

次世代計算基盤のユーザビリティに関する提言

令和7年6月 日

一般社団法人 HPCI コンソーシアム

目次

1. はじめに.....	3
2. 背景:次世代計算基盤の必要性和方向性.....	5
2.1 HPCIシステムのこれまでとこれから.....	5
2.2 次期フラッグシップや基盤センターを含む次世代計算基盤の果たすべき役割.....	6
2.3 国際的な動向.....	8
2.4 アンケートからみる HPCI の現状と課題.....	9
3. アクセラレータを念頭においたユーザビリティの向上.....	10
3.1 アクセラレータの活用状況と課題.....	10
3.2 HPCI 計算資源の多様性確保.....	11
3.3 ソフトウェアエコシステムと標準化.....	11
3.4 レガシーコードの効率的な GPU 移植.....	12
4. 次世代計算基盤の戦略的整備と運用.....	14
4.1 計算資源の効率的な配分とアクセス性向上.....	15
4.2 異なる資源間の連携と統一的 UI/UX.....	17
4.3 大学・研究機関との連携.....	18
5. 新規応用分野の開拓.....	19
5.1 生成 AI と AI for Science.....	19
5.2 量子コンピュータと HPC の統合的活用.....	21
6. 産業界との連携と人材育成.....	22
6.1 産業界との連携強化.....	22
6.2 人材育成とオープンサイエンスの推進.....	24
7. あとがき.....	25
附録.....	25
HPCI システムの今後の運営の在り方に関する調査検討ワーキンググループ委員リスト	26
検討の記録.....	26

1. はじめに

我が国における 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) は、世界トップクラスのスーパーコンピュータ (以下、スパコン) であるフラッグシップ計算機を頂点とし、大学情報基盤センター等が運用するスパコンや大規模共用ストレージを高速ネットワークで接続して、全ての計算資源へのシングルサインオン可能な共用計算環境インフラであり、平成 24 年(2012 年)9 月の運用開始以来、卓越した成果の創出を通じて、我が国の計算科学・計算機科学の振興と発展に貢献してきた。

HPCI の中心的な存在であるフラッグシップ計算機としては、初代「京」の後を受け、二代目である「富岳」の共用が令和 3 年(2021 年度)3 月に開始され、成果創出加速プログラムや政策対応利用課題、Society 5.0 推進利用課題をはじめとした戦略的な活用の成果として、コロナ対策の飛沫シミュレーションや豪雨防災のための数値気象予測の高度化、創薬 DX プラットホームの構築、国産大規模言語モデル(以下、LLM)の構築などの成果が創出されている。さらに、公募による一般・若手課題や産業課題でも様々な成果が創出されている。

HPCI を取り巻く我が国の状況としては、第 6 期科学技術・イノベーション基本計画が策定され、国が推進する長期的な科学技術政策として Society5.0 の実現、持続可能な社会への変革、研究力強化、教育・人材育成などが謳われており、これらの事業に対して「富岳」を中心とした HPCI を活用した計算科学が非常に重要な役割を担うことが期待されている。文部科学省次世代計算基盤検討部会の中間取りまとめ(令和 3 年(2021 年)8 月 27 日)では、ポスト「富岳」時代においては、次期フラッグシップシステム及び大学情報基盤センター等が運用するスパコンをはじめとする国内の主要な計算基盤、データ基盤、ネットワークが一体的に運用され、総体として持続的に機能する基盤となることが望ましいとされている。

上記の中間取りまとめを受けて、文部科学省において「次世代計算基盤に係る調査研究」(以下ではフィージビリティ・スタディの略称で FS)が令和 4 年度(2022 年度)から開始された。この調査研究は、ポスト「富岳」時代の次世代計算基盤の具体的な性能・機能等について、計算科学ロードマップを踏まえつつ、サイエンス・産業・社会のニーズを明確化し、それを実現可能なシステム等の選択肢を提案するためのものである。フラッグシップシステムのアーキテクチャ、システムソフトウェア・ライブラリ、アプリケーションを検討する2つのシステム調査研究チーム(以下ではシステム FS)、量子コンピューティングとスーパーコンピューティングの融合計算を行うための量子スーパーコンピューティングの実現可能性を評価する新計算原理調査研究チーム(以下では新計算原理 FS)、多様なシステムが有機的に結合した持続可能な次世代計算基盤の実現に向けた運用関連技術を調査する運用技術調査研究チーム(以下では運用技術 FS)で構成されている。

近年、生成 AI に係る技術革新などにより、研究開発に必要な計算資源の需要が急拡大するとともに、AI とシミュレーション、さらには自動実験やリアルタイムデータを組み合わせて科学研究分野で活用する AI for Science の重要性が指摘されるなど、求められる計算資源がこれまで以上に多様化している。このため、システムとアプリケーションの協調設計が今後ますます重要になり、その前提として計算基盤の利用側のニーズを具体的に示すロードマップが不可欠である。計算科学ロードマップは、将来計算科学が解決すべき社会的課題・期待される科学的ブレークスルーと、そのために必要となる計算機システム性能等を数年ごとにまとめたものであり、最新版が令和 5 年(2023 年)12 月に公開されており、さらに、各分野での AI for Science に関する取りまとめた内容を加えたものが令和 6 年(2024 年)6 月に公開されている (<https://cs-forum.github.io/roadmap-2023/>)。

他方、令和 4 年(2022 年)から稼働を開始した米国の Frontier やフィンランドの LUMI、令和 5 年(2023 年)から稼働を開始した米国の Aurora やイタリアの Leonardo、令和 6 年(2024 年)から稼働を開始した米国の El Capitan やスイスの Alps など、Top500 スパコンランキングの上位に名を連ねる米国、欧州のスーパーコンピュータにおいては、CPU に加えて GPU などの加速部(アクセラレータ)を有し、計算の一部を加速部で処理することで高速化を図ることが主流となっている。我が国においても、次期フラッグシップであるポスト「富岳」に加速部が導入される予定であり、基盤センターのシステムの多くは既に加速部を導入している。ただし、加速部を導入したシステムにおいては、そのユーザビリティが大きな課題となる。

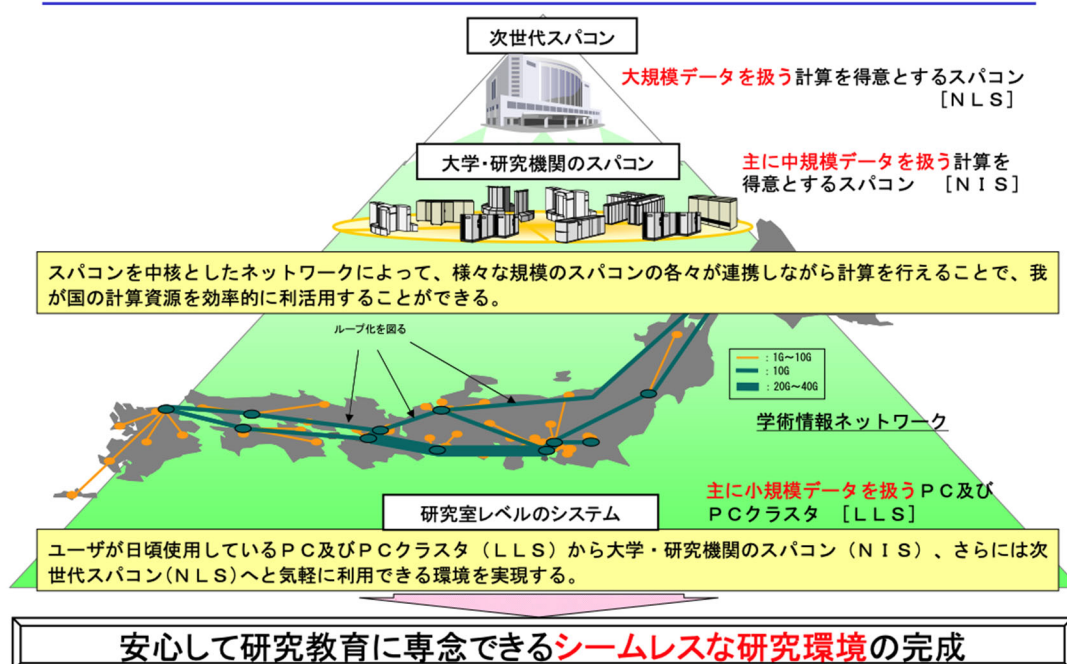
HPCI コンソーシアムは、我が国の計算科学技術振興の中心となり、世界最高水準の成果創出と成果の社会還元を目指して活動することを理念とし、その実現に向けて、計算科学技術に関わるコミュニティの幅広い意見を集約し、国への提言を行ってきた。本提言は、HPCI システムの今後の運営の在り方に関する調査検討ワーキング・グループ(以下、WG)において、次世代計算基盤運用技術調査研究チームの研究代表や「富岳」運用技術チームの代表など、前年度ヒアリング対象者となった方々や計算科学ロードマップの代表などを WG 委員として迎え、HPCI のユーザや HPCI システム構成機関(以下、システム構成機関)の意見を踏まえて、将来、次世代計算基盤を利用することになるユーザ、ならびにシステム構成機関としての立場として、ユーザビリティ向上のためにどのようなことが次期フラッグシップを含む次世代計算基盤の整備・運用の上で求められるかについての議論と提言をまとめたものである。

2. 背景:次世代計算基盤の必要性と方向性

2.1 HPCI システムのこれまでとこれから

平成 18 年(2006 年)に開始された「次世代スーパーコンピュータ」プロジェクトは計算科学技術を発展させ、広汎な分野の科学技術・学術研究及び産業における幅広い利用のための基盤を提供することにより、我が国の競争力強化に資するとともに、材料や医療をはじめとした多様な分野で社会に貢献する研究成果を挙げた。プロジェクト発足当時の資料を以下に示す。

科学技術・学術研究の基盤となるスパコンネットワークの構築



https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu2/006/shiryo/_icsFiles/afldfile/2015/02/12/1264462_002.pdf

この資料の中に「科学技術・学術研究の基盤となるスパコンネットワークの構築」というスライドがあり、次世代スパコンを「ナショナル・リーダーシップ・システム(NLS)」、基盤センターのシステムを「ナショナル・インフラストラクチャー・システム(NIS)」と位置付けている。「第二階層」という呼称はこのピラミッド構造の「一つ下の階層」にあることから生まれたものであるが、ポスト「富岳」時代に我々が目指すべきは下図のようなフラッグシップを中心として特徴あるスパコン、データ基盤、量子コンピュータ、大型実験施設等がフラットにつながり、共通のプラットフォームから利用できる次世代計算基盤の形である。



フラッグシップと基盤センターのスパコンを含む次期計算基盤の在り方を本提言で示していく上で「第二階層」という呼称をまず改める必要があると考える。具体的には、フラッグシップシステムを「ナショナル・フラッグシップ・システム(以下 NFS)」、基盤センターのシステムを「ナショナル・インフラストラクチャー・システム(以下 NIS)」と呼ぶことを提案する。

HPCI は「京」の共用が開始された平成 24 年(2012 年) 9 月からその運用が開始され、運用開始後、既に 13 年あまりの年月が経過している。令和 9 年(2027 年)4 月が特定先端大型研究施設の共用の促進に関する法律の利用者選定業務と利用者支援業務を実施する登録施設利用促進機関の切替えの時期となる。令和 8 年(2026 年)はその選定の公示、公募の時期となるため、令和 7 年(2025 年)から登録機関で選定される機関、利用促進業務を行う機関にどのような課題があるかを考える必要がある。HPCI の運営の改善に向けた検討は HPCI 計画推進委員会の「次世代計算基盤を見据えた今後の HPCI の運営に係る検討ワーキンググループ」で議論されることになるため、本提言では全体的な運営や支援の在り方を検討することとする。

2.2 次期フラッグシップや基盤センターを含む次世代計算基盤の果たすべき役割

令和 3 年(2021 年)8 月、科学技術・学術審議会情報委員会において、次世代計算基盤検討部会「中間取りまとめ」が報告され、次世代計算基盤のあり方やその必要性等について確認されるとともに、その後の検討は HPCI 計画推進委員会において進めることとなった。その後、HPCI 計画推進委員会において、「令和 3 年中間取りまとめ」以降の国内外の情勢変化を確認するとともに次世代計算基盤のあり方について更なる議論を進め、特に次期 NFS の性能や機能等について取りまとめを行った。令和 6 年(2024 年)6 月に HPCI 計画推進委員会から報告された「次世代計算基盤に関する報告書(最終取りまとめ)」には生成 AI の登場などにより、計算科学だけでなく科学技術全体、そして産業競争力等の観点からも、今後、計算基盤の重要性がさらに増し、求められる機能も多様化しつつ大きく移り変わっていくことが予想されることが述べられている。また、このような社会情勢の中であっても時代の要請に

応える計算基盤を常に提供していくために、次世代の NFS には以下のような役割が求められるとされている。

- 科学技術分野において AI を活用する「AI for Science」の取組を通じて、研究サイクルの飛躍的加速や研究探索空間の拡大といった科学研究の革新など、新たな時代を先導し、国際的に卓越した研究成果を創出していくこと
- このためには、計算速度の指標である TOP500 など単一の尺度を対象としたランキングのみを追求するのではなく、AI 性能をはじめ、あらゆる先端分野において世界最高水準の計算能力を提供すること
- また、10年など長期間にわたり同一のシステムで稼働し続けるのではなく、多様化・拡大を続ける需要の変化に柔軟に対応し、時代時代の要請に応じた十分な性能を常に提供し続けること
- 自国の技術を中心にスーパーコンピュータを開発・整備する能力を確保し、人材育成や産業競争力の維持・発展に資するコア技術を特定すること
- 利用が大きく拡大するとともに、その要素技術が世界の情報基盤に採用され、広く普及することで我が国の産業競争力や経済安全保障の強化に資すること

また、この役割を果たすための具体的な提案として、「遅くとも 2030 年頃の運転開始を目指し、電力性能の大幅向上により、既存の『富岳』ユーザに対しては実効性能として現行の 5～10 倍以上の計算能力を提供しつつ、AI 性能については運用開始時点で世界最高水準（実効性能として少なくとも 50 EFLOPS 以上）の利用環境を提供することを目指し、フラッグシップシステムの開発・整備を行うべき」と述べられている。ただし、このような AI 性能を実現するためには加速部の導入が必須であり、これまで「京」や「富岳」で開発されてきたアプリケーションコードを加速部に対応させることが重要な課題となる。「次世代計算基盤に関する報告書(最終取りまとめ)」でも

- アプリケーション開発においては、これまで発展してきたシミュレーションを中心とする計算科学をさらに発展させるため、加速部の導入などに際してもこれまで整備を行ってきたアプリケーションが安定して継続的に利用できるようにするとともに、必要に応じて改良を進める。あわせて、幅広いアプリケーションの利用を促進する。また、近年の動向を踏まえて、生成 AI の利用など、HPC の新たな領域を開拓することが期待されるアプリケーションの開発に取り組む。

と述べられており、これを実施する上で HPCI の果たすべき役割は大きい。「次世代計算基盤に関する報告書(最終取りまとめ)」の内容は次期 NFS に関するものが中心となっているが、NIS の在り方に関しても以下の項目が関連する。

- 計算資源の需要は多様化しており、TOP500 などの単一の尺度を対象としたランキングのみを追求するのではなく、国内の多様な計算資源需要を着実に満たすとともに、新たなシステムによる世界最高水準の利用環境の提供を通じ、我が

国の計算科学を中心とする様々な科学技術・産業分野の研究開発を先導することを期待。更に、科学技術分野における AI 基盤モデルの活用が促進されることを期待。

- フラッグシップシステムの整備においては、システムソフトウェア環境の相互互換性を高めるとともに水平展開し、また、開発されたアプリケーションが国内外のシステムに簡便に移植できることも、利用者の拡大の観点からは重要。

このような NFS と NIS の一体的な運用とアプリケーションソフトウェアのアクセラレータ対応を考える上で諸外国の取り組みは参考になる。

2.3 国際的な動向

近年の HPC を巡る国際動向を見ると、欧州と米国が先導する形でエクサスケール計算基盤の整備が進んでいる。欧州では 2018 年設立の EuroHPC 共同事業体が官民連携の下、パンヨーロッパなスーパーコンピュータ基盤の構築と HPC 技術開発を二本柱として推進している。各国の資金と EU 資金、さらには産業界のリソースをプールし、2020 年代前半に 8 台のプレ・エクサ級(ペタスケール上位)スーパーコンピュータを設置、さらにドイツの JUPITER やフランスの Jules Verne といった欧州初のエクサスケール機を導入した。この欧州の取り組みは、複数国が協調して超大規模機を共有運用する点や、欧州独自の HPC 技術の開発支援、HPC 人材育成など包括的なエコシステム構築を特徴としている。一方米国では、エネルギー省(DOE)が主導する Exascale Computing Project (以下 ECP)が 2010 年代後半から進行し、2022 年に世界初のエクサスケール機 Frontier がオークリッジ国立研究所に導入された。2024 年にはアルゴンヌ国立研究所の Aurora が 6 月の TOP500 でエクサスケールに到達し、ローレンス・リバモア国立研究所の El Capitan が 11 月の TOP500 で約 1.74EFLOPS を達成した。

加えて ECP ではハードウェアだけでなくソフトウェア開発・アプリケーションの性能最適化に巨額の投資を行い、広範な分野のコードを GPU アーキテクチャへ移植してきた。その結果、米国は GPU を活用したエクサスケール計算に必要な包括的ソフトウェアスタック Extreme-scale Scientific Software Stack (以下 E4S) を世界に先駆けて整備した。一方、日本は「富岳」で独自の ARM ベース CPU 路線を追求し大きな成果を上げたが、ポスト「富岳」世代では欧米の動向も踏まえた戦略転換(アクセラレータの本格採用)が避けられない状況である。欧米の事例が示すように、次世代 HPC の競争はハード性能のみならず、ソフトウェア・人材・産業・AI も含めた総合力の競争となる。我が国においても官民連携・国際連携を強化し、国家戦略として HPC 基盤の開発・運用に取り組む必要がある。

2.4 アンケートからみる HPCI の現状と課題

高度情報科学技術研究機構(RIST)では、「特定高速電子計算機施設の登録施設利用促進機関業務および HPCI の運営に関わる業務」の改善のために、終了した HPCI システム利用研究課題の課題代表者を対象に毎年アンケートを行っている。令和 5 年度(2023 年度)には課題代表者 259 名を対象に令和 6 年(2024 年)6 月 10 日～7 月 8 日にアンケートを実施し、151 名の方から回答を得た(回答率 58.3%)。調査内容は、実施研究課題・課題募集、課題の実施環境、利用者サービス、産業利用課題参加者の計算機利活用状況に関するものからなる。このアンケートを通じた HPCI 利用者への調査から、現行の運用に以下のような課題が浮き彫りになった。

- **ユーザビリティに関する課題:** 「富岳」以外の HPCI 資源を利用した課題のうち 40%が「システムの利用経験がある、または使い慣れている」を、計算機の選択理由として挙げている。これは裏を返せば、本来はより性能の高いシステムが別にあるにもかかわらず、使い慣れているシステムを選んでいる可能性があることを表している。NIS のシステム間の移行が容易になるよう、ユーザビリティを改善し、次年度のアンケートでは「アプリケーションの実効性能が高い」ことや「申請可能なノード時間積が大きい」ことなどが選択理由になることを目指すべきである。
- **新規利用者の獲得:** 各課題におけるスーパーコンピュータの利用経験が5年未満のメンバーの割合について「50%以上」が 39%(前回:36%)で増加しているものの、「10%未満」が 28%、「10%以上 25%未満」が 18%も占めており、スパコンの利用経験が少ないメンバーがほとんどいない課題が相当数あることがアンケートの結果から見て取れる。
- **産業利用の現状と課題:** HPCI の産業界による利用は徐々に増加してきたものの、依然として全体に占める割合は限定的であり、多くの企業にとって HPCI は十分活用されていない現状がある。「条件が整えば有償の課題への移行を検討する」と回答したユーザはここ数年 60%程度で推移しており、有償課題への移行を検討しているものの、実際に移行していないユーザがほとんどであることが示唆される。
- **マシンの性能に関する意見:** HPCI で利用したアプリケーションソフトウェアについては、「富岳」利用課題の8割が、また「富岳」以外の HPCI 資源の利用課題の 9 割程度が、申請時に期待した実行性能が得られたと回答している。ただし、この数字には「富岳」政策対応課題で行われた Fugaku-LLM の学習において GPU と比べて 1/100 以下の学習速度しか達成できなかったことなどが含まれておらず、AI 分野の重要性と「富岳」におけるその性能を反映できていない結果となっている。

このような背景に鑑み、HPCI システムの今後の運営の在り方に関する調査検討ワーキング・グループでは次世代計算基盤のユーザビリティに関して以下の通り提言する。

3. アクセラレータを念頭においたユーザビリティの向上

提言

- 戦略分野向け GPU 最適化旗艦プロジェクト創設: 2025 年度に実施
- GPU 利用共通教材の作成: 2026 年度末までに完了
- 基盤ソフトウェア開発プロジェクト創設: 2025 年度に創設
- 生成 AI によるレガシーコードの GPU 移行促進: 2029 年までに完了
- GPU 移行ベストプラクティス共有の旗艦プロジェクト創設: 2025 年度に創設

3.1 アクセラレータの活用状況と課題

近年の最先端スパコンの多くは、計算加速用の GPU やアクセラレータを統合した異種混在型アーキテクチャを採用している。実際、2024 年 6 月時点の TOP500 上位を見ると、1 位の米国 Frontier は AMD 製 GPU (MI250X) を搭載、2 位の Aurora は Intel 製 GPU (Ponte Vecchio)、3 位にはクラウド上の GPU クラスタ (Microsoft Azure 上の NDv5、NVIDIA H100 搭載) がランクインしており、純 CPU 型で上位に残るのは 4 位の「富岳」のみとなっている。ただし、国内の NIS の計算資源においてはアクセラレータの導入は欧米に先行しており、2008 年 12 月にまず東京工業大学が TSUBAME1.2 システムに GPU (NVIDIA S1070 170 台) を導入し、2012 年 3 月には筑波大学が HA-PACS システムに GPU (NVIDIA M2090 1072 台) を導入し、2012 年 4 月には京都大学が Laurel システムに GPU (NVIDIA M2090 64 台) を導入し、2013 年 8 月には九州大学が当時の高性能演算サーバシステムに GPU (NVIDIA K20 240 台/K20X 16 台) を導入し、2017 年 3 月には東京大学が Reedbush-H システムに GPU (NVIDIA P100 240 台) を導入し、同年 12 月には大阪大学が OCTOPUS システムに初めて GPU (NVIDIA P100 148 台) を導入し、2020 年 7 月には名古屋大学の「不老」システムに初めて GPU (NVIDIA V100 884 台) を導入している。

アクセラレータの導入に際してはアプリケーションコードの移植が必要になるが、TSUBAME に GPU が導入された当初は 2009 年 9 月から年に 3 回のペースで CUDA や OpenACC の講習会を実施し、ユーザの GPU 移行を促した。最近の例だと、JCAHPC の Miyabi 導入に際して GPU への円滑な移行のため、東京大学ではベンダーと協力した事前のプレ評価で GPU 機種を選定し、ユーザに十分な準備期間を提供するなどの工夫を行っている。また、主要コミュニティコード 19 種については外部資金で移植作業 (GPU 対応) を実施する「サポート移植」制度を立ち上げ、利用者

自らも情報共有しながら対応できるよう GPU 移行ポータルサイトを公開している。海外では米国の ECP や欧州の HPC Centres of Excellence(以下 HPC CoE)により主要アプリケーションの GPU 対応が体系的に進められており、日本もそれら国際的な知見を取り入れつつ、自国コミュニティのソフトウェア資産の近代化を急がねばならない。特に、気候・創薬・材料など戦略分野のアプリケーションについては、優先的に GPU 最適化・チューニングを支援する旗艦プロジェクトを設定し、メーカーや専門家チームとの協働で対応を進めることを提案する。

3.2 HPCI 計算資源の多様性確保

GPU の重要性が増す一方で、次世代計算基盤では計算資源の多様性も確保すべきである。全ての計算問題が GPU で効率的に解けるわけではなく、アルゴリズムや用途によっては他のアーキテクチャが有効な場合もある。HPCI を構成する大学計算センターでは現在、GPU クラスタに加え、ベクトル計算機(NEC SX Aurora TSUBASA)や FPGA 搭載機、大規模共有メモリ機など多様なシステムが供用されている。このような多様な技術要素の試行と実用化は、次期フラッグシップの設計にもフィードバックされ得る重要な財産である。将来的には、量子コンピュータとのハイブリッド計算も視野に入れる必要がある。量子計算は特定の問題で古典計算を凌駕する可能性を秘めており、欧米では既に量子計算機と HPC を統合して利用する研究が始まっている。日本でも理研が「量子-HPC ハイブリッド」基盤の研究開発を推進しており、「量子計算を実用化するためには HPC との統合が不可欠」との認識の下でソフトウェア基盤構築に取り組んでいる。HPCI としても、GPU に限らず多様な計算資源を柔軟に組み合わせて利用できる環境を整備すべきである。具体的には、ユーザが計算課題に応じて最適な資源(GPU、ベクトル、FPGA、量子シミュレータ等)を選択・併用できるように、ジョブスケジューラやミドルウェアを拡張し、異種計算ノード間のデータ連携をシームレスに行える仕組みを検討する。また、新技術の試行の場として、NIS の基盤センターにプロトタイプ的な先端機を配置し、それを HPCI 経由でコミュニティが利用できるようにするといった工夫も考えられる。多様性確保はリスクヘッジと競争力維持の両面で重要であり、一極集中型ではない持続可能な計算基盤の構築につながる。

3.3 ソフトウェアエコシステムと標準化

アクセラレータ活用と計算資源の多様化を支える鍵が、ソフトウェアエコシステムの整備と標準化である。異なるアーキテクチャ間でアプリケーションを円滑に動作させるには、ポータブルで高性能なプログラミング環境やライブラリ群が不可欠である。米国 ECP では E4S と呼ばれる包括的なソフトウェアスタックを構築し、GPU・CPU いずれのアーキテクチャでも動作する科学技術計算ツール・ライブラリをオープンソースで提供している。E4S は HPC と AI の双方で用いられる主要ライブラリを含み、AWS や

Google Cloud 上でも同様に使えるよう設計されており、クラウドとの親和性も高い。Linux Foundation は 2023 年 11 月に SC23 で、HPC 向けソフトウェアの新基金 HPSF の設立計画が初公表され、2024 年 5 月の ISC24 にて正式な設立が発表された。ECP で開発された Spack (HPC パッケージマネージャ) や Kokkos (性能可搬な C++ 向け並列プログラミングモデル) を含む 10 の HPC 関連プロジェクトが HPSF 傘下となり、知的財産や商標の寄贈・移管が実施されている。

日本でも HPCI の各システムで共通に利用できるソフトウェアスタックの標準化を進めるべきである。現在は各センターごとにモジュール環境や提供ライブラリが異なる場合が多いが、HPCI ポータルからコンテナイメージを配布しどの計算機上でも同一環境を再現できるようにする、あるいは Spack や E4S を参考に主要ソフト群のビルド・配布を統一するといった取り組みが考えられる。また、プログラミングモデルにおいても、CUDA や SYCL、Kokkos などオープンな標準技術に基づく実装を促進し、特定ベンダー依存のプロプライエタリなモデルに頼らない将来にわたる可搬性を確保することが重要である。ソフトウェアの標準化・高度化には継続的な投資が必要である。幸い国内には、数値計算ライブラリやプログラミング言語処理系の開発で世界をリードしてきた専門家が多数存在する。産学官連携の下、HPCI を舞台にこれら人材を集めた次世代 HPC ソフトウェア開発プロジェクトを立ち上げ、基盤ソフトの開発・メンテナンス・人材育成の循環を作ることを提案する。具体的には、ECP における xSDK や欧州の HPC CoE にならい、分野横断的な数値計算ライブラリ群やツールチェーンを整備・統合するほか、機械学習フレームワークの HPC 対応を進める。また、前述のように Linux Foundation の HPSF に日本からも参加し、国際連携を通してソフトウェア開発を行うべきである。東大は既に HPSF に参画しており、理研は準備中である。これにより、HPC 利用者が新旧様々な計算資源を意識せず一貫した開発環境で研究開発できるようになり、HPCI 全体のユーザビリティが飛躍的に向上することが期待される。

3.4 レガシーコードの効率的な GPU 移植

上述の GPU 移植に関する教材が整備できたとしても、膨大な量の科学技術計算のレガシーコードを CUDA や SYCL、Kokkos などに書き換えるのには多大な人的コストを要する。自助努力で事業として組織的に継続できているものもあるが、多くの分野では草の根的な活動としてボランティアベースで支えられ細々と活動している状況である。また、GPU 移植に携わった研究者のキャリアアップをどのように実現するかを考えることが喫緊の課題となる。この際に、どういったスキルセットがどういった企業で必要になるのかを明確にする必要がある。米国 ECP の期間終了後には国立研究所などから NVIDIA に大量に人が流れたが、国内にはこのような受け皿がない。そこ

で、人手をかけずにレガシーコードの GPU 移植を行うための革新的技術が切望される。

HPCI コンソーシアムの人材育成タスクフォースを中心に GPU 利用の共通教材を作成することを提案する。また、この際に NIS の各基盤センターでこれまで蓄積されてきた教材を積極的に取り込むことで教材作成のコストを抑えるとともに、国内の現有資産を集約することで教材の質の向上を図ることができると考えられる。

Fortran を CUDA に自動変換するツールに関する初期の試みとしては F2C-ACC が挙げられる。これは 2010 年から 2012 年にかけて NOAA の Govett らによって開発されたもので、気象分野の Fortran コード NIM を CUDA に書き換えるのに利用された。最近の試みとしては AMD が開発した GPUFORT などがある。GPUFORT は Fortran コードに OpenMP4.5 のディレクティブを自動的に挿入したり、Fortran コードの一部または全てを HIP と C++ のコードに自動変換することができる。Kokkos にも Fortran Language Compatibility Layer (FLCL) があり、Fortran と Kokkos (C++) のコードが混在することを可能にし、デバッグしながら少しずつコードを移植することを可能にする。FLCL は Fortran に合わせた列優先のデータ構造を C++ で実現することもでき、データ構造も徐々に変換することを可能にしている。ただし、このような自動変換ツールは内在するアルゴリズムやデータ構造を大幅に変えることまではできず、元の Fortran コードが高い並列度を既に有している必要がある。

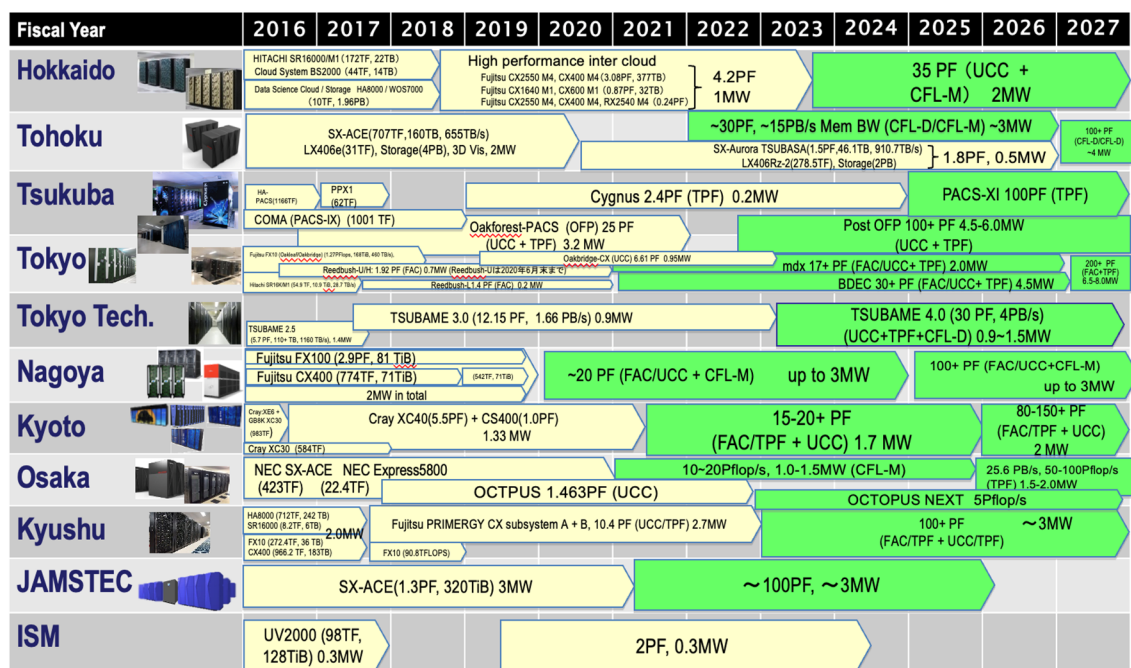
他方、生成 AI のコード生成能力は目覚ましい進歩を遂げており、米国 DOE が主導する S4PST の ChatHPC、アルゴンヌ国立研究所の Code-Scribe、ローレンス・リバモア国立研究所の Fortran2CPP、Google DeepMind の CodeRosetta などは LLM を用いて Fortran コードを CUDA/C++ に自動変換することに成功している。OpenAI の o1 や o3、Google の Gemini 2.0 Flash Thinking、DeepSeek の R1 などの論理推論に強いモデルでは、言語生成よりもコード生成能力のほうに大幅に向上しており、今後 LLM を用いたコード変換の性能は大幅に向上することが予想される。このようなツールを最大限に活用しつつ、レガシーコードを積極的に GPU 移植していくことを提言する。この際に、生成 AI を活用してレガシーコードを書き換えるのはそのコードの開発者を想定しているが、ベストプラクティスやノウハウを収集して開発者に提供するの HPCI に旗艦プロジェクトを作ってやることを提言する。

4. 次世代計算基盤の戦略的整備と運用

提言

- 戦略プログラム枠の刷新: 2026 年度の課題から刷新
- 産学連携・企業利用の促進: 2026 年度に新規枠の開始
- リソースオンデマンド供給: 2026 年度までにオンデマンド枠を提供開始
- コンテナの共通化: 2025 年度にコンテナ環境を運用開始
- UI/UX の改善とポータル統一化: 2026 年度にポータル本格稼働
- データ基盤との統合: 2027 年までに主要データ基盤との連携運用を開始
- NIS センターの機能分担と協調: 2026 年度に役割分担の枠組みを導入
- 広域連携プロジェクト創出と人材交流促進: 2027 年度に制度開始

次世代計算基盤の戦略的整備を実施する上で、NIS のセンター間で導入するシステムの調整を行い、システムの多様性を確保することが重要である。各基盤センターの導入予定のアーキテクチャのマップはこれまで東京大学情報基盤センターを中心に整備が行われてきた。以下に基盤センターに導入予定のシステムの最新のアーキテクチャマップを示す。



このようなアーキテクチャのロードマップの整備は本来であれば、HPCI 連携サービス委員会が担うことが望ましく、今後は連携サービス委員会の活動の一環として NIS の

アーキテクチャマップの整備を行うことを提言する。また、運用 FS 活動も連携サービス委員会に引き継ぎ、継続的な運用改善を行う仕組みを作ることが望ましい。

4.1 計算資源の効率的な配分とアクセス性向上

HPCI 全体として計算資源を最大限に活用するには、効率的な資源配分とアクセス性の向上が重要である。現在、HPCI では公募により計算時間を割り当てる仕組み（課題公募制度）が運用されているが、今後は利用目的に応じた柔軟な割当制度を拡充すべきである。提言として、以下のような施策を挙げる：

- **戦略プログラム枠の創設：**「京」のグランドチャレンジプログラムおよび戦略プログラム、「富岳」の重点課題および成果創出加速プログラムに続き、今後も国の重点分野（気候変動対応、創薬、AI for Science 等）については長期的大規模計算を支援する戦略プログラム枠を設け、安定的かつ集中的に資源提供する。これにより、社会的インパクトの大きいプロジェクトに HPC をテコ入れし、成果の最大化を図る。ただし、生成 AI や量子コンピュータなどの新しい潮流に対応できるよう対象とする分野を刷新し、新しい分野のコミュニティ育成・分野振興を進めるべきである。
- **産学連携・企業利用の促進：**産業利用については従来の公募型（課金あり）に加え、大学との共同研究や業界コンソーシアムによる利用を促進する特別枠を設定する。「富岳」だけなら既にあるが、HPCI にはこのような制度はない。例えば産業界が共同で出資し課題解決に HPC を使う場合にリソースを優先配分する制度など、企業が参画しやすいモデルを構築する。あわせて、スタートアップや中小企業向けには低コストで試行利用できるサンドボックス的環境も提供し、HPC 利活用のすそ野を広げる。
- **リソースオンデマンド供給：**「富岳」のファーストタッチオプションを HPCI 全体に拡張し、従来の年度単位配分に加えて、計算資源の一部をオンデマンド型（随時申し込みで短期間利用）に開放し、小規模なニーズや突発的な計算需要に応える。特に観測データのリアルタイム解析や災害シミュレーション等、即応計算を要する用途のために、迅速な審査とスケジューリングが可能な運用体制を整備する。
- **利用実績の少ない課題に対する対策：**現在、「富岳」以外の NIS の資源提供期間へ支払われる利用負担金はセンターのポリシーが従量制であった場合は定額制とするのは難しい場合もあるため、センターごとに定額制と従量制に分かれている。従量制の場合は実際の利用量に応じた支払いとなるのは当然のことながら、定額制の場合も当該課題の利用実績がゼロの場合に利用負担金は支払われない。しかし、利用実績がゼロの利用者が発生したことにより資源提供機関に利用負担金が支払われないという不利益が生じることは望ましくない。このような

事例の防止策として、HPCI 一般課題で利用実績がゼロの課題、年度末にのみ少量の利用実績がある課題に対しては、次年度の申請時にそのことを考慮した資源配分を行うことを提案する。制度上はこのような対応が可能にも関わらず、これまでは申請時の資源配分の際のオーバーブッキングなどで対応しており、利用実績ゼロ課題に対して実質的にペナルティが課されていなかったことを是正することが狙いである。

- **ユーザーポータルの一元化とシングルサインオン:** アクセス性の向上として、ユーザーポータルの一元化とシングルサインオンを実現する。現在、各センターごとにユーザ登録や認証が必要な場合があるが、将来的には学術認証フェデレーション(学認)を活用して HPCI 全参加機関を横断するシングルサインオンを完備し、利用者が一度の認証で全ての資源にアクセスできるようにする。さらにジョブ管理やストレージも統合的に扱える統合ポータルサイトを構築し、ユーザはブラウザ上で自分の利用状況確認やジョブ投入、データ移動ができるようにする。こうしたシームレス化によって、複数の計算機資源をあたかも一つの巨大クラスタのように扱えるユーザ体験を提供し、利便性を商用クラウド並みに高めることが可能となる。
 - **データアクセスのシームレス化:** HPCI 外部の実験施設や IoT センサ等で得られる大量データを計算に取り込むニーズが高まっており、計算機と遠隔データ源を高速ネットワーク(SINET 等)で直結してリアルタイムにデータ転送できる仕組みを拡充する。例えば東京大学の Aquarius システムでは一部ノードを SINET 経由で外部接続しリアルタイムデータを取得可能にしている。今後は他のセンターでも同様の試みを進め、「計算とデータの垣根」を低くするインフラを目指す。
 - **NIS システムの共通ベンチマーク:** 2017 年 5 月 24 日の HPCI コンソーシアムによる「今後の HPCI 第二階層計算資源の整備とその活用に関する提言」を受け、HPCI 連携サービス委員会(2017/11/29)でベンチマークテスト(BMT)の実施について意見交換が行われた。RIST は 2017-2018 年にかけて検討を行い、委員会の意見を聴取しベンチマーク試験・評価結果を報告した。一方、公開については連携サービス委員会での議論結果、継続審議することとなった。2023 年度の HPCI 連携サービス委員会では課題の洗い出しを行い、検討方針や体制を議論した。その結果を HPCI コンソーシアムで再度議論して今後の方向性・検討内容を定めることとなった。各センターも含めて全体として検討する必要があるため、ベンチマークに関しては次年度の WG の課題として引き続き検討する。
- これら運用上の工夫は、限られた計算資源であっても最大限の成果を引き出すことにつながり、ひいては HPCI 全体の生産性向上に寄与する。

4.2 異なる資源間の連携と統一的 UI/UX

近年、クラウドサービス事業者による HPC サービス(HPC on Cloud)が台頭しており、オンデマンドで大規模 GPU 計算資源を提供するケースが増えている。実際、前述の Microsoft Azure のクラウド GPU システムが TOP500 で 3 位に入るなど、クラウド上で従来の国立スーパーコンピュータに匹敵する計算が可能になってきている。HPCI としても、このクラウド HPC と競合ではなく補完関係を築く戦略が求められる。具体的には以下のような連携策が考えられる:

- **コンテナの共通化:** HPCI 側・クラウド側双方に共通のコンテナ型仮想環境を用意し、ユーザが同一のコンテナを持ち込んでどちらでも動かせるようにする。これにより環境差異を意識せず、必要に応じ計算場所だけを切り替えて利用できる。また、クラウド上で構築した AI モデルや分析パイプラインを HPCI 上に持ち込む(あるいはその逆)ことも容易になる。
- **UI/UX の改善とポータルの一統化:** 複数の計算機センターに跨る HPCI の利用インタフェースを統合し、統一的で直感的なユーザーエクスペリエンスを提供する。具体的には、ウェブブラウザ経由で各計算資源にアクセス・ジョブ管理ができるポータルを構築し、従来のコマンドライン中心の操作に不慣れなユーザでも扱いやすい環境を用意する。現在、「富岳」への Open OnDemand 導入など対話的利用環境整備が進みつつあるが、ユーザ自身が自由に環境構築できるわけではないなど制約も残っている。今後、コンテナ技術の活用等により利用者ごとのカスタム環境構築を可能にするなど、一層の UI/UX 改善を図る。
- **データ基盤との統合:** クラウドはデータ格納・分析基盤としても優れるため、大規模データを扱う HPC ワークロードではクラウドストレージと HPCI 計算資源を組み合わせるのが効果的な場合がある。HPCI とクラウド間で高速かつセキュアにデータ連携できる仕組みを整え、ユーザが意識せずに最適配置のストレージにアクセスできるようにする。具体例として、学術分野向けデータ基盤 mdx との連携が挙げられる。mdx は国内の仮想化基盤であり、HPCI の計算結果データを mdx に蓄積・公開したり、逆に mdx 上のオープンデータセットを HPCI 計算に利用したりといったシナジーが期待できる。

このためには、システム構成機関、データプラットフォーム、研究データ基盤、商用クラウド等の認証の連携、複数のスパコンのストレージや共用ストレージの間でのデータ連携、リアルタイムなデータの取得を含めた HPCI 外部データへの自由かつセキュアなアクセスが求められる。これに向けて、「富岳」のクラウド的利用に向けた共同研究プロジェクト民間サービスプロバイダを活用したクラウド的利用に関する理研 R-CCS での共同研究プロジェクトが 2020 年から 2 年間実施され、ユーザ100名以上、民間企業 35 社以上、「富岳」利用資源 400 万ノード時間以上と、その利便性の実証がなされている。その後、「富岳」を円滑に利用するための付加サービスとし

て、サービスプロバイダによる「富岳」のクラウド的利用が開始されている。支援を担うサービスプロバイダは増加しつつあるが、ユーザやサービスプロバイダの要望やニーズを踏まえ、利用拡大を図るべきである。

4.3 大学・研究機関との連携

HPCIを強化する上で、日本全国の大学・研究機関との連携も引き続き重要である。HPCIは元来、9大学と理研・NIIの協力により構築された経緯があり、各大学が計算センターを拠点に人材育成や地域の計算科学支援を担ってきた。今後、このネットワークを活かしつつ国内HPCコミュニティ全体の底上げを図るべきである。具体的提案として、国内の大学計算センター・研究所をいくつかの機能別に分類し、それぞれCenter of Excellence的な役割を与えることが考えられる。例えば、あるセンターは気候・環境分野のHPC研究拠点、他のセンターはAIとデータ科学の融合拠点、といった具合に専門性を打ち出し、人材・ソフトウェアの集積を促す。欧州ではEuroHPCの下で各国に「国民コンピテンスセンター(以下NCC)」を設置し、教育訓練や産業支援のハブとするEuroCCプロジェクトを展開した。33か国のNCCが国内のHPCリソース・スキルをマッピングし不足を分析、産業界や行政を含む多様なユーザへの窓口となることで、欧州全体のHPC利用拡大に貢献している。日本でも、HPCI参加機関それぞれが専門とする分野や得意技術を明確化し、国内のHPC相談窓口やトレーニング提供者として機能するよう支援するのが望ましい。例えば、「量子計算×HPC」の相談は理研R-CCS、「研究データ管理基盤」はNII、「分子シミュレーション応用」は特定大学、といったように役割分担と連携を進めることで、ユーザは目的に応じた支援を受けやすくなる。加えて、大学間・機関連携を促す資金制度も有効である。欧州のHPC研究プログラムでは、複数国の大学・企業が連携するコンソーシアム形式で公募課題に応募することが一般的であり、自然と広域連携が生まれている。日本でも、複数大学・企業の連名によるHPC共同研究課題に対し重点的に計算資源や助成金を配分する仕組みを作ることで、オールジャパンでのHPC活用プロジェクトを創出できる。人材の流動性促進も重要であり、計算科学やHPC技術の専門人材が大学・国研・企業の間を行き来しやすいように、クロスアポイントメント制度の活用や、大型プロジェクトへの産学混成チーム編成を推進する。以上のように、HPCIをハブとして国内の研究リソース・知見を結集すれば、単一機関では成し得ないような革新的成果の創出につながると期待できる。また、これを実現するために統合的な枠組みを実現するべきであり、計算資源の全体最適化を行うべきである。

欧州のEuroHPC研究プログラムには、多数のHPCアプリケーション高度化プロジェクトや、人材育成プログラム(例:EUMaster4HPC)が含まれており、日本にとっても参考になる。例えばEUMaster4HPCでは大学・スーパーコンピューティングセンター・企業が連携し欧州全域で統一カリキュラムのHPC大学院課程を創設している。日本

でも産学連携の高度人材育成スキームを検討すべきである。これについては後述の人材育成章で詳述する。

5. 新規応用分野の開拓

提言

- 学際・国際連携の推進: 2025 年度に学際・国際連携の枠組みを整備
- 小型で論理推論能力の高い AI の開発: 2025 年度に開発を完了
- 量子-HPC ハイブリッド環境の整備: 2028 年度までにプロトタイプ環境を構築
- 量子 x HPC 人材育成と交流促進: 2025 年度に初回イベント開始

5.1 生成 AI と AI for Science

2024 年ノーベル物理学賞は、「人工ニューラルネットワークによる機械学習を可能にした基礎的発見と発明」に対する業績により John Hopfield (プリンストン大学) と Geoffrey Hinton (トロント大学) が受賞した。また、2024 年ノーベル化学賞は、「計算によるタンパク質の構造予測及び設計」に対する業績により David Baker (ワシントン大学)、Demis Hassabis (DeepMind)、John Jumper (DeepMind) が受賞した。AI 分野の基礎研究がノーベル物理学賞、応用研究がノーベル化学賞を受賞したことは、AI のインパクトが情報分野にとどまらず科学全体に波及したことを表している。AI 技術の進歩はますます加速しており、ノーベル化学賞の受賞対象となった AlphaFold のような技術が創薬以外の科学分野においても今後次々と生まれることが予想される。

高速ネットワークで相互接続されたアクセラレータと大容量ストレージが具備された NFS と NIS を含む HPCI の計算資源は、我が国が AI やデータ科学の分野においても国際的なリーダーシップを発揮するためにも必須の研究基盤となることが期待されている。このことに関連して「特定高速電子計算機施設の共用の促進に関する基本的な方針」(平成 23 年(2011 年)文部科学省告示第 120 号(令和元年(2019 年)9 月 17 日改正))においても、フラッグシップ計算機に関して、「従来の研究開発に加え、人工知能、ビッグデータに関する研究開発を実施するための不可欠な計算基盤として活用され、それを通じて Society5.0 の実現に大きく貢献する。」と言及されている。

このような状況の中、令和 6 年度(2024 年度)の HPCI 一般課題において「AI やデータサイエンスを活用して、科学的・社会的課題の解決に資する研究開発」を重点分野として設定し、多くの計算資源が AI for Science に向けた取り組みに割り当てられた。HPCI コンソーシアム 計算科学フォーラムがまとめる令和 6 年度(2024 年度)版の計算科学ロードマップにも AI for Science 章が追加された。AI for Science のロードマップは素粒子・原子核、ナノサイエンス・デバイス、エネルギー・材料、生命科学、脳・神経科学、創薬・医療、設計・製造、社会科学、地震・津波、気象・気候、宇宙・天

文の 11 分野の専門家が AI をどのように活用しているかがまとめられている。これらの先進的な取り組みから国内の AI for Science における課題がいくつか浮かび上がっている。

- **「富岳」に GPU が搭載されていない:** 「富岳」は CPU(富士通 A64FX)のみで GPU を搭載していないシステムであり、AI の計算に向いていない。AI 向けに半精度型も用意されているが、加算を単精度で行う GPU の Tensor Core などとは異なり加算も半精度で行うため AI の学習にこれは利用できない。A64FX の単精度の理論ピーク性能は 6.8TFLOPS であり、現在 AI の学習に多く用いられている NVIDIA H100 の半精度の理論ピーク性能 989TFLOPS と比べて約 145 倍遅い。昨今、中国の DeepSeek 社が劣化版の GPU で安価に AI が学習できることを示したと話題であるが、彼らの用いた NVIDIA H800 の半精度の理論ピーク性能は 756TFLOPS であり H100 の 1.3 倍程度しか遅くない。AI for Science を含めたあらゆる AI の研究を遂行する上で、日本のフラッグシップシステムに GPU が搭載されていないことは致命的である。
- **NIS における GPU 資源の不足:** NIS のシステムの多くには GPU が導入されているが「富岳」の 158,976 ノードと比べると NIS の中では比較的規模の大きい Miyabi (JCAHPC)でも 1,120 ノード、ABCI3.0 でも 766 ノードであり、圧倒的に規模が小さい。このような規模の NIS の GPU 資源で本来フラッグシップが提供するべき AI for Science のための GPU 資源を補うことはできない。
- **LLM のスケール則:** LLM はモデルの規模や学習するデータ量を大きくするだけで性能がべき乗則に従って向上し、計算資源への投資に対する効果が保証されていることが知られている。米国では OpenAI や大手 IT 企業主導で、「スターゲート(Stargate)」計画と呼ばれる巨額投資が行われており、総額 5,000 億ドル(約 70 兆~80 兆円)を投じ、全米各地に巨大データセンターを建設して AI モデル開発を支えるインフラを構築する計画である。このような計算資源で開発される先端の AI 技術に国内の計算資源だけで対抗することは困難である。

ただし、生成 AI 技術は目覚ましい勢いで進歩しており、DeepSeek の例に見られるような大幅なコスト削減がこれまでも何度も起きている。また、OpenAI などの一部の企業が持っているとされる独自技術は次々と他の研究チームによって再現されており、技術の格差は縮まっていることが示唆される。このような状況において、生成 AI 及び AI for Science を新規応用分野として取り込む上で以下の内容を提言する。

- **学際・国際連携の推進:** AI for Science の発展のためには、科学の各分野の専門家が最新の AI 技術を取り込むことが不可欠である。このためには、AI 分野の研究者、HPC の研究者、各専門分野の研究者による学際連携を促すような制度設計が必要である。また、学際連携の橋渡し役となる複数分野に跨った専門性を有する人材育成のために大学院教育においても計算科学×AI×各専門領域

のカリキュラムを整備する。また、世界各国で国際連携の機運が高まっており、国境を超えた AI for Science のためのコンソーシアムが立ちがっている。アルゴンヌ国立研究所が主催する Trillion Parameter Consortium (以下、TPC)は世界各国の AI for Science の研究者が参画しており、日本からも理研、産総研、東大、科学大などが参画している。このように、学問分野や国境を超えてデータ・ノウハウ・人材をやりとりできる仕組みを確立することが喫緊の課題となっている。

- **小型で論理推論能力の高い AI の開発:** AI for Science において AI に求められるのは論理推論能力である。昨今の OpenAI の o1 や o3、Google の Gemini 2.0 Flash Thinking、DeepSeek の R1 などに代表される推論時にしばらく考えてから出力するタイプのモデルは論理推論能力が高く、特に数学やコーディングのベンチマークで gpt-4o などと比べても格段に高い性能を達成している。ただし、一定の論理推論能力が獲得できれば知識は Retrieval Augmented Generation (以下、RAG)などを用いて Web や社内データベース上からデータをとってくるができるため、膨大な知識を学習させる必要はない。現在、OpenAI、Google、DeepSeek は 1 兆パラメータ級の大規模なモデルに 10 兆トークンの膨大なデータを学習させたものに強化学習を用いて論理推論能力の高いモデルを作っているが、今後小型のモデルで同程度の論理推論能力を有するモデルが次々と開発されることが予想される。小型で論理推論能力の高い AI が登場すれば、国内の計算資源でも十分競争力のある AI for Science の研究が可能になる。これを見越して、各分野の専門家は o1 などのモデルで自分の研究がどのように加速できるかを試し、計算科学の専門家は推論のさらなる高速化を実現するためのシステム・ソフトウェアを開発、AI の専門家は小型で論理推論能力の高い AI の開発を優先的に行うべきである。また、HPCI としてそれを促す仕組みを構築することを提言する。

5.2 量子コンピュータと HPC の統合的活用

量子コンピュータは将来的に HPC と並ぶ計算プラットフォームとなりうる技術であり、その動向にも注視が必要である。現時点では量子ビット数の制約から限定的な応用に留まるが、量子計算が特定分野で指数関数的な性能向上をもたらす可能性は高い。米国や欧州では、HPC センターに量子計算機を設置し既存のスーパーコンピュータと接続する試みが進んでいる。例えばドイツ・ユーリヒ研究所では超伝導量子計算機をスーパーコンピュータに繋ぎ、ハイブリッド計算を実証しているとの報告がある。日本においても、理研 R-GCS に「量子-HPC 連携プラットフォーム部門」が設立され、ハードウェア接続やソフトウェア統合の研究が始まっている。この状況を踏まえ、量子計算とのハイブリッド HPC 環境を先取り整備することを提言する。具体的には、まず量子計算機(または量子アニーラ)と古典 HPC を接続するネットワーク・インタフ

エース標準を検討し、一部の HPCI センターで実験的に量子計算リソースを導入して
みることが考えられる。ユーザが一つのジョブで量子回路演算と通常の数値演算を
組み合わせて実行できるよう、ジョブスケジューラやプログラミング環境を拡張する。
理研では「実用的な量子計算には HPC との統合が必要」と述べており、大規模シミュ
レーションで量子計算を補完する、あるいは逆に量子計算結果を高速に解析する、と
いった相互補完関係を築くことが鍵となる。また、量子計算機そのもののシミュレー
ション(エミュレーション)も HPC の重要な役割である。量子アルゴリズムの開発や誤り
耐性評価には古典計算機上での大規模シミュレーションが不可欠であり、既に一部
の HPC では量子シミュレータが動作している。HPCI でも高性能な量子計算シミュ
レータを提供し、国内の量子アルゴリズム開発コミュニティを下支えする。また、これは
アカデミア利用の基盤となる。さらに、人材育成面でも量子×HPC の融合領域を担う
専門家を育てる施策が必要である。具体的には、量子計算と数値計算の双方に通じ
た若手研究者を対象にしたハイブリッド計算競技会やハッカソンを開催し、技術交流
を促進する。以上のような量子計算との統合的活用戦略により、将来量子技術が実
用段階に達した際にも日本の HPC コミュニティが先導的立場を取れるよう備えること
が望ましい。

6. 産業界との連携と人材育成

提言

- 産業界との連携強化とイノベーション創出支援: 2026 年度に新制度を運用開始
- HPC 運営への産業界参加: 2025 年度末に産業界の参画を推進
- 産業界向け利用促進策の強化: 2025 年度に主要施策を展開開始
- 企業ニーズに応じた柔軟な利用枠の設定: 2027 年度までに利用枠を整備
- HPC 大学院プログラムの創設: 2029 年度までに HPC 大学院プログラムを開設
- 社会人向け HPC 高度研修講座の開講: 2027 年度に研修講座を開講
- 研究成果の共有と再利用環境の整備: 2026 年度にオープンサイエンス基盤を稼動
- 国際 HPC 連携: 2025 年度に国際連携のプロジェクトを始動

6.1 産業界との連携強化

HPCI を持続的に発展させるためには、産業界との連携強化と、新たなビジネスモデルの模索が欠かせない。従来、日本の HPC は官公庁主導の色合いが強く、産業利用は一部を除き限られた枠内で行われてきた。しかし Society5.0 の実現やデジタルトランスフォーメーション(DX)、ChatGPT をはじめとする生成 AI の潮流の中で、産業界こそ HPC の主要な利用者・ステークホルダーとなりつつある。欧州の EuroHPC が官民共同出資の PPP として運営されているように、日本も産業界の資金・技術を巻き

込みつつ HPC 基盤を拡充する仕組みを検討すべきである。産業利用課題はこれまで順調に増加し成果も出ているが、潜在的にはより多くの企業が HPCI を活用できる余地がある。Society 5.0 実現には産学連携による技術革新が鍵であり、HPCI はその基盤となり得る。産業界が利用しやすい環境や制度を整備することで、HPCI から新たな価値創出を促すことが可能である。具体的な提言は以下の通りである。

- **産学連携の強化とイノベーション創出支援:** 大学・研究機関と企業の共同研究に HPCI を活用する仕組みを強化する。産学連携プロジェクトに対して計算資源を優先的に配分する制度や、公募課題の評価に「産業応用の可能性」という観点を組み込むことを検討する。さらに、HPCI 上で得られた成果の事業化を支援するため、技術コンサルティングや知的財産面でのサポート体制を整える。HPCI コンソーシアムと産業界の連絡会(産業利用促進 WG 等)を通じて企業からの要望を定期的にヒアリングし、サービス改善や新機能追加に反映させる仕組みも重要である。これらにより、HPCI を通じたオープンイノベーションの創出と我が国全体の産業力強化につなげる。
- **HPC 運営への産業界参加:** 産業利用者からの意見集約の場として、産学連携ワーキンググループやユーザ会を定期開催し、システム改善やソフトウェア整備についてフィードバックを得る。実際、スーパーコンピューティング技術産業応用協議会(以下、産応協)が HPC 運用に活かされた例もある。また、自動車工業会(以下、自工会)と新化学技術推進協会(以下、新化協)には、シミュレーションや AI 活用について調査検討する WG がある。さらに、「富岳」を活用する産学連携コンソーシアムには、自動車次世代 CAE コンソーシアムや FMO 創薬コンソーシアムなどもある。ただし、これらのコンソーシアムは大企業が中心であるが、中小企業への目配りも必要であり、その意味では FOCUS が重要な役割を果たしている。今後はより踏み込んで、HPC 運営への産業界参加(ガバナンス参画)も視野に入れる。例えば HPCI の運営委員会に企業メンバーを加え、投資判断やロードマップ策定に産業的視点を取り入れることは有益である。加えて、HPC と民間クラウド企業との連携もビジネスモデル上重要である。国内クラウド事業者が HPC 分野に進出する動きがある中、HPCI との協業(サービス相互乗り入れ等)により双方の市場を拡大するチャンスがある。総じて、HPCI をオープンイノベーションの場として位置づけ、多様な主体との協働で成長させていく発想が求められる。産業界との共創により、新たな応用分野の開拓や国産 HPC ソフトの事業化など好循環を生み出すことが可能となり、ひいては日本全体の HPC エコシステム強化につながる。
- **産業界向け利用促進策の強化:** 産業利用拡大に向けたプロモーションと支援策を充実させる。例えば、HPCI の産業利用成功事例をまとめて広報し、異業種も含めた企業への周知を図る。加えて、潜在ユーザである企業を対象に HPCI 説

明会やハンズオン研修を定期開催し、HPC リテラシーの底上げを行う。大企業だけでなく中小企業への利用支援策(利用料補助や専門家派遣による技術サポート)も検討し、産業界全体での活用を促進する。

- **企業ニーズに応じた柔軟な利用枠の設定:** 企業が事業計画や研究開発スケジュールに合わせて HPCI を利用できるよう、柔軟な枠組みを導入する。現在は公募期間に合わせた課題申請が主だが、随時利用やオンデマンド利用が可能な制度を整備する。具体的には、急な計算需要に応えるスポット利用枠や、年間を通じて一定量の計算資源を確保できるサブスクリプション型の企業利用契約を検討する。また機密データを扱う企業向けに、セキュアな計算環境(例:VPN 接続や専用ノードの提供)を用意し、安全面の不安を解消する。

6.2 人材育成とオープンサイエンスの推進

最後に、人材育成とオープンサイエンスについて提言する。まず人材育成だが、HPC 分野は高度専門人材の不足が慢性的な課題であり、次世代計算基盤を使いこなすための人材育成を計画的に行う必要がある。欧州では前述の EUMaster4HPC に代表されるように、大学院レベルで統一カリキュラムの HPC 教育プログラムを創設し、人材輩出に力を入れている。日本でも、スーパーコンピューティング技術者やデータサイエンティストを体系的に育成する HPC 大学院プログラムの設置を検討すべきである。物質科学分野における 計算物質科学人材育成コンソーシアム(MPCoMS)の様に、自助努力で人材育成を行っている組織もあるが、本来であればより幅広くより計画的に行うことが望ましい。具体的には、全国の有効大学が連合して HPC 専攻を作り、理論・実習を通じて計算科学の専門修士・博士を育てる仕組みを提案する。また、産業界の研修ニーズにも応えるため、社会人を対象とした HPC 高度研修講座を HPCI が主体となって開講する。オンライン教材や集中講義を組み合わせ、現場のエンジニアが最新 HPC 技術を学べる場を提供する。すでに各センターで年次の HPC 講習会は行われているが、これを全国規模・体系立った形に発展させることで人材層を拡大したい。加えて、次世代を担う中高生・学部生に HPC の魅力を伝える啓蒙活動も重要である。スーパーコンピュータ見学会の開催や計算科学コンテストの実施、科学技術計算の入門書籍・動画の作成などにより、将来の人材の裾野を広げる取り組みを推進する。将来的にはそれぞれの NIS に所属する技術スタッフが連携して効率的に HPCI 全体の技術スタッフとして活躍していただくことも望ましい。

オープンサイエンスの推進については、HPCI が開かれた研究基盤となることを目指す。具体的には、HPCI を利用した研究成果(論文、データ、ソフトウェア)をオープンに共有・再利用できるようにする施策である。前述のデータ公開基盤の整備に加え、HPCI 成果のメタデータカタログを作成し、誰がどのような計算を行いどんな成果を得たかを検索できるようにする。また、計算結果の再現性を高めるため、ジョブの実行

環境(ソフトウェアバージョンやコンパイルオプション等)を保存・公開する仕組みを検討する。幸いコンテナ技術の進展により、計算環境一式をイメージとして保存しておけば後日でも再現が容易になっている。ユーザが希望すれば自分のジョブ実行コンテナをリポジトリに登録し、他の研究者がそれを用いて同じ計算を再現できる、といったサービスがあると理想的である。さらに、国際連携によるオープンサイエンスも推進したい。欧州の ICOS や CERN の WLCG のように、国境を越えたデータ共有・計算協調が進んでいる分野もある。HPCI も将来的には欧州や米国、アジア諸国との HPC 連携プロジェクトに積極参加し、計算資源やデータを融通しあうことでグローバルな課題解決に寄与すべきである。実際、EU と日本は 2023 年にデジタル分野の協力強化で合意し、HPC も重点項目とされている。その一環で日欧の共同研究プロジェクトである HANAMI や JHPCN と NHR(ドイツ)の相互協力協定、日米の共同研究プロジェクトである DoE-MEXT などが開始され、双方の研究チームが計算資源への相互アクセスやワークショップを通じて連携を深めている。こうした国際プロジェクトに HPCI としてコミットすることで、日本の研究者が世界の HPC インフラを活用できる機会が増え、逆に国内 HPC にも海外から優秀な人材・知見が集まるという相乗効果が期待できる。

7. あとがき

本提言は、HPCI システムの今後の運営の在り方に関する調査検討ワーキング・グループ(以下、WG)において、次世代計算基盤運用技術調査研究チームの研究代表や「富岳」運用技術チームの代表など、前年度ヒアリング対象者となった方々を WG 委員として迎え、HPCI のユーザや HPCI システム構成機関(以下、システム構成機関)の意見を踏まえて、将来、次世代計算基盤を利用することになるユーザ、ならびにシステム構成機関としての立場として、ユーザビリティ向上のためにどのようなことが次期 NFS 及び NIS の整備・運用の上で求められるかについての議論と提言をまとめたものである。次世代計算基盤が、様々な科学的・社会的課題の解決を通して、より良い次世代社会の実現に貢献することが期待される。

附録

本報告書は、HPCI 計画推進委員会における決定を受けて、高度情報科学技術研究機構内に設置された「HPCI システムの今後の運営の在り方に関する調査検討ワーキンググループ」において、一般社団法人 HPCI コンソーシアムが中心となり、調査・検討した結果を報告するものであることを付記する。

HPCI システムの今後の運営の在り方に関する調査検討ワーキンググループ委員リスト

主査	横田 理央	東京科学大学 情報基盤センター
副主査	朴 泰祐	筑波大学 計算科学研究センター
委員	合田 憲人	国立情報学研究所 アーキテクチャ科学研究系
委員	佐藤 三久	国立研究開発法人理化学研究所 計算科学研究センター 量子 HPC 連携プラットフォーム部門
委員	塩原 紀行	一般財団法人 高度情報科学技術研究機構
委員	茂本 勇	ダイキン工業株式会社 テクノロジー・イノベーションセンター
委員	庄司 文由	国立研究開発法人理化学研究所 計算科学研究センター 運用技術部門
委員	千葉 滋	東京大学 情報基盤センター
委員	藤堂 眞治	東京大学 大学院理学系研究科
委員	埴 敏博	東京大学 情報基盤センター
委員	福澤 薫	大阪大学 大学院薬学研究科
委員	堀 高峰	国立研究開発法人海洋研究開発機構 海域地震火山部門 地震津波予測研究開発センター

※50 音順

※オブザーバ:文部科学省研究振興局参事官(情報担当)付、伊藤聡(HPCI コンソーシアム理事長)、森雅博(RIST 神戸センター長)、西一成(HPCI コンソーシアム・事務スーパーバイザー)、その他 HPCI コンソーシアムメンバで希望する者

検討の記録

令和 6 年(2024 年)8 月 1 日(木)

第 86 回理事会

- ・令和 6 年度 HPCI コンソーシアム理事の業務分担の決定
- ・調査検討 WG の実施方針(テーマ、検討期間、検討委員)を確認

令和 6 年(2024 年)10 月 28 日(月)

第 1 回調査検討 WG

- ・次世代の計算基盤に関する政府の動向について文部科学省から報告
- ・令和 5 年度 WG で十分検討できなかった課題について前主査から報告
- ・令和 6 年度 WG 計画について検討

令和 6 年(2024 年)11 月 29 日(金)	第 2 回調査検討 WG
・令和 6 年度 WG 報告書案(提言案)の検討(1)	
令和 7 年(2025 年)1 月 7 日(火)	第 3 回調査検討 WG
・令和 6 年度 WG 報告書案(提言案)の検討(2)	
令和 7 年(2025 年)1 月 17 日(金)	第 4 回調査検討 WG
・令和 6 年度 WG 報告書案(提言案)の検討(3)	
令和 7 年(2025 年)1 月 28 日(火)	第 87 回理事会
・調査検討 WG の進捗を確認	
令和 7 年(2025 年)2 月 14 日(金)	第 5 回調査検討 WG
・令和 6 年度 WG 報告書案(提言案)の確認	
令和 7 年(2025 年)2 月 28 日(金)	意見交換会
・令和 6 年度 WG 報告書案(提言案)についての意見聴取	
令和 7 年(2025 年)3 月 10 日(月)	第 6 回調査検討 WG
・令和 6 年度 WG 報告書案(提言案)についての意見交換会の意見反映 ・次年度に向けた検討事項の確認	
令和 7 年(2025 年)3 月 31 日(月)	第 88 回理事会
・令和 6 年度 WG 報告書案(提言案)の確認	

次世代計算基盤の ユーザビリティに関する提言 (案)

HPCIコンソーシアム理事

HPCIシステムの今後の運営の在り方に関する調査検討WG 主査

横田 理央 (東京科学大学)

提言（案）の位置付け

「HPCIシステムの今後の運営の在り方に関する調査検討ワーキング・グループ」において、

HPCIのユーザやHPCIシステム構成機関（以下、構成機関）の意見を踏まえて、

将来、次世代計算基盤を利用することになるユーザ、ならびに構成機関の立場として、

ユーザビリティを向上するために、どのようなことが次期フラグシップを含む次世代計算基盤の整備・運用の上で求められるかについての議論と提言をまとめるもの

提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

1. はじめに

- HPCIの説明

- 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)は、世界トップクラスのスーパーコンピュータ（以下スパコン）であるフラッグシップ計算機を頂点とし、大学情報基盤センター等が運用するスパコンや大規模共用ストレージを高速ネットワークで接続して、全ての計算資源へのシングルサインオン可能な共用計算環境インフラ

- HPCIを取り巻く状況

- 第6期科学技術・イノベーション基本計画が策定
 - Society5.0の実現、持続可能な社会への変革、研究力強化、教育・人材育成
- 文部科学省次世代計算基盤検討部会の中間取りまとめ
 - 国内の主要な計算基盤、データ基盤、ネットワークが一体的に運用
- 次世代計算基盤に係る調査研究（以下ではフーズビリティ・スタディの略称でFS）
 - システム調査研究チーム（2チーム）、新計算原理調査研究チーム、運用技術調査研究チーム
- 技術的な動向
 - ポスト「富岳」へのアクセラレータ導入
 - 生成AI、AI for Science

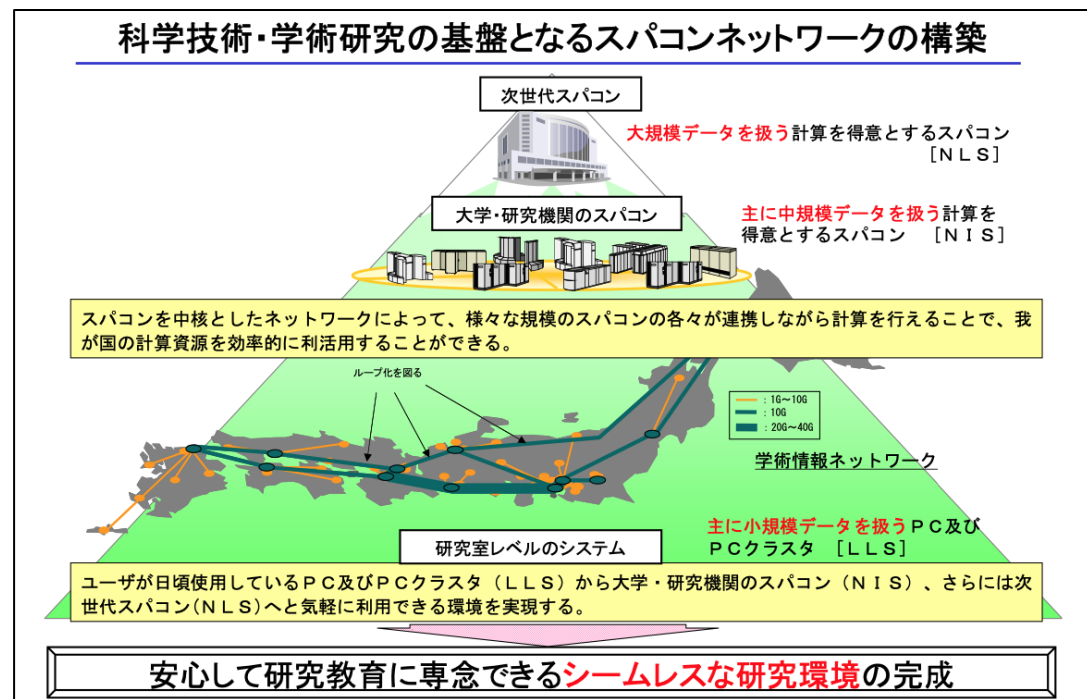
提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

2.1 HPCIシステムのこれまでとこれから

- 「第二階層」という呼び方を改める
 - ナショナル・フラッグシップ・システム(NFS)
 - ナショナル・インフラストラクチャー・システム(NIS)

平成18年の「次世代スーパーコンピュータ」プロジェクトの資料



2.2 次期フラッグシップや基盤センターを含む次世代計算基盤の果たすべき役割

- 次世代計算基盤に関する報告書（最終取りまとめ）
 - TOP500 など単一の尺度を対象としたランキングのみを追求しない
 - 国内の多様な計算資源需要を着実に満たす
 - 10年間の調達サイクルを改め、時代の要請に応じた性能を提供し続ける
 - 自国の技術を中心にスーパーコンピュータを開発・整備する能力を確保
 - 加速部の導入に際して既存のアプリケーションが安定して動作すること
 - システムソフトウェア環境の相互互換性を高めるとともに水平展開
 - 開発されたアプリケーションが国内外のシステムに簡便に移植できること
 - 生成 AI の利用などHPCの新たな領域を開拓する
 - 「AI for Science」の取組を通じて研究サイクルを飛躍的に加速

2.3 国際的な動向

- GPU計算基盤の整備
 - 2022年 Frontier (ORNL), LUMI (CSC)
 - 2023年 Aurora (ANL), Leonardo (CINECA)
 - 2024年 El Capitan (LLNL), ALPS (CSCS)
- GPU向けソフトウェア開発
 - Exascale Computing Project (ECP)
 - EuroHPC
 - 広範な分野のコードをGPUアーキテクチャへ移植
 - Extreme-scale Scientific Software Stack (E4S)

2.4 アンケートからみるHPCIの現状と課題

令和5年度(2023年度)には課題代表者259名を対象に令和6年(2024年)6月10日～7月8日にアンケートを実施し、151名の方から回答を得た(回答率58.3%)

- ユーザビリティに関する意見
 - 「富岳」以外のHPCI資源を利用した課題のうち40%が「システムの利用経験がある、または使い慣れている」を、計算機の選択理由として挙げている
- 新規利用者の獲得
 - スーパーコンピュータの利用経験が5年未満のメンバーの割合が「10%未満」が28%、「10%以上25%未満」が18%
- 産業利用の現状と課題
 - 「条件が整えば有償の課題への移行を検討する」と回答したユーザはここ数年60%程度で推移
- マシンの性能に関する意見
 - 「富岳」利用課題の8割が、また「富岳」以外のHPCI資源の利用課題の9割程度が、申請時に期待した実行性能が得られたと回答
 - Fugaku-LLMなどではGPUスパコンと比べてだいぶ低い性能

提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

3.1 アクセラレータの活用状況と課題

- NISのシステムへのGPUの導入時期
 - 2008年12月：TSUBAME1.2（東工大）NVIDIA S1070 170台
 - 2012年3月：HA-PACS（筑波大）NVIDIA M2090 1072台
 - 2012年4月：Laurel（京大）NVIDIA M2090 64台
 - 2013年8月：九大 NVIDIA K20 240台/K20X 16台
 - 2017年3月：Reedbush-H（東大）NVIDIA P100 240台
 - 2017年12月：OCTOPUS（阪大）NVIDIA P100 148台
 - 2018年8月：ABCI（産総研）NVIDIA V100 4352台
 - 2020年7月：不老（名大）NVIDIA V100 884台
- NISにおけるGPU講習会
 - 2009年：東工大、2012年：筑波大、2016年：東大
 - 2021年：名大、2022年：九大、2023年：京大
 - 東京大学ではベンダーと協力した事前のプレ評価でGPU機種を選定
- HPCIとしてGPU最適化・チューニングを支援する旗艦プロジェクトを立ち上げる

3.2 HPCI計算資源の多様性確保

- 全ての計算問題がGPUで効率的に解けるわけではない
 - ベクトル計算機
 - FPGA搭載機
 - 大規模共有メモリ機
 - 量子コンピュータ
- 多様な計算資源のシームレスな運用・利用
 - ユーザが計算課題に応じて最適な資源を選択・併用
 - ジョブスケジューラやミドルウェアを拡張
 - 異種計算ノード間のデータ連携
 - NISの基盤センターにプロトタイプ的な先端機を配置

3.3 ソフトウェアエコシステムと標準化

- 米国ECPの例
 - E4S: GPU・CPUいずれでも動作する科学技術計算ツール・ライブラリ群
 - Spack : HPCパッケージマネージャ
 - Kokkos : 性能可搬なC++向け並列プログラミングモデル
 - HPCとAIの双方で用いられる主要ライブラリを含む
 - AWSやGoogle Cloud上でも使える
 - 2024年5月からLinux Foundationの新基金HPSFの傘下で持続可能な開発
 - 知的財産や商標の寄贈・移管
- 日本もこのようなエコシステムに参画・貢献すべき
 - 東大は既に参画
 - 理研は準備中

3.4 レガシーコードの効率的なGPU移植

- 背景と課題
 - 膨大な行数のレガシーコードのGPU移植には多大な人的コストを要する
 - GPU移植に携わった研究者のキャリアアップをどのように実現するか？
 - 人手をかけずにレガシーコードのGPU移植を行うための革新的技術が切望される
- ソースコードの自動変換ツール
 - F2C-ACC : 気象分野のFortranコードNIMをCUDA化するのに利用された
 - GPUFORT : OpenMP4.5のディレクティブを自動的に挿入
 - Kokkos FLCL : FortranとKokkos (C++) のコードが混在することが可能に
- 生成AIを用いたGPUコード生成
 - ChatHPC : 米国DOEのS4PST
 - Code-Scribe : アルゴンヌ国立研究所
 - Fortran2CPP : ローレンス・リバモア国立研究所
 - CodeRosetta : Google DeepMind

提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

4.1 計算資源の効率的な配分とアクセス性向上

- 戦略プログラム枠の刷新
 - 生成AIや量子コンピュータなどの新しい潮流に対応できるよう対象とする分野を刷新し、新しい分野のコミュニティ育成・分野振興を進める
- 産学連携・企業利用の促進
 - スタートアップや中小企業向けには低コストで試行利用できるサンドボックス的環境を提供
- リソースオンデマンド供給
 - 「富岳」のファーストタッチオプションをHPCI全体に拡張
 - 観測データのリアルタイム解析や災害シミュレーション等
- 利用実績の少ない課題に対する対策
 - 利用実績がゼロの課題、年度末にのみ少量の利用実績がある課題に対しては、次年度の申請時にそのことを考慮した資源配分を行う
- NISシステムの共通ベンチマーク
 - HPCI連携サービス委員会で引き続き検討し、ベンチマーク結果を公表できるところまでもっていく

4.2 異なる資源間の連携と統一的UI/UX

- コンテナの共通化
 - HPCI側・クラウド側双方に共通のコンテナ型仮想環境を用意
 - クラウド上で構築したAIモデルや分析パイプラインをHPCI上に持ち込むなど
- UI/UXの改善とポータルの一統化
 - 「富岳」へのOpen OnDemand導入など対話的利用環境整備が進みつつある
 - ウェブブラウザ経由で各計算資源にアクセス・ジョブ管理ができるポータルを構築
- データ基盤の統合
 - クラウドストレージとHPCI計算資源を統合
 - HPCIの計算結果データをmdxに蓄積・公開
 - mdx上のオープンデータセットをHPCI計算に利用

4.3 大学・研究機関との連携

- 海外の事例
 - EuroHPCは各国に「国民コンピテンスセンター（NCC）」を設置
 - 33か国のNCCが国内のHPCリソース・スキルをマッピングし不足を分析
 - 産業界や行政を含む多様なユーザへの窓口となっている
 - EUMaster4HPC：欧州全域で統一カリキュラムのHPC大学院課程を創設
- 大学間・機関連携を促す資金制度
 - 複数大学・企業の連名によるHPC共同研究課題に対し重点的に計算資源や助成金を配分する
 - クロスアポイントメント制度の活用や、大型プロジェクトへの産学混成チーム編成を推進
 - 統合的な枠組みを実現するべきであり、計算資源の全体最適化を行うべき

提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

5.1 生成AIとAI for Science

- 背景

- 2024年ノーベル物理学賞：人工ニューラルネットワークによる機械学習を可能にした基礎的発見と発明
- 2024年ノーベル化学賞：計算によるタンパク質の構造予測及び設計
- 令和6年度(2024年度)のHPCI一般課題において「AIやデータサイエンスを活用して、科学的・社会的課題の解決に資する研究開発」を重点分野として設定
- 令和6年度(2024年度)版の計算科学ロードマップにもAI for Science 章が追加された

- 課題

- 「富岳」にGPUが搭載されていない
 - A64FXのFP32ピーク（FP16の加算では学習に使えない）：6.8TFLOPS
 - H100のFP16ピーク（Tensor CoreはFP32で加算）：989TFLOPS
- NISにおけるGPU資源の不足
 - 「富岳」：158,976ノード
 - Miyabi：1,120ノード、ABCI3.0：766ノード
- LLMのスケール則
 - Stargate：総額5,000億ドル
 - DeepSeek：671Bパラメータ、15Tトークン

5.1 生成AIとAI for Science

- 学際・国際連携の推進
 - AI分野の研究者、HPCの研究者、各専門分野の研究者による学際連携
 - 学際連携の橋渡し役となる複数分野に跨った専門性を有する人材育成
 - Trillion Parameter Consortium : 理研、産総研、東大、科学大などが参画
- 小型で論理推論能力の高いAIの開発支援
 - Reasoning Model : OpenAI o3、Google Gemini 2.0 Flash Thinking、DeepSeek-R1
 - Deep Research : 外部データにアクセスし、それを元に論理推論を行う
 - 国内の計算資源でも十分競争力のあるAI for Scienceの研究が可能
 - 計算科学 : 推論のさらなる高速化を実現するためのシステム・ソフトウェアを開発
 - AI : 小型で論理推論能力の高いAIの開発

5.2 量子コンピュータとHPCの統合的活用

- 海外の動向
 - ユーリツヒ研究所：超伝導量子計算機とスーパーコンピュータを繋ぐ
- 国内の動向
 - 理研R-CCSに「量子-HPC連携プラットフォーム部門」が設立
- 多方面からのアプローチ
 - 量子回路演算と通常の数値演算を組み合わせて実行
 - それができるよう、ジョブスケジューラやプログラミング環境を拡張
 - 量子回路シミュレータの活用
- 人材育成
 - 量子×HPCの融合領域を担う専門家を育てる施策が必要
 - ハイブリッド計算競技会やハッカソンを開催し、技術交流を促進

提言（案）の内容

1. はじめに
2. 背景：次世代計算基盤の必要性と方向性
3. アクセラレータを念頭においたユーザビリティの向上
4. 次世代計算基盤の戦略的整備と運用
5. 新規応用分野の開拓
6. 産業界との連携と人材育成

6.1 産業界との連携強化

Society5.0の実現やデジタルトランスフォーメーション（DX）、ChatGPTをはじめとする生成AIの潮流の中で、産業界こそHPCの主要な利用者・ステークホルダーとなりつつある

- 産学連携の強化とイノベーション創出支援
 - 産学連携プロジェクトに対して計算資源を優先的に配分する制度
 - 公募課題の評価に「産業応用の可能性」という観点を組み込むことを検討
 - 技術コンサルティングや知的財産面でのサポート体制を整える
 - 企業からの要望を定期的にヒアリングし、サービス改善や新機能追加に反映させる
- HPC運営への産業界参加
 - スーパーコンピューティング技術産業応用協議会(産応協)
 - 自動車工業会（自工会）
 - 新化学技術推進協会（新化協）
 - 製薬工業協会（製薬協）
 - 自動車次世代CAEコンソーシアム
 - FMO創薬コンソーシアム
- 企業ニーズに応じた柔軟な利用枠の設定
 - 随時利用やオンデマンド利用が可能な制度を整備
 - 機密データを扱う企業向けに、セキュアな計算環境を用意

6.2 人材育成とオープンサイエンスの推進

- HPC人材育成プログラムの設置を検討
 - 例：計算物質科学人材育成コンソーシアム(MPCoMS)
 - 社会人を対象としたHPC高度研修講座
 - 現場のエンジニアが最新HPC技術を学べる場を提供する
 - 中高生・学部生を対象とした見学会・コンテスト
 - 例：スーパーコン（科学大・阪大）
- オープンサイエンスの推進
 - データ公開基盤の整備
 - ジョブの実行環境を保存・公開する仕組みを検討
- 国際連携
 - DoE-MEXT：日米の共同研究プロジェクト
 - HANAMI：日欧の共同研究プロジェクト
 - JHPCNとNHR（ドイツ）の相互協力協定

おわりに

- 本提言は、HPCI システムの今後の運営の在り方に関する調査検討ワーキング・グループ（以下、WG）において、次世代計算基盤運用技術調査研究チームの研究代表や「富岳」運用技術チームの代表など、前年度ヒアリング対象者となった方々をWG委員として迎え、**HPCIのユーザやHPCIシステム構成機関（以下、システム構成機関）の意見を踏まえて、将来、次世代計算基盤を利用することになるユーザ、ならびにシステム構成機関としての立場として、ユーザビリティ向上のためにどのようなことが次期NFS及びNISの整備・運用の上で求められるか**についての議論と提言をまとめたものである。
- 次世代計算基盤が、様々な科学的・社会的課題の解決を通して、より良い次世代社会の実現に貢献することが期待される。

令和6年度 HPCI システムの今後の運営の在り方に関する 調査検討ワーキンググループ委員

主査	横田 理央	東京科学大学 総合研究院 スーパーコンピューティング研究センター
副主査	朴 泰祐	筑波大学 計算科学研究センター
委員	合田 憲人	国立情報学研究所 アーキテクチャ科学研究系
委員	佐藤 三久	国立研究開発法人理化学研究所 計算科学研究センター 量子HPC連携プラットフォーム部門
委員	塩原 紀行	一般財団法人 高度情報科学技術研究機構 神戸センター
委員	茂本 勇	ダイキン工業株式会社 テクノロジー・イノベーションセンター
委員	庄司 文由	国立研究開発法人理化学研究所 計算科学研究センター 運用技術部門
委員	千葉 滋	東京大学 情報基盤センター
委員	藤堂眞治	東京大学 大学院理学系研究科
委員	埴 敏博	東京大学 情報基盤センター
委員	福澤 薫	大阪大学 大学院薬学研究科
委員	堀 高峰	国立研究開発法人海洋研究開発機構 海域地震火山部門 地震津波予測研究開発センター