

松岡 聡

東京工業大学 学術国際情報センター
教授

「EBD:次世代の年ヨッタバイト処理に向けたエクストリームビッグデータの基盤技術」

§ 1. 研究実施体制

(1) 松岡グループ

- ① 研究代表者:松岡 聡 (東京工業大学学術国際情報センター、教授)
- ② 研究項目
EBD システムアーキテクチャの設計及びシステムソフトウェアの研究開発
 - ・EBD システム・アーキテクチャの基本設計
 - ・EBD システムソフトウェアの基本設計
 - ・アプリケーションによる評価・高度化

(2) 建部グループ

- ① 主たる共同研究者:建部 修見 (筑波大学システム情報系、准教授)
- ② 研究項目
EBD 分散オブジェクトストアの研究
 - ・分散オブジェクトストアの設計
 - ・プロトタイプ実装・性能評価
 - ・アプリケーションによる評価・高度化

(3) 鯉淵グループ

- ① 主たる共同研究者:鯉淵 道紘 (国立情報学研究所アーキテクチャ科学研究系、准教授)
- ② 研究項目
EBD インターコネクトの研究開発
 - ・低遅延トポロジとルーティング

- ・ストレージへの直接通信機構

(4) 秋山グループ

- ① 主たる共同研究者: 秋山 泰 (東京工業大学大学院情報理工学研究科、教授)
- ② 研究項目
 - EBD データ処理 API の開発と、大規模ゲノム解析等での評価
 - ・EBD データ処理 API の開発
 - ・大規模ゲノム解析での実応用評価

(5) 鈴木グループ

- ① 主たる共同研究者: 鈴木 豊太郎 (University College Dublin, School of Computer Science and Informatics, Visiting Associate Professor)
- ② 研究項目
 - 大規模データ駆動型による社会動向解析基盤の研究・開発
 - ・社会現象モデルの構築
 - ・EBD システムアーキテクチャ上の予測シミュレーション基盤の設計・実装
 - ・予測シミュレーション基盤の性能最適化技術
 - ・実証実験による精度検証

(5) 三好グループ

- ① 主たる共同研究者: 三好 建正 (理化学研究所データ同化研究チーム・チームリーダー)
- ② 研究項目
 - 次世代ビッグデータ同化アプリケーション
 - ・フェールセーフの EBD ワークフローの開発
 - ・Geographical Search アルゴリズムの最適化
 - ・EBD のプラットフォームの設計・開発のコ・デザイン

§ 2. 研究実施の概要

将来 Zeta(10^{21})Byte/日(あるいは Yotta(10^{24})Byte/年)という、今の Google/Amazon の個々の IDC に代表される 10 万ノード級のクラウドのデータ処理能力の、最大で 10 万倍に至る処理能力を達成するための EBD(Extreme Big Data)システム基礎技術の確立を達成することを目標とし、そのためにスパコン技術、特にメニーコア超並列処理と広帯域低遅延ネットワーク技術・不揮発性メモリ技術・及び高性能データベース技術を融合し、単なる「ビッグデータ」から「EBD」への相転移的な技術革新をはかる。本年度は、EBD を実現する際の技術的課題の特定のためのベンチマーク実行基盤の整備と、EBD を支えるシステム・ソフトウェアと、アプリケーション・インスタンスのプロトタイプの開発、初期の性能評価を中心に行った。

EBD を支えるシステム・ソフトウェアに関しては、次世代のストレージ装置として期待されるフラッシュデバイスや不揮発性メモリを想定した高速な並列データアクセスを実現するためのローカルオブジェクトストアの設計や、EBD アプリケーション向けのインターコネクトのトポロジとルーティングの設計、ソートやグラフ処理などの EBD アプリケーションを最大限支援するためのソフトウェアカーネルの研究開発、また、GPU 仮想化・マイグレーションによる、計算とデータの co-location による全体最適化などを進めた。

また、これらの総合的な「EBD システム」を基盤に、大規模ゲノム相関、社会シミュレーション(特に超大規模非構造グラフ構造)、超高速センサ・エクサスケール気象データ同化、などを代表的な「EBD アプリケーション」のインスタンスとして類型し、EBD アーキテクチャ上での実行に向けた設計、及び、初期の評価を進めた。具体的には、大規模ゲノム相関に関しては、EBD 対 EBD の相関計算や相同性計算を実現するために必要となる API 群の概念設計、社会シミュレーションに関しては、エージェントベースの大規模交通シミュレーションの TSUBAME2 上での初期の性能評価、超高速センサ・エクサスケール気象データ同化に関しては、気象庁非静力学モデル NHM を用いた局所アンサンブル変換カルマンフィルタ LETKF(Local Ensemble Transform Kalman Filter)のプロトタイプの初期の性能評価を行った。

また、EBD システム・ソフトウェアとアプリケーション・インスタンス、ハードウェアアーキテクチャの設計・開発の妥当性を示すためのランドチャレンジとして、2013 年 11 月に、スーパーコンピュータのビッグデータ処理の省エネルギー性を競うランキングである Green Graph500 に計測結果を投稿した結果、SCALE30 の問題で 4.35MTEPS/W となり、Big Data Category で 4 位(EBD-RH5885v2)の成績となり、シングルノードの性能では世界第 1 位となり、EBD アーキテクチャの設計の正しさを示した。

§ 3. 成果発表等

(3-1) 原著論文発表

【松岡グループ】

論文詳細情報(国内)

- A-1.岩渕 圭太, 佐藤 仁, 溝手 竜, 安井 雄一郎, 藤澤 克樹, 松岡 聡, 不揮発性メモリを考慮した Hybrid-BFS アルゴリズムの省メモリ化, 第 143 回ハイパフォーマンスコンピューティング研究発表会, Mar 2014
- A-2.三浦 信一, 滝澤 真一郎, 松岡 聡, 棟朝 雅晴, 實本 英之, 小林 泰三, ” HPCI 先端ソフトウェア運用基盤の構築と運用”, 第 143 回ハイパフォーマンスコンピューティング研究発表会, Vol.2014-HPC-143(30), Mar 2014
- A-3.Jin Guanghao, Endo Toshio, Matsuoka Satoshi, “Multi-level Temporal Blocking for Stencil Computation for Memory Hierarchy on TSUBAME2.5”, 第 143 回ハイパフォーマンスコンピューティング研究発表会, Mar 2014
- A-4.遠藤 敏夫, 額田 彰, 松岡 聡. “ウルトラグリーンスパコン TSUBAME2.5/TSUBAME-KFC”, 大学 ICT 推進協議会 2013 年度年次大会講演論文集, T5G-4, 2013 年 12 月
- A-5.遠藤 敏夫, 額田 彰, 松岡 聡.” TSUBAME-KFC: 液浸冷却を用いたウルトラグリーンスパコン研究設備”, ハイパフォーマンスコンピューティングとアーキテクチャの評価に関する北海道ワークショップ(HOKKE), 情報処理学会研究報告, 2012-ARC-199/HPC-142, 2013 年 12 月
- A-6.佐藤仁, 松岡聡 “Extreme Big Data 時代に向けた TSUBAME スーパーコンピュータ上での取り組み”, AXIES 大学 ICT 推進協議会 (AXIES2013), 幕張, 2013 年 12 月(accepted to be published)
- A-7.Chih-Song Kuo, Akihiro Nomura, Satoshi Matsuoka, Aamer Shah, Felix Wolf, Ilya Zhukov, “Environment Matters: How Competition for I/O among Applications Degrades their Performance”, IPSJ SIG Technical Reports 2013-ARC-199 2013-HPC-142 (HOKKE-21), Hokkaido, Nov 2013 (in print) (unpublished)
- A-8.Aleksandr Drozd, Satoshi Matsuoka. “MSD Radix String Sort on GPU: Longer Keys,

Shorter Alphabets, IPSJ SIG Technical Reports 2013-ARC-199 2013-HPC-142
(HOKKE-21), Hokkaido, Nov 2013 (unpublished)

A-9. 松岡聡、佐藤賢斗、遠藤 敏夫「エクサスケールスパコンに向けた耐故障性の評価—
TSUBAME2.0 を例にして —」、情報処理学会研究報告 HPC-141、2013 年 10 月

論文詳細情報(国際)

A-10. Kento Sato, Kathryn Mohror, Adam Moody, Todd Gamblin, Bronis R de
Supinski, Naoya Maruyama and Satoshi Matsuoka, "A User-level InfiniBand-based
File System and Checkpoint Strategy for Burst Buffers", 2014 14th IEEE/ACM
International Symposium on Cluster, Cloud, and Grid Computing (CCGrid), May
2014(to appear)(DOI: 未定)

A-11. Keita Iwabuchi, Hitoshi Sato, Ryo Mizote, Yuichiro Yasui, Katsuki Fujisawa
and Satoshi Matsuoka, "Hybrid BFS Approach Using Semi-External Memory"
IPDPS-HPDIC2014 3rd High Performance Data Intensive Computing Workshop ,
Phoenix, Arizona, USA, May 23, 2014(to appear)(DOI: 未定)

A-12. Kento Sato, Adam Moody, Kathryn Mohror, Todd Gamblin, Bronis R. de
Supinski, Naoya Maruyama and Satoshi Matsuoka, "FMI: Fault Tolerant
Messaging Interface for Fast and Transparent Recovery", In Proceedings of the
International Conference on Parallel and Distributed Processing Symposium 2014
(IPDPS2014), Phoenix, USA, May, 2014(to appear)(DOI: 未定)

A-13. Miquel Pericàs, Abdelhalim Amer, Kenjiro Taura and Satoshi Matsuoka
"Analysis of Data Reuse in Task-Parallel Runtimes"
4th International Workshop on Performance Modeling, Benchmarking and
Simulation of High Performance Computer Systems (PMBS13), Denver, November
2013
http://www.dcs.warwick.ac.uk/~sdh/pmbs13/PMBS13/Workshop_Schedule.html
(to be published)(DOI: 未定)

A-14. Satoshi Matsuoka, "The TSUBAME2.5 Evolution" TSUBAME E-Science
Journal Vol.10, pp.2-8, Nov., 2013,
http://www.gsic.titech.ac.jp/sites/default/files/TSUBAME_ESJ_10en_0.pdf

【鯉渕グループ】

論文詳細情報(国際)

C-1. Ikki Fujiwara, Michihiro Koibuchi, Hiroki Matsutani, Henri Casanova, “ Skywalk: a Topology for HPC Networks with Low-delay Switches ”, The 28th IEEE International Parallel & Distributed Processing Symposium (IPDPS '14)(米国, フェニックス)(10 pages) May 2014(to appear)(未発表のため現時点では DOI 不明)

【秋山グループ】

論文詳細情報(国際)

D-1. Yuri Matsuzaki, Nobuyuki Uchikoga, Masahiro Ohue, Takahiro Shimoda, Toshiyuki Sato, Takashi Ishida, Yutaka Akiyama, “ MEGADOCK 3.0: a high-performance protein-protein interaction prediction software using hybrid parallel computing for petascale supercomputing environments”, Source Code for Biology and Medicine, vol. 3 No. 8(1), pp.18, 2013 (DOI: 10.1186/1751-0473-8-18)

D-2. Nobuyuki Uchikoga, Yuri Matsuzaki, Masahito Ohue, Takatsugu Hirokawa, Yutaka Akiyama. “Re-Docking Scheme for Generating Near-Native Protein Complexes by Assembling Residue Interaction Fingerprints”, PLOS ONE, Vol. 8, No. 7, e69365, 2013. (DOI: 10.1371/journal.pone.0069365)

D-3. Takehiro Shimoda, Takashi Ishida, Shuji Suzuki, Masahito Ohue, Yutaka Akiyama. “MEGADOCK-GPU: Acceleration of Protein-Protein Docking Calculation on GPUs”, ACM Conference on Bioinformatics, Computational Biology and Biomedicine 2013 (ACM-BCB 2013), 2nd International Workshop on Parallel and Cloud-based Bioinformatics and Biomedicine (ParBio2013), Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine 2013 (ACM-BCB 2013), pp. 884-890, Sep. 2013. (DOI: 10.1145/2506583.2506693)

【鈴木グループ】

論文詳細情報(国内)

E-1. 岡田瑞穂、鈴木豊太郎「大規模ソーシャル・ネットワークシミュレーション基盤の設計と実装」日本データベース学会年次大会 DEIM 2014 第5回データ工学と情報マネジメントに関するフォーラム(査読付き) 2014/3 (DOI:未定)