

「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

H28 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트의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、昨年度に引き続きグラフデータストアや、疎行列ベクトル演算などの基本アルゴリズム、加えて EBD アプリケーション向けのプログラミングフレームワークの大規模実行環境下における性能解析と最適化を行った。また、巨大なデータに効率よくアクセスするためのバーストバッファ技術 CloudBB の開発を継続するとともに、大規模並列実行環境下における EBD アプリケーションのファイルアクセスの特性の精査を行った。また、EBD アプリケーションの資源スケジューリングについても引き続き検討をすすめ、ワークロードをより高密度・高効率に実行できるようになった。さらに、交通シミュレーションに加えて新たな EBD アプリケーションとして昨年度より研究を継続している機械学習アプリケーションについては、性能モデリングとその応用、実アプリケーション最適化について検討をすすめた。また、GPU 以外のデバイスを用いた EBD 実行環境として、FPGA におけるアプリケーションの性能特性の解析を行った。

これまで、ストレージ装置としてフラッシュデバイスや不揮発性メモリを想定し、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の最適化と評価をすすめた。本年度は、数多くのアプリケーションがオブジェクトストアを利用できるようにするため、POSIX インターフェースの設計を行った。並列アクセスによる性能向上を得るため、分散メタデータサーバを用いた。オブジェクトはジャンプコンシステントハッシングを用いて分散オブジェクトストアに格納している。これらにより、並列アクセスにおいて、分散オブジェクトストアのノード数にスケーラブルな性能を達成した。

これまでの EBD トラフィックモデリングに基づく最適化技術を発展させ、かつ、低遅延通信を達成するために配線遅延を抑える cable-geometric 相互結合網の設計手法を開発した。また、昨年度に開発した EBD 向け通信シミュレータ SimGrid において、MPI を利用した評価環境を構築するとともに、ネットワークポロジの構成について、与えられたスイッチ数、そのポート数に応じて、現時点で知られた最適なグラフをデータベースなどから収集し、SimGrid 実行のネットワークポロジとして採用できるように拡張を行った。

EBD×EBD 型の相同性解析を可能とするため、ゲノム解析分野で世界的に用いられている BLASTX ソフトウェアに比べて 185~261 倍の高速な相同性検索を達成できる GHOSTZ を開発した。さらに、GHOSTZ において計算時間のかかるステップを GPU 上で行う GHOSTZ-GPU を開発し、論文発表およびソフトウェア公開を行った。GHOSTZ-GPU では、例えば 12CPU スレッドと

3GPU を併用することにより、既存の GHOSTZ を 12CPU スレッドで使用した場合よりもさらに 5.8～7.7 倍高速な相同性検索を実現できる。これにより大量の微生物ゲノム配列データの解析が可能となり、東京歯科大学との共同研究においては、歯周病患者・健常者由来のデータへの応用を開始し、解析結果が歯周病リスク予測に使える可能性が出てきた。

また、パブリッククラウド上での実応用評価として、大規模並列相同性解析ソフトウェア GHOST-MP およびタンパク質間相互作用解析ソフトウェア MEGADOCK について、Microsoft Azure 上での評価を実施した。GHOST-MP は 30 virtual machine まで、MEGADOCK は 70 virtual machine までノード間並列化性能を計測し、データの相互移動能力が十分ではないクラウド環境上にも関わらず、70%以上の並列化性能が確認でき、パブリッククラウドを用いた並列計算が実用に耐えうることを示した。

気象における次世代ビッグデータ同化アプリケーションに関して、突然の計算機の異常などにより計算機が 15 分間停止した場合を想定したフェールセーフの EBD ワークフローの基本設計を完了した。基本的に計算時間と予報精度はトレードオフの関係にあるが、予報精度を比較的保ちつつ、計算時間を大きく短縮できる同化ウィンドウの長さが存在することが分かった。さらに、松岡グループ及び建部グループと協働して Geographical Search アルゴリズムの最適化に着手した。モデル格子点の近傍にある観測データ点を探索する順方法について Tree を用いた高速化を図り、いくつかの観測シナリオ下での探索時間の計測を行った。さらに観測データ点から近傍にあるモデル格子点を探索する逆アプローチのアルゴリズムも開発した。

#### 代表的な原著論文

Yosuke Oyama, Akihiro Nomura, Ikuro Sato, Hiroki Nishimura, Yukimasa Tamatsu, Satoshi Matsuoka, "Predicting Statistics of Asynchronous SGD Parameters for a Large-Scale Distributed Deep Learning System on GPU Supercomputers", IEEE Big Data 2016, Washington D.C., USA, 2016/12/5

Suzuki S, Kakuta M, Ishida T, Akiyama Y. Acceleration of sequence homology searches by means of graphics processing units (GPUs) and database subsequence clustering, PLoS ONE, 11(8): e0157338, 2016.

Takemasa Miyoshi, Guo-Yuan Lien, Shinsuke Satoh, Tomoo Ushio, Kotaro Bessho, Hirofumi Tomita, Seiya Nishizawa, Ryuji Yoshida, Sachiko A. Adachi, Jianwei Liao, Balazs Gerofi, Yutaka Ishikawa, Masaru Kunii, Juan Ruiz, Yasumitsu Maejima, Shigenori Otsuka, Michiko Otsuka, Kozo Okamoto, and Hiromu Seko, "Big Data Assimilation" Toward Post-Petascale Severe Weather Prediction: An Overview and Progress," in Proceedings of the IEEE, vol. 104, no. 11, pp. 2155-2179, 2016.