き込みが起るまでに参照されるであろうデータのうち、90%以上のデータは冗長であることを意味する。このように、同一のデータ値を有するラインが多数存在する性質をデータ値の局所性と定義する。このデータ値の局所性が高い場合、キャッシュ内には同一の値を有するラインが数多く存在しており、キャッシュ容量を無駄に浪費していることを意味する。

そこで本研究では、キャッシュライン単位でのデータ値の局所性を活用した新しいキャッシュアーキテクチャとして、ライン共有キャッシュ(LSC: line sharing cache)を提案した。図 7 で示すように、従来型キャッシュとは異なり、LSC では複数のタグが単一のキャッシュ・ラインに対応することを許す。そして、アドレスは異なるが同一のデータ値を有するラインを 1 カ所の格納場所に割り当てることで、冗長なラインの記憶を回避する。これにより、従来型キャッシュと同じミス率を維持しつつ、実装すべきキャッシュラインのエントリ数を削減できる。これに加

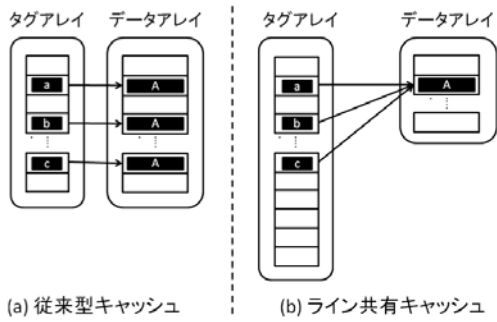


図 7: ライン共有キャッシュの概念

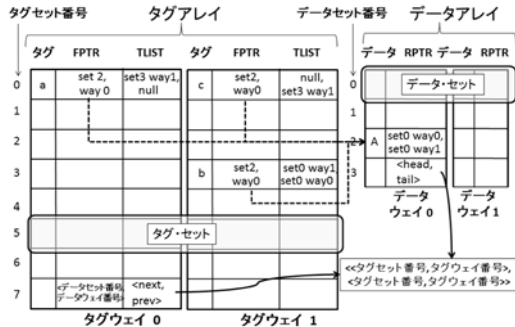


図 8: ライン共有キャッシュの内部構造

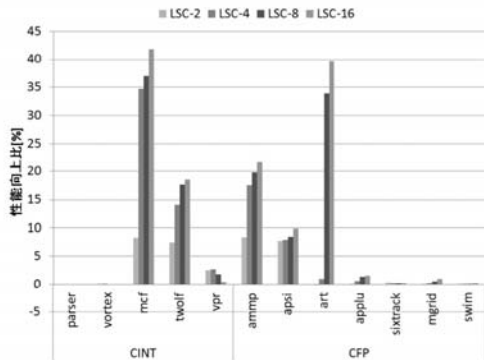


図 9: ライン共有キャッシュの性能

え、節約したトランジスタ資源を用いてタグエントリ数を増やす。これにより、同一面積制約下において従来型キャッシュよりも多くのタグエントリを格納することが可能となり、大幅な容量性ミスの削減を実現する。LSC の内部構造を図 8 に示す。LSC では、タグとデータの関連をポインタによって管理することで、柔軟なデータ管理が可能となる。ベンチマーク・プログラムを用いた定量的な評価を行った結果、図 9 に示すように、最大で 40%以上の性能向上を実現できることが分かった。

(5) スーパーコンピューティングのためのインターコネクト・シミュレータ開発

大規模並列計算機ではより多くの計算ノ

ードを相互結合網により接続することで総演算性能を高めており、近い将来、数 10 万ノードを接続する大規模並列計算機が主流となると予測される。このような大規模並列計算機向けにアプリケーションの性能チューニングや通信ライブラリ、通信アルゴリズム、および、相互結合網の制御方式の開発を効率的に進めるためには、相互結合網における通信の性能を予測する必要がある。しかしながら、相互結合網の性能は、輻輳やプロセスのロードインバランスなどに影響されるため、静的および定性的に予測することは困難である。したがって、通信関数やスイッチの振る舞い、およびノード間でのパケット転送を的確にシミュレーションできるツール(シミュレータ)が必要となる。その他、通信パターンの記述インターフェースの柔軟性、対応する通信パターンおよび取得可能な情報の多様性、結果の精度、シミュレーション速度なども、シミュレータに求められる要件である。シミュレーション対象の大規模化に伴う処理量の増加を抑えることは難しいため、シミュレーションを現実的な時間内で完了させるためには単位処理量あたりのシミュレーション時間を低減しなければならない。このためには既存のシミュレーションモデルの見直しが必要である。

一方、シミュレータの実行環境としてマルチコア・プロセッサを搭載した並列計算機が普及している。特徴は、物理メモリを共有するプロセッサ・コア間では共有メモリを介して低いレイテンシで通信可能であること、および旧来の大規模メモリ共有並列計算機と比較して容易かつ安価に入手できることである。ただし、物理メモリを共有するプロセッサ・コア数は限られているため、低いレイテンシでの通信が可能ない並列数に制約があり、利用可能メモリ容量も制限される。また、高い並列性や多くのメモリを利用する場合にはネットワークを介した通信が必要となり、通信レイテンシが高まる。既存のシミュレータは、並列性もしくは低レイテンシ通信のいずれか片方の特徴を利用するように設計されている。並列性と低い通信レイテンシとを両方利用するためには、大規模並列計算機だけでなく小規模メモリ共有並列計算機での動作を考慮してシミュレータを設計することが望ましい。

そこで本研究では、相互結合網シミュレータ NSIM を開発した。本シミュレータでは、数 10 万ノードを接続する将来の相互結合網を対象として、様々な通信パターンを現実的な時間内でシミュレーションすることが可能である。シミュレーション処理量およびメモリ消費量が減るようにシミュレータを設計した。これにより、小規模メモリ共有並列計算機上での大規模シミュレーションも可能と

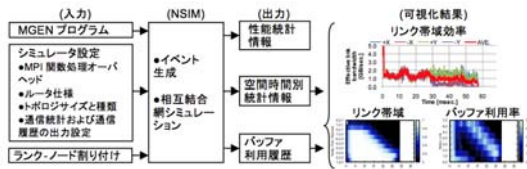


図 10 : NSIM の入出力

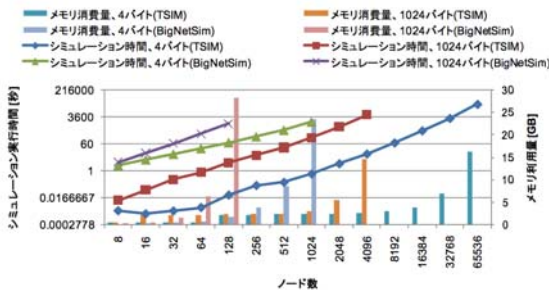


図 11 : NSIM の性能

なる。シミュレーション対象通信パターンは、図 10 に示すように、NSIM が提供する MPI 互換インタフェース (MGEN API) を利用しメッセージパッシング方式の並列プログラム (MGEN プログラム) として記述する (C 言語のみ可能、FORTRAN は未対応)。MGEN プログラムは MPI プログラムと同様の要領で記述できる。MGEN API には MPI には定義されていない拡張関数が含まれており、複数の NIC を同時に利用した通信、ゼロ・コピー通信などを指定したシミュレーションが可能である。

図 11 に、NSIM ならびに関連研究である BigNetSim のシミュレーション実行時間およびメモリ消費量を示す。横軸はノード数、左縦軸はシミュレーション実行時間、右縦軸はメモリ消費量を示す。線グラフがシミュレーション実行時間、棒グラフが消費メモリ量に対応する。今回のシミュレーション実行環境では、NSIM は BigNetSim と比較して 2700 倍 (1024 ノード、4 バイトメッセージ)、370 倍 (128 ノード、1024 バイトメッセージ) 高速に動作することが確認できる。

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 37 件)

- ① 岡慶太郎, 佐々木広, 阿部祐希, 井上弘土, 村上和彰, “データ値の局所性を利用したライン共有キャッシュ,” 情報処理学会論文誌 コンピューティングシステム (ACS), Vol. 5, No. 4, pp.36-47, 2012 年, 査読有り
- ② Hideki MIWA Ryutaro SUSUKITA
Hidetomo SHIBAMURA Tomoya
HIRAO Jun MAKI Makoto
YOSHIDA Takayuki KANDO

Yuichiro AJIMA Ikuo MIYOSHI
Toshiyuki SHIMIZU Yuji OINAGA
Hisashige ANDO Yuichi INADOMI
Koji INOUE Mutsumi AOYAGI
Kazuaki MURAKAMI, “NSIM: An Interconnection Network Simulator for Extreme-Scale Parallel Computers,” IEICE TRANSACTIONS on Information and Systems, Vol.E94-D, No. 12, pp.2298-2308, 2011, 査読有り

- ③ 井上弘土, 安藤 壽茂, 薄田 竜太郎, 山村 周史, 柴村 英智, 三輪 英樹, 本田 宏明, 稲富 雄一, 眞木 淳, 平尾 智也, 青柳 睦, 村上 和彰, 石附 茂, 小松 秀実, 安島 雄一郎, 三吉 郁夫, 清水 俊幸, 黒川 原佳, “大規模スーパーコンピュータ向けシステム性能評価環境の構築,” 情報処理学会論文誌 コンピューティングシステム, Vol.4, No.1, pp.1-23, 2011, 査読有り

[学会発表] (計 8 件)

- ① Yuki Abe, Hiroshi Sasaki, Koji Inoue, Kazuaki Murakami, and Shinpei Kato, “On the Power and Performance Analysis of GPU-Accelerated Systems,” Poster Session, USENIX Annual Technical Conference, 2012, 査読有り
- ② Koji Inoue, “3D memory architecture,” D43D: 3rd Design for 3D Silicon Integration Workshop, 2011, 招待講演
- ③ Koji Inoue, “Adaptive Execution on 3D Microprocessors,” 11th International Forum on Embedded MPSoC and Multicore, 2011, 招待講演

[図書] (計 0 件)

[産業財産権]

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

[その他]

なし

6. 研究組織

(1) 研究代表者

井上弘土 (INOUE KOJI)

九州大学・大学院システム情報科学研究所・准教授
研究者番号: 80341410

(2) 研究分担者

なし

(3) 連携研究者

なし