

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

深谷 猛¹⁾, 岩下 武史¹⁾, 金子 修己²⁾, 折野 神恵²⁾, 更科 高広²⁾

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 new supercomputer system in Information Initiative Center, Hokkaido University

Takeshi Fukaya¹⁾, Takeshi Iwashita¹⁾, Naoki Kaneko²⁾, Kamie Orino²⁾, Takahiro Sarashina²⁾

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月より、新学際大規模計算機システム（北海道大学ハイパフォーマンスインテークラウド）によるサービスを開始する。本稿では、新システムにおけるスーパーコンピュータシステムと関連サービスの概要を紹介する。

1 はじめに

北海道大学情報基盤センターでは、新学際大規模計算機システム（北海道大学ハイパフォーマンスインテークラウド）によるサービスを2018年12月から開始する [1]。新システムはスーパーコンピュータシステム（スパコンシステム）とクラウドシステムから主に構成され、システム全体としての総合演算性能は約4PFLOPSとなる。本稿では、新システムにおけるスパコンシステムと関連するサービスの概要を紹介する。

2 新スパコンシステムの構成

新スパコンシステムは、サブシステム A (Grand Chariot^{*1})、サブシステム B (Polaire^{*2})、およびストレージシステムから構成される (図 1)。二つのサブシステムはどちらも Intel 製の CPU と Linux OS を搭載しており、旧スパコンシステム (HITACHI SR16000 M1, IBM Power プロセッサと AIX を搭載) に比べて、一般的な研究室の計算機環境との親和性という点においては向上が期待される。サブシステム A と B を合わせた、新スパコンシステム全体の理論演算性能は 3.96PFLOPS である。

2.1 ハードウェア

以下、サブシステム A, B, ストレージシステムの概要を述べる。

- サブシステム A (Grand Chariot) : 1004 ノードで構成されるシステムで、各ノードは 2 基の Intel 製マルチコア CPU (Intel Xeon Gold 6148) と 384GB のメモリを搭載。ノード間は Intel Omni-Path ネットワーク (ノード当たり 2 ポート) で結合。
- サブシステム B (Polaire) : 288 ノードで構成されるシステムで、各ノードは 1 基の Intel 製メニーコア CPU (Intel Xeon Phi 7250) と 96GB のメモリを搭載。ノード間は Intel Omni-Path ネットワーク (ノード当たり 1 ポート) で結合。
- ストレージシステム : 全体で 16PB の物理容量を有するストレージシステムで、Lustre ベースの並列ファイルシステムを採用。

なお、サブシステム A, B およびストレージシステムの間ネットワークも Intel Omni-Path である。

表 1 に新旧スパコンシステムの主な性能諸元の比較を示す。新スパコンシステムは、システム全体の演算性能、総ノード数、電力効率等を重視して設計されたため、ノード当たりのメモリバンド幅 (と BF 値) は旧スパコンシステムから低下することになった。しかし、新スパコンシステムでは、利用料金算出の根拠と

*1 「グラン・シャリオ」, フランス語で北斗七星。

*2 「ポレール」, フランス語で北極星。

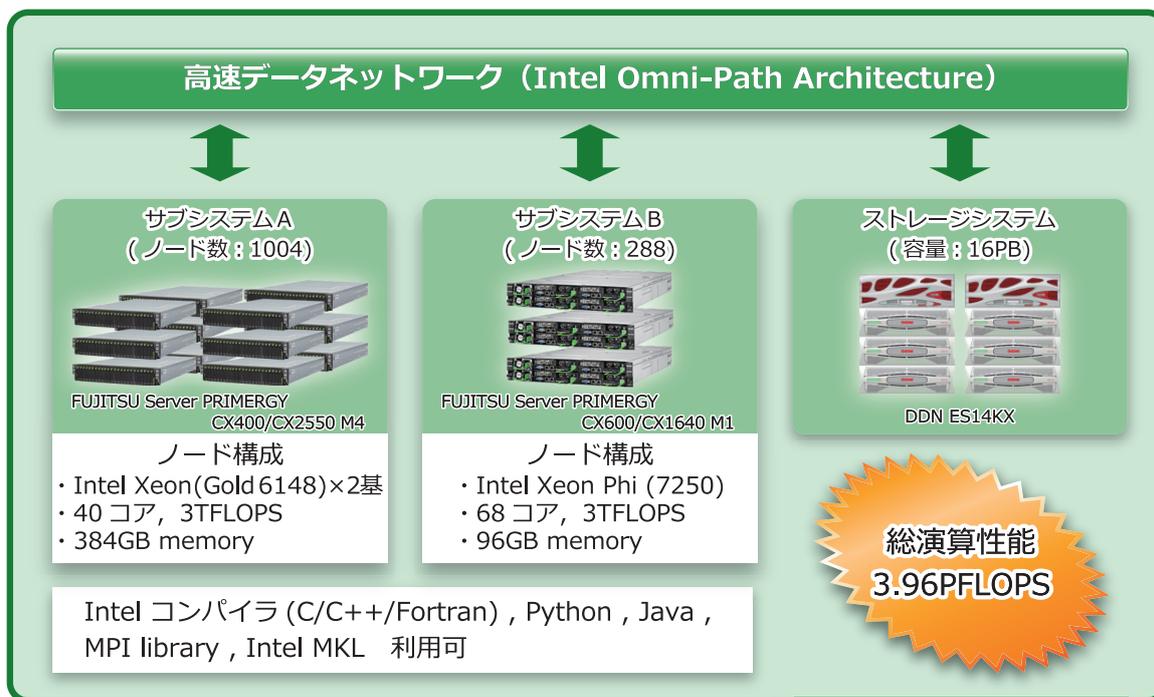


図 1: 新スパコンシステムの全体像.

表 1: 新旧スパコンシステムの性能諸元: サブシステム B の括弧内の数値は高速メモリ (MCDRAM) の場合の値.

| 項目 | | 旧スパコン | 新スパコン | |
|--------|-----------|--------------|----------------------|-------------------------|
| | | | サブシステム A | サブシステム B |
| システム全体 | ノード数 | 176 | 1004 | 288 |
| | 理論演算性能 | 0.172 PFLOPS | 3.08 PFLOPS | 0.87 PFLOPS |
| プロセッサ | 種類 | IBM POWER7 | Intel Xeon Gold 6148 | Intel Xeon Phi 7250 |
| | 周波数 | 3.83 GHz | 2.4 GHz | 1.4 GHz |
| | コア数 | 8 | 20 | 68 |
| ノード | 搭載プロセッサ数 | 4 | 2 | 1 |
| | 搭載メモリ容量 | 128 GB | 384 GB | 96 GB (16 GB) |
| | 理論演算性能 | 980 GFLOPS | 3072 GFLOPS | 3046 GFLOPS |
| | 理論メモリバンド幅 | 512 GB/s | 255.9 GB/s | 115.2 GB/s (921.6 GB/s) |
| BF 値 | | 0.52 | 0.083 | 0.038 (0.30) |

なる, ノード当たりの消費電力が大幅に下がる見込みであるため, 複数のノードを用いたジョブが安価に実行でき, ジョブ全体として利用可能なメモリバンド幅を向上させることができる。したがって, 新スパコンシステムのユーザに対して, 「プログラムを適切に並列化し, 多数のノードを利用することで, 旧スパコンシステムよりも効率的にシミュレーション等を行ってもらう」ように推奨することが重要となる。

2.2 ソフトウェア

サブシステム A および B において, Intel 製のコンパイラ (Fortran, C, C++), 数値計算ライブラリ (MKL),

MPI ライブラリ等を提供する。また, Java や Python も利用可能である。有償のアプリケーションソフトウェアとしては, 分子解析ソフトウェア (Gaussian) と流体解析ソフトウェア (V-FaSTAR) を提供する (サブシステム A のみ)。また, フリーソフトとしては, 数値計算ライブラリ (PETc, Trilinos, PLASMA, EigenExa, z-Pares 等) や計算科学分野のアプリケーション (OpenFOAM, FrontFlow, GAMESS, PHASE 等) に加えて, 機械学習分野のソフト (Chainer, Tensorflow, Caffe) も整備する。

2.3 運用方針

サブシステム A および B の計算ノードは、「共用ノード」と「占有ノード」に分けて運用する予定である。共用ノードと占有ノードの概要は以下の通りである。

- **共用ノード**：演算時間（ノード時間積）に応じて利用料金が発生し、ユーザ間で計算資源を共同利用する。計算科学分野における、チャンピオンデータを取得するために行う超大規模シミュレーションにも対応できるサービスとして位置づけている。
- **占有ノード**：年間定額制を採用し、申請者（と申請者が許可したユーザ）のみが計算資源を利用できる。年間を通じて一定の資源の利用が可能となるため、安定的な研究の推進を希望するユーザのニーズに対応できるサービスとして位置づけている。

ストレージシステムは、「ホーム領域」と「ワーク領域」に分けて運用予定である。ホーム領域とワーク領域では、データのバックアップ等の点において差別化をする予定である。ホーム領域には、プログラムやソースコードなどの重要性の高いデータを置き、ワーク領域には、プログラムの出力結果など一時的な大容量のデータを置くことを想定している。

3 提供サービス

北海道大学情報基盤センターでは、新スパコンシステムを含む学際大規模計算機システムに関して、大学や研究所等の研究者向けの「一般利用コース」と、民間企業等の利用者向けの「民間企業等利用コース（成果公表または成果非公表）」を提供している。なお、民間企業等利用コースでは審査を行い、利用目的等が適切であるか、事前に確認を行っている。

システムの利用を希望する場合、最初にユーザ単位で利用者登録（基本サービス経費が必要）を行い、その後、必要なスパコンシステム（あるいはクラウドシステム）の付加サービスを申請する手順となる。

以下では、一般利用コースについて、新スパコンシステムで提供するサービスの概要を紹介する。

- **基本サービス**：サブシステム A および B における「試用・デバッグ用の共用ノード」と一定量のストレージ領域（ホーム領域）が利用可能。
- **付加サービス**：サブシステム A および B の共用

ノードにおける演算時間（トークン）、占有ノード（1 ノード単位、一定容量のワーク領域付き）、ストレージの追加領域（ホーム領域およびワーク領域）を申請。

なお、付加サービスは申請者だけでなく、申請者が許可した他のユーザも申請資源を利用することが可能である。

また、上記の一般利用だけでなく、新スパコンシステムの計算資源は旧スパコンシステムと同様に、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）、JHPCN（学際大規模情報基盤共同利用・共同研究拠点）、北海道大学情報基盤センター萌芽型共同研究で公募する課題においても利用可能となる。

4 プログラムの移行・チューニング支援

新スパコンシステムのアーキテクチャは旧スパコンシステム（HITACHI SR16000 M1）のアーキテクチャと大きく異なる。そこで、旧スパコンシステムのユーザのプログラム資源を円滑に新スパコンシステムに意向するための「移行支援」、更に新スパコンシステムにおいてユーザが有するプログラムを高性能化するための「チューニング支援」を公募し [2]、表 2 にあるプログラムを選定した [3]。採択されたプログラムは、旧スパコンシステムで実行されていた大型シミュレーションプログラムで、新スパコンシステムでも実行予定であり、移行の遅れが今後の研究に支障をきたす恐れがあるものである。また、新スパコンシステムにおいて、高性能化を図ることにより、一層の研究成果が期待されるものである。

採択された各プログラムについては、入力データを含むプログラム一式と旧スパコンシステムにおける出力結果、性能情報（使用ノード数、使用プロセッサ数、計算時間等）をユーザから受領し、新スパコンシステムの稼動開始前および後に以下の内容で支援を実施する。

- **移行支援**：新スパコンシステムの稼動開始までに、新スパコンシステムにおけるプログラムの動作確認とそれに必要なプログラムの改修を行う。外部のライブラリや特殊な関数の利用等により、新スパコンシステムにおける動作が確認できない場合は、問題点を記述したレポートをユーザに送付する。また、その場合には、可能な限り新スパコンシステムにおける代替手段を提示する。
- **チューニング支援**：新スパコンシステム稼動後に、

表 2: 移行・チューニング支援 採択課題.

| プログラム名 | 移行 | チューニング |
|---|----|--------|
| 全球気候モデルと領域気象モデルの結合システムによる地球環境・領域気象予測 | ○ | ○ |
| 海洋植物プランクトン多様性モデル | ○ | ○ |
| sph3D (3次元 Smoothed Particle Hydrodynamics) | ○ | - |
| 次世代ワイヤレス接続サービスの回線設計・評価支援プログラム | ○ | ○ |
| 海洋非静水圧モデル Kinaco | ○ | - |
| 海洋熱塩・栄養物質循環モデル | ○ | ○ |

各対象プログラムの高性能化に関する支援を行う。各プログラムについて、ユーザ、新システムの落札ベンダ、北大情報基盤センターの三者が協議し、約3年間に亘り、継続的にプログラムの高性能化に取り組む予定である。プロファイラによるホットスポットの同定やループレベルのチューニングなど微視的なプログラムチューニングに留まらず、解法や並列化アルゴリズム、実装方式などについても検討を行う予定である。即ち、情報基盤センターの教職員が有する高性能計算や応用数学等に関する知見を実際のユーザプログラムにおいて活用し、計算科学を主体とする様々な応用分野の発展に寄与することを目指している。

5 おわりに

本稿では、2018年12月にサービスを開始する、北海道大学情報基盤センター 新学際大規模計算機システムのスーパーコンピュータシステムおよび関連サービスの概要を紹介した。新スパコンシステムは、旧スパコンシステムに対して、全体あるいはノード当たりの理論演算性能やノード数が飛躍的に増加する。また、OSに、計算サーバやクラウドシステム等のコンピューティングプラットフォームとして広く普及しているLinuxを採用し、一般的な研究室の計算環境との親和性を高めている。そのため、様々な研究分野において、今まで以上にスパコンが活用されることが期待される。一方で、旧スパコンシステムからは、プロセッサのアーキテクチャが大幅に変わるなど、プログラム資産の移行コストが少なくないことが予想されるため、移行支援・移植支援を行い、ユーザを出来る限りサポートする。新スパコンシステム稼働後は、システムの利用状況やユーザのニーズ等を踏まえながら、より使いやすいシステムを目指して、運用に努めたいと考えている。

謝辞

本稿を作成するにあたり、図の作成等でご協力頂いた、北海道大学 総務企画部 情報企画課の職員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のご案内, <https://www.iic.hokudai.ac.jp/newsys2018/>
- [2] 北海道大学情報基盤センター 新学際大規模計算機システム 移行支援・チューニング支援要領, https://www.hucc.hokudai.ac.jp/pdf/ikousien_h28.pdf
- [3] 北海道大学情報基盤センター 新学際大規模計算機システムのユーザプログラム移行支援及びチューニング支援制度の採択結果について, https://www.hucc.hokudai.ac.jp/pdf/ikousien_kekka_h28.pdf