

# 東南アジア地域HPCを牽引するNational Supercomputing Centre Singapore

## National Supercomputing Centre Singapore: The HPC Leader in Southeast Asia Region

● 宇津井 峻

● 山田基弘

---

### あらまし

National Supercomputing Centre Singapore(NSCC)はシンガポールの国家スーパーコンピュータセンターであり、「国家研究開発構想支援」「産業分野での研究協業招聘」「国家研究開発能力の向上」を目標として掲げている。このシンガポール国家の研究を支えるHPC(High Performance Computing)システムの構築において、富士通はシステム基盤提供および運用業務を担った。NSCCに最適なシステムとして、PRIMERGYシリーズで構成された計算ノードをはじめ、ストレージ・高速ネットワークを提供した。システムの運用については、これまで培ってきたHPCのノウハウと富士通シンガポールの既存のシステム運用系サービス(マネージドサービス)を組み合わせることで最適なサポート体制を整えた。

本稿では、システム導入から運用にかけて直面した課題と、それらに対する対応策を紹介する。

### Abstract

The National Supercomputing Centre Singapore (NSCC) is Singapore's national supercomputer center, and it has the objectives of supporting national R&D initiatives, attracting industrial research collaboration, and enhancing Singapore's research capabilities. In assisting this national-level research initiative for high performance computing (HPC) in Singapore, Fujitsu provided the system platform and undertook its system operation. As an optimal system platform for NSCC, we selected the PRIMERGY Series for the computing nodes, and integrated storage and a high-speed network provided by third-party vendors. As for the system operation, Fujitsu developed an optimized support system through a combination of its proprietary know-how in HPC and the existing Managed Services of Fujitsu Singapore. This paper describes the challenges and solutions we found during system introduction and operation.

---

## まえがき

2016年3月28日、シンガポールの国家スーパーコンピュータセンター「National Supercomputing Centre Singapore (NSCC)」のシステムが稼働した。富士通はシステム基盤およびシステム運用業務の提供を担うことになった。<sup>(1)</sup>

NSCCシステムは、シンガポール科学技術研究庁 (A\*STAR) を筆頭に、シンガポール国立大学 (NUS)、南洋工科大学 (NTU)、シンガポール工科大デザイン大学 (SUTD) の協力により立ち上がった。本システムは1 P (ペタ) FLOPSの計算能力を誇る。言い換えると、毎秒1,000兆回の計算能力であり、これは日本最高で世界5位 (2016年6月「TOP500」による) の「京」<sup>(注1)</sup> に比べ約10分の1の性能である。NSCCでは、約1,300台のサーバを高速ネットワークでつなぎ合わせることで、この性能を実現している。2016年6月に発表された世界のHPC (High Performance Computing) システムの性能ランキング「TOP500」では93位となった。これは東南アジアでは最高順位である。日米欧が先行するHPC市場においてシンガポールを含む東南アジア諸国はまだ後発ではあるが、NSCCの発足をきっかけに、HPCに対する機運は高まりと急成長の兆しを見せている。付け加えると、NSCCのミッションの一つはHPCの活用をシンガポールで促進することである。世界中で多数の大規模HPCシステムを運用する富士通のノウハウをNSCCにも共有することが、本プロジェクトにおける富士通の重要な役目でもある。

本稿では、NSCCプロジェクトにおける課題およびそれらに対する富士通の果たす役割について論ずる。

## 課 題

冒頭で述べたように、本プロジェクトにおける富士通の役割は「システム基盤提供」と「システム運用業務提供」である。これらの実施に当たり、直面した課題は下記のとおりである。

まず、システム基盤提供に関しては三つの技術的課題があった。一つ目は、データセンターの設

計である。NSCCのデータセンターはConnexis North TowerというA\*STARの複合オフィスビルの17階に位置しており、それに付随する様々な課題が存在した。また、これに加え、気候が一年中暖かいシンガポールで高い電力効率の実現が求められた。

二つ目は、各ステークホルダーそれぞれのロケーションとNSCCをつなげる高速ネットワークの整備である。NSCCの母体は前述の複合オフィスビル内に位置するが、NUS、NTUにも一部、機器が配置された。これらを高速ネットワークでつなぎ、各大学からユーザーがNSCCシステムにアクセスする際にデータ転送が簡単に行えることや、スムーズなシステム利用が目標とされた。

三つ目は、階層型ストレージシステムおよび高速なストレージ性能要件であった。HPCの実用化においては計算処理能力だけではなく、ストレージ性能も同等に重要である。ファイルへの書き込みや読み取りが多いアプリケーションは、ストレージ性能がボトルネックとなる可能性がある。NSCCシステムではその点が考慮され、500 Gバイト/秒という高い性能要件が求められた。また、ユーザー視点で利用しやすいシステムとするべく、階層型のシステムが必要とされた。

次にシステム運用業務提供については、シンガポールでは日米欧のHPC先進国に比べ、この規模のシステム運用を経験している人材が容易に見つかる環境ではないという背景がある。そこで、富士通がこの人材の不足を補うために完全にHPCの運用業務をこなせるチームを顧客先に派遣する必要があった。

次章以降で上記の課題に対して富士通がどのようなソリューションを提供しているかを述べる。

## システム基盤提供

本章ではまず、NSCCシステム概要から説明する。システムは大きく分けて、計算ノード、ストレージ、高速ネットワークの三つの要素から成り立っている (図-1)。

計算ノードは富士通のPCサーバPRIMERGYシリーズによって構成される。計1,298ノード中、1,160台がFUJITSU Server PRIMERGY CX2550 M1 (スタンダードノード)、128台が各GPUカード

(注1) 理化学研究所と富士通が共同開発したスーパーコンピュータ。「京」は理化学研究所の登録商標。

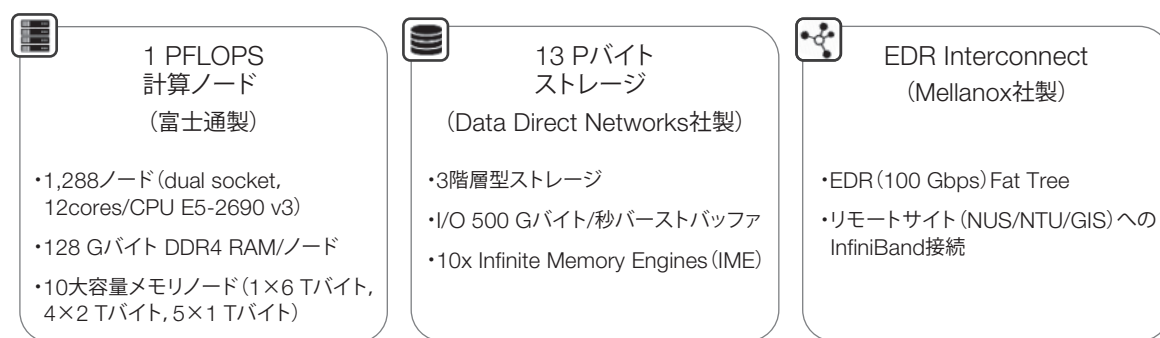


図-1 NSCCシステムを形成する3要素

1枚を搭載しているFUJITSU Server PRIMERGY CX2570 M1 (GPUノード)、10台がメモリ容量が他ノードより大きいFUJITSU Server RX4770 M2 (大容量メモリノード) としている。CPUはIntel社のXeon Processor E5-2690 v3およびXeon Processor E7-4850 v3を搭載している。本構成で、TOP500の性能指標となるHPC ChallengeのLINPACKベンチマークテストでは、1.01 PFLOPSの性能を実現した。<sup>(注2)</sup>

以下では、各課題に対して富士通が提供するソリューションを紹介する。

### ● 課題1：データセンターの設計

今回のプロジェクトにおいて、データセンターのロケーションがいくつかの課題を発生させた。前述のとおり、NSCCのデータセンターは複合オフィスビルの17階に設置された。ここで、17階まで機器を運び込むためのエレベーター寸法によるラックの高さ制限や、荷重の物理的な制限がいくつか生じたため、提供するハードウェアもそれらの制限に抵触しないように設計をする必要があった。更に、シンガポールのデータセンター平均値はPUE (Power Usage Effectiveness, データセンターの効率を表す指標値で、低ければ低いほど効率は良い) 2.5以上と言われる中 (NSCC調べ)、NSCCではPUE 1.4という目標値が設定された。これに対し、富士通は水冷、空冷、リアドア水冷式ヒートエクスチェンジャー (RDHx) の三つの冷却方式の組み合わせで対応した。

CX2550 M1/2570 M1については富士通の

PRIMERGYシリーズ初の水冷機を納品した。この水冷機の提案は消費電力値を抑えるためには非常に重要であった。約40℃の温水を利用して、CPU、メモリ、およびGPUを直接水冷することにより、計算ノードシャーシ1台 (4台のCX2550 M1、もしくは2台のGPU付CX2570 M1が収納される) あたり、約140 ~ 160 Wの消費電力が削減できる。システム全体に換算すると、これだけで約50 kW以上の削減になる。また、CX2550 M1/2570 M1搭載ラック以外は、全てRDHxが実装されている。更に、そのほかの部分 (計算ノードラック搭載のネットワーク機器など) は、空冷方式で冷却している。

### ● 課題2：リモートサイト接続

各ロケーションとの高速ネットワーク接続を理解するためには、全体のネットワーク設計の概要を説明する必要がある。NSCCには、ユーザーネットワーク、管理ネットワーク、高速ネットワークの三つのネットワークが存在する。ユーザー・管理ネットワークはTCP/IPを用いて通信を行うイーサネットである。高速ネットワークは計算ノード間の通信やストレージシステムへのデータアクセス時に利用される。これには、Mellanox社のInfiniBandが採用された。NSCCシステムには一つのリンクあたり、100 Gbps InfiniBand EDRという現時点での最新テクノロジーが搭載されている。トポロジはQuasi Fat-treeでノンブロッキング、すなわち、どのスイッチを経由しても、そのスイッチで本来得られるリンクあたりの性能が失われないように設計されている。

ここで課題となるのは、NTU/NUS、更にはA\*STARの一部であるGIS (Genome Institute of Singapore)のリモートサイト接続である。ユーザー

(注2) 本試験はCX2550 M1およびCX2570 M1に搭載されたCPUのみを利用している。CX2570 M1搭載のGPU、およびRX4770 M2のCPUは利用していない。

がシステムにアクセスするためのゲートウェイであるログインノードは物理的に複数拠点に設置されることとなった。現時点ではNSCCに配置されている4台のほかに、NUSのキャンパス内に4台およびNTUキャンパス内にも4台が設置されている。また、GISにも1台の大容量メモリノードが設置された（図-2）。

解決策としては、これらのノードをそれぞれ、Mellanox社のMetroXという装置を経由し、光ケーブルで直接NSCCのInfiniBandネットワークに接続することとした。NSCCからGIS/NTU/NUSへの物理配線の距離はそれぞれ2 km、22.4 km、42 kmである。この接続があることにより、各リモートサイトからNSCCまでのシステムアクセスには最低でも40 Gbpsのリンクが確保されている。また、このネットワーク上ではTCP/IPではなく、RDMAプロトコルが利用される。これによって、それぞれの大学のキャンパス内からでもインターネットを経由してシステムを利用するよりもデータ転送などが快適に行えるような設計となっている。

### ● 課題3：ストレージシステム

NSCCシステムにおけるもう一つの特徴となったのが、三つ目の課題の階層型ストレージシステムおよび高速なストレージ性能要件に対するソリューションである。NSCCプロジェクトにて、富士通が提案したストレージはData Direct Networks社のシステムであった。このストレージシステムは四つの階層に分けている。最上位階層

のバーストバッファテクノロジーに加え、Lustre、GPFS、アーカイブがある。中でも、最上位の階層は500 GB/sという高い性能が求められ、これに対してはData Direct Networks社のInfinite Memory Engine（IME）という製品を導入した。IMEはこれまでのHDDディスクとは違い、キャッシュのような働きをするNVMe SSDを活用した共有プールであり、<sup>(2)</sup> I/Oを頻繁に繰り返すアプリケーションをHPCシステム上で動作させるユーザーが特に恩恵を受けることになる。第2階層は4 PバイトのLustreファイルシステム（一時領域）、第3階層は4 PバイトのGPFSファイルシステム（恒久領域）、第4階層はアーカイブとある。GPFSはユーザーが主にファイルやバイナリをアップロードする領域である。Lustreファイルシステムはユーザーが計算資源を利用する際に、アップロードしたファイルやバイナリを移動し、ジョブを実行する領域である。LustreとGPFSの平行ファイルシステムに関しては、それぞれの利用用途に応じてチューニングが行われている。

### システム運用業務提供

NSCCシステム運用に向けて、前述のとおり、最適な運用チームのチームビルディングが課題であった。NSCCでは、富士通シンガポールに既に存在するシステム運用系サービス（マネージドサービス）に、日本および外部からHPCの運用経験を持ち込み、これらを融合することによってユ



図-2 NSCCのリモート接続

ニークかつ最適な運用体制を整備することとした(図-3)。

まず、顧客先には富士通のHPCシステム運用チームが常駐することとした。現状、このチームはパートナー会社からのメンバーも含めて、6名で構成されている。この中に、HPCシステム運用者、ストレージシステム運用者、HPCアプリケーション最適化支援要員、データセンター運用者の役割を担うメンバーがそろっている。いわば、各専門分野のプロフェッショナルを結集したチームである。また、チームの国際色は非常に豊かで、筆者(日本人)をはじめ、オーストラリア人、インド人、インドネシア人、英国人が在籍している。シンガポール人がいないが、そもそもシンガポールには人口約554万人のうち、シンガポール人・永住者は約390万人しかいない<sup>(3)</sup>。この中からHPCのノウハウを持ち合わせた人材を探すのは困難である。そこで、富士通がHPCシステム運用のプロフェッショナルを派遣し、その常駐チームがシンガポールにおける人材不足を補っている。

このほかに、ユーザーからの一時受け窓口機能

を果たすサービスデスク(SD)チームと、システムアラートをモニタリングするリモートモニタリングセンター(RMC)チームが存在する。これらのチームはNSCCプロジェクトが始動する以前より、シンガポールで他顧客に対しても同様のサービスを提供している。また、SDとRMCはそれぞれ顧客サイトではなく、富士通のオペレーションセンターで活動している。

具体的なユーザーに対するサービス提供の流れを説明すると、以下のとおりとなる。まず、ユーザーがNSCCシステムにおいてジョブ実行中に何かしらのシステム起因の問題に気づいたとする。これをレポートする先は、一時受け窓口であるSDになる。これを受けたSDは、問題やその影響範囲を確認後、必要があればNSCC常駐のHPCシステム運用チームにエスカレーションする。ここから、オンサイトのHPCエンジニアはSDおよびユーザーと頻繁に情報共有しながら問題解決を目指す。問題が解決したら、SDよりユーザーに最終確認を行い、問題案件をクローズする。

更に、複雑な問題に備えて、サードパーティベ

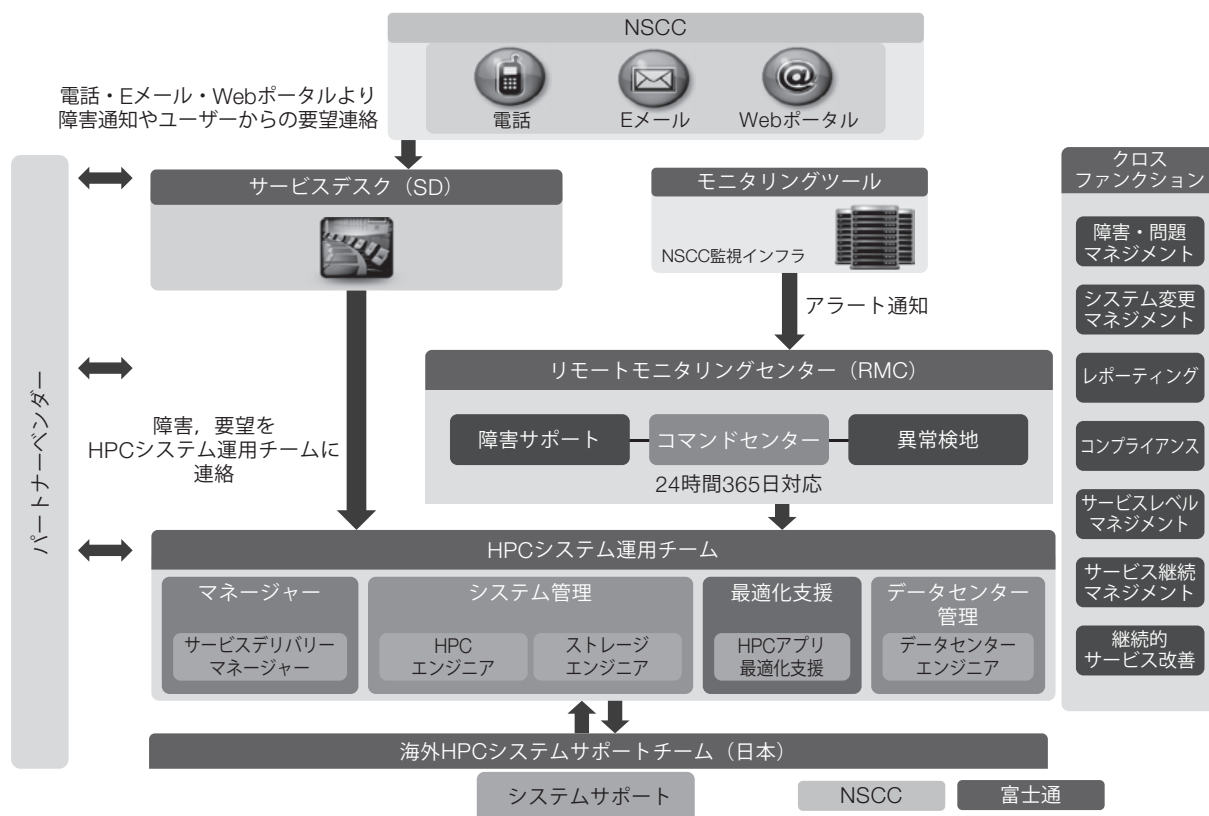


図-3 NSCC運用体制

ンダのサポートチームや富士通本社のHPCチームがバックエンドで富士通シンガポールのマネージドサービスチームをサポートする体制としている。

NSCC以外のHPCサイトの多くでは、顧客側にITチームがいて一次対応をするケースが多く、富士通で一次対応からクロージングまで担当するケースは稀である。これをNSCCでは富士通が一手に引き受けている。これが実現できたのは、一次受けからフォローアップまで、全てを体系化して担う、強いマネージドサービスチームが富士通シンガポールに存在したからである。NSCCプロジェクトの発足により、このマネージドサービスのノウハウに、HPCのノウハウが加わった。これまであまり交わることのなかった二つのノウハウではあるが、これらが手を組むことによって最適な運用体制ができあがった。

#### NSCCに見る今後のHPC動向

NSCCでシステム運用を継続していると、あることに気づく。HPCにおけるシミュレーションを代表するような流体解析や構造解析はシステム上で数多くのユーザーに利用されているが、CaffeやTheanoなどのDeep Learning系のアプリケーション利用が増えている。これは数年前までは多くのHPCセンターでは見られなかった現象である。この傾向には、NSCCがHPC初心者にもオープンな環境を構築していることが大きく寄与していると考えられる。

Deep Learningの世界ではデータが蓄積されればされるほど、コンピュータが正確な情報を予測できるとされている。そこに絡んでくるのが、InfiniCortexというシンガポールと世界各国を高速ネットワークでつなげるプロジェクトである。NUSやNTUとNSCC間が光ファイバーでつながれていると同様に、A\*STARが主導してシンガポールは各国に対して海底の光ファイバーを設置している。例えば、シンガポールから日本の東京工業大学に直接光ファイバー回線がつながっている。そのほか、アメリカやヨーロッパにも複数の接続拠点を持つ。InfiniCortexプロジェクトのリーダーであるA\*STARの発表によると、この高速リンクは「各国のHPC拠点をつなぎ、各システムを一つの大きな分散型のHPCシステムのように利用し、

重要課題を解決していくための基盤づくり」であるという。これらの高速リンクは各国システムをつなぐだけでなく、データの転送がより容易にできるという観点では、非常に重要な意味を持つ。国土が狭く、資源が少ないシンガポールは外国の資源に頼らざるを得ない<sup>(4)</sup>。InfiniCortexプロジェクトにより、今では「データ」という資源をシンガポール国内に取り込みやすい環境までできあがりつつある。

#### む す び

NSCCが運用開始して半年以上が経過した。これから続くシステム運用についてもこれまでと同様に社内外のノウハウを最大限に活かし、その場に応じて臨機応変なソリューションを見つけ、NSCCに貢献するのが富士通に与えられた役割である。それだけでなく、前述のとおり、今後のHPC市場は従来のものとは異なり、その利用目的は多岐にわたり始めている。各ユーザーの隠れたニーズを見つけ出し、NSCCとともにHPCの利用を促進させていくことも重要である。

また、シンガポールを筆頭に、東南アジア地域全体にHPCを徐々に浸透させていくことに貢献することも、これからの重要課題である。技術によって課題を解決するのはもちろんのこと、既存のマネージドサービスチームをHPCの運用に活用できたように、技術以外の部分でも場面に応じての柔軟な発想を常に持ち合わせていたい。

#### 参考文献

- (1) 日本経済新聞「富士通、スパコンでアジア開拓 シンガポールと契約」2015年11月19日。
- (2) DDNホームページ。  
<http://www.ddn.com/>
- (3) 外務省ホームページ「シンガポール共和国」。  
<http://www.mofa.go.jp/mofaj/area/singapore/data.html>
- (4) 岩崎育夫：物語 シンガポールの歴史（中公新書）。中央公論新社，2013。

## 著者紹介

---



**宇津井 峻** (うつい しゅん)

Fujitsu Asia Pte. Ltd.

NSCCプロジェクトをはじめ、シンガポール・東南アジアにおけるHPCビジネスに従事。



**山田基弘** (やまだ もとひろ)

テクニカルコンピューティング・ソリューション事業本部

計算科学ソリューション統括部

グローバルHPCビジネスの企画・推進に従事。