

「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」  
平成 25 年度採択研究代表者

H29 年度  
実績報告書

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
  - ・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

### (2) 建部グループ

- ① 主たる共同研究者:建部 修見 (筑波大学計算科学研究センター 教授)
- ② 研究項目
  - ・EBD 分散オブジェクトストアの研究

### (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

### (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

### (5) 三好グループ

① 主たる共同研究者:三好 建正

(理化学研究所計算科学研究機構データ同化研究チーム チームリーダー)

② 研究項目

・ゲリラ豪雨予測を可能にする次世代ビッグデータ同化アプリケーションの EBD 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、そのためにスーパーコンピューティング技術、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、昨年度までと引き続き各グループにおける EBD 要素技術の研究開発を推し進めるとともに、グループ間のコ・デザインに基づく開発を加速し、CREST ビッグデータ領域全体で利用可能なシステムソフトウェア・アプリケーションとするべく積極的にグループ間の共同研究・開発を推進した。

EBD アプリケーションのための要素技術として、近年の多様化する EBD 向けアーキテクチャに向けた高性能・高生産性を達成するため、ノードローカルの NVMe デバイス、GPU、Intel Xeon Phi、FPGA のそれぞれの特性を生かしたミドルウェアの開発を継続して行った。また、高並列実行環境で問題となる耐故障性や、ジョブ間の干渉による性能劣化についての研究もを行い、EBD アーキテクチャ全体での実行効率の向上を行った。さらに、クラウド環境での EBD アプリケーションの高効率実行のための性能特性の調査を行った。

EBD ストレージについては、これまで研究開発してきたオブジェクトストアに対し、秋山チームのメタゲノムアプリケーションとのコデザインを進めた。計算ノードのローカルストレージを用いたメタゲノムアプリケーションのワークフローの設計を行い、性能評価を進めた。プロトタイプ実装において、7 ノードまでの評価でスケールすることが分かった。また、ゲノムアプリケーションなどをより高速に実行するため、計算ノードのローカルストレージに高速なフラッシュデバイスや不揮発性メモリを想定した分散ファイルシステムの研究開発を進めた。

ネットワークに関しては、これまでの EBD トラフィックによる最適化技術を発展させ、EBD アプリケーションとのコデザインを行った。その結果、データ転送の一部において、多少の誤差を許容しても結果の大勢に影響しないことを 256 スイッチ規模の EBD システムを対象としたイベントドリブンシミュレーションにより確認し、ある程度の転送時の誤りを許容することで 400Gbps~1Tbps の広帯域かつ低遅延通信を実現できることを示した。その結果、アプリケーションの実行時間を最大 1/5 以下に短縮することができた。

EBD ターゲットアプリケーションについては、当 CREST 事業で開発した相同性解析ソフトウェアの大規模並列版 GHOST-MP に関して国際論文誌 IJMS に発表した。既存ソフトウェア BLAST の並列版である mpiBLAST に比べて、メタゲノム解析等に必要な精度をほぼ保った上で、同じコア数当たりで 80 倍以上高速に計算が実施でき、また「京」コンピュータ上の計測では約 20 万コアまで効率的に動作可能であった。さらに、最新のアルゴリズムを搭載した GHOSTZ-GPU についても大規模並列化を行い、動的スケジューリングを採用することによって、従来の実装に比べて 64 プロセスを越えた領域での実行効率を改善した。また、気象アプリケーションとして、4D-LETKF (4-Dimensional Local Ensemble Transform Kalman Filter)ワークフローを対象に観測データサーチャアルゴリズムの改良を行い、性能分析とその向上を行った。

秋山グループを中心として行ったパブリッククラウドを利用した環境構築については、昨年度までの Azure 上での実装に加えて、本年度は Apache Spark による MapReduce を利用したプロトタイプ実装を行った。クラウド上のデータ消失回避機構である HDFS を利用した場合でも、比較的高い効率を保てることを確認した。

軽量かつ性能に優れた仮想化技術として注目されるコンテナ型仮想化を用いた分散計算環境への対応について、昨年度に実施した Docker に加えて、Singularity および Shifter での実装と評価も実施した。

海外への技術移転の一環として、当事業で開発した3種のソフトウェアを、シンガポールの National Supercomputing Center Singapore が所有する Aspire-1 スーパーコンピューター上、および同国 NTU 大学の SCELSE 研究所サーバに移植し、現地の研究者に開放した。

#### 代表的な原著論文

Kakuta M, Suzuki S, Izawa K, Ishida T, Akiyama Y. A massively parallel sequence similarity search for metagenomic sequencing data, *International Journal of Molecular Sciences*, 18(10): 2124, 2017.

Jian Guo, Kun Qian, Björn Schuller, Satoshi Matsuoka, "GPU-based Training of Autoencoders for Bird Sound Data Processing", *IEEE International Conference on Consumer Electronics-Taiwan 2017 (ICCE-TW 2017)*, Taipei, Taiwan, Jun, 2017

Yanagisawa K, Komine S, Suzuki SD, Ohue M, Ishida T, Akiyama Y. Spresso: An ultrafast compound pre-screening method based on compound decomposition, *Bioinformatics*, 33(23): 3836-3843, 2017.