

SINETを介したデータベース基盤とHPC基盤の連携による医療画像解析基盤実現に関する研究

Research on realization of medical image analysis platform by linking database platform and HPC platform via SINET

研究代表者：村尾 晃平 (国立情報学研究所) / 副代表者：森 健策 (名古屋大学)

共同研究者：佐藤 真一・合田 憲人・大江 和一・二宮 洋一郎 (国立情報学研究所) / 明石 敏昭 (順天堂大学) / 大竹 義人・崇風 まあぜん (奈良先端科学技術大学院大学)

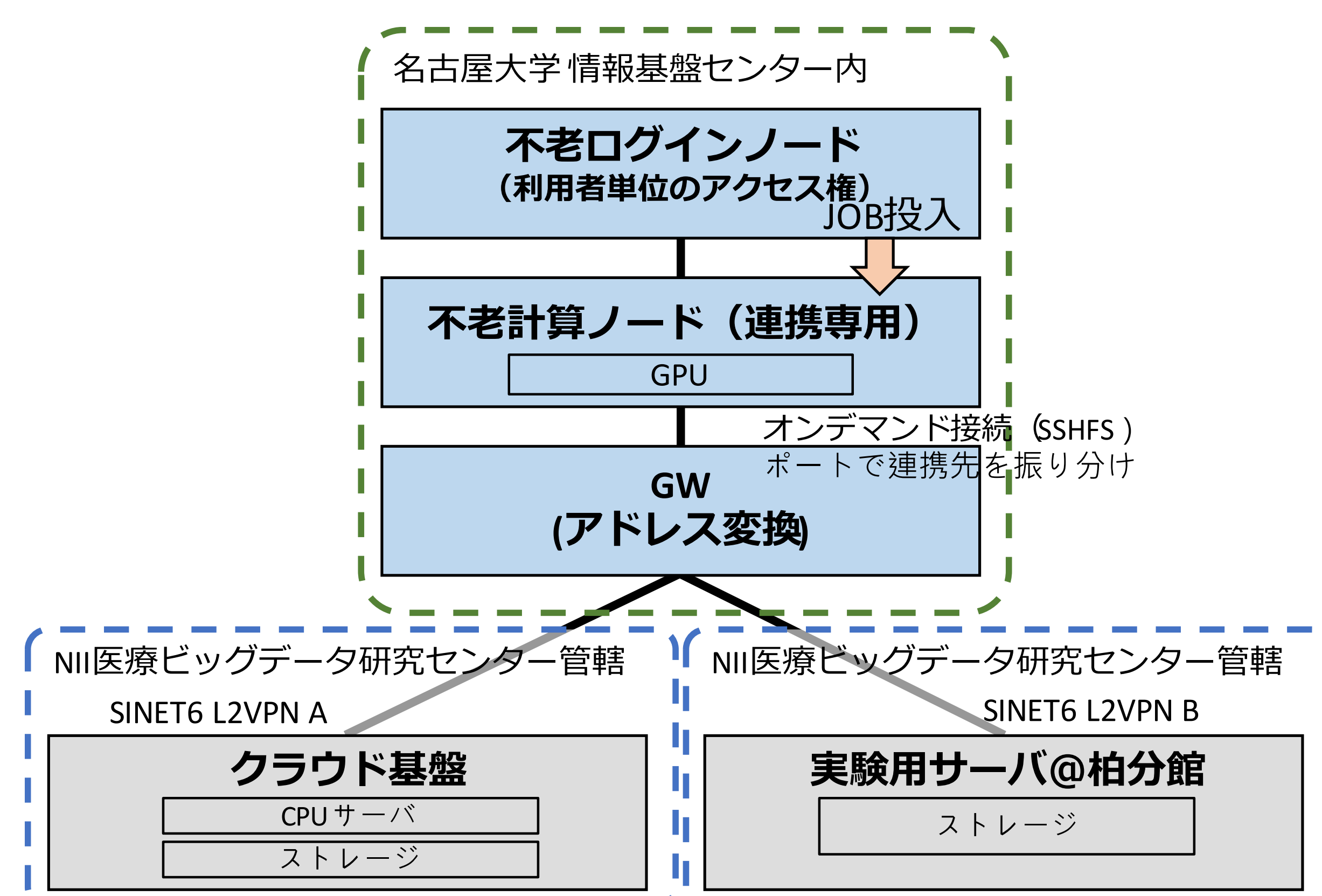
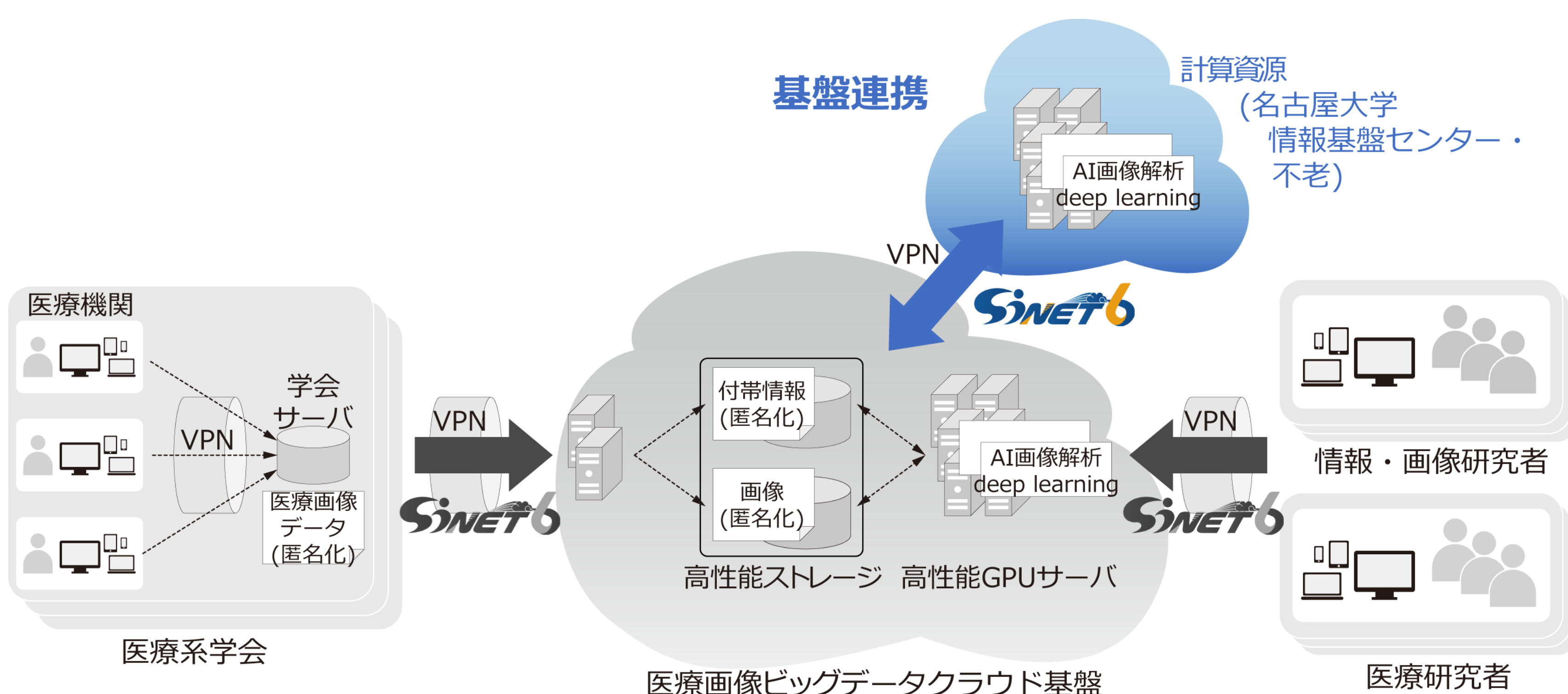
研究の背景と目的

画像診断支援AIを実現する学習データには多様性が求められ、異なった撮影装置・撮影条件でのデータが必要となる。我々は医療系学会を通して全国の医療機関から悉皆的に多彩な医療画像を収集し、AI開発が可能な医療画像ビッグデータクラウド基盤（以降クラウド基盤と記す）を構築した。2018年から2024年5月末までで6つの医療系学会を通じて約5.5億枚の医療画像を蓄積している。

このクラウド基盤には小規模であるがGPUサーバが接続され、NIIを含む国内14の情報系研究チームが累計40以上の医療AI研究開発タスクに取り組んで

いる。しかし、データ・タスク・利用者の増加に伴って計算資源に対する需要は急激に増大し、単一拠点の計算・データベース資源だけで対応することは困難となりつつある。

そこで、データを保有するクラウド基盤とHPC資源を提供する大学のセンターとをSINETを介して接続することで分散計算基盤を構成することを考えた。しかし、医療AI研究が求めるセキュリティと性能を確保するための分散計算基盤を実現する方式は確立されていない。本研究では、医療AI研究開発のために、データベース基盤とHPC基盤を安全かつ実運用可能な速度性能で連携するシステムの構築を目的とする。



2023年度の成果

以下の結果により、SINETを介したデータベース基盤とHPC基盤を連携する仕組みが医療画像解析に適用できる見込みが得られた。

(1) クラウド基盤と名古屋大学の不老との連携方式の策定と実装

クラウド基盤はインターネットから隔離されており、利用許可のある施設・メンバーのみがアクセスできるようになっている。その仕組みを踏襲した上で、基盤連携に特有の基本方針として「クラウド基盤内の医療情報を外に残さない」こととした。そのための体系的な要件を次のように設定した。

- ① 利用者は、計算に必要な時のみクラウド基盤に接続する
- ② 計算終了後に医療情報をクラウド基盤の外部に残さない
- ③ 外部の計算環境にて、クラウド基盤にアカウントの無いユーザはクラウド基盤にアクセスできない

実際に、名古屋大学の不老から専用の計算ノードを切り出し、GWを介してJOB管理によりオンデマンドsshfs接続をすることにより、上記の要件を満たすシステムが実現できた。

(2) AI開発の速度性能検証・ネットワーク一時遮断に対する頑健性実験

オープンソースのAI-benchmarkを使い、画像をローカル環境に置いた場合と基盤連携環境に置いた場合で処理時間を比較した所、物理的距離に比

例して画像読込時間のみが長くなることがわかった(a)。

次にCT画像を用いてSegmentationの学習・テストの処理時間を測定した所、基盤連携環境でもローカルで処理した場合とほとんど変わらなかった(b)。通信状況を見ると、スパイク状になっていて通信が律速になっていない(c)。これは、CT画像がシリーズごとに1ファイルにまとめられ、GPUメモリにバッチサイズ分のデータが保持できるからと考えられる。

さらに、柏分館との連携環境を使い、ネットワーク一時遮断に対する頑健性の実験を行った。Segmentationテスト中に5分ほどの遮断を設けても、回復後は最後まで問題なく計算できることを確認できた(d)。

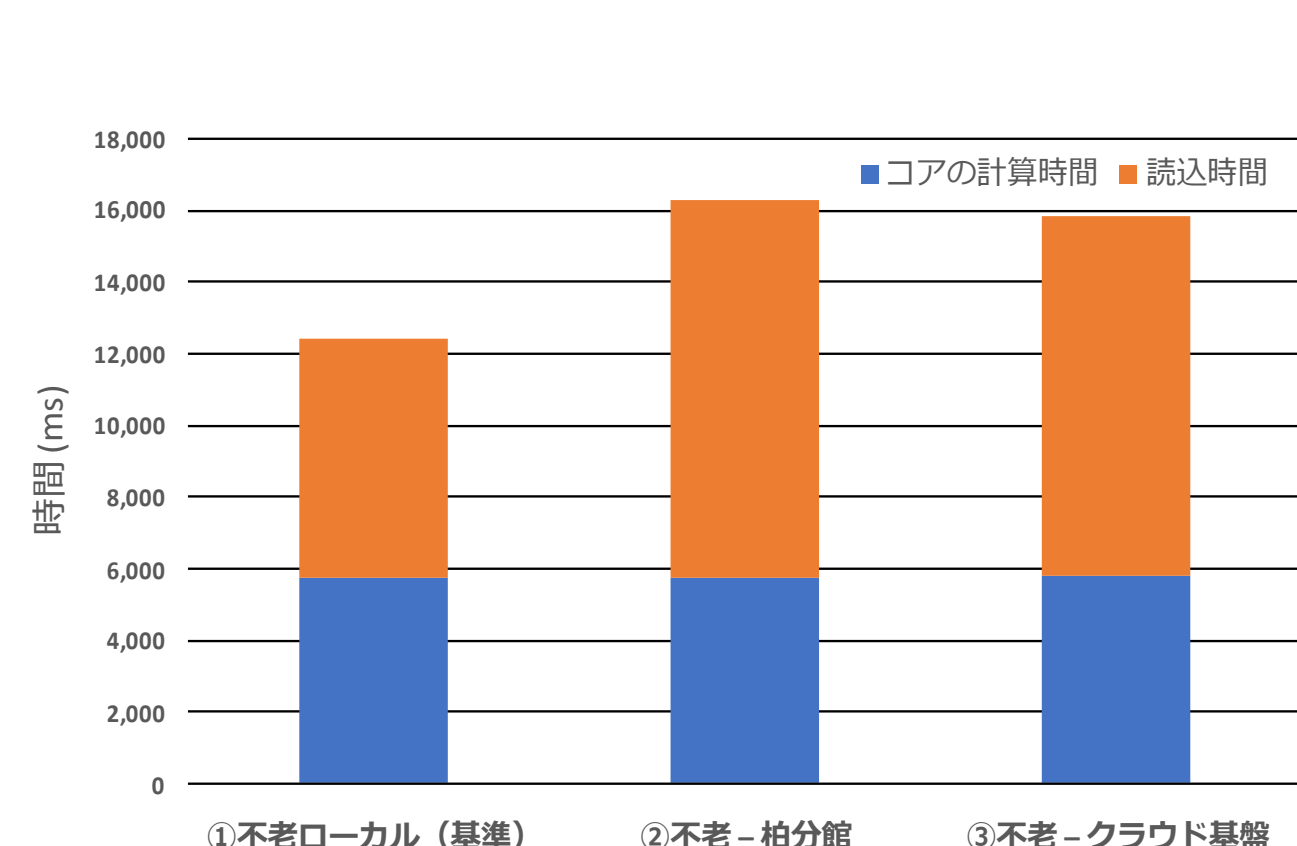
(3) 基盤連携を利用した医療画像AIモデル構築 (2024年度継続)

医療画像としては超大規模であるデータベースを活用して、AIモデルの構築を進める。

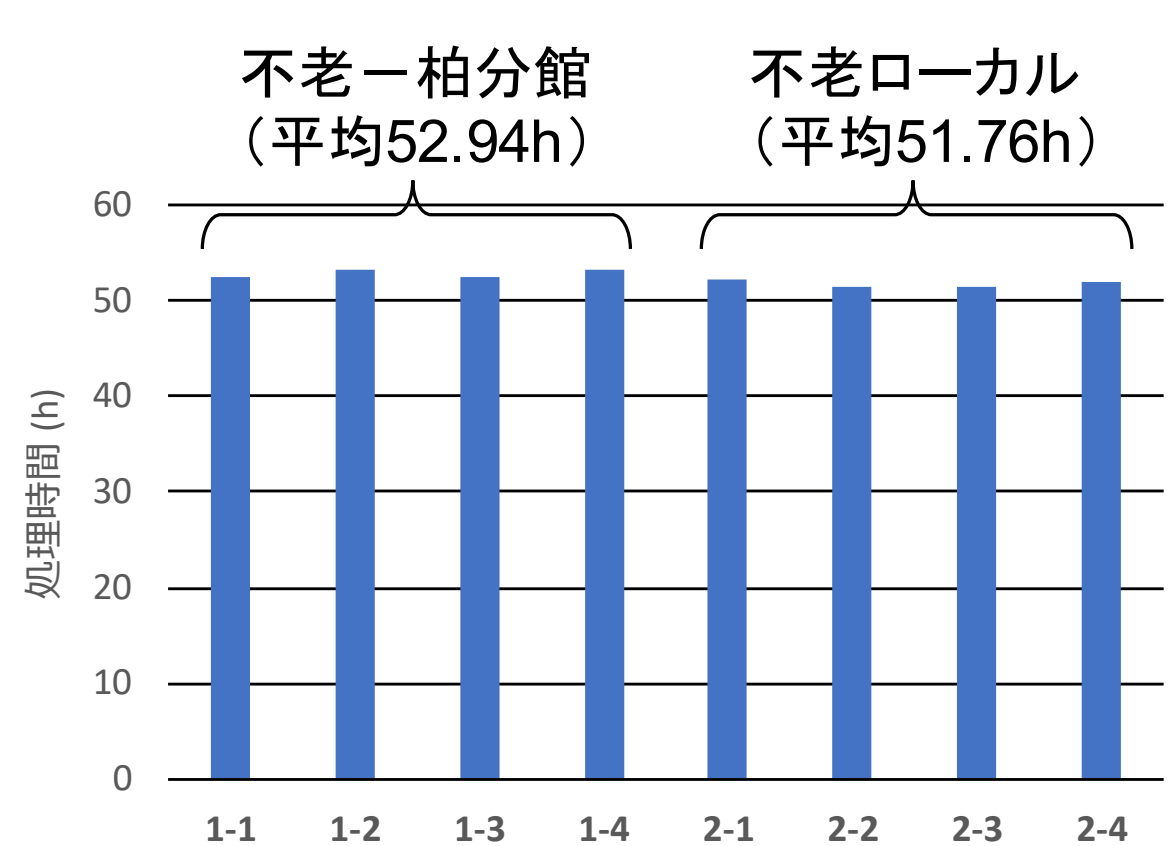
- ・CT腹部画像3万件以上のデータでFoundation Model作成
- ・CT画像4万件の筋骨格系セグメンテーション・解析

上記Foundation Modelは計算中で、安定して1.3Gbpsのパフォーマンスで動作している(e)。

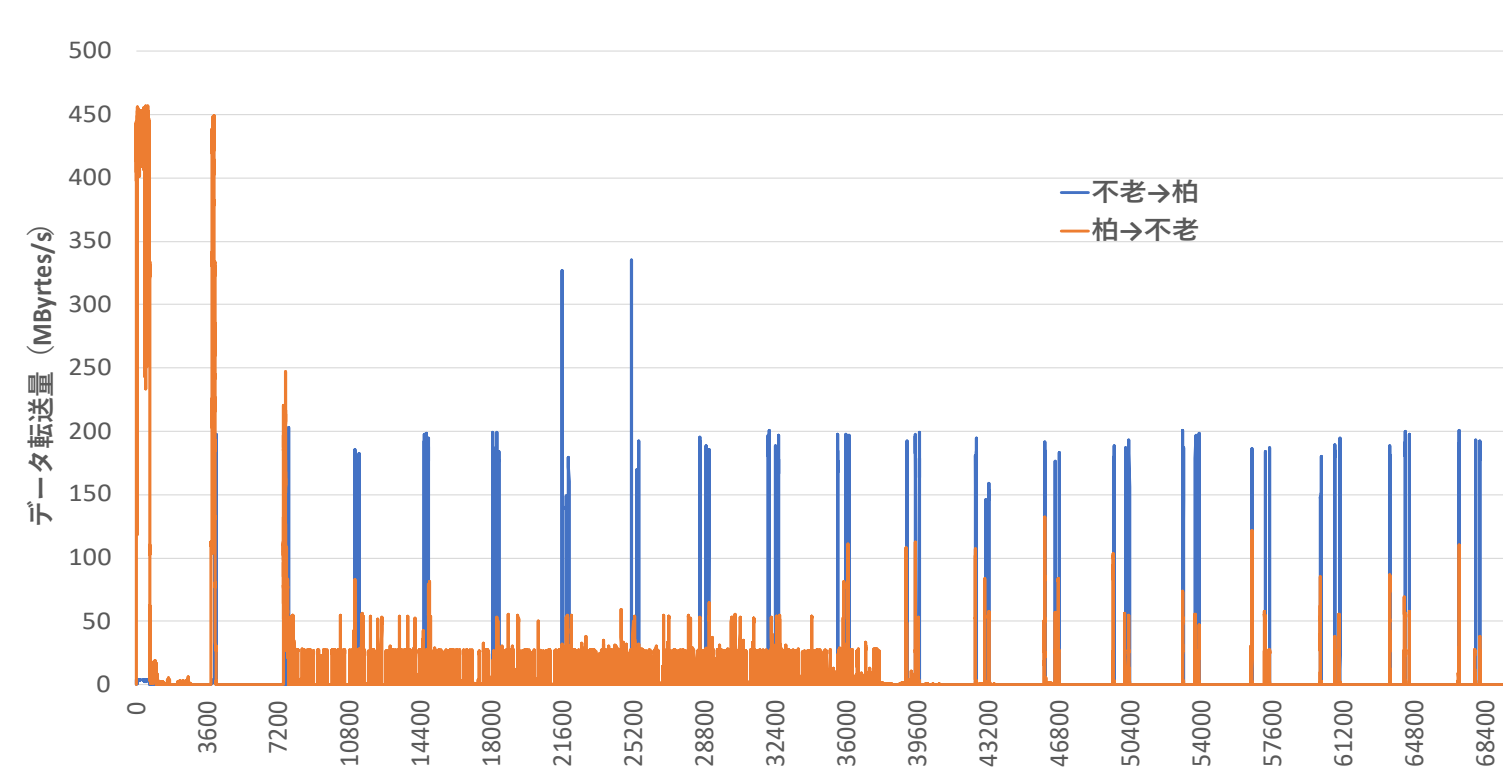
今後、筋骨格系の計算もできるように名古屋大学の不老のGPUを追加で切り出し、インフラ強化を進める。またmdxとの連携についても、セキュリティを考慮したシステムを設計・構築していきたい。



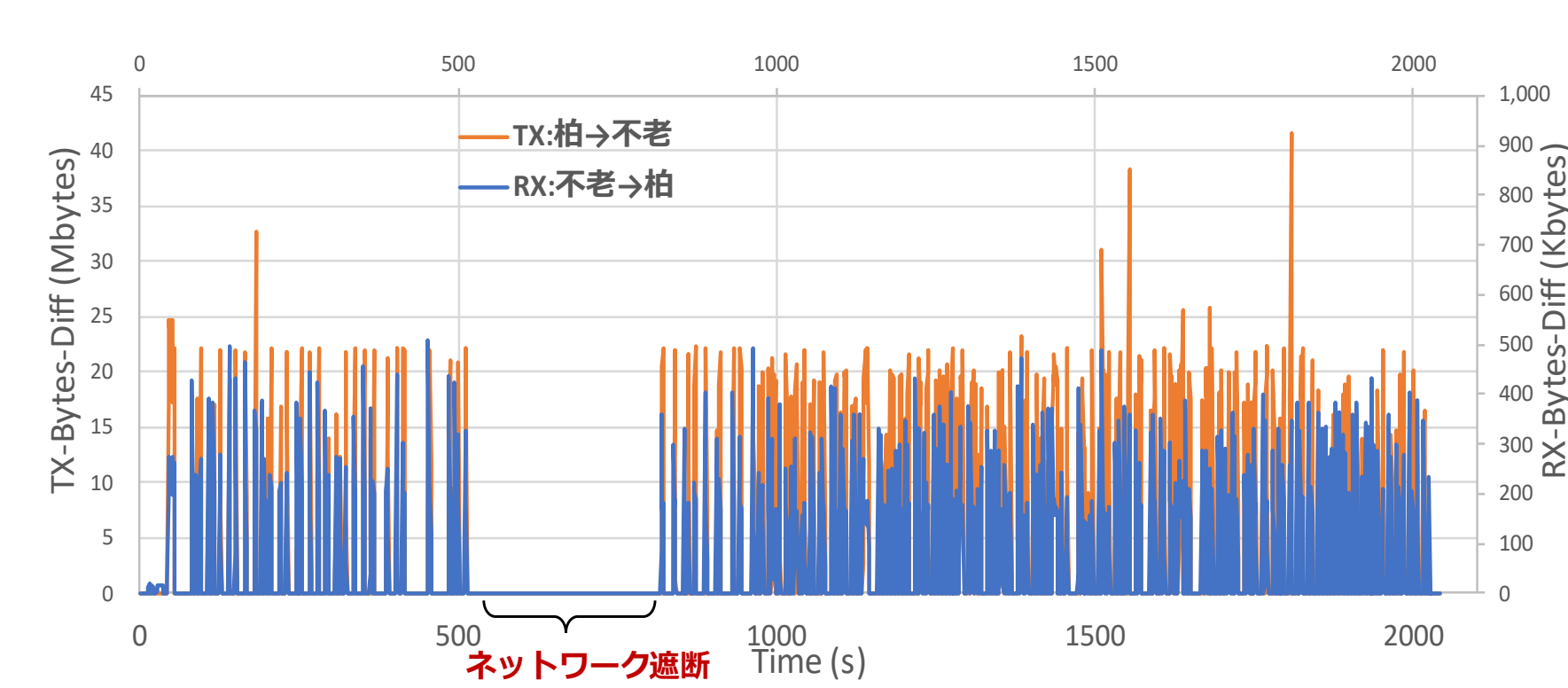
(a) AI-benchmarkを使った処理時間測定



(b) CT画像を使った並列処理時間測定

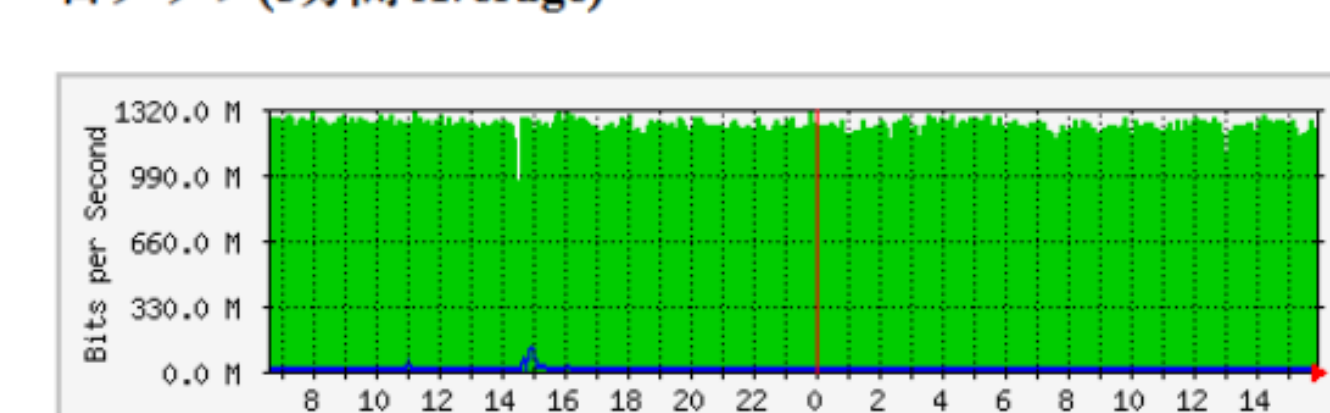


(c) CT画像を使った単体処理時のGWでの通信量測定



(d) 頑健性の実験時の柏分館での通信量記録

日グラフ(5分間 Average)



最大 1306.3 Mb/s (1.3%) 平均 1229.6 Mb/s (1.2%) 最小 1227.1 Mb/s (1.2%)
送信 109.8 Mb/s (0.1%) 11.8 Mb/s (0.0%) 10.8 Mb/s (0.0%)

(e) Foundation Model作成時のスイッチ上での通信量測定

謝辞

クラウド基盤は、戦略的イノベーション創造プログラム (SIP)「統合型ヘルスケアシステムの構築」JPJ012425の補助を受けて運用している。名古屋大学情報基盤センターの片桐孝洋教授と大島聡史准教授 (現在、九州大学情報基盤研究開発センター) には基盤連携の設計で助言をいただいた。名古屋大学情報基盤センターの山田一成氏には不老からの計算ノードの切り出しやネットワーク設定などでお世話になった。国立情報学研究所の陳明康氏にはAI-benchmarkを今回の評価実験に使えるよう手を加えて頂いた。