

北海道大学情報基盤センター次期スーパーコンピュータシステムの紹介

深谷 猛¹⁾, 梅田 隆行¹⁾, 佐竹 祐樹¹⁾, 岩下 武史¹⁾, 更科 高広²⁾, 吉川 浩²⁾, 吉川 潤²⁾, 角鹿 千枝²⁾, 高口 智美²⁾, 折谷 智咲²⁾, 齋藤 珠紀²⁾, 村田 欽正¹⁾

1) 北海道大学 情報基盤センター

2) 北海道大学 総務企画部情報企画課

fukaya@iic.hokudai.ac.jp

An overview of the next supercomputer system at Information Initiative Center, Hokkaido University

Takeshi Fukaya¹⁾, Takayuki Umeda¹⁾, Yuki Satake¹⁾, Takeshi Iwashita¹⁾, Takahiro Sarashina²⁾, Hiroshi Yoshikawa²⁾, Jun Yoshikawa²⁾, Chie Tsunoka²⁾, Tomomi Takaguchi²⁾, Chisaki Oritani²⁾, Tamaki Saito²⁾, Yoshimasa Murata¹⁾

1) Information Initiative Center, Hokkaido University

2) ICT Planning Division, General Affairs and Planning Department, Hokkaido University

概要

北海道大学情報基盤センターでは、2018年12月から運用を行っている現在のスーパーコンピュータシステムを2025年4月に次期システムに更新する予定である。そこで、本稿では、現スーパーコンピュータシステムのこれまでの運用状況を簡単に振り返り、それを踏まえた上で設計を行った次期システムの概要について紹介する。

1 はじめに

北海道大学情報基盤センター（以下、本センター）では、スーパーコンピュータシステム（スパコン）とクラウドシステムを中心に構成される学際大規模計算機システム（学際システム）を運用している [1]。現在の学際システムは、2018年12月に運用を開始し、2025年4月に次期学際システムに更新予定である。本稿では、次期学際システムの一部として導入予定の次期スパコンの概要について紹介する。次期スパコンの仕様やサービス体系は、現スパコンの運用状況を踏まえた上で検討を行った。そこで、以下では、最初に現スパコンの運用状況を簡単に振り返り、その後、次期スパコンの概要を述べる。

2 現スパコンの振り返り

2018年12月から運用が開始された本センターの現学際システムの全体像を図1に示す。図1から分かるように、現スパコンは Grand Chariot（サブシステム A）と Polaire（サブシステム B）の二つの演算サブシステムとストレージシステムから構成されている。また、利用者に提供しているサービス形態は、占有ノー

ド利用（年間定額制）と共用ノード利用（従量課金制）の二種類である。

本センターで運用していた前スパコンは Power アーキテクチャ（HITACHI SR16000 M1, IBM Power プロセッサと AIX を搭載）であり、現スパコンを導入するに当たり、x86 アーキテクチャの採用を含む大きな変更を行った [3]。その背景の一つとして、一般的な研究室等における計算機環境との親和性の向上、幅広い科学技術アプリケーション（特にオープンソースソフトウェア）への対応、という狙いがあり、現スパコンの Grand Chariot にその方針が強く出ている。一方で、Polaire は、現学際システムの仕様作成当時の HPC システムのトレンド（メニーコア、高バンド幅、国内の他のスパコンの状況など）を踏まえて、構成を決定した。また、前スパコンでは従量課金制のみのサービス体系であったが、利用者のニーズを考慮した結果、占有ノード利用のサービス形態も導入した。更に、利用者番号の管理方法の変更（プロジェクト単位から利用者単位）やそれに伴うグループ利用方法の変更なども併せて実施した。

現スパコンの運用は2025年2月末で終了予定であり、2018年12月の稼働開始から約6年間（当初の

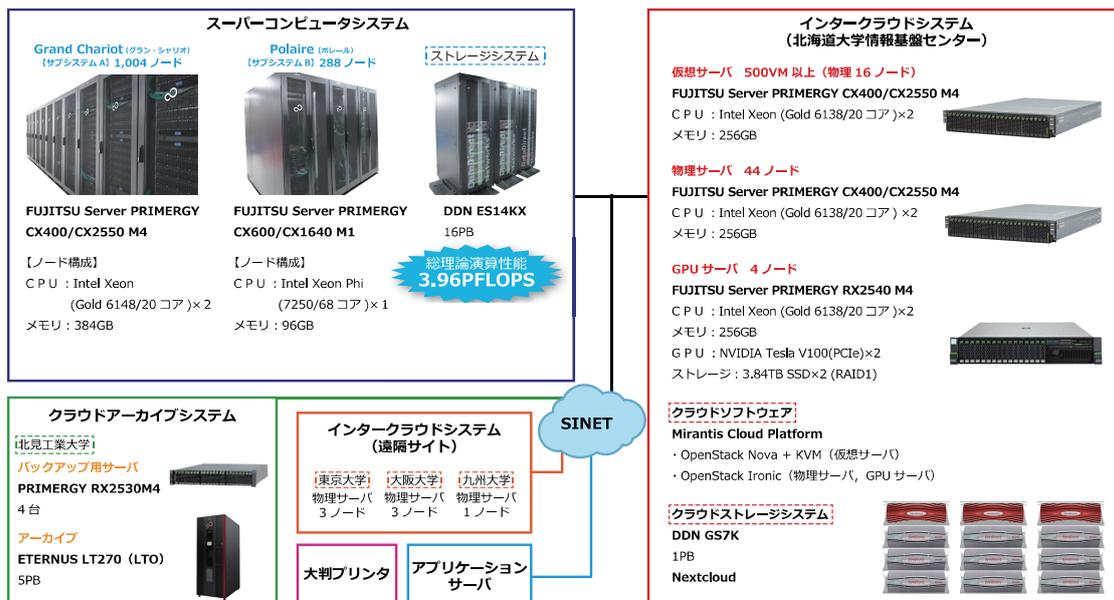


図 1: 現学際システムの全体像。

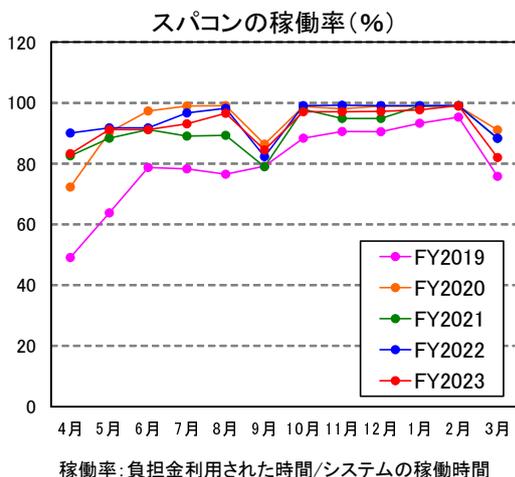


図 2: 現スパコンシステムの稼働状況。

5年から約1年の延長)の運用となる[2]。様々な理由により、システムの規模や演算性能が制限されるなか、上述のように、利用者のニーズを踏まえた「使いやすいスパコン」の実現が目標であった。これに対して、現スパコンの稼働率は図2に示した結果となっている。図2が示す通り、運用期間全体を通して、非常に高い稼働率となっており、利用者に十分に活用されたと言える状況である。また、HPCI課題における利用に関しても、毎年、提供している資源量に対して多くの希望があり[4]、多くの研究者に受け入れられたシステムであったと言える。北海道大学においても、化学反応創成研究拠点(WPI-ICReDD)や前田化学反応創成知能プロジェクトのような学内有数の大型プロジェクトでの利用や人獣共通感染症国際共同研究所と

いった部局単位での利用実績がある。さらに、民間企業における利用も活発で、本センターで想定している提供可能資源量(全体の10%程度)に到達する年度が大半であった。このように、現スパコンの利用状況は良好であり、その結果として、利用負担金収入の面でも余裕が生じたため、運用期間中の電気料金の高騰に対して利用負担金の値上げを回避することができ、次期スパコンのための設備等の整備もスムーズに行うことができた。

現スパコンは、運用期間中に大きなシステムトラブルが生じることもなく、上述のように順調な運用状況のまま、2025年2月末の運用終了に近づいている。利用者にとって使いやすいスパコンという目標の下で、前スパコンからの大きな方針転換は、これまでの結果を見る限り、十分な成功であったと判断できる。次期スパコンの導入に際しては、現スパコンの長所を維持しつつ、新しいニーズや周囲の状況に対応することが求められる。

3 次期スパコンの概要

まず、次期スパコンの設計方針を簡単に紹介する。現スパコンの設計思想とこれまでの利用状況を踏まえると、現スパコンの設計思想の踏襲は妥当な方針の一つである。一方で、HPCシステムのトレンドは現スパコンの設計時とは少なからず異なっており、また、AI・機械学習分野のニーズへの対応が新たな課題として生じている。以上の点を踏まえて検討を行った結果、次期スパコンは「標準的なCPU+一部にGPU搭載」と

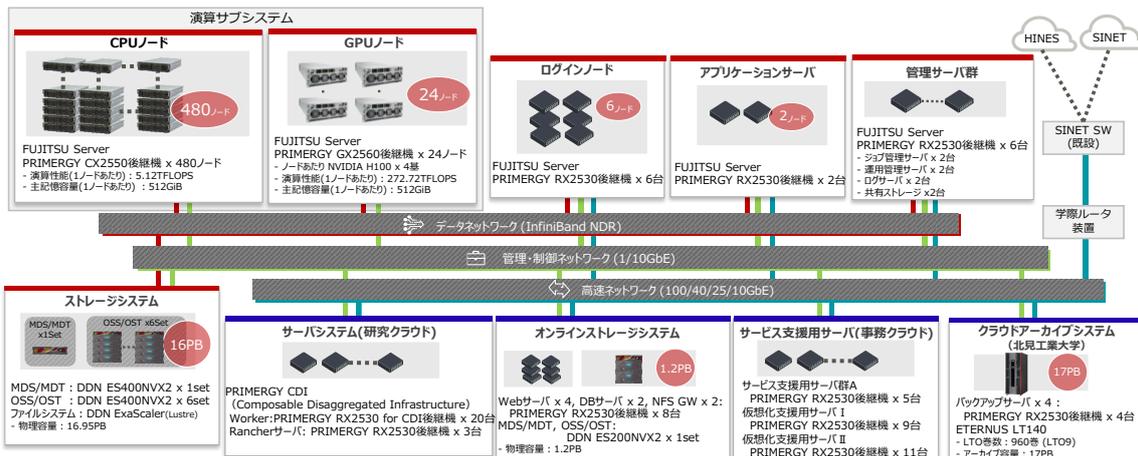


図 3: 次期学際システムの全体像。

演算システム		ストレージシステム
CPUノード群	GPUノード群	
FUJITSU PRIMERGY CX2550後継機 PRIMERGY CX400シャノン後継機	FUJITSU PRIMERGY GX2560後継機	DDN ES400NVX2
480ノード (2.457PFLOPS)	24ノード (6.55PFLOPS)	16.95PB
ノード構成: ・ CPU: Intel Xeon Gold 6548Y+ x 2 (Emerald Rapids, 32コア, 2.5GHz) ・ メモリ: 512GB	ノード構成: ・ CPU: Intel Xeon Gold 6548Y+ x 2 (Emerald Rapids, 32コア, 2.5GHz) ・ NVIDIA H100 x 4 ・ メモリ: 512GB	・ DDN ExaScaler (Lustre) ・ オールフラッシュ(フルSSD)
ネットワーク: InfiniBand NDR (400Gbps, Fat Tree)		

図 4: 次期スパコンの主な構成。

いう方針で設計することとなった。また、サービス形態に関しては、現スパコンの占有ノード利用が人気であることを踏まえて、現スパコンのサービス形態を原則維持し、細かい部分を調整する方針とした。

スパコンに関しては上述の方針で仕様を策定して、調達を実施した結果、導入予定となった次期学際システム全体の概要は図3の通りである。なお、現学際システムに続いて、富士通株式会社が落札した。次期学際システムは、現システムと同様に、スパコンとクラウドを中心に構成される。次期スパコンは、一部ノードにGPUが搭載され、CPUノード群とGPUノード群から成る一つの演算サブシステムとストレージシステムで構成される。

次期スパコンの主たる構成は図4に示した通りである。CPUは二つのノード群で共通で、Emerald Rapids世代のIntel Xeonプロセッサを搭載する。搭載するCPUの仕様は、演算性能だけでなく、消費電力も考慮した上で選定した。前述の通り、現スパコンでは占有ノード利用への人気が高く、次期スパコンでも同様のサービス形態を継続することを考えている。その場合、ノード当たりの消費電力が利用負担金のベースとなり、また、ノード当たりの演算性能(価格)とノード数がトレードオフの関係となる。この点を考慮した

上で、次期スパコンのCPUの仕様の選定を行った。GPUノード群は、NVIDIAのGPU(H100)をノード当たり4枚搭載する構成となった。また、システム全体で、ノード当たり512GBのメモリを搭載する構成である。現スパコン(Grand Chariot)では、ノード当たり384GBのメモリを搭載しており、多くの利用者に好評であった。次期スパコンでは、より大きい512GBのメモリを搭載する構成として、この点を強化した。

次期スパコンのインターコネクトはInfini Band NDRを採用した。また、ストレージシステムは、オールフラッシュ(フルSSD)で、容量は現スパコンと同程度の16.95PBとなる。現スパコンのストレージシステムの利用状況に関しては、容量的には余裕があり、その点を踏まえて、次期スパコンのストレージシステムでは、容量の増強よりも、転送速度や省電力性能の点で有利なオールフラッシュのシステムの採用を優先した。オールフラッシュのストレージシステムの採用により、ストレージシステムに関する省電力化とIOが支配的なアプリケーションの高速化が期待される。また、現スパコンは、当初の運用期間の5年を超えて運用を行うことになり、延長後の運用において、ストレージシステムのハードウェア障害が増加する傾向となっている。次期スパコンに関しても、運用期間を延長する可能性は十分に考えられるため、その場合のストレージシステムの安定性という観点も、今回、オールフラッシュのシステムを採用するに至った理由の一つである。

次期スパコンのCPUノード単体の理論演算性能(倍精度)は5.12TFLOPSで、現スパコンのGrand Chariotのノード単体の演算性能の約1.66倍である。

また、GPU ノード単体の理論演算性能（倍精度）は 272.72TFLOPS である。次期スパコンシステム全体の理論演算性能（倍精度）は 9.0PFLOPS で、現スパコンに対して約 2.27 倍となる。次期スパコンの演算性能は当初の想定を下回る結果となったが、その理由として、為替相場の影響、価格の高騰（特に GPU）、設置場所に関する制約などが挙げられる。次期スパコンは、現スパコンが設置されている情報基盤センター北館の将来的な改修を考慮し、改修が済んだ南館に設置する。これに伴い、設置スペースや冷却設備等の整備のための予算の確保などが制約となった。

次期スパコンに関するサービス形態の詳細は、本稿執筆時は検討段階である。既に述べた通り、現スパコンのサービス形態を原則維持する方針である。一方で、占有ノード利用に関して、ノード当たりの消費電力が増加（特に GPU ノード）したため、CPU ソケット単位、GPU ボード単位での占有利用の提供を検討している。これらの点に関する進展に関しては、適宜、発表当日に紹介する予定である。

現スパコンと同様、一般利用（利用負担金利用）に加えて、HPCI、JHPCN、北海道大学情報基盤センター萌芽型共同研究といった公募利用も行う予定である。また、民間企業向けの利用（有償、審査・制限あり）も継続する予定である。次期スパコンでは、北海道大学情報基盤センターのスパコンとしては初めて GPU を搭載する。そこで、GPU 向けの利用者支援にも取り組む予定であり、他大学（例：東大、筑波大、東京科学大、名大、九大）とも積極的に連携 [5] を図ることを考えている。また、現スパコンにおいて、試験的に講義等における教育目的の利用を実施しており、十分な効果が確認できたため、他大学での事例 [6] を参考にしながら、次期スパコンでは教育目的の利用制度を整備する予定である。

4 おわりに

本センターでは、学際システムの一部として、2018 年 12 月から、現スパコンシステムの運用し、2025 年 4 月に次期スパコンに更新予定である。現スパコンは、高いシステム稼働率などが示すように、多くの利用者に活用され、使いやすいスパコンとして、その役割を終えようとしている。次期スパコンは、現スパコンの長所を引き継ぎつつ、新しいニーズや状況に対応することが求められ、その点を踏まえて、仕様を策定し、導入システムが決定した。次期スパコンは CPU ノード群と GPU ノード群で演算システムを構成し、オール

フラッシュのストレージシステムを備える。現スパコンの利用状況や利用者の声を踏まえた上で、現在、次期スパコンの具体的なサービス形態を検討中である。発表当日は、本稿の執筆以降に決定したサービス形態の詳細等についても、適宜、報告したい。次期スパコンも、引き続き、利用者にとって使いやすいシステムとなることを目指し、正式運用開始（現時点では 2025 年 6 月もしくは 7 月が濃厚）に向けた準備を進めたい。

謝辞

日頃からスパコンの運用に携わっている北海道大学情報基盤センターおよび情報企画課の関係者の皆様に感謝いたします。また、現スパコンの運用にご尽力いただくとともに、次期スパコンに関する図等をご提供いただいた富士通株式会社の関係者の皆様に感謝いたします。

参考文献

- [1] 北海道大学情報基盤センター，学際大規模計算機システム，<https://www.hucc.hokudai.ac.jp/>，（参照日：2024 年 10 月 21 日）。
- [2] 更科 高広，吉川 浩，角鹿 千枝，吉川 潤，高口 智美，折谷 智咲，齋藤 珠紀，村田 欽正，深谷 猛，岩下 武史，スーパーコンピュータシステムの運用状況について，大学 ICT 推進協議会 2023 年度年次大会論文集，pp. 126–132，2023。
- [3] 深谷 猛，岩下 武史，金子 修己，折野 神恵，更科 高広，北海道大学情報基盤センター新スーパーコンピュータシステムの概要，大学 ICT 推進協議会 2018 年度年次大会論文集，WE2-1，2018。
- [4] HPCI，これまでの応募状況分析（定期募集），https://www.hpci-office.jp/using_hpci/project_application_stats，（参照日：2024 年 10 月 21 日）。
- [5] ，東京大学情報基盤センター，第 12 回 GPU ミニキャンプ，<https://www.cc.u-tokyo.ac.jp/events/lectures/233/>，（参照日：2024 年 10 月 21 日）。
- [6] 東京大学情報基盤センター，教育利用，<https://www.cc.u-tokyo.ac.jp/guide/education/>，（参照日：2024 年 10 月 21 日）。