

スーパーコンピュータシステムの運用状況について

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

1) 北海道大学 総務企画部情報企画課

2) 北海道大学 情報基盤センター

sarasina@jimuhokudai.ac.jp

Operational Status of the Supercomputer System

Takahiro Sarashina¹⁾, Hiroshi Yoshikawa¹⁾, Chie Tsunoka¹⁾, Jun Yoshikawa¹⁾,
Tomomi Takaguchi¹⁾, Chisaki Oritani¹⁾, Tamaki Saito¹⁾, Yoshimasa Murata²⁾,
Takeshi Fukaya²⁾, Takeshi Iwashita²⁾

1) ICT Planning Division, General Affairs Planning Department, Hokkaido University

2) Information Initiative Center, Hokkaido University

概要

北海道大学情報基盤センターではスーパーコンピュータシステム及びクラウドシステムを一体とした学際大規模計算機システムを運用している。現在の学際大規模計算機システムは2018年12月に運用が開始され、約5年が経過した。本稿では、スーパーコンピュータシステムに焦点を当て、その運用状況について報告する。

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（以下、スパコン）とクラウドシステムを中核とする学際大規模計算機システム（以下、本システム）を運用している[1]。2018年9月に発生した北海道胆振東部地震の影響により、システムの導入スケジュールの遅延が懸念されたが、導入ベンダおよび北海道大学の関係部署の努力により、予定通り2018年12月に本システムの運用を開始することができた。本システムの運用開始にあたって、想定される電気料金に相当する形で利用負担金を設定した。しかし、運用開始時の想定を超えて、徐々に電気料金が値上がり、本システムの運用（および本センターの運営）に大きな影響を及ぼすことになった。それでも本システムのサービス停止などの大きなトラブルが生じることはなく、運用開始から約5年が経過するに至った。

本稿では、本システム運用開始後の約5年間におけるスパコン[2]の運用状況について報告する。スパコンの消費電力量は、本システム全体の消費電力量の約9割を占めており、スパコンの運用状況は本システム全体の運用状況に大きく影響する。そこで、運用開始から約5年が経過するこのタイミングで、これまでのスパコンの運用状況を検証し、成果と課題を明らかにすることで、次期システムの設計・運用に役立てる。

以下、2節で本システムおよびスパコンの構成を簡単に紹介する。次に、3節でスパコンのサービスの概要を説明するとともに、これまでの利用状況を報告する。その後、4節で現状の課題を述べて、最後に5節でまとめを述べる。

2 学際大規模計算機システムとスーパーコンピュータ

2.1 学際大規模計算機システムの構成

本システムの構成を図1に示す。スパコンと



図 1：学際大規模計算機システム構成図

クラウドシステムが中心で、これらに、商用アプリケーションを実行するためのアプリケーションサーバや、学会等でのポスター発表の際に利用される大判プリンタが加わる。また、本システムでは、近年発生頻度が増大している地震等の災害への備えとして、北見工業大学に設置されたクラウドアーカイブシステムに遠隔バックアップを行うシステムを初めて導入した。

2.2 スーパーコンピュータシステムの構成

図 2 にスパコンシステムの構成を示す。主力機となるサブシステム A (通称: Grand Chariot) は、CPU として、ノード当たり 2 基の Intel 社製 Xeon (Skylake 世代) と 384GB のメモリを搭載する。前スパコンシステム (HITACHI SR16000) は IBM 社製 Power7 を搭載していたが、現スパコンでは x86 アーキテクチャの CPU に転換した。サブシステム A 全体は 1,004 ノードで構成され、前スパコンの 176 ノード (論理ノード数) から大幅な増加となった。加えて、Intel 社製 Xeon Phi (Knights Landing 世代) を搭載した 288 台の計算ノードで構成されるサブシステム B (通称: Polaire) も導入し、2 種類のサブシステムを利用者が選択可能とした。ストレージシステム (16PB) は、2 つのサブシステムで共通化されており、Intel Omni-Path ネットワークで結ばれている。

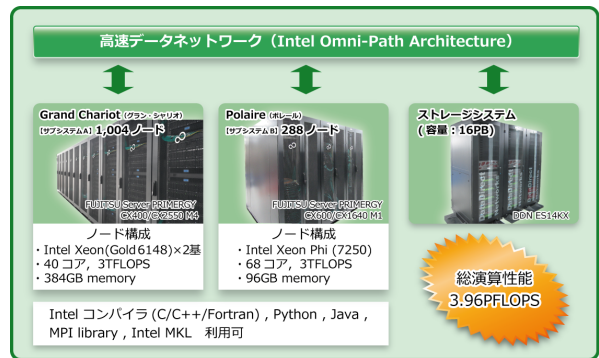


図 2：スーパーコンピュータ構成図

3 スパコンサービスの概要と利用状況

3.1 基本サービス

スパコンサービスをはじめとする、本システムで提供されているサービスを利用するためには、アカウント (利用者番号) を取得する必要がある。アカウントの取得には、基本サービス経費 (1 アカウント毎に、一般: 年額 12,960 円、学生: 年額 2,160 円) が必要である。

基本サービスを申し込むことで、以下のサービスが利用可能となる。

- スパコンの試用・デバッグ用リソースグループ
- スパコンのストレージ: home 領域 100GB
- アプリケーションサーバ
- WebDAV ストレージ (Nextcloud): 1TB (学生は 100GB)

スパコンの試用・デバッグ用リソースグループでは、各サブシステムにおいて、テスト用のジョブ (最大ノード数: 4、最大実行時間: 1 時間) が実行できる。これにより、プログラムの動作確認や実行時間の目安の把握が可能であり、本格利用に向けた検討に役立てることができる。前スパコンでは、試用・デバッグ用のリソース提供できていなかったため、初めてスパコンを利用する場合に、費用の目安が分からないなど、利用の障害になっていた可能性があると考えられる。それに対して、現スパコンでは、基本サービスの一部として、試用・デバッグ用のリソースを提供することで、スパコンサービスの利用拡大

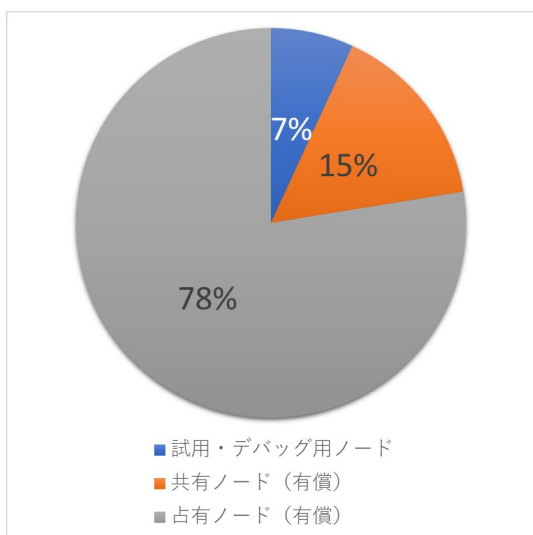


図 3 : 2022 年度にスパコン上で実行されたジョブ件数の内訳

や利便性向上につながったと考えている。

図 3 は 2022 年度にスパコン上で実行された総ジョブ数の内訳を示したグラフである。このグラフから分かるように、試用・デバッグ用リソースで実行されたジョブは全体の 7% であり、十分に活用されていたと考える。

3.2 付加サービス

スパコンを利用した本格的な計算は、基本サービスの申込後、別途、付加サービスを申請することで可能となる。スパコンを利用した演算に関する付加サービスとしては、以下で紹介するように、共用ノード利用と占有ノード利用の 2 種類が提供されている。また、スパコン用のストレージの領域を追加するための付加サービスも利用可能である。いずれの付加サービスも、申請した利用者を代表とするグループに別の利用者を登録し、申請したリソース（演算時間やストレージ領域）をグループで共有することが可能である。例えば、研究室やプロジェクト単位でスパコンを利用する場合、利用者全員が基本サービスを申請してアカウントを取得し、別途、代表者が付加サービスを申請後、研究室の学生やプロジェクトの参加メンバーをグループに追加することで、申請した付加サービスをシ

表 1 : 利用負担金
(サブシステム A・一般利用コース)

コース	配分演算時間	負担金
A3	3,000,000 ノード秒 (約 35 ノード日)	24,000 円
A15	15,000,000 ノード秒 (約 174 ノード日)	81,000 円
A100	100,000,000 ノード秒 (約 1,157 ノード日)	405,000 円
A250	250,000,000 ノード秒 (約 2,894 ノード日)	810,000 円

ェアすることができる。

3.2.1 共用ノード利用

共用ノード利用のサービスでは、利用者は事前に演算時間（トークン）を申請し、それを消費する形でジョブを実行する。共用ノード利用において使用するリソース（計算ノード）は利用者間で共有されており、一般的なジョブスケジューラの管理の下で各利用者のジョブが実行される。サブシステム A は、最大 256 ノード、サブシステム B は最大 128 ノードのジョブが実行可能である。

共用ノード利用の付加サービスの申請は、各サブシステムにおいて設定されている 4 種類のコースから選択する形式となっている。例えば、サブシステム A（一般利用）で提供されているコースは表 1 の通りである。共用ノード利用の場合、利用者のジョブが実行されるタイミングが事前に分からず、それを踏まえた上で計算ノードの管理などを行う必要があり、より演算時間の多いコースを申請してもらう方が、スパコンの運用計画を検討する上で都合がよい。そのような点を踏まえて、本センターでは、過去のスパコンサービスから、表 1 にあるような、割引があるコース体系を採用している。

2022 年度にスパコン上で共用ノード利用として実行されたジョブの内訳を調べてみると、ジョブの件数としては、1 から 5 ノードを使用したジョブが多くを占めていた。一方、総演算時間（総ノード時間積）の観点で内訳を

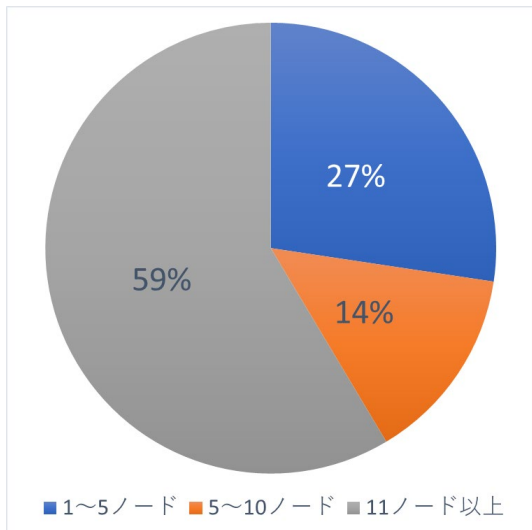


図 4：2022 年度に共用ノード利用として実行されたジョブの総演算時間（総ノード時間積）の内訳

調べると、図 4 が示す通り、11 ノード以上を使用したジョブが約 60%を占めていた。

3.2.2 占有ノード利用

占有ノード利用のサービスでは、利用者（とそのグループのメンバー）は申請したノードを占有してジョブを実行できる。つまり、電気設備の法定点検等でサービス全体が停止する場合を除いて、グループ外の他の利用者のジョブの状況を気にすることなく、スパコンを利用することが可能である。

占有ノード利用は、各サブシステムにおいて 1 ノードを単位として申請する形式である。利用負担金は年額であり、四半期を単位として、申請時期により減額される。また、1 ノード毎にスパコンストレージの work 領域 3TB が付与される。

この数年、サブシステム A の占有ノード利用の希望が非常に多く、実際に提供するノード数を調整せざるを得ない状況になっている。現在の運用の方針としては、サブシステム A 全体（1,004 ノード）の約半数である 500 ノードを占有ノード利用に配分するノード数の上限の目安としている。これに対して、前年度の利用者が

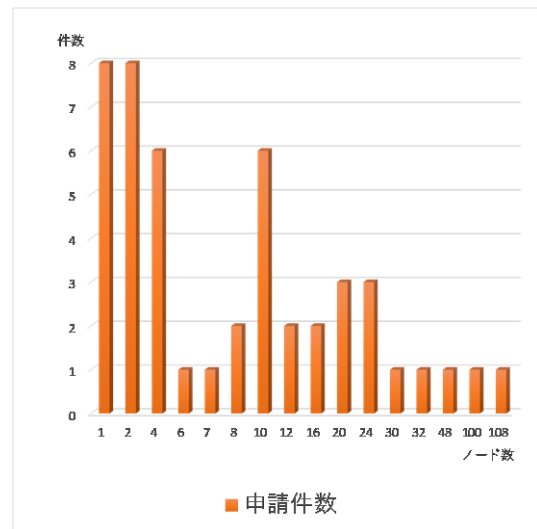


図 5：占有ノード利用における提供ノード数ごとの申請件数（2022 年度）

翌年度への継続申請の際に希望する占有ノード利用のノード数の合計が 500 ノードを超えており、センター側で調整を行う必要が生じている。具体的には、新年度開始後に新規利用者からの占有ノード利用の申請を一定数受け付けることができるように、年度当初の提供ノード数を調整している。一方、年度開始後の新規利用者の申請については、最大受付可能ノード数を 2 と制限し、申請時期が遅くなった場合には受付不可とする場合もある。

図 5 は 2022 年の占有ノード利用における提供ノード数ごとの申請件数を示している。上述した通り、年度途中の新規利用者に対する提供ノード数の上限を 2 と制限せざるを得ない状況であるため、1 から 2 ノードの申請が多い。一方で、大規模な研究プロジェクトでも利用されており、100 ノードを超える申請も存在している。

3.3 スパコン全体の稼働状況について

まず、2022 年度の消費電力量の状況を図 6 に示す。図 6 の右側のグラフは、学際大規模計算機システムにおけるスパコンとスパコン以外の IT 機器の消費電力量の割合を図示している。このグラフが示すように、スパコンの消費電力量が全体の 96%を占めている。また、左側のグラフは、学際大規模計算機システムに関連する空

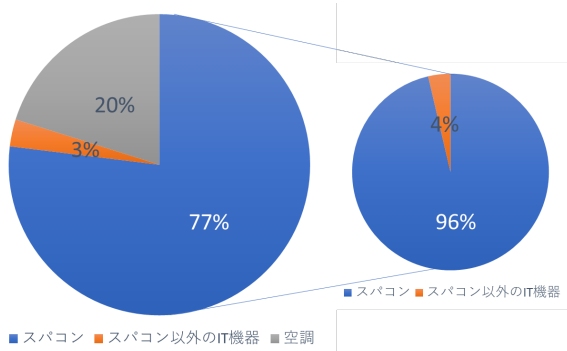


図6：消費電力量の比率（2022年度）

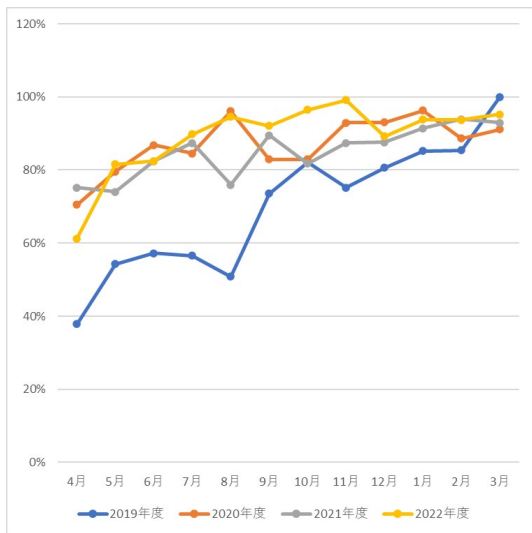


図7：スパコンの稼働率の推移

調機器の消費電力量も含めた場合の内訳を図示している。空調機器を含めた場合でも、スパコンの消費電力量は全体の77%と大部分を占めている。

上述したように、学際大規模計算機システムの消費電力量の大部分をスパコンが占めており、そのための電気料金をカバーできるように、スパコンサービスの利用負担金を設定することが求められる。スパコンの消費電力量は、計算ノードの稼働率に応じて変化する部分と、ストレージやネットワークのように稼働率に関わらず定常的に発生する部分があるため、これらを総合的に考慮して、適切な負担金の体系を設計することが重要となる。

図7はスパコンの稼働率の推移を示したグラフである。現スパコンの運用開始時、本センターの運営に支障が生じることがないよう、多少

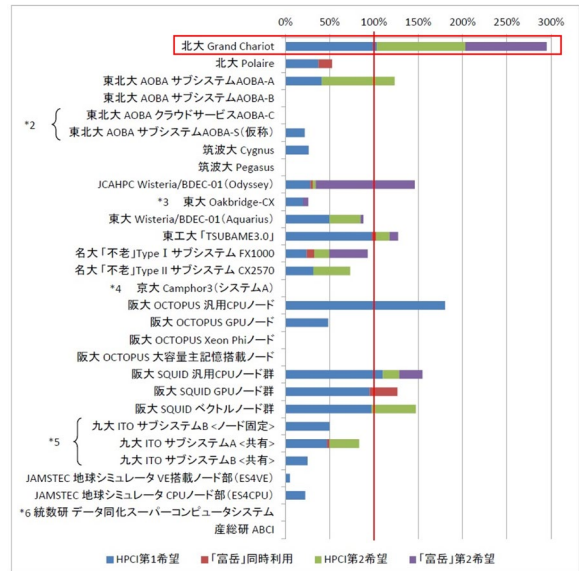


図8：2023年度 HPCI 共用計算資源ごとの要求資源量の割合（[3]の図3-1より引用）

控えめに想定稼働率を設定し、それに基づいて負担金を決定した。しかし、結果的には、初年度（2019年度）前半を除いて、想定を大きく上回る稼働率で運用ができています。そのため、昨今の電気料金の高騰に対して、2023年度までは、運用開始時の負担金を値上げすることなく、対処ができています。ただし、昨今の電気料金の高騰に対して、北海道大学では全学的な補助があり、その結果として、学際大規模計算機システムの電気料金を含む本センター全体の電気料金をやりくりできている状況です。したがって、負担金の値上げに関しては、予断を許さない状況です。

図8は、資源提供している HPCI（革新的ハイパフォーマンス・コンピューティング・インフラ）の2023年度A期の研究課題の応募状況である[3]。図8から分かるように、サブシステムAに関して、応募課題の要求資源量の総和が提供資源量の3倍に迫る状況です。この HPCI 課題における積極的な利用は、図7に示した高い稼働率の大きな一因になっていると考える。

4 課題

前節で報告した通り、高い稼働率をはじめとして、概ね順調にスパコンを運用できているが、

一方で、対処が困難である課題もいくつか生じているので、以下で簡単に述べる。

4.1 年度末の待ち時間の長時間化

共用ノード利用では、通常時でも数百件程度の実行待ちジョブが発生しているが、年度末が近づくと、その数がより多数になり、ジョブの実行開始までの時間が非常に長くなる傾向が顕著になっている。そのため、年度末近くの新たな共用ノード利用の申請については、ジョブの実行待ち時間の長時間化により、配分された演算時間が使いきれない可能性がある旨を利用者に確認した上で、申請の受付を行っている。また、HPCIやJHPCNの課題利用者に対して、早い時期からの計画的な利用を促している。

4.2 講習会用のリソース不足

共用ノード利用における実行待ちジョブの増加と占有ノード利用の希望多数のため、講習会を開催するためのリソースの確保が難しい状況になっており、近年は講習会の開催頻度が低くなっている。ただし、現状のリソースの状況を考えると、新規利用者を増やしても、リソースの限界に近い状況のため、満足に利用してもらえない可能性が高いとは言い難い。講習会の開催頻度が下がった反面、既存の利用者に対するサポートにより注力できており、また、利用の必要性が高い場合には、個別の対応を行い、新規利用につながっている。スパコンの利用者を拡大することは重要な課題であるが、一方で、提供可能な資源には限界があり、利用者数がそれを超えるような場合には、利用者の満足度が低下する可能性がある。その点を踏まえると、結果的に、現在提供可能なスパコンのリソースに見合う数の利用者に対して、適切なサポートが提供できているのではないかと考える。

4.3 設置建屋の老朽化

システムの設置場所である、情報基盤センター北館は、本センターの前身である計算機センターが1962年に発足した後、1970年の情報基

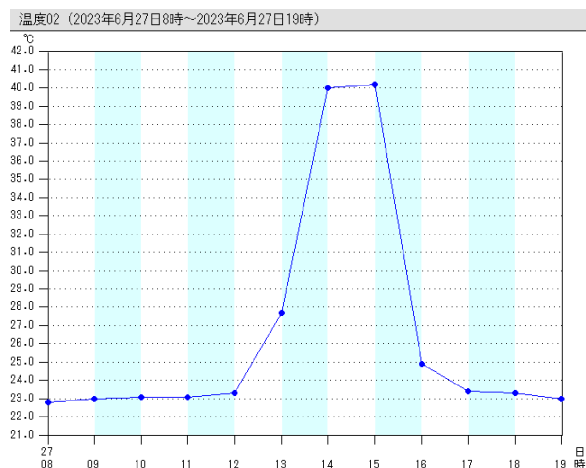


図9：チラー障害時のスパコン室の温度変化の例

盤センター設立時に竣工し、以降増築や耐震改修等を行ってきたものの、建物自体が非常に老朽化しており、対策が必要となっている。現在、北館と同様に老朽化が問題となっていた南館の改修を行っており、これが完了した後、次期システムを南館に設置する予定である。その後、北館の改修を行い、次々期システムを再び北館に設置する「式年遷宮方式」の運用を計画している。

4.4 冷却設備（チラー）の老朽化

スパコンの冷却設備の一つであるチラーは、前システムの導入開始時（2011年）に導入したものであり、2022年中から障害発生頻度が増加傾向にある。障害によりチラーが停止すると、スパコンが設置しているサーバ室の温度が上昇し、スパコンの計算ノードがダウンするリスクが高くなる。

図9は過去のチラーの障害時のスパコン室の室温の推移を示すグラフである。この障害時、スパコンの計算ノードは停止寸前の状況になったものの、復旧作業が間に合い、ノードの停止は避けることができた。なお、チラーの障害の影響を受けた部屋には、スパコンの計算ノードのみが設置されており、スパコンストレージやネットワーク機器は、外気冷却および空冷空調により冷却している別室に設定していたため、このチラーの障害の影響は受けなかった。

5 おわりに

現在運用中のスパコンは、2018年12月から5年の賃貸借契約により導入した学際大規模計算機システムの一部であり、契約満了（運用開始から5年）が近づいているが、その間、安定的な運用ができ、多くの研究者に利用されてきた。次期システムは、4.3節で紹介したように、現在改修中の南館に設置予定であるが、半導体不足による資材供給の不安定な情勢に起因する竣工時期の遅延を想定し、2025年4月の運用開始を目指すこととした。そのため、現在運用中のスパコンを含むシステムについて、1年4か月の運用の延長を予定している。電気料金の高騰の影響、設置建屋の改修、そして冷却設備の経年劣化による故障など、懸案が多い状況ではあるが、引き続き、安定運用に注力し、多く利用者の研究に貢献できるように努めたいと思う。

謝辞

運用開始から現在に至るまでのスパコン運用に対するサポートについて、北海道大学情報基盤センターおよび総務企画部情報企画課、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/>
2. 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,
<https://axies.jp/report/publications/papers/papers2018/>
3. 令和5年度A期HPCIシステム利用研究課題の応募状況,
https://www.hpci-office.jp/using_hpci/project_application_stats