

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：合田 憲人・佐藤 真一・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)／黒瀬優介（東京大学）

研究の背景と目的

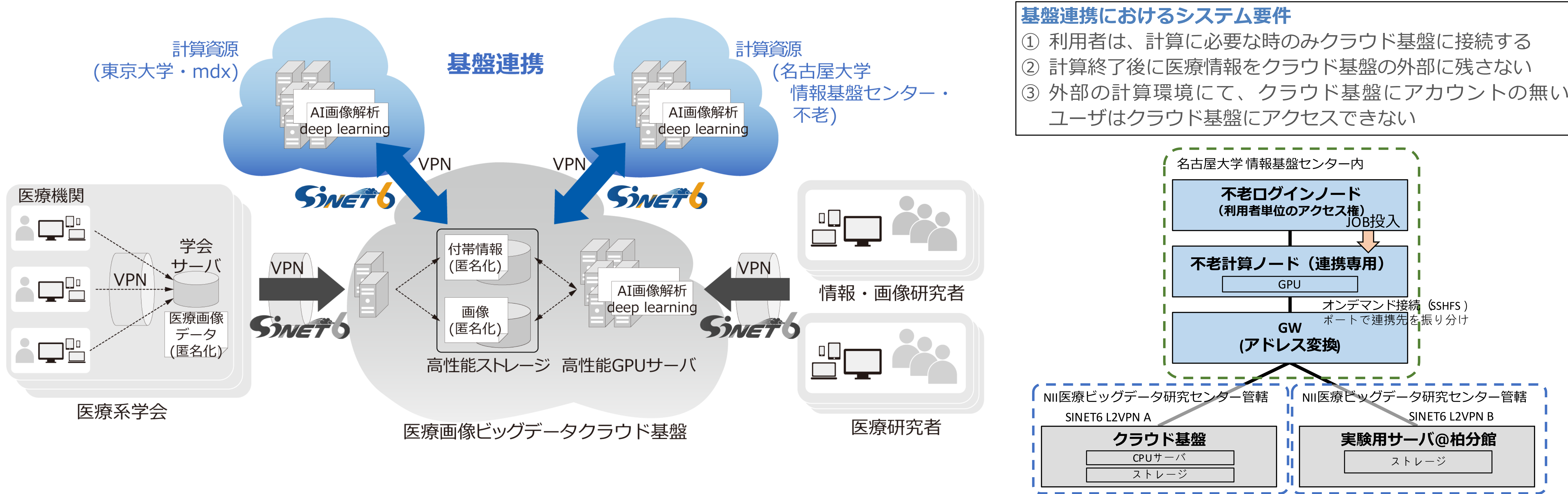
画像診断支援AIを実現する学習データには多様性が求められ、異なった撮影装置・撮影条件でのデータが必要となる。我々は医療系学会を通して全国の医療機関から悉皆的に多彩な医療画像を収集し、AI開発が可能な医療画像ビッグデータクラウド基盤（以降クラウド基盤と記す）を構築した。2018年から2025年5月末までで6つの医療系学会を通じて約7.8億枚の医療画像を蓄積している。

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医療AI研究開発に取り組んでいる。しかし、データやモデルの大規模化、利用者の増加に伴って計算資源に対する需要は増大しており、単一拠点の計算資源だけで対応することは困難となりつつある。

そこで我々は、データを保有するクラウド基盤とHPC資源を保有する大学のセンターとを連携して分散計算基盤を実現することを目指している。利用者や利用目的が限定されるような機微データを扱うため、特にセキュリティの確保を重要視している。

2023年度には、基盤連携におけるシステム要件を策定し、名古屋大学の情報基盤センターの不老との連携を実現した。実験の結果、遠隔に起因する遅延はあるものの、AI学習においてはトランザクション回数が多くなくて実用上の問題なく、並列計算に対しても耐性があることがわかった。

2024年度にはHPC連携を強化するため、SINET L2VPNのネットワーク構成を見直し、同時にセキュリティ強化を行った。



2024年度の成果

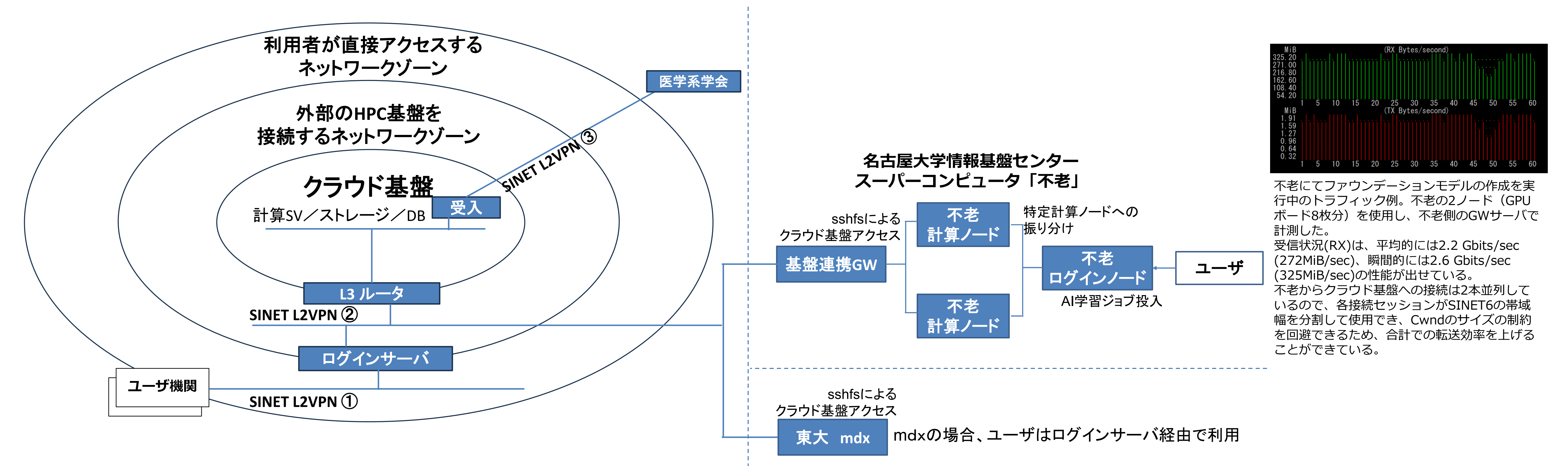
以下の取り組みにより、基盤連携のセキュリティを強化すると共に、新たなHPC基盤を接続できるネットワーク環境を整えた。

(1) ログインサーバの導入

これまでのクラウド基盤では、ユーザ機関はSINET L2VPN経由ではあるが、クラウド基盤の中のサーバに直接接続していた。これでは輻輳のようなトラブル時に影響が広がってしまう。そこで、新たにログインサーバを設置し、ネットワーク境界による防波堤のような仕組みを導入した。

(2) 新たなSINET L2VPNの緩衝地帯の設定

ログインサーバとクラウド基盤の間に新たなSINET L2VPNを設定し、外部のHPC基盤をそこに接続することとした。名古屋大学の「不老」のように独自にログインノードを持ち、ジョブ管理ができている計算基盤はそのまま接続でよいが、東京大学のmdxのように計算ノードに直にログインし、ジョブ管理なしで使えるようなHPC基盤の場合には、NIIが用意したログインサーバを経由して利用することとした。



2025年度の予定

これまで、一か所のデータ基盤に対してHPC基盤を接続することを実験してきたが、複数のデータ基盤を利用したAI開発することを想定した実験を行いたい。

また、多数のAI開発が同時進行することを想定し、データ基盤側の読み書きのボトルネックがどのようなものになるか、ディスクの物理環境を変えるなどの実験をしたいと考えている。

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ている。
基盤連携の設計時には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）に助言をいただいた。