

VI. 高性能計算システム研究部門

1. メンバー

教授	佐藤 三久、朴 泰祐、児玉 祐悦、高橋 大介
准教授	建部 修見、埴 敏博
講師	川島 英之
助教	多田野 寛人
研究員	田中 昌宏、梅田 宏明、中尾 昌広、奴賀 秀男、松本 和也、Jabri Mohamed Amin

2. 概要

本研究部門では、高性能計算システムアーキテクチャ、並列プログラミング環境、GPU利用技術、並列数値処理の高速化研究、広域分散環境におけるデータ共有を中心とするグリッド計算技術等の研究を行っている。

3. 研究成果

【演算加速機構を持つ将来の HPCI システムに関する調査研究】(佐藤、朴、児玉、高橋)

本調査研究においては、将来の HPCI システムのあり方の調査研究に対して科学技術計算のアプリケーションに適合した演算加速機構による並列大規模システムを提案し、調査研究を行った。

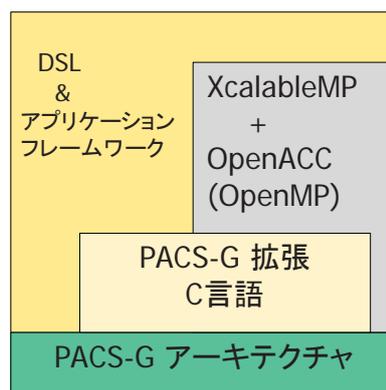
本システムのサイクルベースの評価に関して、当該年度においては、昨年度開発したシングルチップ向けサイクルベースシミュレータをベースとして、マルチチップによるシミュレーションを可能とする開発を行った。主な改良点は、(ア) グローバルメモリ (GM) 機能、(イ) チップ内通信機構の 3D/4D メッシュへの対応、(ウ) チップ間通信機構の 3 点である。GM 機能はオフチップの高バンド幅メモリをシミュレートするもので、各ブロードキャストメモリ (BM) に GM の各パーティションが一对一に対応する形で同時転送を行う。チップ内通信機構は 3D/4D メッシュへの対応するために 3 ホップ先のコアと直接接続するバイパスポートをサポートした。チップ間通信機構は、4 次元トーラスをサポートし、チップ間は 1 対 1 の接続を行う。チップ間通信のためのバッファとして BM と同様に各コアと接続された CM を 2 次元メッシュの 4 方向にもち、コアから CM への転送および CM からコアへのデータ転送はコア命令で行い、チップ間の通信は DMA 起動で行う。本サイクルベースシミュレータを用いて GM 利用および複数チップ利用の場合の性能評価を行った。アプリケーションとしては地震波計算コードである Seism3D を利用した。これにより、ピーク性能を用いたアプリケーション性能見積もり方法が妥当であることを確認した。

また、本研究において、演算加速プロセッサ PACS-G 向けプログラミング・モデルの検討を行った。PACS-G のプログラミング・モデルを検討するために、まず、ソフトウェアからのアーキテクチャを明らかにする命令レベルシミュレータを開発した。次に、アーキテクチャの詳細を記述できる、低レベルプログラミング向けの記述言語 PACS-G 拡張 C 言語を設計し、PACS-G の命令レベ

ルシミュレータ用にコンパイラを実装した。この知見をもとに、一般プログラマ向けに指示文ベースでプログラミングをするための並列言語 PACS-G 拡張 XcalableMP を設計した。また、プロセッサ間の通信ライブラリ API についても設計を行い、シミュレータに実装した。

想定されるプログラミング環境のソフトウェアスタックを図VI-1 に示す。

- (1) 低レベルプログラミングのための、PACS-G 拡張 C 言語。PE の制御構文、Intrinsic 命令により命令レベルの記述が可能
- (2) XcalableMP 拡張 (+ OpenACC) による指示文ベースのプログラミング環境を提供。
- (3) 既存のコード (主にステンシル計算) の移植を容易する。Domain-Specific Language (DSL) とアプリケーションフレームワーク。例えば、粒子法のための DSL 等、が考えられる。
- (4) さらに、ホストと協調してプログラミングする場合には、ホストからのオフロードのための指示文ベースのプログラミング環境をサポートする。低レベルではホストインタフェースライブラリを提供する。
- (5) プロセッサ間については、MPI もしくは専用メッセージ通信ライブラリを想定。



図VI-1 PACS-G システムのプログラミング・ソフトウェアスタックの構成図

拡張 C 言語は、低レベルの詳細なプログラミングのためのもので、一般のアプリケーションユーザーには煩雑である。そのために、指示文ベースで、分散メモリ向けの PGAS 言語である XcalableMP をベースに、OpenACC と組み合わせて、Fortran のアプリケーションの PACS-G 向け移植を容易にするために設計した。

以下に、特徴を挙げる。

- (1) 基本的には、OpenACC の拡張

XcalableMP の枠組みを用いて、OpenACC で PE にオフロードした時のデータ配置を指定する。template を拡張して、PE/LM を仮想化する。

- (2) template 指示文

PE 上の仮想 index 空間を定義する。分散方法については、block 分散のみを想定する。

(3) data 指示文

OpenACC の拡張で、PE にオフロードされたデータについて PE でのデータ環境を設定する。データは当初は、GM またはマスタプロセッサ側にあることを前提とする。

copy/copyin/copyout/create/present/present_or_copy/present_of_copyin/present_or_copyout/prpresent_or_create により、PE 側にセットアップするデータを指定することができる。

(4) align 指示文 / shadow 指示文

align 指示文 により、template に従い、分散配置を指定。align で指定されない場合には、それぞれの PE で重複。align で分散された配列に関しては、shadow 指示文で、袖領域を確保を指定できる。

(5) parallel 指示文

各 PE で実行するコードを指定。

(6) loop 指示文

ループを分割実行する。on 節で、並列の index スペースを指定。reduction 節で、reduction する変数を指定する。

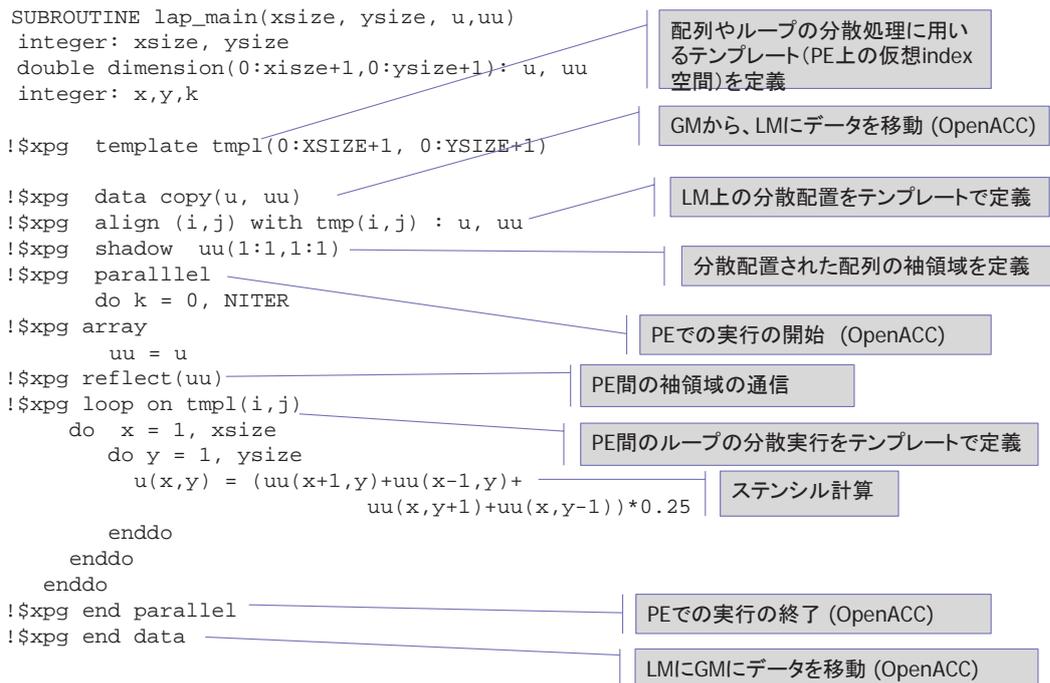
(7) reflect 指示文

shadow 領域の update を実行する。

(8) update directive

data 環境中で、GM/マスタプロセッサと PE のデータをやり取りする。

以下に、laplace 方程式の陽解法のステンシル計算の例を示す。



【PGAS モデル並列プログラミング言語 XcalableMP の開発と HPC Class2 Award の受賞】 (佐藤)

XcalableMP は、PC クラスタコンソーシアム XcalableMP 規格部会でコミュニティで議論されている PGAS モデルの並列プログラミング言語である。筑波大と理研でレファレンス実装 Omni XcalableMP コンパイラを開発している。

当該年度においては、11 月に開催される SC で行われている HPC Class2 ベンチマークに京コンピュータでの XcalableMP の結果を提出し、Award を獲得することができた。

【Xeon Phi 向け XcalableMP 実行時システムの設計】 (佐藤)

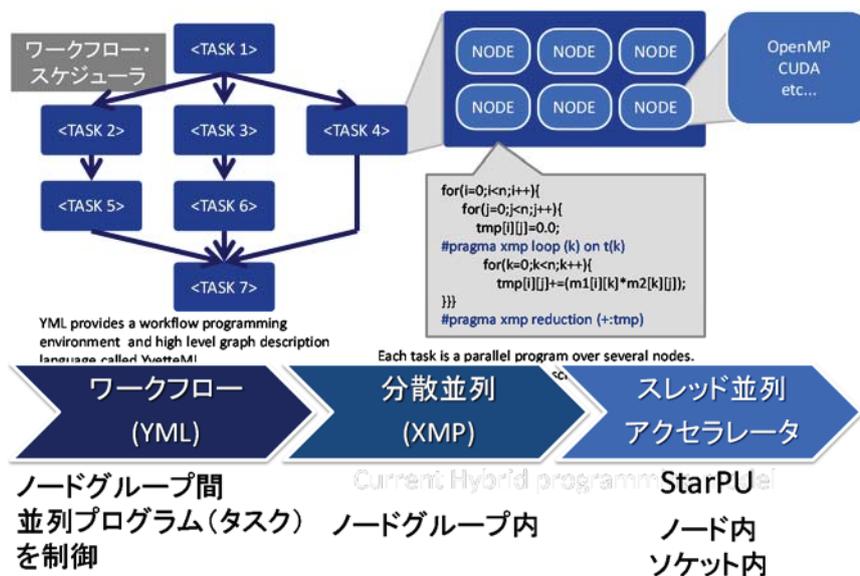
PGAS 言語 XcalableMP で記述したステンシル計算をマルチノードのメニーコアプロセッサ Xeon Phi 上で効率よく実行するための実行時システムを試作・評価した。この実行時システムでは、プログラムの並列実行にはプロセスを用い、ステンシルコード上のグローバル配列は並列数でブロック分割し、それぞれに袖領域を加えた上で共有メモリ上に置く。配列のブロッキングはメモリ参照の局所性を高め、さらに袖領域の追加により、ブロック間の境界領域を異なるコアがアクセスすることによるキャッシュラインのシェアリングを削減できるため、演算性能が向上する。また、袖領域交換のために発生するノード内のプロセス間通信を共有メモリ上での規則的なメモリ転送に置き換えることができるため、この通信時間も削減することができる利点がある。評価は、よる Laplace 方程式の計算と姫野ベンチマークの 2 つのベンチマークを用いて行った。この結果、姫野ベンチマークにおいては、2 ノードから 16 ノードのマルチノード上で、従来プロセッサ向きに実装された実行時システムを用いた場合の 2 倍程度の性能が得られることが確認

できた。

【階層的並列コンポーネントベース FP2C プログラミングモデルの研究】(佐藤)

JST 国際科学技術共同研究推進事業（戦略的国際共同研究プログラム）課題名「ポストペタスケール・コンピューティングのためのフレームワークとプログラミング」において、フランスとの共同研究を実施した。このなかで、ポストペタスケール・コンピューティングのための並列プログラミング基盤として、並列プログラミング言語拡張仕様 XcalableMP (XMP) を用いて、他のグループで研究開発されるコンポーネントへのインターフェースを与えるとともに、上位のプログラミングパラダイムとして、ワークフロー言語である YAML を用いる。XMP で記述されたコンポーネントからは、OpenMP もしくは、後述する GPU への拡張仕様 XMP-dev を用いて、GPU を利用することができる。XMP-dev からは、実行時スケジューラ StarPU を用いることにより計算を GPU と CPU に自動的に振り分けることができる。その階層的なプログラミングモデル FP2C (Framework for Post-Petascale Computing) を図 1 に示す。

当該年度においては、YML の開発を担当するベルサイユ大学グループと INRIA サクレグループ、理化学研究所のグループと協力し、京コンピュータ上で性能評価を行った。この FP2C を用いて、階層構造を持つブロックガウスジョルダン法による密行列解法プログラムをこのプログラミングモデルで記述し、その有効性について評価した。具体的には、XMP で並列行列積のコンポーネントを記述し、そのコンポーネントを用いたガウスジョルダン法を YML で記述する。評価の結果、逆行列を計算するベンチマークにおいて FP2C で記述したプログラムが、通常の MPI で記述したプログラムよりも性能が良いことを確認し、その有効性について検証した。



図VI-2 FP2C による階層的プログラミングモデル

【エクサスケールの気候シミュレーションのための耐故障技術およびプログラミングモデルの研究】 (佐藤)

多国間国際研究協力事業 (G8 Research Councils Initiative) の研究課題「エクサスケールの気候シミュレーションを実現する技術(Enabling Climate Simulation at Extreme Scale)」において、エクサスケールの気候シミュレーションのための耐故障技術およびプログラミングモデルの研究を行った。

耐故障技術については、共同研究者である INRIA の Cappello 博士のグループが研究を進めている、ハイブリッドなチェックポイント手法(Franck)を利用する言語機能について、言語サポートとクラスタに分割するための高速アルゴリズムを設計・実装・評価を行った。気象コードをはじめとする典型的な科学技術計算のアプリケーションでは、時間発展等のメインループを持ち、ループごとに同じ通信パターンになることが多い。ここに注目し、メインループを指示文を用いて指定することにより、ここにおいて、最初の繰り返しで通信パターンを検出し、次回以降のチェックポイントにおいて、これを利用する。ハイブリッドチェックポイント手法は、局所的なクラスタ間でチェックポイントするメッセージロギングと、クラスタで協調チェックポイントを組み合わせたもので、その分割がチェックポイントの効率を左右する。ここでは、最初に通信パターンを検出し、実行時にクラスタ分割を計算しなくてはならないために、クラスタ分割自体の時間も重要である。本グループでは、通信する各プロセスが実行されているノードを分割の単位とし(node-based)、各ノードは正規的なグリッド・トポロジーで構成されている(grid-based)とすることにより、短時間に最適化分割を行うことができる。これを、気象のミニアプリである GC-POP や他の科学技術計算コード(GeoFEM)に適用した結果、一般的なグラフ分割アルゴリズムよりも大幅に短時間でグラフ分割が可能で、得られた分割も同程度の最適化(エッジカット等)が行えることを示した。

また、プログラミングについては、理研計算科学機構にある京コンピュータにおける気象コードの性能評価と最適化に取り組んだ。特に、京コンピュータにおいては片方向通信を用いて気象コードの通信を行い、性能の改善を行った。本グループでは、GAS(Partitioned Global Address Space)モデルに基づく並列プログラミング言語 XcalableMP を開発しており、片方向通信にはこの言語の機能の1つである Coarray を用いることにより、従来の MPI 通信よりも簡便に記述することができる。その結果、GG-POP ではハードウェアによるリダクション通信の利用と合わせて、84%(7500 ノード)の性能改善、NICAM においては 7%(640 ノード)、CICE においては 15%(1024 ノード)の性能改善をすることができた。特に、ノード数が多い大規模なプログラムで性能が改善でき、将来のエクサスケール計算においては、片方向通信による通信の最適化が重要であることが分かった。

【HA-PACS/TCA の研究開発】 (朴、児玉、埴)

文部科学省特別経費「エクサスケール計算技術開拓による先端学際計算科学教育研究拠点の充

実」(H23～H25 年度の 3 年間)において、密結合演算加速装置 (Tightly Coupled Accelerators: TCA) を提案し、中核となる PEACH2 (PCI Express Adaptive Communication Hub ver.2)チップを FPGA により実装し、これを汎用 GPU クラスタノードに実装するための PCI Express ボードのハードウェアを完成させ、これを量産したバージョンを全計算ノードに搭載した HA-PACS/TCA を構築した。

図VI-3 に量産型の PEACH2 ボードの写真を示す。ボード中央に放熱フィンと共に FPGA (Altera Stratix-IV) が設置され、DRAM メモリ、他のノードとの外部 PCIe リンク接続のためのポート (計 3 ポート) とマザーボードに接続される PCIe スロットを備えている。ボードの右側 1/3 程度のエリアは安定化電源のためのレギュレータ類である。



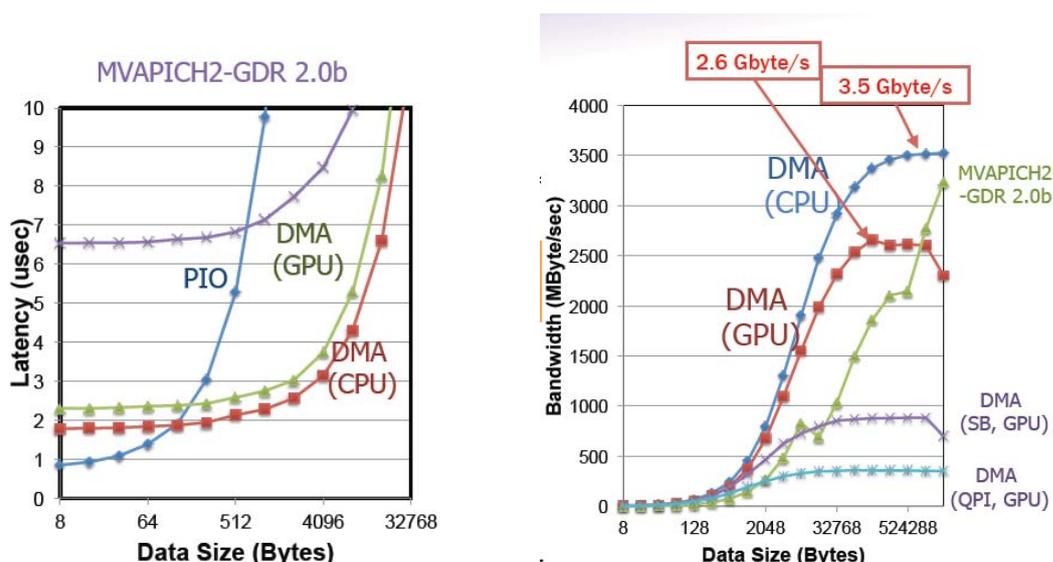
図VI-3 PEACH2 ボード (量産型)

HA-PACS/TCA (以下、TCA 部) は平成 25 年 10 月に導入され、平成 24 年 2 月から稼働開始した HA-PACS ベースクラスタの拡張部分として接続された。ベースクラスタと TCA 部は共通の InfiniBand ネットワークスイッチに接続され、ファイルサーバ等を共有する。TCA 部がベースクラスタと大きく違う点は、全ノード (64 ノード) に PEACH2 ボードを装着し、TCA 技術による低レイテンシの GPU 間直接通信が可能であることと、GPU として NVIDIA 社の最高性能製品である Tesla K20X を搭載している点である。なお、K20X を用いた Linpack 性能は効率が 70%以上を達成し、TCA 部のみでの Linpack 性能は 277 TFLOPS、電力当たりの Linpack 性能は 3518MFLOPS/W で、2013 年 11 月の Green500 リストで世界第三位にランクされた。

TCA 部の完成により、これまでの 2 ノード間での基本通信性能評価に加え、多数ノードでの集合通信ライブラリやアプリケーションへの適用を開始した。TCA 部完成後の本格的なシステム及びアプリケーションソフトウェアの開発は JST-CREST 研究領域「ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出」における「ポストペタスケール時代に向けた演算加速機構・通信機構統合環境の研究開発」(研究代表者: 朴泰祐) に引き継がれている。PEACH2 による TCA 通信は基本通信性能のチューニングが進み、平成 25 年度末時点において、最小レイテンシを約 2 μ 秒まで縮め、PC クラスタにおける事実上の標準通信技術である InfiniBand FDR

の約 1/3 という極めて短いレイテンシを実現している。また、256KB 程度までのデータ転送においては同ネットワークを上回るバンド幅を達成しており、加えて DMA チェイニングやブロックストライド転送等、典型的な科学技術計算における様々な通信パターンの最適化を実現している。

図VI-4 に point-to-point での通信性能を示す。TCA 部において Intel の IvyBridge アーキテクチャプロセッサの CPU を導入したことにより、CPU 内蔵の PCIe スイッチの性能が改善され、ノードを跨ぐ GPU 間通信でも十分なバンド幅が達成された。



図VI-4 最適化された TCA 機構による通信基本性能 (左：レイテンシ、右：バンド幅)

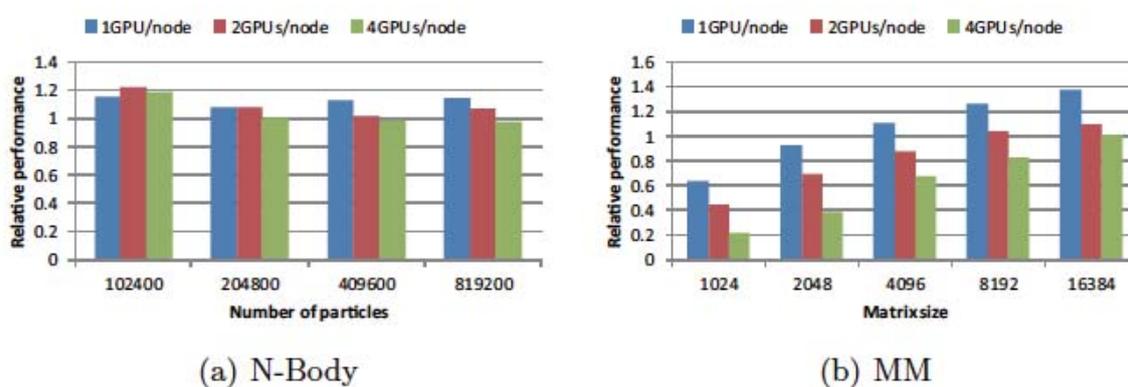
TCA を用いたアプリケーションに関しては、NVIDIA 社で開発が続けられている GPU 向け QCD ライブラリである QUDA を、同社との共同研究により TCA に移植した。オリジナルは MPI によって記述されているが、これを TCA に移植するために、片方向通信をベースとした RMA (Remove Memory Access) ライブラリを新たに定義し、通信部分をこれによって書き換えた。その上で、RMA ライブラリの TCA 実装と、比較対象となる MPI-3 の片方向通信を用いた実装を行い、比較した。通信がボトルネックとなるような小さな計算モデルにおいて、TCA を用いた strong scaling が有効であることを確認している。また、より一般的な数値計算ライブラリとして、並列 CG 法の GPU 実装に TCA を適用した。NAS-PB 等の CG 法実装を参考に、各 GPU 上での疎行列ベクトル積ルーチンの他、GPU 上の計算経過ベクトルの reduction、及び各部分ベクトルの alltoall 通信という 2 種類の集合通信を TCA により実装した。本研究に関しては進行中であり、今後性能チューニングを順次行っていく。

【GPU クラスタにおける並列言語ベースでの GPU/CPU ワークシェアリング】(佐藤、朴)

平成 25 年度末をもって終了した戦略的国際科学技術協力推進事業（日仏共同研究）「ポストペタスケールコンピューティングのためのフレームワークとプログラミング」において、高レベル並列プログラミング環境での CPU/GPU ワークシェアリングに関する研究を行った。理化学研究所と共同開発している XcalableMP の GPU 対応版である XMP-dev を対象に、並列化されたプログラムに対し、各計算ノード上のランタイムシステムによって GPU と CPU コアの全計算リソースを有効利用する動的負荷分散システムを開発した。本研究ではフランス INRIA との共同研究に基づき、彼らが開発した StarPU ランタイムシステムを我々の XMP-dev コンパイラと協調動作させることにより、本システムを実現した

StarPU 本体はユーザにとって極めて扱いやすく、負荷分散の対象となる並列タスクの定義やデータの入出力を煩雑な記述によって表さなければならず、実応用の障壁となっていた。XMP-dev とのカップリングにより、ユーザは通常の GPU を対象とした XMP-dev 言語での記述を行うだけで、StarPU の GPU/CPU 動的負荷分散機能を利用できる。その際、GPU と CPU コアという演算性能が大幅に異なるリソースにどのようなサイズのタスクを割り当てるかという、タスク分割方法が性能上の問題となることを突き止め、この分割自体を動的に変更する機能を追加した。例えば、反復計算において GPU と CPU コアの負荷状況を観察し、簡単なアルゴリズムによってその比率を変更することが可能である。

図VI-5 に、重力多体問題（N-Body）及び行列行列積（MM）において、GPU のみを用いた XMP-dev プログラム実行に対し、本システムによる GPU/CPU ワークシェアリングを最適に適用した場合の性能向上を示す。アプリケーションや問題規模、使用 GPU 数に依存するが、最大で 1.2～1.4 倍の性能向上が実現された。



図VI-5 XMP-dev/StarPU による GPU/CPU ワークシェアリングの性能向上

【核融合シミュレーションコードの並列 GPU 化に関する研究】（朴）

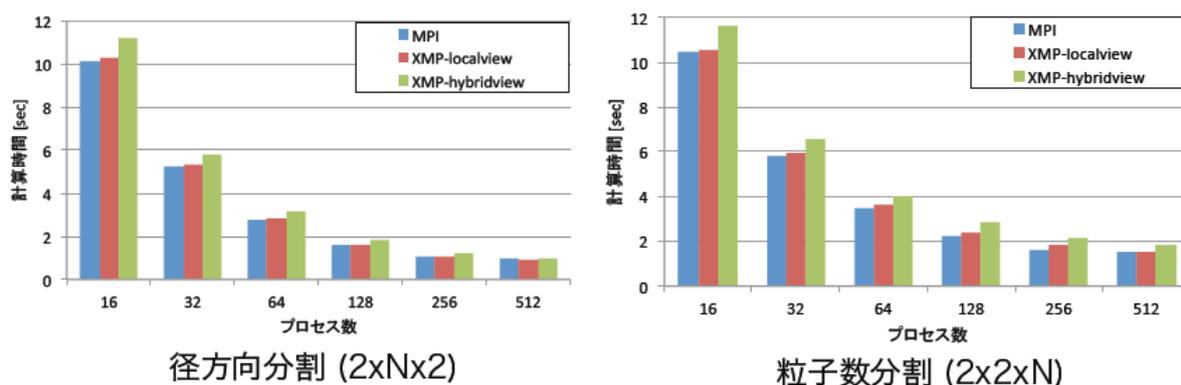
JSPS 多国間国際研究協力事業（G8 Research Councils Initiative）における研究課題「エクサスケ

ール規模の核融合シミュレーション」(H23～H25 年度、研究代表者：朴泰祐)において、日本原子力研究開発機構及び米国プリンストン大学との共同研究により、次世代核融合シミュレーションコードの並列 GPU 化を行った。前者においては「京」等で開発中の GT5D コードを、後者においては BG/Q 等で開発中の GTC-P コードを対象とする。

GT5D コードについては、PGI CUDA Fortran を用いて HA-PACS ベースクラスタへの移植を行い、全コードを GPU 化し、データに関しては全てを GPU メモリに保持するようにコードを作成した。MPI 通信と GPU・CPU 間データ転送のオーバーラップ、データ再配置を伴う MPI 通信における CPU と GPU での処理の分散、各カーネルの最適化等の性能向上を実現した。最終的に、HA-PACS ベースクラスタにおいて最大 16 ノードを用いた場合、CPU のみを用いた場合に対して 1.91 倍の性能が得られた。

GTC-P コードについては並列記述言語 XscalableMP (XMP)の応用例としてコーディングを行った。ここでは、XMP による実アプリケーションとして初めての試みである、global view と local view の hybrid view 記述による PIC (Particle In Cell)コードの記述を実現した。GTC-P は PIC コードの一種であるが、静的な空間(場)の物理の記述と、電子等の粒子が動的に動き回るという二種類の振る舞いに対し、データ及び処理の分散の記述のし易さとプロセス間の負荷の変動を同時に記述する必要がある。従来の MPI 実装では全てを詳細に記述する必要があったが、XMP の hybrid view 記述では、静的な空間の分割は XMP の宣言文だけでデータ分割・処理分割・分散データの同期を簡潔に記述でき、プログラムの生産性が向上する。一方、動的要素が強い部分については coarray 記法による片方向通信で記述し、同様に MPI における煩雑な配列記述のようなプログラム生産性低下要因が排除できる。

図VI-6 に HA-PACS ベースクラスタ(CPU のみ使用)における strong scaling の性能評価を示す。オリジナルの MPI 実装、XMP による coarray 記述で単純に MPI 通信を置き換えた場合、hybrid view による記述の場合について、並列化による処理時間の短縮を示している。hybrid view では他の 2 方式に比べ最大で若干処理時間が延びているが、これは配列のインデックス計算を自動化している部分でのオーバヘッド等が原因で、今後コンパイラの改良を行う予定である。hybrid view による生産性の向上を考慮すると十分十分許容可能な性能差と考えられる。

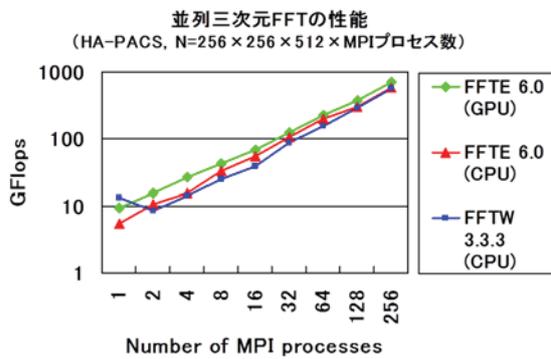


図VI-6 XMP hybrid-view による GTC-P コードの strong scaling 性能（左右は並列化における分割方向を変えた場合で、どの分割方法でもほぼ同じ傾向の高速化が行えている）

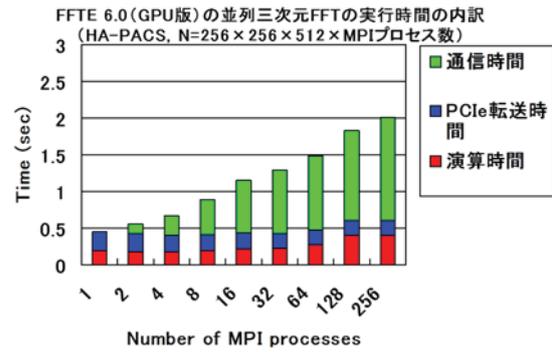
【並列高速フーリエ変換に関する研究】（高橋）

科学技術計算で広く用いられている並列高速フーリエ変換（FFT）の性能を改善するために、高速化手法に関する研究を行った。ポストペタスケール計算環境の一つと考えられる GPU クラスタにおいて並列 FFT を実行する際に、どのようなアルゴリズムや実装が望ましいかについて検討を行った。GPU クラスタにおいて並列 FFT を行う際には、全対全通信が複数回行われることから、計算時間の大部分が全対全通信によって占められることになる。さらに CPU と GPU 間を接続するインターフェースである PCI Express バスの理論ピークバンド幅は PCI Express Gen 2 x 16 レーンの場合には一方向あたり 8GB/sec となっていることから、CPU と GPU 間のデータ転送時間を削減することも重要になる。そこで、CPU と GPU 間のデータ転送と計算ノード間の MPI 通信をパイプライン化してオーバーラップすることができる MPI ライブラリである MVAPICH2 を用いることで、この問題を解決した。さらに、FFT の処理において出現する行列の転置の処理を GPU 上で行うなどの工夫も行った。

その結果、GPU クラスタにおける並列 FFT が PC クラスタにおける並列 FFT よりも高速に実行できることを確認した。



HA-PACS における並列三次元 FFT の性能

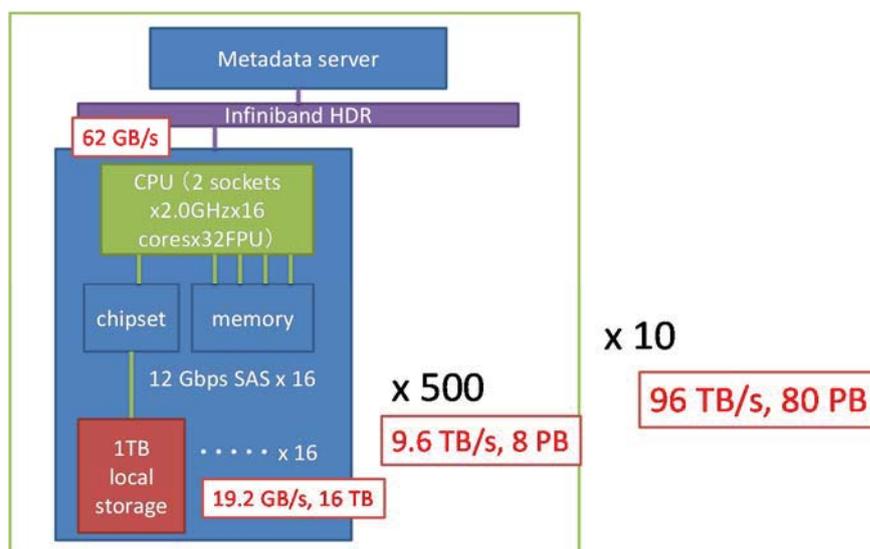


HA-PACS における並列三次元 FFT の実行時間の内訳

図VI-7 高性能 FFT の性能評価

【データインテンシブサイエンスのためのシステムソフトウェア】(建部、川島)

ビッグデータ解析が必要となるデータインテンシブサイエンスをポストペタスケールシステムで実施するため、分散ファイルシステム、計算ノード OS、実行時システムの設計を行い、プロトタイプ実装による評価を行った。現行の分散ファイルシステムは、メタデータサーバの性能、データアクセス性能のいずれにおいてもポストペタスケールシステム、エクサスケールシステムにおいては十分な性能を発揮することができない。ポストペタスケール以降のシステムにおいて必要となるメタデータサーバの性能、データアクセス性能を達成するためには、アーキテクチャ、システムソフトウェアの転換が必要となる。本研究では、ポストペタスケールシステムにおいて必要とされる性能を達成するため、以下のようなアーキテクチャを想定し、システムソフトウェアの研究開発を実施している。



図VI-8 ポストペタスケールデータサーバのアーキテクチャモデル

各計算ノードは高速なローカルストレージを備える。2015 年を想定すると、たとえば 12 Gbps SAS を 16 本用い、フラッシュストレージあるいは不揮発性 RAM を利用することで 19.2 GB/s の I/O バンド幅を実現可能と考えられる。この計算ノードを 5,000 ノード、ファイルシステムを管理するメタデータサーバを 10 ノード利用することにより、96 TB/s の I/O バンドを実現することが可能となる。IOPS については、理論的な性能を出すのは難しいが、経験的に 1 メタデータサーバでは 100K IOPS ほどは処理可能であるため、スケールアウトできれば 10 ノードで 1M IOPS が達成可能となる。

上記の性能はハードウェアの限界性能であるが、実際にシステムソフトウェアでこの性能を達成するためにはいくつかの問題を解決する必要がある。分散ファイルシステムについては、1M IOPS、100 TB/s を目指してアーキテクチャ設計、システム設計を実施した。途中結果であるが、分散 KVS に対してノンブロッキング分散トランザクション処理を行うことにより、15 サーバで 250K IOPS を達成した。また、フラッシュストレージ、不揮発性 RAM を効率的に利用して高い I/O 性能を達成するためのローカルストレージ、ローカル KV ストアの設計、評価を行った。更には、遠隔の計算ノードのローカルストレージに格納されているデータを高速にアクセスするためのネットワーク、システムの設計、計算ノードに障害が発生した場合にデータを消失しないための複数ノードのストレージへの冗長書込手法の設計、評価を行った。

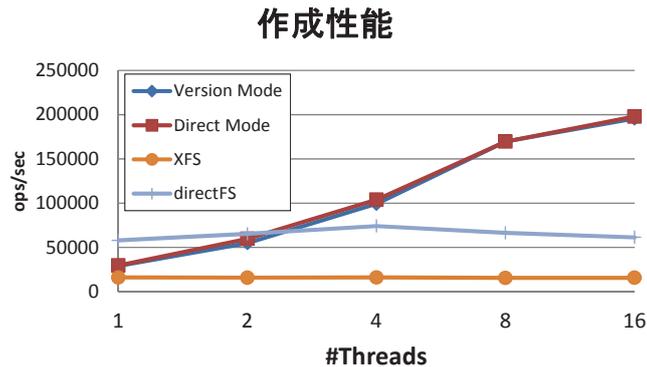
【エクストリームビッグデータの基盤技術】(建部、川島)

エクストリームビッグデータ (EBD) アプリケーションの実行に求められる、数万~数十万プロセスからの並列アクセスを想定した IOPS、プロセス数に比例した読込、書込アクセスバンド幅性能を目標として、分散オブジェクトストアの設計を行った。

本年度は、設計において重要な部分となる、数十万プロセスからの並列アクセスを可能とする分散メタデータサーバ、高速なアクセスを実現するローカルオブジェクトストア、大規模データに対する典型的な問合せを行う並列問合せ手法の概念設計を行った。分散メタデータの設計においてはこれまで分散ファイルシステムの分散メタデータサーバの研究開発をベースとし、より高効率な分散処理が可能となるように設計を行った。現在 EBD アプリケーションの実行に求められるアクセスパターン、API を他グループとのコ・デザインにより設計中であり、EBD アプリケーションで効率的な処理が可能となるよう進めている。

ローカルオブジェクトストアの設計においては、flash および不揮発性メモリにおいて可能となる操作をベースに、より効率的なデータアクセスが可能となるように設計を行った。具体的には OpenNVM を用いて、ローカルオブジェクトストアと範囲検索可能な Key Value ストアの設計を行った。ローカルオブジェクトストアでは、OpenNVM で可能となるスパースアドレス空間と、不可分書込を用い、スレッドセーフな設計を行った。従来の i ノードをベースにする設計では多

段の間接参照が性能を落とす原因となっていたが、スパースアドレスを用いたリージョンベースの設計によりこの間接参照を回避した。この設計を元に、現在プロトタイプ実装を進め、検証を行っている。図VI-9 にこれまでの性能結果を示す。提案手法である Version Mode、Direct Mode では並列クライアント数を増やすと性能が向上し、16 スレッドで 190,000 ops/sec を達成した。一方、従来手法である directFS や XFS では、並列クライアント数を増やしても性能は向上していない。



図VI-9 エクストリームビッグデータ処理基盤のプロトタイプにおける性能評価

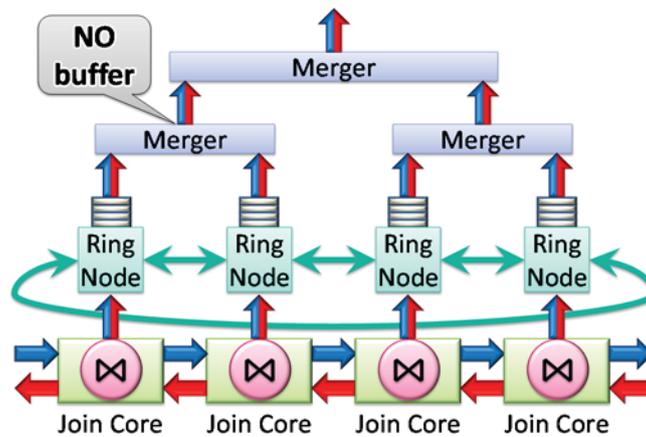
範囲検索可能な Key Value ストアの設計では、OpenNVM で提供される Key Value ストアに対して、範囲検索を可能とするためインメモリの B+木を利用する。並列アクセスを可能とするためには、スレッドセーフなノンブロッキング B+木が必要であるが、その設計がほぼ完了し、現在設計の検証段階にある。

【大規模広域分散ファイルシステム及びグリッド/クラウド技術に関する研究】(建部)

文部科学省が進める革新的ハイパフォーマンスコンピューティングインフラ (HPCI) の HPCI 共用ストレージ、素粒子物理学データ共有システム JLDG のシステムソフトウェアとして利用される Gfarm ファイルシステムの整備を行った。具体的には、それぞれのファイルについてチェックサムを自動的に計算することによるファイル損傷検知機能、運用監視システムにおける監視項目の整備、不具合修正、ドキュメントの更新などである。この結果として、Gfarm バージョン 2.5.8.6 をリリースした。

【高性能ストリームデータ処理に関する研究】(川島)

2つのストリームデータを高速に結合する従来研究としてハンドシェイクジョインがある。ハンドシェイクジョインの結果回収パイプラインにはデータがオーバーフローしやすいという欠点が存在する。本研究ではその欠点を克服するために適応的マージ機構を提案した。FPGA 上に提案システムを実装し、オリジナルのハンドシェイクジョインと比較した結果、5倍程度の性能向上を得た。

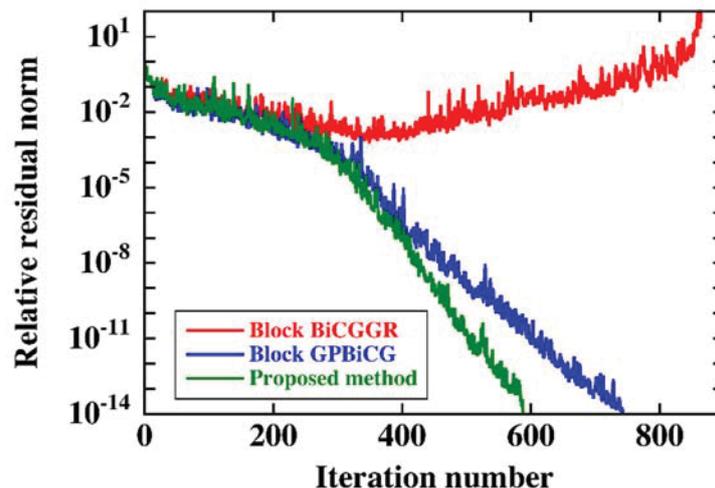


図VI-10 ストリーム結合

【Block Krylov 部分空間反復法に関する研究】(多田野)

複数右辺ベクトルをもつ連立一次方程式の数値解法である、Block Krylov 部分空間反復法の研究を行った。Block Krylov 部分空間反復法では、右辺ベクトル数が増加すると数値的に不安定な状況に陥り、残差の発散や停滞を引き起こすことがある。また、計算過程に発生する誤差の影響で、最終的に得られる近似解の精度が劣化する場合もある。

本研究では、残差の収束性向上、及び近似解の精度劣化を回避するアルゴリズムを構築した。格子 QCD 計算で現れる連立一次方程式（行列サイズ：1,572,864、右辺ベクトル数：12）に対して実験を行った。図VI-11 に示すように、提案手法は従来手法である Block BiCGGR 法、Block GPBiCG 法よりも少ない反復回数で残差が収束しており、高い収束性を示している。また、近似解の精度の指標である真の相対残差の値は、Block GPBiCG 法は 2.3×10^{-7} であるのに対し、提案法は 1.2×10^{-14} となり、提案法は高精度の近似解を生成することが分かった。



図VI-11 Block BiCGGR 法、Block GPBiCG 法、及び提案法の相対残差履歴

4. 外部資金の獲得状況

1. 文科省委託研究「次世代 IT 基盤構築のための研究開発将来のスーパーコンピューティングのための要素技術の研究開発」 課題名「演算加速機構を持つ将来の HPCI システムに関する調査研究の研究開発」 H24～H25 年度, 42,929 千円(H25) (代表:佐藤三久)
2. JST 国際科学技術共同研究推進事業 (戦略的国際共同研究プログラム) 課題名「ポストペタスケール・コンピューティングのためのフレームワークとプログラミング」 H22～H25 年度, 8,446 千円 (H25) (代表:佐藤三久)
3. 多国間国際研究協力事業 (G8 Research Councils Initiative) 研究課題「エクサスケールの気候シミュレーションを実現する技術(Enabling Climate Simulation at Extreme Scale)」 H23～H25 年度, 2,090 千円 (H25) (代表:佐藤三久)
4. JST CREST 研究領域「ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出」, 「ポストペタスケール時代に向けた演算加速機構・通信機構統合環境の研究開発」, H24～H29 年度, 77,180 千円 (H25) (代表:朴泰祐)
5. 多国間国際研究協力事業 (G8 Research Councils Initiative) 研究課題「エクサスケール規模の核融合シミュレーション (Nuclear Fusion Simulation at Exascale)」 H23～H25 年度, 8,640 千円 (H25) (代表:朴泰祐)
6. JST CREST 研究領域「ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出」, 「数値計算ライブラリによる超並列複合システムの階層的抽象化に関する研究 H24～28 年度, H25 14,820 千円 (共同研究者:高橋大介)
7. 科学研究費補助金 新学術領域研究 (研究領域提案型)「大規模並列環境における数値計算アルゴリズム」 H22～26 年度, H25 3,770 千円 (分担:高橋大介)
8. 科学研究費補助金 基盤研究 (C)「エクサスケール計算環境に向けた高速フーリエ変換のアルゴリズムに関する研究」 H24～26 年度, H25 1,300 千円 (代表:高橋大介)
9. JST CREST 研究領域「ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出」, 「ポストペタスケールデータインテンシブサイエンスのためのシステムソフトウェア」, H23 年度～H27 年度, 40,245 千円 (H25) (代表:建部修見)
10. JST CREST 研究領域「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」, 「数値計算ライブラリによる超並列複合システムの階層的抽象化に関する研究 H25～30 年度, 2,000 千円 (H25) (共同研究者:建部修見)

11. 科学研究費補助金 基盤研究 (C)「リアルタイム事象検知基盤に関する研究」H24～26 年度, 1,100 千円 (H25) (代表: 川島英之)
12. 科学研究費補助金 基盤研究 (B)「高性能計算のための抽象化に基づく資源管理システムソフトウェア」H25～27 年度, 1,000 千円 (H25) (分担: 川島英之)
13. 科学研究費補助金 基盤研究 (A)「大規模・異種の時空間データ統合で生じる矛盾を許容するサイエンスクラウド基盤」H24～26 年度, 800 千円 (H25) (分担: 川島英之)
14. 共同研究,「ストリームデータ処理における異常検知基盤の構築 (H25 年度)」日立製作所, 300 万円 (H25 年度) (代表: 川島英之)

5. 研究業績

(1)研究論文

1. 埜 敏博, 児玉 祐悦, 朴 泰祐, 佐藤 三久, "Tightly Coupled Accelerators アーキテクチャに基づく GPU クラスタの構築と性能予備評価", 情報処理学会論文誌コンピューティングシステム, Vol.6, No.4, pp.14-25, 2013.
2. 梅田 宏明, 埜 敏博, 庄司 光男, 朴 泰祐, 稲富 雄一, "フラグメント分子軌道法に現れる Fock 行列の GPGPU 化", 情報処理学会論文誌コンピューティングシステム, Vol.6, No.4, pp.26-37, 2013.
3. 扇谷 豪, 三木 洋平, 朴 泰祐, 森 正夫, 中里 直人, "計算宇宙物理のための GPU クラスタ向け並列 Tree Code の開発と性能評価", 情報処理学会論文誌コンピューティングシステム, Vol.6, No.3, pp.58-70, 2013.
4. Y. Miki, D. Takahashi and M. Mori, "Highly scalable implementation of an N-body code on a GPU cluster", Computer Physics Communications, Vol. 184, No. 9, pp. 2159--2168 (2013).
5. Ken T. Murata, Hidenobu Watanabe, Kazunori Yamamoto, Eizen Kimura, Masahiro Tanaka, Osamu Tatebe, Kentaro Ukawa, Kazuya Muranaga, Yutaka Suzuki, Hirotsugu Kojima, "A high-speed data processing technique for time-sequential satellite observation data", IEICE Communications Express, Vol.3, No.2, pp.74-49, 2014 (doi: 10.1587/comex.3.74)
6. Eric S. Fukuda, Hideyuki Kawashima, Taro Fujii, Koichiro Furuta, Tetsuya Asai and Masato Motomura, "C-based Design of Window Join for Dynamically Reconfigurable Hardware", Journal of Computer Science and Engineering, Volume 20, Issue 2, November 2013.

7. Masafumi Oyamada, Hideyuki Kawashima, and Hiroyuki Kitagawa. 2013. “Data Stream Processing with Concurrency Control”, SIGAPP Appl. Comput. Rev. 13, 2 (June 2013), 54-65.
8. T. Sakurai, Y. Futamura and H. Tadano, “Efficient parameter estimation and implementation of a contour integral-based eigensolver”, J. Alg. Comput. Tech., Vol. 7, No 3, pp. 249—269, 2013.
9. 山崎 育朗, 今倉 暁, 多田野 寛人, 櫻井 鉄也, “残差最小性に基づく Krylov 部分空間反復解法に対する疎行列用直接解法を用いた前処理のパラメータ推定”, 日本応用数学会論文誌, Vol. 23, No. 3, pp. 381—404, 2013.
10. A. Imakura, L. Du and H. Tadano, “A weighted Block GMRES method for solving linear systems with multiple right-hand sides”, JSIAM Letters, Vol. 5, pp. 65—68, 2013.

(2) 学会発表

(A) 招待講演

1. Mitsuhsa Sato, Issues for Exascale Accelerated Computing - system architecture and programming, 7th Int'l. Conf. on PGAS Programming Models.
2. T. Boku, “Nu-FuSE Exascale Simulations and the XMP programming language”, Exascale Applications and Software Conference, Edinburgh, April 2013.
3. T. Boku, “Toward Exa-scale Accelerated Computing”, Int. HPC Forum China, Changsha, May 2013.
4. 朴泰祐, “G8 NuFuSE プロジェクトにおけるプラズマコードの大規模並列化/GPU化”, プラズマシミュレーションシンポジウム2013, 土岐, 2013年9月.
5. T. Boku, “Accelerated Computing Unified with Communication Towards Exascale”, CODESIGN China 2013, Guilin, October 2013.
6. 児玉 祐悦, “将来の演算加速機構について”, 計算分子化学研究拠点第4回研究会, 2013年9月.
7. 児玉祐悦, 朴泰祐, 埴敏博, 佐藤三久, 梅村雅之, “超並列GPUクラスタHA-PACSにおけるGPU間直接通信機構”, 大学ICT推進推進協議会年次大会, 2013年12月.

8. 児玉 祐悦, ”スーパーコンピュータにおけるアクセラレータ技術”, 情報処理学会第 76 回全国大会「エクサスケールを目指す技術開発最前線」, 2014年3月.
9. D. Takahashi, “Implementation of Parallel FFTs on GPU Clusters”, 2014 Conference on Advanced Topics and Auto Tuning in High Performance and Scientific Computing (2014 ATAT in HPSC), National Taiwan University, Taipei, Taiwan, March 2014.
10. Osamu Tatebe, Gfarm: Present Status and Future Evolution, OpenSFS APAC Lustre User Group, Tokyo, October 17, 2013.
11. 建部修見, Gfarmファイルシステムの実装と応用事例、第4回地域間インタークラウドワークショップ、沖縄、2014年3月.
12. Hideyuki Kawashima, “Taming Big Data Streams”, The International Symposium on Embedded Multicore/Manycore System-on-Chip (MCSoc-13), Tokyo, Japan, September 26-28, 2013, Keynote talk.

(B) その他の学会発表(査読付)

1. Hitoshi Murai and Mitsuhsa Sato: An Efficient Implementation of Stencil Communication for the XcalableMP PGAS Parallel Programming Language, In Proc. 7th Int'l. Conf. on PGAS Programming Models (2013)
2. Masahiro Nakao, Hitoshi Murai, Takenori Shimosaka and Mitsuhsa Sato: Productivity and Performance of the HPC Challenge Benchmarks with the XcalableMP PGAS language, In Proc. 7th Int'l. Conf. on PGAS Programming Models (2013).
3. Akihiro Tabuchi, Masahiro Nakao, Mitsuhsa Sato: A Source-to-Source OpenACC Compiler for CUDA. Euro-Par Workshops 2013: 178-187
4. Miwako Tsuji, Mitsuhsa Sato, Maxime R. Hugues, Serge G. Petiton: Multiple-SPMD Programming Environment Based on PGAS and Workflow toward Post-petascale Computing. ICPP 2013: 480-485

5. Tatsuya Abe, Toshiyuki Maeda, Mitsuhsa Sato: Model Checking Stencil Computations Written in a Partitioned Global Address Space Language. IPDPS Workshops 2013: 365-374
6. Tomotake Nakamura, Mitsuhsa Sato: XMP-IO function and its application to MapReduce on the K computer. PARCO 2013: 35-42
7. 小田嶋 哲哉, 朴 泰祐, 佐藤 三久, 塙 敏博, 児玉 祐悦, Raymond Namyst, Samuel Thibault, Olivier Aumage, "GPU クラスタ向け並列言語 XMP-dev における GPU/CPU 動的負荷分散機能", 2014 年ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2013 論文集, 2014.
8. 塙 敏博, 児玉 祐悦, 朴 泰祐, 佐藤 三久, "Tightly Coupled Accelerators アーキテクチャに基づく GPU クラスタの構築", 先進的計算基盤システムシンポジウム SACSIS2013 論文集, 2013.
9. T. Odajima, T. Boku, M. Sato, T. Hanawa, Y. Kodama, R. Namyst, S. Thibault, O. Aumage, "Adaptive Task Size Control on High Level Programming for GPU/CPU Work Sharing", Proc. of Int. Workshop on Advances of Distributed and Parallel Processing 2013 (ADPC-2013, with ICA3PP-2013), Vietri sul Mare, LNCS-8286 Part II, pp.59-68, 2013.
10. T. Hanawa, Y. Kodama, T. Boku, M. Sato, "Interconnect for Tightly Coupled Accelerators Architecture", Proc. of HotInterconnect? 2013, San Jose, 2013.
11. T. Hanawa, Y. Kodama, T. Boku, M. Sato, "Tightly Coupled Accelerators Architecture for Minimizing Communication Latency among Accelerators", Proc. of 3rd Int. Workshop on Accelerators and Hybrid Exascale Systems (AsHES 2013, with IPDPS2013), Boston, CD-ROM, 2013.
12. Y. Yamaguchi, K. Fujinami, A. Sugiura and Y. Kodama, "A Study of a Three-Dimensional Multiphase-Flow Simulator", International Conference on Field Programmable Logic and Applications, 1-4, Sep 2013.
13. Y. Kodama, T. Boku, T. Hanawa, M. Sato, "HA-PACS/TCA: Tightly Coupled Accelerators for Low-Latency Communication", The 4th AICS International Symposium, poster, Dec. 2013.

14. S. Ochiai, Y. Yamaguchi and Y. Kodama, “The Flexible Sound Synthesizer on an FPGA”, International Symposium on Computing and Networking, 104-111, Dec. 2013.
15. D. Mukunoki and D. Takahashi, “Using Quadruple Precision Arithmetic to Accelerate Krylov Subspace Methods on GPUs”, Proc. 10th International Conference on Parallel Processing and Applied Mathematics (PPAM 2013), Part I, Workshop on Numerical Algorithms on Hybrid Architectures, Lecture Notes in Computer Science, Vol. 8384, pp. 632--642, Springer-Verlag, 2014.
16. T. Hiragushi and D. Takahashi, “Efficient Hybrid Breadth-First Search on GPUs”, Proc. 13th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2013), Part II, 2013 International Symposium on Advances of Distributed and Parallel Computing (ADPC 2013), Lecture Notes in Computer Science, Vol. 8286, pp. 40--50, Springer International Publishing, 2013.
17. D. Takahashi, “Implementation of Parallel 1-D FFT on GPU Clusters”, Proc. 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE 2013), pp. 174—180, 2013.
18. Y. Sato, M. Miwa, S. Takeuchi and D. Takahashi, “Optimizing Objective Function Parameters for Strength in Computer Game-Playing”, Proc. 27th AAAI Conference on Artificial Intelligence (AAAI-13), pp. 869—875, 2013.
19. D. Mukunoki and D. Takahashi, “Optimization of Sparse Matrix-vector Multiplication for CRS Format on NVIDIA Kepler Architecture GPUs”, Proc. 13th International Conference on Computational Science and Its Applications (ICCSA 2013), Part V, Lecture Notes in Computer Science, Vol. 7975, pp. 211--223, Springer-Verlag, 2013.
20. Hiroki Ohtsuji, Osamu Tatebe, "Optimized Communication Layer for Exa-scale Storage Systems", PRAGMA Students Workshop, Beijing, October 16, 2013
21. Li Xieming, Osamu Tatebe. Design of Data-Aware Scheduler, PRAGMA Students Workshop, October 16, 2013
22. Fuyumasa Takatsu, Kohei Hiraga, and Osamu Tatebe, Object Storage for High Speed Storage Device, PRAGMA Students Workshop, October 16 2013

23. Naoko Kosugi, Naoki Kodama, Sachiko Shimizu, Shunsuke Saruwatari, Tsutomu Terada, Hiroaki Kazui, Koichi Yamashita, Hideyuki Kawashima, Masayuki Hata, "A Prototype System of Remote Music Therapy Using the Latest Communication Technology in Japan", The 15th International Conference on Information Integration and Web-based Applications & Services (iiWAS2013) , pp. 671-676.
24. Naotaka Nishimura, Hideyuki Kawashima, Hiroyuki Kitagawa, "A High Throughput Complex Event Detection Technique with Bulk Evaluation", 5th International Workshop on Streaming Media Delivery and Management Systems In Conjunction with 7th 3PGCIC-2013 Conference, pp. 624-629.
25. Yasin Oge, Masato Yoshimi, Takefumi Miyoshi, Hideyuki Kawashima, Hidetsugu Irie and Tsutomu Yoshinaga, "An Efficient and Scalable Implementation of Sliding-Window Aggregate Operator on FPGA", The First International Symposium on Computing and Networking (CANDAR), pp. 112--121.
26. Masahiro Oke, Hideyuki Kawashima, "A Multiple Query Optimization Scheme for Change Point Detection on Stream Processing System", SHORT PAPER, The 7th International Workshop on Business Intelligence for the Real Time Enterprise (BIRTE 2013) in conjunction with International Conference on Very Large Databases (VLDB).
27. Yasin Oge, Masato Yoshimi, Takefumi Miyoshi, Hideyuki Kawashima, Hidetsugu Irie, Tsutomu Yoshinaga (2013): Wire-Speed Implementation of Sliding-Window Aggregate Operator over Out-of-Order Data Streams. In: IEEE 7th International Symposium on Embedded Multicore/Many-core SoCs (MCSoc-13), pp. 55-60, 2013.
28. H. Tadano, Improvement of numerical stability of a Block Krylov subspace method for computing high accuracy solutions, International Conference on Simulation Technology (JSST2013), 2013.

(C) その他の学会発表(査読無)

1. Takenori Shimosaka, Hitoshi Murai, Mitsuhsa Sato: A communication library between multiple sets of MPI processes for a MPMD model. EuroMPI 2013: 147-148
2. 中尾昌広, 佐藤三久. “京速コンピュータ「京」における CGPOP Miniapp の性能評価“, 第 139 回 HPC 研究会, 2013 年 5 月, 東京大学 柏の葉キャンパス (千葉)

3. 中尾昌広, 佐藤三久. “京速コンピュータ「京」における PGAS モデルによる 気象コード NICAM の実装”, 2013 年並列/分散/協調処理に関する『北九州』サマー・ワークショップ (SWoPP 北九州 2013), 研究報告 ハイパフォーマンスコンピューティング (HPC), Vol.2013-HPC-140, pp.1-5, 2013 年 8 月, 北九州 国際会議場.
4. 藤井 久史, 埜 敏博, 児玉 祐悦, 朴 泰祐, 佐藤 三久, 藏増 嘉伸, Mike Clark, "GPU 向け QCD ライブラリ QUDA の TCA アーキテクチャによる実装", 2014-HPC-143, 2014.
5. 奴賀 秀男, 朴 泰祐, 藤田 典久, 中尾 昌広, 佐藤 三久, William Tang, "並列言語 XcalableMP による核融合シミュレーションコードの開発", 2013-HPC-142, 2013.
6. 藤井 久史, 埜 敏博, 児玉 祐悦, 朴 泰祐, 佐藤 三久, "TCA アーキテクチャによる並列 GPU アプリケーションの性能評価", 2013-HPC-140, 2013.
7. 藤田 典久, 奴賀 秀男, 朴 泰祐, 井戸村 泰宏, "GPU クラスタ HA-PACS における核融合シミュレーションコードの性能評価", 2013-HPC-140, 2013.
8. 小田嶋 哲哉, 朴 泰祐, 佐藤 三久, 埜 敏博, 児玉 祐悦, Raymond Namyst, Samuel Thibault, Olivier Aumage, "並列言語 XMP-dev における GPU/CPU 動的負荷分散機能", 2013-HPC-140, 2013.
9. 児玉 祐悦, 山口 佳樹, 中里 直人, 牧野 淳一郎, 朴 泰祐, 佐藤 三久, "大規模 SIMD 型アクセラレータの検討", 2013-HPC-140, 2013.
10. 下坂 健則, 佐藤 三久, 朴 泰祐, William Tang, "京速コンピュータ「京」における核融合シミュレーションコード GTC-P の評価", 2013-HPC-139, 2013.
11. 高橋 航平, 埜 敏博, 朴 泰祐, 児玉 祐悦, 扇谷 豪, 佐藤 三久, "各種アプリケーションにおける GPGPU 対 Many Core Processor の性能比較", 2013-HPC-139, 2013.
12. 丸山裕士, 矢葺徹, 山口佳樹, 児玉祐悦, “FPGA を用いた実時間動画像手ブレ補正システムの構築”, 信学技報, Vol.113, No.22, pp.13-18, 2013 年 4 月.
13. 藤浪健太, 杉浦彰, 山口佳樹, 児玉祐悦, “3D 流体解析用 FPGA システムの構築”, 信学技報, Vol.113, No.22, pp.37-42, 2013 年 4 月.

14. 藤浪健太, 山口佳樹, 児玉祐悦, “FPGA による三次元混相流シミュレータの設計と検証”, 信学技報, Vol.113, No.52, p.103-108, 2013 年 5 月.
15. 落合優, 山口佳樹, 児玉祐悦, “FPGA サウンドシンセサイザとそのユーザインタフェースについて”, 信学技報, Vol.113, No.52, pp.85-90, 2013 年 5 月.
16. 金紅坤, 山口佳樹, 児玉祐悦, “FPGA 用マルチコアプロセッサをターゲットとしたマルチポート共有キャッシュシステムの提案”, 信学技報, Vol.113, No.324, pp.65-70, 2013 年 11 月.
17. 藤浪健太, 山口佳樹, 児玉祐悦, “格子法に基づく 3 次元流体シミュレータの高速化”, 数値流体力学シンポジウム講演論文集, D06-2, 1-4, 2013 年 12 月.
18. T. Nagashima, Y. Yamaguchi and Y. Kodama, “A study of SuperH processor cores on an FPGA”, International Workshop on Innovative Architecture for Future Generation High-Performance Processors and Systems, Mar. 2014.
19. 棕木大地, 高橋大介 : GPU における 4 倍精度浮動小数点演算を用いたクリロフ部分空間法の高速化, 情報処理学会研究報告, 2013-HPC-140, No. 35, 2013.
20. 平櫛貴章, 高橋大介 : GPU クラスタにおける幅優先探索の高速化, 情報処理学会研究報告, 2013-HPC-139, No. 12, 2013.
21. Hiroki Ohtsuji, Osamu Tatebe, "High-throughput Remote File Access for Exa-scale Storage Systems", IWCST2013, Hangzhou, October 26, 2013
22. 渡邊英伸, 黒澤隆, 一岡翔太郎, 木村映善, 村田健史, 建部修見, UDT を用いた並列ファイル転送技術, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-139(8),1-8 (2013-05-22)
23. 村田 健史, 渡邊 英伸, 山本 和憲, 久保田 康文, 建部 修見, 田中 昌宏, 深沢 圭一郎, 木村 映善, 鶴川 健太郎, 村永 和哉, 鈴木 豊, 磯田 総子, Gfarm/Pwrake による NICT サイエンスクラウドの並列分散処理技法, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-139(9),1-6 (2013-05-22)

24. Marilia Melo, Osamu Tatebe, "Fault Tolerance Design for Hadoop MapReduce on Gfarm Distributed Filesystem",研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-140(14),1-5 (2013-07-24)
25. 鷹津冬将 , 平賀弘平 , 建部修見, 不揮発性デバイス向けの Object Storage の設計, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-140(12),1-7 (2013-07-24)
26. 李燮鳴 , 建部修見, データ配置を考慮したタスクスケジューリング, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-140(19),1-5 (2013-07-24)
27. 渡邊英伸 , 亀澤祐一 , 高杉英利 , 平野一樹 , 今井潔 , 村田健史 , 建部修見, 広域分散ファイルシステム Gfarm の SLA 評価手法, インターネットと運用技術シンポジウム 2013 論文集,2013,9-16 (2013-12-05)
28. 建部修見, 原田浩, 實本英之, 佐藤仁, 平川学, HPCI 共用ストレージの性能評価, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-142(8),1-6 (2013-12-09)
29. 大辻弘貴, 建部修見, 分散ストレージシステムに対する低オーバーヘッド冗長化書き込み手法の提案と評価, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2013-HPC-142(10),1-6 (2013-12-09)
30. 鷹津冬将, 平賀弘平, 建部修見. 不揮発性デバイス向けの Object Storage の実装と評価.研究報告ハイパフォーマンスコンピューティング (HPC) ,2014-HPC-143(1),1-7 (2014-02-24)
31. 田中昌宏, 建部修見, ワークフローシステム Pwrake における I/O 性能を考慮したタスクスケジューリング, 研究報告ハイパフォーマンスコンピューティング (HPC) ,2014-HPC-143(3),1-10 (2014-02-24)
32. 川島英之 : In-DSMS 分析へ向けて. 電子情報通信学会技術研究報告. 知的環境とセンサネットワーク 113(38), 13-18, 2013 年 5 月.
33. 西村直孝, 川島英之 : AIS を用いた複合イベント処理の効率化. 情報処理学会研究報告. UBI, [ユビキタスコンピューティングシステム]/2013(13)/pp.1-6. 2013 年 5 月.
34. 川島英之, 黒川茂莉 : 推論処理を有するストリーム処理基盤. 情報処理学会研究報告. UBI, [ユビキタスコンピューティングシステム]/2014(38)/pp.1-6. 2014 年 3 月.
35. 川島英之, 黒川茂莉 : ストリームイベントの効率的なオンライン名寄せ処理. 電子情報通信学会総合大会. 2014 年 3 月.

36. 多田野 寛人, 石川 陽一, 今倉 暁, “双共役残差型反復解法の複数右辺ベクトル対応版への拡張と性能評価”, 日本応用数学会 2013 年度年会, 2013.
37. 齊藤 周作, 多田野 寛人, 今倉 暁, “Block BiCGSTAB(l)法の構築と安定化”, 第 10 回日本応用数学会 研究部会連合発表会, 2013.

(D) 著書

6. 異分野間連携・国際連携等

1. 戦略的国際科学技術協力推進事業（日仏共同研究）「ポストペタスケールコンピューティングのためのフレームワークとプログラミング」（佐藤）
2. 多国間国際研究協力事業 G8 Research Councils Initiative “Nuclear Fusion Simulation for Exascale” 日本代表 PI（朴）
3. 多国間国際研究協力事業 G8 Research Councils Initiative “Exascale Climate Simulation” 日本代表 PI（佐藤）

7. 国際活動

1. M. Sato: Program Committee, CCGRID 2013
2. M. Sato: Program Committee, Cluster 2013
3. M. Sato: Program Committee, EASC 2013
4. M. Sato: Program Committee, EuroMPI 2013
5. M. Sato: Program Committee, FTXS 2013
6. M. Sato: Program Committee, HeteroPar 2013
7. M. Sato: Program Committee, IWOMP 2013
8. M. Sato: Program Committee, PGAS 2013
9. M. Sato: Track Chair, HiPC 2013
10. T. Boku: Program Committee, CASS 2013
11. T. Boku: Program Committee, HiPC 2013
12. T. Boku: Organizing Chair, HPC in Asia Session at ISC2013

13. T. Boku: Program Committee, ICS2013
14. T. Boku: Program Committee, Int. Conf. HPCS 2013
15. T. Boku: Program Committee on Tutorial, ISC2013
16. T. Boku: Program Committee, PMAM2013
17. T. Boku: Area Chair on System Software Area, SC13
18. T. Boku: Committee member of ACM Gordon Bell Prize 2013
19. D. Takahashi: Program Committee, The International Conference for High Performance Computing, Networking, Storage and Analysis (SC13)
20. D. Takahashi: Program Committee, The International Conference on Computational Science (ICCS 2013)
21. D. Takahashi: Review Committee, The International Conference for High Performance Computing, Networking, Storage and Analysis (SC13) Doctoral Showcase
22. D. Takahashi: Program Committee, 15th IEEE International Conference on High Performance Computing and Communications (HPCC 2013)
23. D. Takahashi: General Chair, IEEE 7th International Symposium on Embedded Multicore/Many-core SoCs (MCSoc-13) Special Session on Legacy HPC Application Migration (LHAM 2013)
24. O. Tatebe: Program Committee, IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC13)
25. O. Tatebe: Program Committee, 5th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2013)
26. O. Tatebe: Program Committee, International Supercomputing Conference
27. O. Tatebe: Program Committee, The 42nd International Conference on Parallel Processing (ICPP)
28. O. Tatebe: Program Committee, International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC 2013)
29. O. Tatebe: Program Committee, Architecture, Languages, Compilation and Hardware support for Emerging ManYcore systems (ALCHEMY) workshop
30. O. Tatebe: Program Committee, Workshop on Big Data Management in Clouds (BigDataCloud)
31. O. Tatebe: Program Committee, The Sixth International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS-2013)

32. O. Tatebe: Scientific Committee, 11th International Meeting on High-Performance Computing for Computational Science (VECPAR 2013)
33. O. Tatebe: Program Committee, International Conference on Grid and Pervasive Computing (GPC)
34. O. Tatebe: Program Committee, IEEE International Conference on High Performance Computing and Communications (HPCC)
35. H. Kawashima: Program Committee, 4th International Workshop on Streaming Media Delivery and Management Systems, In Conjunction with 8th 3PGCIC-2013 Conference

8. 教育

・学位論文（博士課程）

1. 椋木大地, 博士（工学）, A study on linear algebra operations using extended precision floating-point arithmetic on GPUs (GPU における拡張精度浮動小数点演算を用いた線形計算の研究), 筑波大学大学院システム情報工学研究科博士論文, 2013 年 11 月（指導：高橋大介）
2. 佐藤佳州, 博士（工学）, ゲームにおける棋譜の性質と強さの関係に基づいた学習, 筑波大学大学院システム情報工学研究科博士論文, 2014 年 3 月（指導：高橋大介）

・学位論文（修士課程）

1. Alamri Hussain Ali A, 修士（工学）, Evaluation of Many-core processor with Analysis of Sparse Matrix Vector Multiplication（疎行列ベクトル積の解析によるメニーコアプロセッサの評価）, 筑波大学大学院システム情報工学研究科修士論文, 2013 年 3 月（指導：佐藤三久）
2. 石川 奏太, 修士（工学）, Non-Homogeneous 置換モデルを実装した分子系統解析プログラムの並列化, 筑波大学大学院システム情報工学研究科修士論文, 2013 年 3 月（指導：佐藤三久）（計算科学ディグリープログラム）
3. 金 紅坤, 修士（工学）, FPGA 用マルチコアプロセッサの提案と実装, 筑波大学大学院システム情報工学研究科修士論文, 2013 年 3 月（指導：児玉祐悦）
4. 鷹津冬将, 修士（工学）, 高性能計算機向けのストレージシステムに関する研究, 筑波大学大学院システム情報工学研究科修士論文, 2014 年 3 月（指導：建部修見）

5. 李燮鳴, 修士 (工学), 大規模データ処理基盤のタスクスケジューリング, 筑波大学大学院システム情報工学研究科修士論文, 2014 年 3 月 (指導: 建部修見)
6. Marilia Rodrigues de Melo e Silva, 修士 (工学), Fault Tolerance Design for Hadoop MapReduce on Gfarm Distributed File System (Gfarm 分散ファイルシステム上での Hadoop MapReduce の耐故障性設計に関する研究), 筑波大学大学院システム情報工学研究科修士論文, 2014 年 3 月 (指導: 建部修見)

・ 学外教育

1. 佐藤三久: 神戸大学システム情報学研究科・客員教授

9. 社会貢献

1. 佐藤三久: PC クラスタコンソーシアム 理事 XMP 規格部会会長
2. 佐藤三久: 情報処理学会フェロー
3. 佐藤三久: 大川出版賞 「岩波講座 計算科学 別巻 スーパーコンピュータ」(共著)
4. 朴泰祐: 先進的計算基盤システムシンポジウム SACSYS 運営委員会委員
5. 朴泰祐: 情報処理学会ハイパフォーマンスコンピューティング研究会運営委員
6. 朴泰祐: Annual Meeting on Advanced Computing System and Infrastructure (ACSI) 2015 組織委員長
7. 児玉祐悦: 情報処理学会 デジタルプラクティス査読委員
8. 児玉祐悦: 電子情報通信学会 論文誌査読委員
9. 児玉祐悦: 情報処理学会 計算機アーキテクチャ研究会運営委員
10. 児玉祐悦: Cool Chips プログラム委員
11. 高橋大介: 先進的計算基盤システムシンポジウム SACSYS2013 プログラム委員
12. 建部修見: 特定非営利団体つくば OSS 技術支援センター理事長
13. 建部修見: インターネットカンファレンス 2013 プログラム委員
14. 建部修見: 情報処理学会ハイパフォーマンスコンピューティング研究会運営委員
15. 川島英之: 情報処理学会 データベースシステム研究会 幹事
16. 川島英之: 情報処理学会論文誌編集委員会委員

17. 川島英之: 電子情報通信学会 知的環境とセンサネットワーク研究会 委員
18. 多田野寛人: 日本応用数学会 「行列・固有値問題の解法とその応用」 研究部会 幹事
19. 多田野寛人: 日本応用数学会 「若手の会」 研究部会 運営委員
20. 多田野寛人: 日本応用数学会 「応用数理」 編集委員
21. 多田野寛人: 情報処理学会ハイパフォーマンスコンピューティングと計算科学シンポジウム
HPCS2014 実行委員
22. 多田野寛人: 情報処理学会ハイパフォーマンスコンピューティングと計算科学シンポジウム
HPCS2014 プログラム委員
23. 多田野寛人: 先進的計算基盤システムシンポジウム SACSIS2013 プログラム委員

10. その他

1. 筑波大学計算科学研究センター・HPC サマーセミナー開催