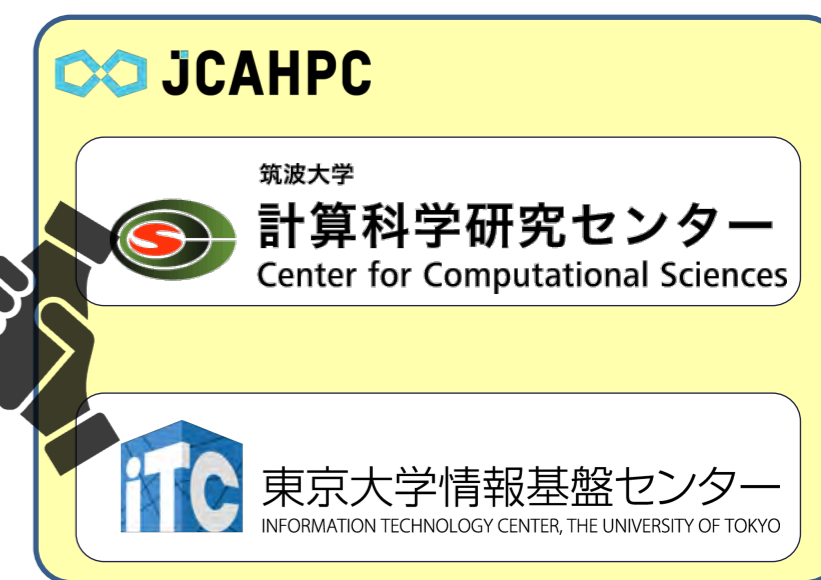


次世代演算加速装置とそのファイルIOに関する研究



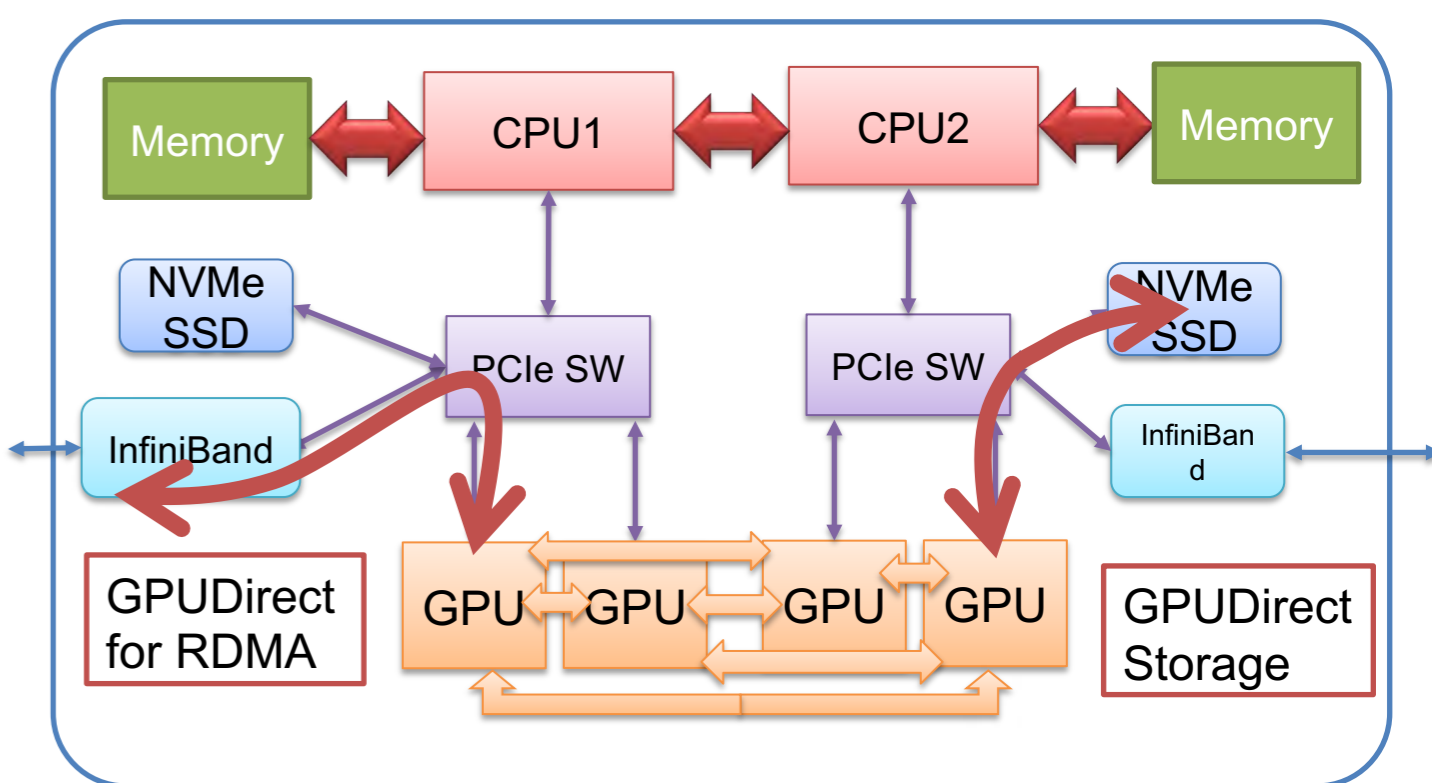
概要

- 実アプリケーションでGPUを利用するには、GPUに対するデータ入出力の考慮が必要であり、データ転送と演算のオーバラップ、転送レイテンシの短縮を工夫する必要がある。
- 従来GPUからファイル IO を行うには、ホスト CPU を介して行う必要があったが、NVIDIAによってGPU Direct for RDMA をファイル IO に拡張した **GPU Direct Storage (GDS)** が提供されている。しかし国内にはGDSが使える大規模システムはまだ存在しない。ストレージ側の適切な構成についても検討が十分ではない。

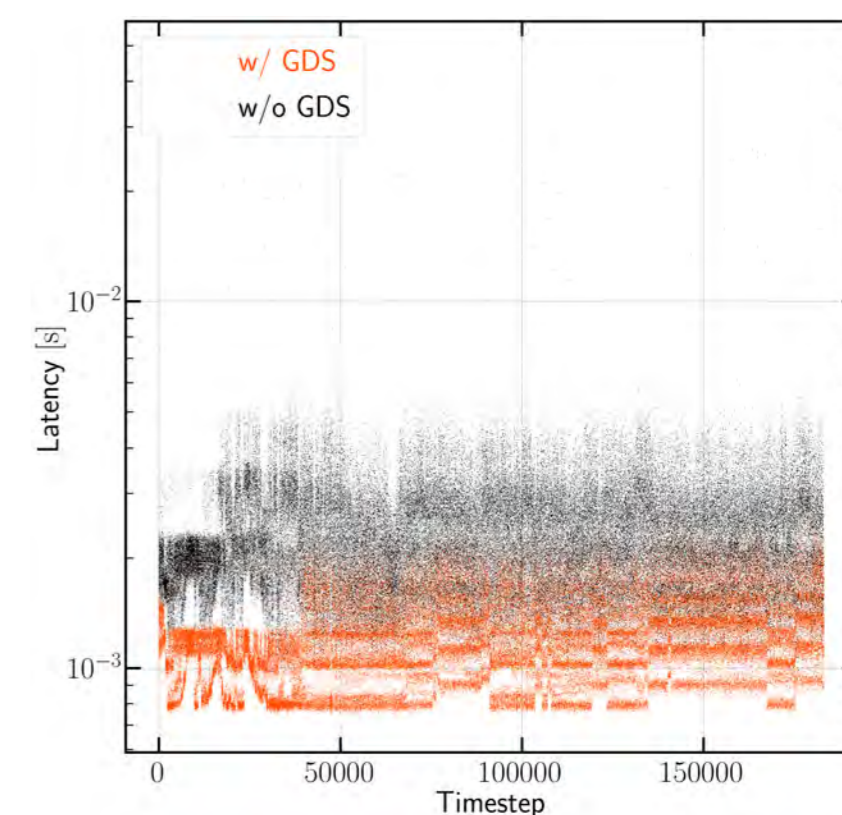


研究目的

- GPU上データの直接ファイル IO, あるいは計算とファイル IO のオーバラップを容易に取り扱い可能にする手法を確立し、様々なファイル入出力特性を持つ実アプリケーションにおいてGPU-ファイル IO 間の処理を効率化する。
- 最先端共同 HPC 基盤施設(JCAHPC)の次期システム“Miyabi (OFP-II)”で採用するGH200 (Grace Hopper)での性能評価を実施する。
- この成果を、Miyabi (OFP-II)の運用に反映する。TSUBAME4.0でも実証実験を行う。



GPUノード構成イメージ

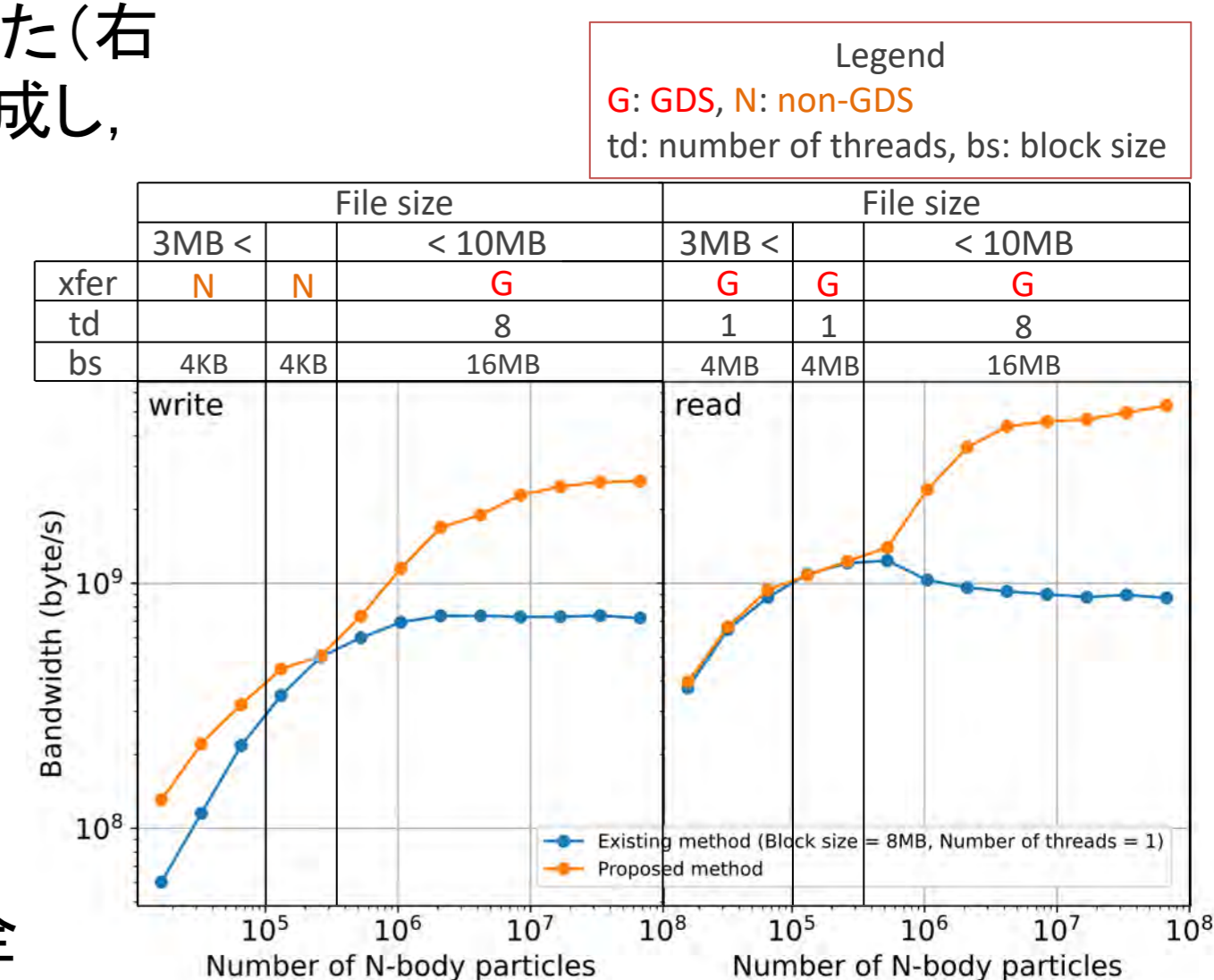


GDS使用の有無によるデータ出力時間の比較

対象アプリケーション

元矮小銀河の中心ブラックホールがまとう星団の力学進化

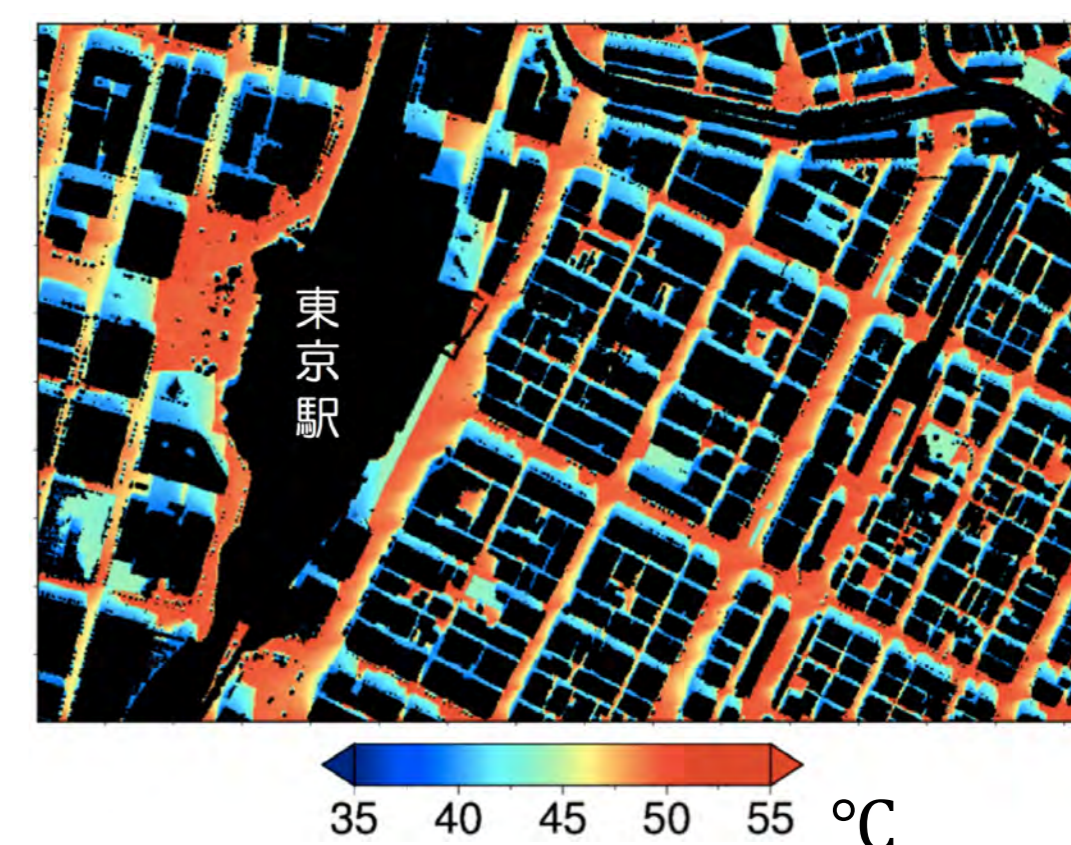
- かつて矮小銀河の中心に位置していたブラックホールが比較的小さな星団をまとった状態でアンドロメダ銀河の北東領域を軌道運動していると示唆されている [1]。銀河衝突に付随して形成されているため、N体シミュレーションを用いてこの星団の力学的性質を調べる必要がある。元矮小銀河を含めた全体の進化に比べ、星団を構成するブラックホール周辺粒子の挙動を調べるには高頻度の解析データ取得が望ましい。
- 昨年度までの成果と今年度の展望
 - 完全GPU化されたツリーコードGOTHIC [2]に、HDF5を介してGDSを利用する機能を追加した。実験サーバ上での性能評価においては、GDS使用(赤)によるファイル出力の高速化が確認できた(右図)。HDF5のAPIを用いたGDSによるファイル入出力性能を測定するベンチマークコードを作成し、それを元にHDF5向けGDSプラグイン(vfd-gds)の性能を最大6倍改善した[3]。
 - 今年度はさらなるHDF5+GDSの性能改善、およびGH200環境、TSUBAME4.0での測定を行う。



LustreFSIにおけるGDSの性能改善

都市街区気象LESを用いた高解像度熱環境シミュレーション

- 近年の都市の暑熱環境悪化への対策のため、街路樹やドライミストの設置、風通しの良い都市の設計などのアイデアが提案されている。これらの効果を精度良く評価するため、都市を街区レベルで解像し、熱環境を評価できるシミュレーションコードの開発・利用が重要となる。
- **使用コード**: 都市街区気象LESモデルであるCity-LES [4]、暑熱計算にかかる箇所をGPU化
 - 高解像度シミュレーションを行う場合ファイルサイズが大きくなるため、出力の際GDSを用いることでI/Oにかかる時間の削減が見込まれる。また、時間詳細な出力を必要とする場合、全体の実行時間に対するファイル出力の時間の割合が大きいため、計算と出力をオーバラップする等の工夫を施すことでより高効率な計算が可能となる。
 - 最適化後の性能を測定するためのCity-LESをベースとしたベンチマークソフトにより評価、最適化: 気流計算部のうち移流拡散過程, Poisson方程式の求解部+計算結果の出力部
 - Parallel netCDFを用いたノンブロッキングIO



機械学習におけるトレーニング

- MLPerf ベンチマーク(Training)を例に、巨大データセットを読み込みながら複数ノードのGPUを用いて訓練を行うワークロードについて、GDSの有無に関する評価や、ストレージの構成に合わせた最適化を実施
- **使用コード**: MLPerf, PyTorch, DALI, KvikIO等

文献

[1] Y. Miki, M. Mori, T. Kawaguchi, and Y. Saito: "Hunting a Wandering Supermassive Black Hole in the M31 Halo Hermitage", The Astrophysical Journal, 783, 87, Mar. 2014

[2] Y. Miki, M. Umemura, "GOTHIC: Gravitational oct-tree code accelerated by hierarchical time step controlling", New Astronomy, vol. 52, Apr. 2017, pp. 65-81.

[3] M. Tominaga, T. Hanawa, Y. Miki, "Toward Optimizing File IO on GPU Clusters", GTC 2024 poster

[4] R. Ikeda, H. Kusaka, S. Iizuka, and T. Boku: Development of Urban Meteorological LES Model for Thermal Environment at City Scale. Proc. of the 9th international conference for Urban Climate, Jul. 2015.