



## 国内最大規模 50PB の高速大容量 ストレージシステム DDN EXAScaler が ゲノム医学研究を支える

～スパコンシステムを Phase 3 システムへ更新～

### DDN ソリューション

#### 導入システム

・ DDN ES400NVX 13 台

#### 利用用途

- ・ ヒト全ゲノム解析
- ・ スーパーコンピューティング

東北大学 東北メディカル・メガバンク機構 (ToMMo) は、2012年2月に、未来型医療を築いて東日本大震災の復興に取り組むために設置され、被災地の地域医療再建と健康支援に取り組みながら、医療情報とゲノム情報を複合させたバイオバンクを構築している。その要となるのが、スーパーコンピュータシステムである。東北メディカル・メガバンク機構が運用するスーパーコンピュータシステムは、2014年7月に Phase 1 システムの本格運用が開始され、2018年に第2世代となる Phase 2 システムに更新された。このスーパーコンピュータシステムは、ヒト全ゲノム解析やメタボローム解析、脳画像解析などさまざまな用途に使われている。中でも、リソースを必要とするのがヒト全ゲノム解析である。ゲノム解析には、演算性能だけでなく、多量のストレージが要求される。東北メディカル・メガバンク機構が進めている大規

模ゲノムコホート調査では、15万人という日本最大規模のゲノム解析を行う予定だが、生のゲノムデータだけでも1人あたり1TBにもなるため、単純計算で15PBものストレージが必要になる。Phase 2 システムには、並列ファイルシステムの Lustre ベースで構築された22PBもの DDN EXAScaler 高速ストレージが搭載されていたが、ゲノム解析と研究が進んだことにより22PBでもストレージ容量が足りなくなってきた。

そこで、東北メディカル・メガバンク機構は、さらなる大容量化と高性能化を求めて、再びスーパーコンピュータシステムの大規模更新に取り組み、2022年4月に Phase 3 システムの稼働を開始した。システム更新の責任者であり、東北メディカル・メガバンク機構 副機構長、ゲノムプラットフォーム連携センター長の木下 賢吾 教授にその経緯を伺った。

### チャレンジ。システム更新における4つのポイント

木下教授が今回のスーパーコンピュータシステム更新において、重視したポイントは以下の4つである。

- セキュリティと利便性のバランスを重視する。
- CPU 負荷に対してデータ量が多いため、ストレージの容量とパフォーマンスを重視する。
- セキュリティレベルに応じて3つのユニットを柔軟に運用できるようにする。
- 同じデータのコピーを持たないことで、ストレージの利用効率を高める。

「まず、セキュリティと利便性のバランスを取るようなシステムを目指しました。ゲノム解析結果を含む健康調査情報は参加者のプライバシーに関わるデータですので、セキュリティを最大限担保する必要があります。セキュリティを担保するなら金庫に入れて誰からもアクセスできないようにするのが一番安全ですが、そうすると解析ができないので意味がない。そのバランスの落とし所を探るのが非常に難しいと思って

います」と木下教授。

ストレージの容量については、Phase 2 システムの約2倍の50PBを目標としたが、単に容量を増やすだけでなく、より柔軟な使い方ができるようにしたかったと木下教授は語る。「我々は扱うデータのセキュリティレベルに応じて、ストレージをユニットA、B、Cという3つのユニットに分けて利用していますが、Phase2システムでは、そのユニット間でリソースの配分を変えるのが難しかった。Phase 3 システムではユニット間でのリソースを柔軟に配分できるようなシステムを目指しました。また、Phase 2 システムでは、ユニットBとユニットCに同じゲノムデータをコピーして持つ場合があったため、冗長に容量を占有することがありました。今5万人の全ゲノム解析が終わって、10万人の目処がついてきた状態で、この規模だとコピーを持つのは限界だと判断して、Phase 3 システムではコピーをしないですむ新たな手法を考えました」



東北メディカル・メガバンク機構 副機構長  
ゲノムプラットフォーム連携センター長  
木下 賢吾 教授

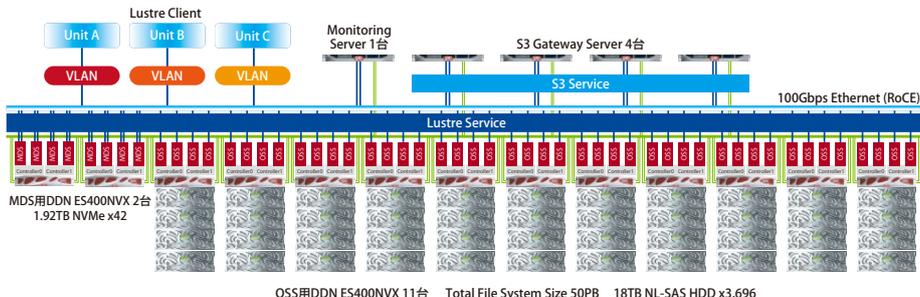


## Phase 3システムの導入効果

Phase 3システムは、AMD EPYC 7713搭載サーバー110台、AMD EPYC 7713およびNVIDIA A100を8基搭載したサーバー3台から構成される計算ノードと、DDN EXAScaler約50PBで構成されるストレージを備えている(Phase3システム以外の連携システムのDDN Lustreを含めると60.8PBとなる)。チャレンジで挙げた重要ポイントについての、Phase

3システムでの達成状況は以下の通りだ。

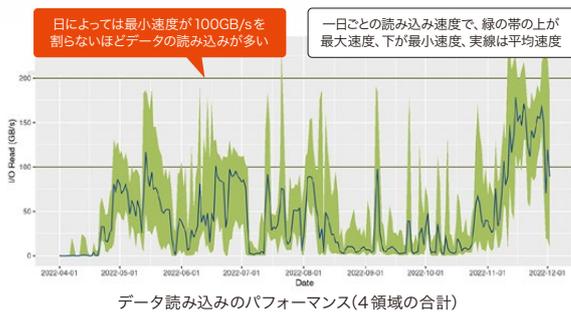
- 限られた予算の中でシステムを更新し、セキュリティと利便性、パフォーマンスのすべてのバランスが取れたシステムを実現
- 大規模ゲノムコホート調査の要となるストレージは、Phase 2システムに引き続きDDNストレージを中心に構成し、合計60PBという大容量を実現



- データ使用量の大きいBとCのユニット間のストレージ容量の配分を柔軟にできるシステムを構築
- ハードリンクとサブディレクトリマウントを活用することで、データをコピーせずにBとCのユニットからデータにアクセスできるストレージシステムを構築

木下教授はPhase 3システムについて、以下のように評価している。「ストレージについては、容量だけでなくスピード面も非常に重要です。解析では同時に同じファイルにアクセスすることが多いので、並列なファイルシステムが大きな役割を担います。そのような中でDDNのLustreベースのストレージは期待通りの性能

を発揮しました。Phase 3システムでは、BとCのユニットへのストレージリソース配分が柔軟にできるようになっただけでなく、ハードリンクとサブディレクトリマウントを活用することで、同じデータをコピーせずに別のユニットからアクセスできるようになり、ストレージをより無駄なく使えるようになりました。こうした使い方はあまり他ではされてないと思いますが、大きな魅力でした。またPhase 2システムまではInfinibandとイーサネットを組み合わせていましたが、今回は柔軟性を重視してすべてイーサネットで構築しました。それでもパフォーマンスは十分満足できるものでした。Phase 1システムからPhase 3システムまで継続してDDNのストレージを使い続けていますが、一度もデータロスがなく、高い信頼性を担保していると認識しています」



## これからのインフラについて: 安定性とチャレンジのバランスを追求

木下教授は、今後の目標やベンダーへの期待について次のように語った。「現在他の研究機関とゲノムデータを共有しています。ジオレプリケーション的なイメージで、磁気テープでデータを持ち合っています。今後はそれをディスクにしたいのですが、理想をいえばやはりネットワーク越しですね。技術的な検討は継続的に実施していますが、さすがにペタバイト級の容量だと難しいです。このスパコンは基盤、インフラです。

インフラというのは動いていて当たり前で、理想をいえば4.5年ごとに切り替えるときも、ユーザーはそれを切り替えたことにさえ気付かないくらいになればいいと思います。ただ、それは非常に難しい。サステナビリティをこの規模のシステムで実現するのはどうすればいいかというのが今後のチャレンジだと思

ます。一般論としてベンダー側に期待したいのは、先のインフラという観点では何より安定性です。その一方で、安定性を誤って解釈し、同じモノを常に維持し続けることだと思っていると、陳腐化する。だから、安定性とは一見背反することにはなりますが、常にチャレンジをし続けたいと聞いています。チャレンジと安定性のバランスを取ることが非常に重要です。簡単なことのように思えるかもしれませんが、とても難しいことですよ。技術力を保ちつつ、開発をしながら、バグが発生したときにはすぐに対応できる技術者も揃えておかなければいけない。それをきちんと理解して成長していくようなベンダーと今後もお付き合いして、最先端のシステムを検討していきたいと考えています」



東北大学 東北メディカル・メガバンク機構 (ToMMo)

※本事例は2023年3月2日、東北大学 東北メディカル・メガバンク機構におけるインタビューに基づいて作成しました。

### 株式会社データダイレクト・ネットワークス・ジャパン

- 本社 〒102-0081 東京都千代田区四番町6-2 東急番町ビル8F
- 大阪オフィス 〒530-0001 大阪府大阪市北区梅田3-3-45 マルイト西梅田ビル5F
- 名古屋オフィス 〒451-6040 愛知県名古屋市中区牛島町6-1 名古屋ルーセントタワー40F

Team-JPSales@ddn.com  
https://ddn.co.jp/contact

© DataDirect Networks. All Rights Reserved. DataDirect Networks, DataDirect Networks ロゴ, DDNは、DataDirect Networks の商標です。その他の名称およびブランド名は、各社の登録商標または商標です。202304