

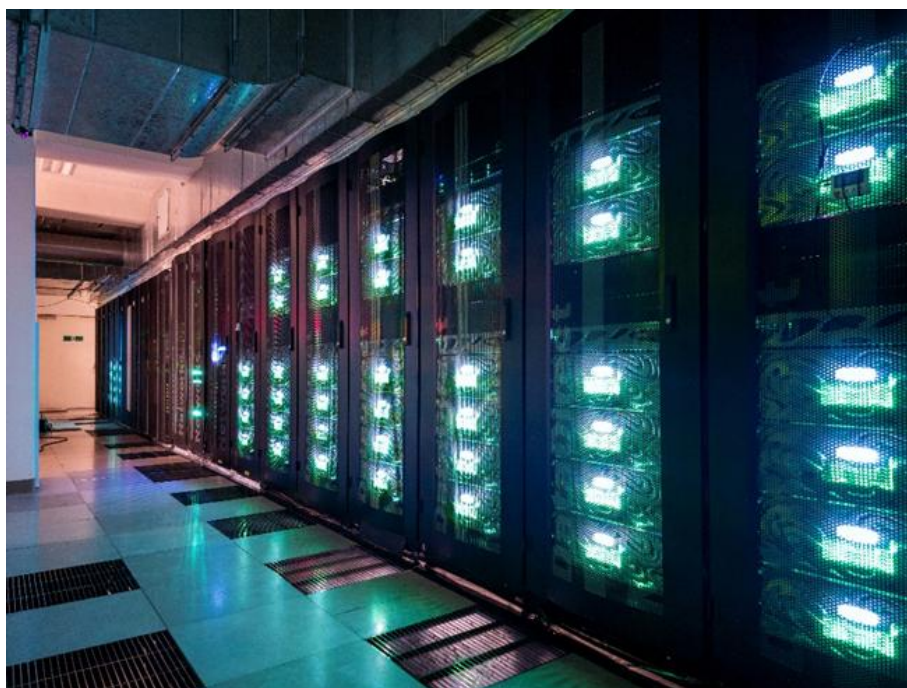
2026年4月1日
東京大学医科学研究所

ゲノム解析に特化した データサイエンスコンピューティングシステム Shirokane8 を稼働

——運用実績に基づくリソースの最適化と堅牢な解析基盤の提供を実現——

発表のポイント

- ◆ 最新型のデータサイエンスコンピューティングシステム Shirokane8（以下、Shirokane8）を2026年4月に稼働開始しました。
- ◆ Shirokane8 の稼働により、東大医科研附属ヒトゲノム解析センター（以下、HGC）の生命科学データサイエンス用スーパーコンピュータシステム SHIROKANE はCPU総コア数31,180コア、ストレージ容量41PBとなり、より効率的かつ安定運用が可能なデータ解析基盤の構築を実現しました。
- ◆ 令和7年11月21日に閣議決定された「ゲノム医療施策に関する基本的な計画」に示されている通り、ゲノム情報を用いた研究開発が社会に果たす役割はより一層重要なものとなっています。HGCは今後も利用者目線にたったデータ解析基盤の提供を通し生命医科学研究の発展に貢献していきます。



スーパーコンピュータシステム SHIROKANE

概要

東京大学医科学研究所 附属ヒトゲノム解析センター（以下、HGC）（センター長：井元 清哉）は、ゲノム情報をはじめとするデータ解析基盤を強化するため、最新型データサイエンス用スーパーコンピュータシステム SHIROKANE8 を 2026 年 4 月 1 日に稼働し、同日から利用者へ提供を開始しました。Shirokane8 は株式会社日立製作所が構築を担当し、運用期間中の安定稼働を支援します。

Shirokane8 は、これまでの運用実績から計算リソースを最適化した構成を採用しています。また、GPU 搭載サーバにはバイオインフォマティクス解析を実行するうえでコストバランスの優れた NVIDIA L4 GPU を採用しました。稼働にあたり、前世代システムに格納されていた約 30PB のデータの移行を、ユーザへの影響を最小限に抑えつつ、約 110 日間で完遂しました。また、新たに整備した専用ストレージ室の運用も開始し、研究データの堅牢な管理を実現します。

背景・課題

ヒト一人分のゲノム配列を決定したヒトゲノム計画の完了からおおよそ四半世紀が過ぎ、ゲノム解析の技術は現在も目覚ましく進歩し続けています。近年では、2019 年の新型コロナウイルス感染症パンデミックにおける重症化リスクとなるゲノム多型を特定するための世界的なデータ収集や、本邦において 2019 年から始まったがん・難病等における国家プロジェクト「全ゲノム解析等実行計画」の推進により、ゲノム解析および産出される膨大なゲノムデータの解析の重要性は上昇の一途をたどっています。東大医科研は、全ゲノム解析等実行計画において、2025 年度末までにがん領域約 17,000 症例の全ゲノムデータ解析を完了しました。

こうした背景から、SHIROKANE（注）への期待とニーズは高まり続けており、学術機関だけでなく民間機関も含めて利用者は、前システム Shirokane6 稼働開始時の 2022 年 4 月時点から約 1.7 倍の 4,180 アカウントを突破しました。加えて、HGC では次世代の研究者育成を支援し、国内の大学・研究機関が実施するスーパーコンピュータ（スパコン）利用の講習会や実習用のアカウント提供も積極的に行っています。その申請数は、Shirokane6 稼働開始時と比べ 4 倍の年間約 900 アカウントまで増加しています。このように急増する解析ニーズとデータ量に対応するため、計算リソースの拡充による処理能力向上と、膨大な研究データを長期間安定して保存できる強靱な基盤の整備を目指しました。

今回の取り組み

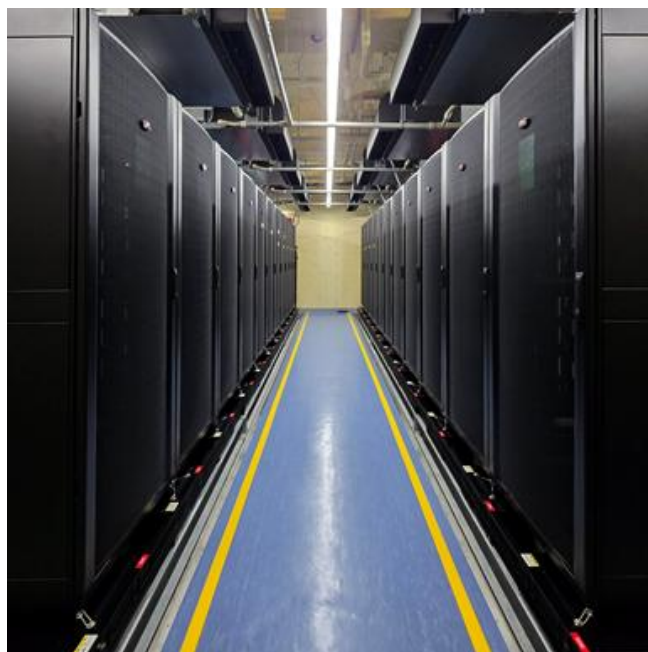
2026 年 4 月に稼働を開始した Shirokane8 は 6,432CPU コアの計算サーバ群と 41PB の大容量高速ストレージを有し、将来的なデータ増加に備え容量の増強が可能な構成を採用しました。

計算サーバ群には、これまでのジョブ実行状況と運用実績に基づき、1 ノード当たり 192 コアを搭載できる最新世代の Intel CPU を採用し、コア当たり 4GB のメモリ構成とすることで、ジョブ実行リソースの最適化を図っています。また、GPU 搭載サーバを 8 台導入し、バイオインフォマティクス解析や AI を用いた解析を実行するうえでコストバランスの優れた NVIDIA L4 GPU 1 台あたり 8 基の計 64 基採用しました。これによりゲノムデータ解析を高速化しつつ、コストパフォーマンスに優れた環境を実現しています。

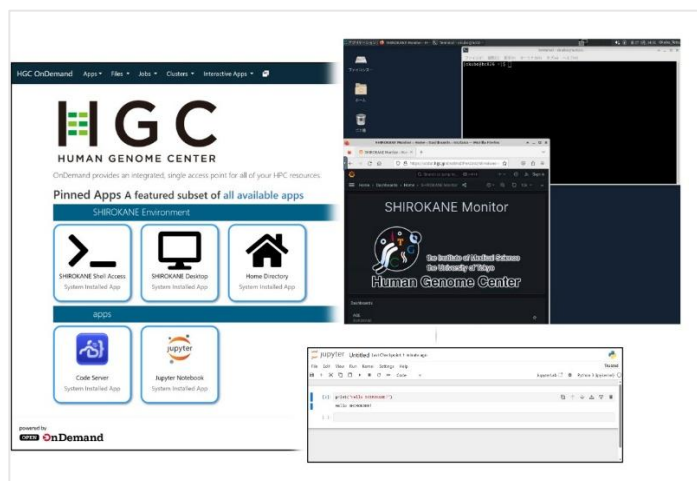
ストレージは、前世代システムで安定稼働した分散ファイルシステムストレージを継続して採用しました。年々増加する解析データに対応するため、初期容量の 41PB を確保しつつ、将来的には 80PB まで増強可能な構成とすることで、大量のデータを蓄積・高速処理し続けることが可能なデータ基盤を提供します。また、Shirokane6 の大容量ストレージからのデータ移行に際しては、SHIROKANE ユーザが実行する計算ジョブやデータ I/O への影響を最低限に抑え、シス

テム停止期間を短縮するべく、移行グループの細分化や計画的なアナウンスを実施し、約 30PB の研究データ移行をおよそ 110 日間かけて完遂しました。

ファシリティ面では、年間を通じてサーバの入気温度を一定に保つことが可能な空調システムと、全てのラックに免震台を設置したストレージ専用室を整備しました。システムの排熱量や耐熱性能に応じた「設置場所の最適化」を徹底し、高発熱な計算・GPU サーバ群は冷却能力に優れた既存のサーバ室へ、安定した温度環境が不可欠なストレージ群はストレージ専用室へとすみわけることで、強靱で安定した稼働環境を実現しています。加えて、Shirokane8 では新たな試みとして、Open OnDemand を使用した Web ポータルを導入しました。これにより Web ブラウザ上でジョブ管理や GUI アプリケーションを利用できるようになり、ユーザの利便性の向上が期待されます。研究者の環境構築の負担を可能な限り軽減し、研究者が即座に解析へ着手できるよう運用改善を推し進めることで、ゲノム研究を高速化し、さらなる発展を支援します。



ヒトゲノム解析センター棟に新設したストレージ専用室



Open OnDemand を使用した Web ポータル

今後の取り組み

HGCは、SHIROKANEを最先端ゲノム研究の基盤とし、全ゲノムシーケンスデータをはじめとするマルチオミクスデータ解析や生命科学分野におけるAIを含むデータ解析が大規模かつ高速に実現可能なデータ解析環境と、利用者目線に立った質の高いサービスを提供します。これにより、公共・民間問わず国内の生命科学研究を大きく加速させ、医学の発展と社会に貢献していきます。

商標に関する表示

記載の会社・組織名、製品名は、それぞれの会社・組織の商標もしくは登録商標です。

用語解説

(注) 生命科学データサイエンス用スーパーコンピュータシステム SHIROKANE
2026年4月からのSHIROKANEは、Shirokane6(2022年4月から運用開始)、Shirokane7(2024年4月から運用開始)そしてShirokane8で構成される。

問合せ先

<本件について>

東京大学医科学研究所 附属ヒトゲノム解析センター 健康医療インテリジェンス分野
センター長・教授 井元 清哉 (いもとせいや)

https://www.ims.u-tokyo.ac.jp/imsut/jp/lab/hgclink/page_00072.html

<機関窓口>

東京大学医科学研究所 プロジェクトコーディネーター室 (広報)

<https://www.ims.u-tokyo.ac.jp/>