

jh220025

## センター間連携による柔軟な計算資源提供に関する研究

滝沢寛之（東北大学）

### 概要

各情報基盤センターの計算資源の利用率は時々刻々と変化しており、ある機関で計算資源が足りない時に、他の機関では同じ種類の計算資源が遊休状態にある可能性がある。このため、より柔軟に計算資源を融通しあう環境を整備することができれば、JHPCN 全体で計算資源のより効率的な活用を期待できる。そのような複数機関の連携による運用を実現するためには、ストレージシステムやバッチジョブスケジューラの連携という技術的検討に加えて、課金体系や利用者管理などの運用方針の観点からも様々な検討が必要となる。本研究では、技術的視点からの検討によって実現可能性を示すとともに、複数のセンター間連携による運用の実現に向けた運用上の課題の明確化を行った。その結果、スーパーコンピュータ AOBA のサブシステム間での連携を題材として、地理的に距離が離れているサブシステム間がお互いにジョブを投入しあう機能を実現した。そのような機能を実運用でも利用するために必要な検討事項を、東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターとの間で今後も協議していく。

### 1. 共同研究に関する情報

#### (1) 共同利用・共同研究を実施している拠点名

東北大学 サイバーサイエンスセンター  
大阪大学 サイバーメディアセンター

下村 陽一 ジョブ転送機能の設定  
高橋 慧智 システム利用支援  
川本 裕樹 混相流の解析と評価  
遠藤 新 クラウド利用方式の検討

#### (2) 課題分野

大規模計算科学課題分野

#### (3) 共同研究分野

超大規模数値計算系応用分野  
超大規模情報システム関連研究分野

#### (4) 参加研究者の役割分担

滝沢 寛之 全体統括とジョブ転送設計  
佐々木大輔 大規模ジョブ実行時の課題調査  
伊達 進 大阪大学側統括  
高橋 俊 大規模ジョブ実行時の課題調査  
小川泰一郎 大規模ジョブ実行時の課題調査  
江川 隆輔 運用方針の検討

### 2. 研究の目的と意義

JHPCN は 8 つの機関から構成されるネットワーク型の共同利用・共同研究拠点であり、それぞれの機関が独自の計算資源を独自の運用方針に基づいて提供している。本研究の目的は、JHPCN 計算資源の有効活用を実現する資源管理機構とその運用体制を検討し、実現可能性を明らかにすることである。

各計算資源の利用率は時々刻々と変化しており、ある機関で計算資源が足りない時に、他の機関では同じ種類の計算資源が遊休状態にある可能性がある。このため、より柔軟に計算資源を融通しあう環境を整備することができれば、JHPCN 全体で計算資源のよ

り効率的な活用を期待できる。そのような複数機関の連携による運用を実現するためには、ストレージシステムやバッチジョブスケジューラの連携という技術的検討に加えて、課金体系や利用者管理などの運用方針の観点からも様々な検討が必要となる。本研究では、技術的視点からの検討によって実現可能性を示すとともに、複数のセンター間連携による運用の実現に向けた運用上の課題の明確化を行う。

### 3. 当拠点の公募型研究として実施した意義

本研究課題は、複数の JHPCN 資源提供機関が連携してより柔軟な資源提供の可能性を検討するものであり、技術的課題の解決に向けた実証実験や、拠点間連携の運用体制構築に向けた課題の明確化のために JHPCN 資源を実際に利用する必要がある。両センターの代表的なアプリケーションを他方のセンターの資源に転送して実行する実証実験を、条件を変えて繰り返し行うことを計画しており、両センターの計算資源利用や担当者の参画が必要不可欠である。

本研究は、課題募集要項に記載されている「本公募型共同研究が重視する研究内容」のほとんどの項目に該当する。まず、計算科学者と

計算科学者から構成される学際的研究体制であることは言うに及ばず、ソフトウェアおよびデータ活用推進の項目にも該当し、実証実験によって構築されたジョブ実行の連携機構に関しては、明らかに正式運用によって成果の幅広い活用を期待するものである。また、IT 基盤技術開発や拠点間連携の項目に該当することも明らかである。大規模データ・大容量ネットワーク利用の項目にも該当し、両センター間で大量のデータ転送を行うことが想定されている。

以上の理由から、本研究課題は JHPCN が重要視する記載している研究内容と合致しており、JHPCN の課題として極めて意義深い。

東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターはともにベクトル型システムを提供する点などで提供資源に共通性があるものの、例えば大阪大学側では GPU クラスタや大規模ストレージを提供している点などの相違も見られる。このことから、混雑状況やジョブの種類に応じて両センターの計算資源を柔軟かつ適切に利用する技術を研究開発することは、計算機科学分野での学術的成果も期待できる。

### 4. 前年度までに得られた研究成果の概要

新規課題のため該当しない。

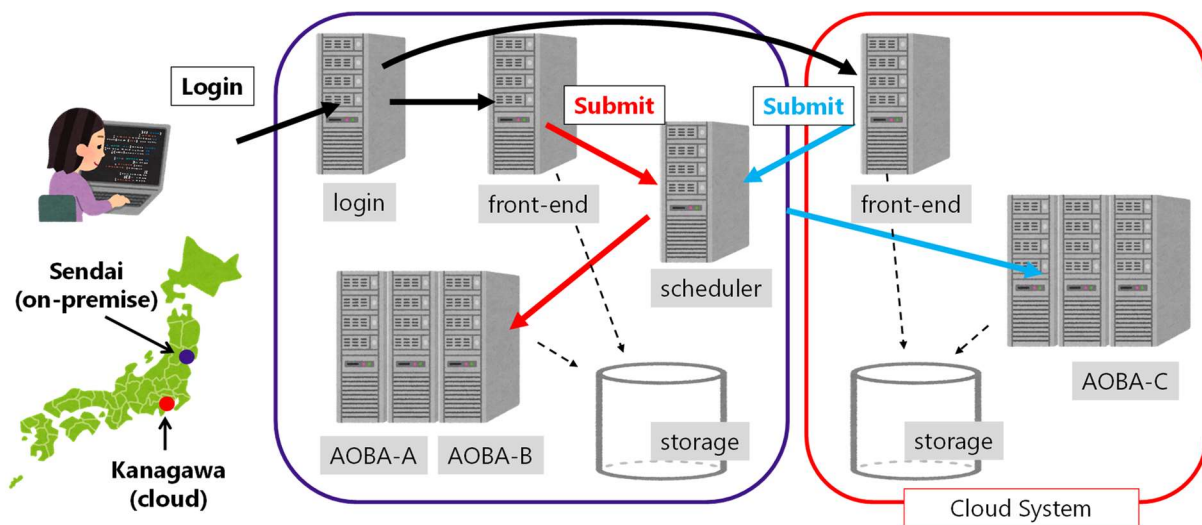


図 1 : スーパーコンピュータ Aoba におけるクラウド連携の実現

## 5. 今年度の研究成果の詳細

### (1) センター間ジョブ転送の仕組みの構築

本研究では、東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターのそれぞれのセンターにおいて専用のバッチキュー(以下、転送キューと呼ぶ)を設け、そこに投入されたジョブは他方のセンターの計算資源を使って実行されるように設定することを目指してきた。

東北大学サイバーサイエンスセンターでは令和 4 年 10 月 3 日からクラウド利用を開始し、サブシステム AOBA-C として運用されている。AOBA-C 導入後のスーパーコンピュータ AOBA の構成を図 1 に示す。東北大学のクラウド利用では、オンプレミス(AOBA-A および AOBA-B)とデータセンター(AOBA-C)との間でジョブスケジューラの連携とユーザ認証情報の共有を実現しており、そのためのソフトウェア環境を整備することができた。

スーパーコンピュータ AOBA の場合には、オンプレミスシステムとクラウドシステムの両方が東北大学サイバーサイエンスセンターによって運用されている。上述のソフトウェア環境整備の結果として、両者が別の組織によって運用されている場合にも同様の連携が技術的には可能となったと考えている。ジョブスケジューラがオンプレミス側でもデータセンター側でも動作しているが、利用者はその構成を意識する必要はない。このことから、技術的には転送キューに当たる機能を実現できたといえる。ただし、異なる組織で運用するシステム間の連携には後述の通り様々な課題が残されており、本共同研究終了後も引き続き東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンター間で検討を続けていくことを予定している。

なお、総理論演算性能 21Pflop/s という世界最大規模のベクトル型スーパーコンピュ

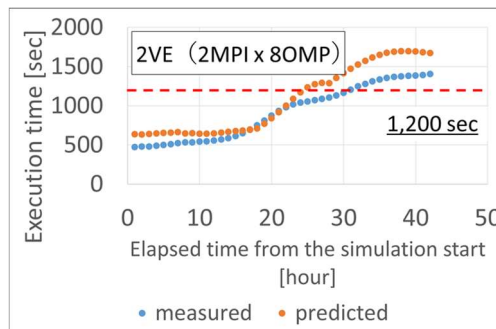


図 2:2VE 実行では×切に間に合わないことを正しく予測

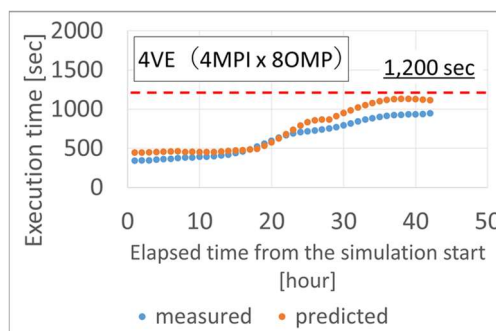


図 3:4VE 実行では×切に間に合うことを正しく予測

ータとして、オンプレミスのサブシステム AOBA-S が令和 5 年 8 月 1 日より稼働する。AOBA-C は運用開始時から常に安定的に運用されているが、AOBA-S の稼働によりオンプレミスの計算資源が大幅に増強されることから、令和 5 年 7 月末で AOBA-C の運用を終了することが決まっている。

### (2) ジョブスケジューリングの検討

東北大学サイバーサイエンスセンターでは、Slurm シミュレータに独自の拡張を加えて同センターが運用するスーパーコンピュータ AOBA のジョブスケジューリングを再現する研究開発を行っている。現在では AOBA 内でのジョブスケジューリングを対象とした研究開発を行っているが、パブリッククラウドや他のセンターへのジョブの転送も可能になった場合を想定して、その適切なジョブスケジューリングを実現するための検討を進めている。

本年度の主な成果として、三井共同建設コンサルタント株式会社と共同研究を行っているリアルタイム洪水シミュレーションを対象とし降雨データに適応して必要な計算資源量を予測する手法を提案したことが挙げられる[1][8][11][14]。図2および図3に予測結果の例を示す。この降雨データの条件下では、2基のベクトルエンジン(VE)を使う場合には、図2に示す通り〆切時間である1,200秒以内に洪水シミュレーションを実行することはできない。一方、4基のVEを使う場合には、図3に示す通り〆切時間までにシミュレーション実行を完了することができる。提案手法によって、〆切時間までにシミュレーションを完了するために必要な計算資源量を予測できている。事前に必要な計算資源量を予測することで、降雨データに合わせて必要最小限の資源のみを洪水シミュレーションに割り当てることが可能となった。

他にも令和4年度には、強化学習によって電力制御に関するジョブスケジューラのパラメータを自動調整する研究[4][5]や、緊急ジョブ実行時の効率化のためのジョブスケジューリング手法を検討した[13]。外部との

データ連携によってより複雑なジョブスケジューリングが求められる場合を想定した研究開発であり、ジョブスケジューラの現在の機能と限界、および将来の開発の方向性を明確化することができた。

また、民間クラウドベンダの提供するIaaS型クラウド資源を計算機センター保有のオンプレミス資源を併用する場合の運用面での課題について、性能およびコスト面から評価を行った[3]。具体的には、本研究では、民間クラウドベンダの提供するクラウド資源、および、オンプレミス計算資源の性能差および利用時コスト差を考慮した、スケジューラキューの構成方法および運用方法についての提案を行い、利用者視点でクラウドバースト機能の利用促進を行うための検討を行った。本研究で得られた知見は、複数のセンター間連携での柔軟な計算資源提供を実用的かつ実践的な視点で実現するに際しても、コストパフォーマンス面での検討は不可避であり、今後の研究に生かしていく予定である。

### (3) 運用面での課題の明確化

本研究課題で実際の運用を念頭において

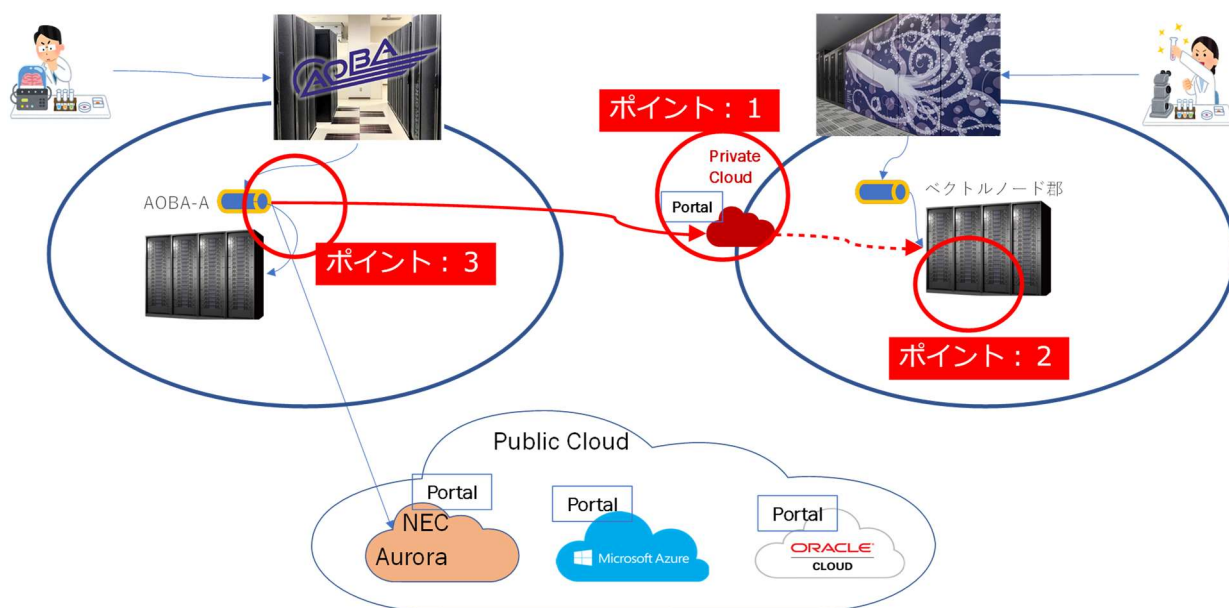


図4：プライベートクラウド構築イメージ



課題を明確化するとともに、その解決に向けた運用方針についても具体的に検討してきた。

令和4年度には、東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターの運用関係者がミーティングを行い、実証実験の実施に向けた課題の洗い出しと整理を行った。Software as a Service (SaaS) ではライセンス等の問題が生じる可能性があるために Infrastructure as a Service (IaaS) を前提とし、本課題用のプライベートクラウドを構築するための実装方式について検討を進めた。クラウド提供する資源の管理、監視や保守のために運用コストの増大が懸念されることから、現実的な実現案を検討した。

まず、「柔軟な計算資源提供の前提条件」を整理し、任意のセンターにおける計算需要が、急激かつ極端に高まったバースト時に、任意の他センターに余裕があれば迅速に他センターのリソースに切り替えることを可能にする IaaS サービスの実現を最終目標と定義した。そのための実現方式案としてはいくつか考えられるが、各センターにプライベートクラウドを構築する方式から検討を始めるのが望ましいとの結論に至った。大阪大学のベクトル資源を、プライベートクラウド方式によって東北大学へ提供するイメージを図4に示す。実現に必要な要素として、図中のポイント1～3を以下に説明する。

#### ポイント1：Cloud ポータル

フルセットのWebサービス全てを必要としないが、資源要求や資源要求に応じた IaaS 環境構築の「プロセスのほぼ全てが自動化されていることが望ましい。

#### ポイント2：クラウド資源管理機能

資源提供要求に従って要求された計算資源を自センターの運用から（事前または、オンデマンドに）切り離し、IaaS サービスとし

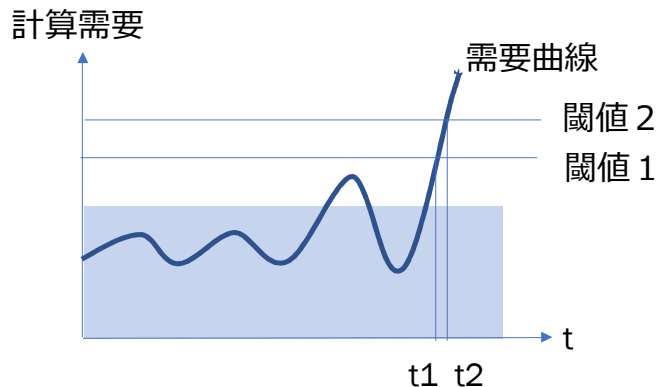


図5：閾値設定と計算需要予測

て提供する計算資源が自動的に構築されることが望ましい。また、ネットワーク管理、および物理サーバ管理が自動化されることが望ましい。

#### ポイント3：提供リソースの組込み機構

組み込みオンプレ実行、クラウド実行（マルチクラウドとなる場合はクラウドの選択）の判断、および優先選択機能等の機能は必須である。

ポイント1および2はセンター間の調整で固定的な環境を事前構築することで実現可能であると言えるが、技術的にはポイント3の実現には課題があり、今後の「計算資源連携」の実現に向けた研究課題となる。どの時点で以下の選択肢を判断するか、およびどれくらいのジョブをバースト対象とするかの検討が必要である。

- ・オンプレ資源で実行
- ・パブリッククラウドへバーストして実行
- ・プライベートクラウドへバーストして実行

例えば、それぞれの選択肢に閾値1と閾値2を定義したとしても、図5に示すとおり、閾値1を超えてどれくらいの時間で閾値2を超えるか、など色々な需要予測が技術的には必要となる。また、判断基準は要求資源の総量なのか、ジョブ数、実行待ち時間平均など複数の因子を複合的に判断すれば良いの

か等も検討する必要がある、それぞれのセンターの運用ポリシーを策定して決める必要がある。

なお、図 1 で示したようにジョブスケジューラの機能による連携も技術的には可能であるが、運用では少なくとも以下の課題を解決する必要がある、本研究における検討の結果として、「柔軟な計算資源提供の前提条件」を満たすのは難しいとの結論に至っている。

- 異なるユーザアカウント
- 異なるジョブ記述
- ジョブ操作権限
- ジョブデータへのアクセス

また、プライベートクラウド方式を採用したとしても、セキュリティ、システム更新計画、従量課金額、ストレージの観点からの運用課題があることがすでに分かっている。運用の安定性と利用者の利便性を考慮しつつ、東北大学サイバーサイエンスセンターと大阪大学サイバーメディアセンターとの間でこれらの課題を今後も継続的に議論していく予定となっている。

## 6. 進捗状況の自己評価と今後の展望

### (1) センター間ジョブ転送の仕組みの構築

図 1 で示したように、地理的に離れた複数の計算資源間でジョブ転送する仕組みを構築し、実運用もできたことから進捗状況としては良好であったと判断している。ただし、運用主体のことなる運用を考えるとジョブスケジューラの連携では運用上の課題を解決するのは難しいという結論になり、プライベートクラウド方式での実現方法の検討を始めている。

### (2) ジョブスケジューリングの検討

より高度なジョブスケジューリングに関してはいくつかの具体的な用例を設定して研究開発することができ、その成果発表も行うことができた。このことから、進捗状況と

しては良好であったと判断している。

### (3) 運用面での課題の明確化

本研究課題で実際の運用を念頭において課題を明確化するとともに、その解決に向けた運用方針についても具体的に検討してきた。研究計画時にはジョブスケジューラ間の連携によって転送キューを実現することを考え、その運用面での課題の明確化を考えていた。しかし、ジョブスケジューラ間の連携によって生じる運用上の課題を解決するよりは、双方にプライベートクラウドを構築して IaaS サービスとして提供の方が運用上の課題が少ないという結論に至り、当初予定とは異なる実装方式での課題検討を進めた。

## 7. 研究業績

### 国際会議プロシーディングス（査読あり）

- [1] [Yoichi Shimomura](#), Akihiro Musa, Yoshihiko Sato, Atsuhiko Konja, Guoqing Cui, Rei Aoyagi, [Keichi Takahashi](#), and [Hiroyuki Takizawa](#), “A Real-time Flood Inundation Prediction on SX-Aurora TSUBASA,” The 29th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC), pp. 192–197, Bengaluru, 2022.  
DOI: 10.1109/HiPC56025.2022.00035.
- [2] Kazuki Ide, [Keichi Takahashi](#), [Yoichi Shimomura](#), and [Hiroyuki Takizawa](#), “A task-parallel runtime for heterogeneous multi-node vector systems,” The 23rd International Conference on Parallel and Distributed Computing, Applications and Technologies, 2022.  
DOI: 10.1007/978-3-031-29927-8\_26.
- [3] Arata Endo, Shinji Yoshida, Shuichi Gojuki, Hiroaki Kataoka, Yoshihiko Sato, Akihiro Musa, [Susumu Date](#), “Consideration of a Supercomputing System with Cloud

Bursting Functionality from an Operational Perspective,” 13th IEEE International Conference on Cloud Computing Technology and Science (CloudCom2022), pp. 154-161, Bangkok, 2022.

DOI: 10.1109/CloudCom55334.2022.00031.

- [4] Fitra Rahmani Khasyah(+), Kadek Gemilang Santiyuda(+), Gabriel Kaunang(+), Faizal Makhrus(+), Muhammad Alfian Amrizal(+), and Hiroyuki Takizawa, “An Advantage Actor-Critic Deep Reinforcement Learning Method for Power Management in HPC Systems,” The 23rd International Conference on Parallel and Distributed Computing, Applications and Technologies, 2022.

DOI: 10.1007/978-3-031-29927-8\_8.

- [5] Muhammad Alfian Amrizal(+), Kadek Gemilang Santiyuda(+), Faizal Makhrus(+), Reza Pulungan(+), and Hiroyuki Takizawa, “SleepAgent: A Reinforcement Learning-Based Power Management System for Heterogeneous Clusters,” Research Poster Presentation at ISC23, 2023 (accepted).

#### 国際会議発表 (査読なし)

- [6] Hiroyuki Takizawa, Keichi Takahashi, Yoichi Shimomura, Ryusuke Egawa, Kenji Oizumi, Satoshi Ono, Takeshi Yamashita, and Atsuko Saito, “AOBA: The most powerful vector supercomputer in the world,” 34th Workshop on Sustained Simulation Performance, Sendai, 2022.
- [7] Susumu Date, “Osaka University’s Data Aggregation Infrastructure for Supporting Data-intensive Science,” 34th Workshop on Sustained Simulation Performance, Sendai, 2022.
- [8] Hiroyuki Takizawa, Yoichi Shimomura,

Akihiro Musa, Yoshihiko Sato, Atsuhiko Konja, Guoqing Cui, Rei Aoyagi, and Keichi Takahashi, “Real-time flood inundation simulation on SX-Aurora TSUBASA,” 35th Workshop on Sustained Simulation Performance, Stuttgart, 2023.

- [9] Hiroyuki Takizawa, “Job scheduling and resource management,” 2023 Conference on Advanced Topics and Auto Tuning in High-Performance Scientific Computing, Special Panel Discussion: Next supercomputer system requirement from Japan and Taiwan, Taoyuan, 2023.
- [10] Shinichiro Ogawa, and Daisuke Sasaki, “Development of Wall Injection Model for Solid Rocket Motor Internal Flow by Block-Structure Cartesian Mesh,” AIAA SCITECH 2023 Forum, Maryland & Online, 2023.

#### 国内会議発表 (査読なし)

- [11] 青柳嶺, 高橋慧智, 下村陽一, 滝沢寛之, “計算特性に着目した実行時間予測に基づくリアルタイム洪水シミュレーションの動的資源割当,” 第185回HPC研究会(SWoPP), 下関, 2022.
- [12] 石井翔, 高橋慧智, 下村陽一, 滝沢寛之, “機械学習に基づくジョブスケジューリングのためのGANによるデータ拡張,” 第185回HPC研究会(SWoPP), 下関, 2022.
- [13] 中井大貴, 大村竜義, 高橋慧智, 下村陽一, 滝沢寛之, “緊急ジョブの優先実行を考慮したジョブスケジューリングに関する一検討,” 第187回HPC研究会, 沖縄, 2022.
- [14] 滝沢寛之, “ジョブスケジューリングのデジタルツイン化と自動チューニング,” 第14回自動チューニング技術の現状と応用に関するシンポジウム(ATA2022), 東京, 2022.
- [15] 江川隆輔, “使える高性能計算機システムの実現に向けて,” 京都大学学術情報メデイ

アセンターセミナー「時代に合った HPC の活用」, 京都, 2023.

[16] 小川秦一郎, 大内健太郎, 佐々木大輔, “直交格子積み上げ法を用いたキャビティ付き固体燃料ロケット内部流れの流体解析手法の構築,” 第 40 回航空宇宙数値シミュレーション技術シンポジウム, 盛岡, 2022.

[17] 小川秦一郎, 佐々木大輔, “埋め込み境界法を適用した固体燃料ロケット推進剤壁面噴出モデルの構築, 第 36 回数値流体力学シンポジウム, オンライン, 2022.

#### その他 (特許, プレスリリース, 著書等)

[18] プレスリリース (2022年4月25日), “NEC、東北大学サイバーサイエンスセンターからベクトル型スーパーコンピュータによる大規模クラウドサービスを受注 ～東北大学 AOBA と連携した研究者支援環境を提供～,” [https://jpn.nec.com/press/202204/20220425\\_01.html](https://jpn.nec.com/press/202204/20220425_01.html).