

「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

H27 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

(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

(3) 鯉渕グループ

- ① 主たる共同研究者:鯉渕 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) 鈴村グループ

- ① 主たる共同研究者:鈴村 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

(6) 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、昨年度に引き続き、EBD システム処理基盤としての基本アルゴリズム、ストレージ装置およびインターコネクトと、大規模ゲノム・社会シミュレーション・気象を含むアプリケーションに関するコ・デザイン体制の研究を推進することに加えて、新たにディープラーニングフレームワークの高速化に着手するとともに、EBD 基本アルゴリズムを実現するソフトウェアのオープンソース化にむけたハードニングを行った。

EBD 基本アルゴリズムに関しては、大規模グラフ処理および NVM を用いたソートアルゴリズムの性能最適化をすすめた。また、EBD オブジェクトの資源スケジューリングに関する研究開発を加速し、アクセラレータ仮想化による高効率実行およびアプリケーションマイグレーションや、電力消費最適化による省電力実行など、EBD アプリケーションにおけるデータおよび計算の移動とコロケーションのトレードオフによる性能の最適化に向けた道筋をつけた。

EBD に向けた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。これまでのハードディスクドライブでは、並列にアクセスすると逆に性能が落ちてしまうが、フラッシュデバイスなどでは、並列にアクセスすることにより性能を上げることが可能となる。そのため、並列アクセス可能なローカルオブジェクトストアの設計は重要である。並列アクセスを可能とするためには通常ロックが利用されるが、ロックを用いると並列効率が下がってしまう。そのため、ロックを用いないノンブロッキングな設計を行っている。評価の結果、オブジェクト作成性能について従来システムに比べ 4.3 倍から 10 倍の性能向上が得られた。

EBD インターコネクトについては、リング構造を含むネットワークポロジであればトーラスのような規則性を保持する必要がないことに注目し、これまでに開発したスモールワールド性による低通信遅延と規則性を生かした高スループットを両立する EBD トポロジのソートなどの性能と初期導入コストの優位性を示した。

EBD×EBD 型の相同性解析を可能とするため、超高速なメタゲノム解析ソフトウェア GHOST-MP の改良を進めた。本年度はクラスタリングに基づく新規アルゴリズムと、データ構造の改良により高速化を達成した GHOSTZ ソフトウェアを開発した。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 倍の高速化を達成。大量の微生物ゲノム配列データの解析が可能となり、ヒト腸内や口腔内の細菌集団の比較解析を実現した。さらに、東京歯科大学との共同により、歯周病患者・健康者のヒト由来データへの適用を開始し、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、膨大な計算結果のアーカイブ化や可視化についても、本年度にクライアントソフトウェアの第一版を作成した。成果普及のため、MEGADOCK と GHOST-MP をパブリック・クラウド上に移植し、前者では 1120 コア、後者では 480 コアまで、MPI 上での良好なスケーラビリティを確認した。これらをさらに大量に分散動作させれば、EBD 解析も

可能となる。

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 ワークフローとして、4 次元アンサンブルカルマンフィルタを実装し、検証した。その結果、計算時間と予報精度はトレードオフの関係にあることが判明した。また、ビッグデータ時代のデータ同化を改善するための応用調査として、10240 個のアンサンブルカルマンフィルタで現実大気の観測データと解像度 112km の全球大気モデルを使ったデータ同化実験に成功し、大気の誤差構造などに関する理解を深めた[2]。さらに、Geographical Search アルゴリズムの最適化に着手した。

社会シミュレーションについては、道路交通シミュレーションを題材に、大規模グラフ上で動作するエージェントベースのアプリケーションの負荷平準化を行い、SUBAME2.5 を 16 ノード用いて行ったダブリン市の道路交通網シミュレーション実験では、従来手法と比較して 4 倍の速度での実行を達成した。

また、新たな EBD アプリケーションとして、ディープラーニングアプリケーションの性能解析と高速化に着手した。ディープラーニングアプリケーション SPRINT の GPU 向け最適化および処理の改善を行い、この分野では例のなかった 32 ノード、96 GPU を用いた大規模並列実行を実現した。また、SPRINT の性能改善のために、実行時間およびミニバッチサイズ(1 回のパラメータ更新時に利用される学習データ量)を算出する性能モデルの作成を行った。性能モデルを用いることで認識性能を左右するミニバッチサイズを固定した状態で最も効率の良い実行パターンを探索することができ、結果として TSUBAME-KFC/DL において TSUBAME2.5 で最高の認識率を達成した条件と同じ学習が約 70%の時間で実行できること示せた。

代表的な原著論文

- [1] Suzuki S, Kakuta M, Ishida T, Akiyama Y. Faster sequence homology searches by clustering subsequences, PLoS ONE, 31(8): 1183-1190, 2015, DOI: 10.1093/bioinformatics/btu780
- [2] Takemasa Miyoshi, Keiichi Kondo and Koji Terasaki, "Big Ensemble Data Assimilation in Numerical Weather Prediction", Computer, vol.48, no. 11, pp. 15-21, 2015, DOI: 10.1109/MC.2015.332
- [3] Jens Domke, Torsten Hoefler, Satoshi Matsuoka "Routing on the Dependency Graph: A New Approach to Deadlock-Free High-Performance Routing", HPDC'16, Kyoto, Japan, June 2016 (to appear)