

【提言】次世代計算基盤を利用した成果の最大化に向けて

令和6年6月25日

一般社団法人 HPCI コンソーシアム

目次

1. はじめに.....	3
2. 次世代計算基盤の整備・運用の方向性.....	5
2.1 背景・課題.....	5
2.2 シームレス化を軸としたユーザビリティ向上	6
2.3 次世代計算基盤の戦略的な整備・運用	8
2.4 産業利用の強化.....	9
2.5 分野拡大(AI・データサイエンス、量子コンピュータ関係、その他応用分野)	10
2.6 人材育成、分野振興.....	11
3. 次期フラッグシップシステム検討にあたって.....	14
3.1 アーキテクチャの方向性.....	14
3.2 コデザイン	14
3.3 アプリケーション・ソフトウェアの開発について.....	16
4. あとがき.....	17
附録.....	18
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	18
検討の記録.....	19

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(第二階層計算機)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 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、二代目である「富岳」の共用が令和 3 年 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などが創出されている。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。

近年、AI・機械学習を用いた大量かつ多様なデータの収集・活用が進展し、データ駆動科学においてより大規模なデータ処理能力が求められるとともに、大規模シミュレーション結果を学習データとして活用する、AI・機械学習を使ってシミュレーションを加速する等、大規模シミュレーションとデータ駆動科学が両輪となって、新たなサイエンスの成果を創出するとともに、社会的な課題を解決する時代になってきている。

このようなデータサイエンスや AI 技術とシミュレーションの融合、リアルタイムデータ処理といった先進的な処理方法等やデータ駆動型科学、さらには研究スタイルの革新を目指す研究 DX(研究活動のデジタルトランスフォーメーション)の進展を踏まえ、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 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」及び第二階層計算機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。また、次期「フラッグシップシステム」開発の方向性として、システムとアプリケーションの協調設計が必須であり、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とされている。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 2023 年末に公開されており(<https://cs-forum.github.io/roadmap-2023/>)、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 2024 年度はじめには公開される予定である。

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 年度から開始されている。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、前述の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。

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、国への提言を行なってきた。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、FS の研究代表や「富岳」開発関係の方々、WG 委員等へのヒアリングや 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、成果を最大化するために、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。第2章では、次世代計算基盤全体の整備・運用の方向性に関して提言する。ここでは、運用技術FS 関連の内容に加えて、産業利用の強化や AI・データサイエンス、量子コンピュータ関係の分野拡大についても扱う。第3章では、次世代計算基盤の中で、特に次期フラッグシップシステムの方向性に関して提言する。

2. 次世代計算基盤の整備・運用の方向性

2.1 背景・課題

「はじめに」で触れた通り、次世代計算基盤検討部会の中間取りまとめで、次世代計算基盤の整備にあたっては、次期「フラッグシップシステム」及び第二階層計算機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。これを受けて運用技術 FS では、フラッグシップシステムに加えて、第二階層計算機群、データ活用社会創成プラットフォーム mdx(大学の情報基盤センター群等の 11 機関で共同運営しているクラウドライクなデータサイエンス向け研究基盤:<https://mdx.jp/>)をはじめとする各種データプラットフォーム、NII Research Data Cloud (RDC)といった研究データ基盤(NII が開発・運営している研究データの管理・公開・検索に必要なプラットフォーム:<https://rcos.nii.ac.jp/service/>)、これらを高速に接続する学術情報ネットワーク SINET6(<https://www.sinet.ad.jp/>)、このようなシステムがより有機的に結合した、持続可能な次世代計算基盤の実現を検討するとしている。その目的は、様々なユーザのニーズに対し、求められる適切な資源を提供し、平易で柔軟な利用を実現することであり、これらによって、次世代計算基盤の使い勝手の良さ、すなわち、ユーザビリティを向上し、ユーザの裾野を広げることによって、我が国の科学技術の水準を押し上げることを目指している。

運用技術 FS で扱われているユーザビリティ向上の具体的な中身は、従来のバッチ型のジョブ実行だけでなく、クラウドのような多様な利用方法が提供されていること、その上で、複数の計算資源や複数のストレージを、GakuNin RDM(<https://support.rdm.nii.ac.jp/>)や Open OnDemand(<https://joss.theoj.org/papers/10.21105/joss.00622>)等による web ベースのユーザポータルを通してシングルサインオンかつ GUI で使えること、データ解析や機械学習でデファクトスタンダードとなっている Python の実行環境である Jupyter Lab(<https://jupyter.org/>)が mdx などのクラウドライクなスパコンシステムでも使えること、クラウド環境からスパコンにジョブを投入できること、などが挙げられる。また、それらを実現するためのシームレス化として、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、などが挙げられる。これらは、我々が求める次世代計算基盤の利用成果の最大化にとって不可欠な要素であり、2.2 節のユーザビリティ向上やシームレス化で取り上げる。

また、運用技術 FS では、上記のようなセンター間での協調・連携や外部データ利用等を進めていくための運用体制や「富岳」から次期フラッグシップへの移行期間、いわゆる端境期の資源不足対策等の検討も行なっている。これらの検討は、HPCI を形成する第二階層のシステム構成機関の運用や計算機導入計画などにも関わるものであり、各大学や国研における独自のミッションや予算措置のもとで行われているため調整は難しいが、次世代計算基盤の利用成果の最大化にとっては避けられない課題であり、2.3 節の次世代計算基盤の戦略的な整備・運用で扱う。

次世代計算基盤の利用成果の最大化においては、ユーザのニーズをいかに把握して、今後の整備・運用に反映させるかが重要な課題である。その意味で、産業界からの意見・提言(例えば、

産応協 <http://www.icscp.jp/>)を踏まえた、産業利用促進の観点での次世代計算基盤の整備・運用のニーズを 2.4 節の産業利用の強化で取り上げる。また、2.5 節の分野拡大では、従来の HPC の計算科学分野の枠を超えて、AI・データサイエンス、量子コンピュータ関係、その他新規応用分野で次世代計算基盤が利用されるための課題・提言を述べる。そして、成果最大化にとって不可欠であり、以前からの課題である人材育成・分野振興を 2.6 節で扱う。

2.2 シームレス化を軸としたユーザビリティ向上

Society 5.0 の実現や産業競争力の一層の強化のため、AI・データサイエンスや産業利用のニーズの多様化に応え、HPCIの利用分野を拡大していくためには、AI やデータサイエンスを活用している人たちが現在享受している、あるいは本当に求めているユーザビリティを、次期フラッグシップシステムを含む次世代計算基盤として実現することが必要不可欠である。そのためには、認証の連携、web ブラウザ経由などによる対話的利用に適した環境の提供、大規模データへのスムーズなアクセス、外部からのリアルタイムでのデータ収集・集積等のための連携といった、様々なレベルのシームレス化が必要である。また、AI・データサイエンス分野等で進んでいるクラウド上でのプログラムの実行方法やライブラリの共通化等のシームレス化のノウハウを導入する観点も重要である。2.1 節で触れたように、こうしたユーザビリティ向上やシームレス化に向けた取り組みは始まっており、運用技術 FS でも検討されている。ここでは、これらの取り組みの課題を取り上げるとともに、次世代計算基盤の利用成果最大化の観点で必要と考えられる、他のユーザビリティ向上やシームレス化の課題についても取り上げ、それらに対する提言を述べる。

ユーザビリティ向上の取り組みは、データ活用社会創成プラットフォーム mdx のような仮想化された基盤で、Jupyter Lab を含めた自由な構成を構築できる環境の導入や、「富岳」への Open OnDemand 導入などにみられる。しかし、スパコンを使う場合、通常は管理者権限がユーザにないので、管理者権限が必要なパッケージが手軽には導入できない。Jupyter Lab などが使えるのは、あくまでスパコン管理者の側で用意した環境を使うということであって、ユーザが自由に利用環境を構築できるわけではない。また多くの場合、スパコンはログインノードを経由して計算ノードにアクセスするので、直接外部ストレージと計算ノード間のデータ転送を行うことができない運用となっており、データサイエンス的な研究を行う上で制約となっている。これらは、技術的な問題だけでなく、スパコンの運用やセキュリティポリシーの問題であり、これらの考え方に関する議論や検討が必要である。

一方、「富岳」の利活用拡大検討の一環として、「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・ニーズを踏まえ、利用拡大を図るべきである。

上記のユーザビリティ向上は、どちらかといえば、並列計算等の経験のない非 HPC ユーザ向けといえる。一方で、従来の HPC ユーザ向けのユーザビリティ向上も当然重要である。アプリケーションを利用する側として並列計算の経験がある HPC ユーザの場合は、HPCI 関係で開発されたアプリケーション、商用アプリケーション、オープン・ソース・ソフトウェア(以下、OSS)などがスムーズに次世代計算基盤上に移植されていて、コンパイル等も問題なく実行できることなどが保証されていること、さらには、アプリケーションのバージョンアップに応じて継続的にアップデートされることが求められる。このような移植や継続的な支援が、「富岳」向けのアプリケーションから、次世代計算基盤の次期フラッグシップはもちろんのこと、第二階層計算機や企業の自社クラスタのアーキテクチャに対して実現すれば、「京」や「富岳」の産業利用でこれまで活用してきたアプリケーションを、次世代計算基盤でもシームレスに実行できるようになり、その利用の成果最大化に貢献することになる。これらは、次期フラッグシップ開発におけるシステムソフトウェアやコンパイラ開発とその継続に位置づけられる重要な課題である。なお、アプリケーション開発者にとってのユーザビリティ向上は、フラッグシップシステムにおける協調設計(コデザイン)の充実が必要であり、3章で取り上げる。

ユーザビリティ向上に向けた様々な取り組みは、現時点では個別的・ボランティア的なものであり、統合的な取り組みに強化する必要がある。例えば、ユーザビリティ向上の実証実験の鍵となる mdx には運用費の担保はなく、利用負担金だけでは不足するため、運営機関による一部経費負担で運用されている(令和 5 年 12 月 18 日 HPCI 計画推進委員会)。このような状況は、実は HPCI 設立前の状況と類似している。HPCI の運用開始は平成 24 年だが、その前の平成 22 年度・23 年度に、HPCI の仕様設計や詳細仕様設計に関わる 2 つのプロジェクトが立ち上がり、理研と基盤センター等で設計・構築が行われた。そのさらに前には、ボランティア的な取り組みとして、9 大学の基盤センター+NII によるグリッド実証実験、すなわち今の HPCI の原型である、基盤センターの計算機群を連携させたシングルサインオン・共用ストレージ・申込みワンストップ等の実証実験が行われていたのである。この実証実験が、現在の mdx 等を活用したボランティア的な取り組みに相当するだろう。次世代計算基盤の構築は、HPCI 全体のグランドデザインの要とも位置付けられるため、次期フラッグシップ開発プロジェクトにおいても、これら mdx 等の有効利用を念頭に、ユーザビリティ向上に必要なフラッグシップ周辺環境等の仕様の概念設計・詳細設計・整備等に取り組む必要があると言える。

当面の課題として、ユーザビリティ向上の取組を強化するためには、mdx 等の実証試験的な取り組みを継続・発展させ、それらを活用して様々な分野のニーズを把握するとともに、国内外のユーザビリティ向上の取組を調査し、それらの知見を次期フラッグシップの開発にフィードバックすることが不可欠である。すでに運用技術 FS では、システム FS と連携して、ユーザビリティ向上やシームレス化を支えるツールの研究開発について検討している。したがって、次期フラッグシップ開発プロジェクトの中で、運用技術 FS の活動を継続することで、ユーザビリティ向上の取組を継続・発展させることが必要である。

2.3 次世代計算基盤の戦略的な整備・運用

2.1 節で触れたように、運用技術 FS では、次世代計算基盤のユーザビリティ向上やシームレス化の実現に必要な、システム構成機関間での協調・連携や外部データ利用等を進めていくための運用体制の検討もなされている。また、「富岳」から次期フラッグシップへの移行期間、いわゆる端境期の資源不足への対策も検討されており、これはこれまでの提言でも繰り返し課題として挙げていた問題である。いずれも、次世代計算基盤の利用成果の最大化にとって不可避な課題であるとともに、第二階層計算機群の整備・運用を担う大学や国研のシステム構成機関の理解と協力無しには解決することができない。一方、第二階層のシステム構成機関は、機関毎の独自のミッションのために、限られた予算・体制のもとで運用を行なっている。そのため、各機関にとって次世代計算基盤の一翼を担うことが、当該機関にとって、メリットが明確であることが必要となる。

例えば、2.2 節のアプリケーションのシームレス化で述べたような、「富岳」の産業利用向けのアプリケーションを次世代計算基盤の第二階層で実行する課題では、移植や利用環境の移行、移行後の継続的な更新等への支援が必要になる。移植については、HPCI 内に開発者のいるアプリケーションと商用アプリケーションや OSS とで状況は異なるが、いずれにしても、アプリケーション開発側の自助努力や第二階層のシステム構成機関の独自の支援体制に頼るのは、それぞれに負担が大きいだけでなく、共通する技術や知見が多い移植や支援であっても個別に行うことになり、効率が悪くなる。そこで、当該機関のメリットとして、産業利用ニーズの高いアプリケーション群の移植性を高めることで当該機関における産業利用が広がるということがあり、また運用側からのメリットとして、利用支援を当該機関単独で負担するのではなく、同じ「富岳」向けアプリケーション群を支援してきた機関や複数の類似アーキテクチャのシステム構成機関で協力して支援するような体制をつくることできれば、効率的な支援が可能となる。なお、産業利用が広がることで、産学官連携を加速させるための柱の一つとして HPC が貢献できる可能性をコミュニティとしてアピールしていくことは重要であろう。

次世代計算基盤全体の利用の成果を最大化するためには、アカデミアから産業界まで、分野としても広範囲に渡るニーズを満たすことが求められるが、次期フラッグシップシステムだけですべてのニーズを満たすことができる訳ではない。そこを補うためには、次世代計算基盤を構成する第二階層計算機の多様性が不可欠である。第二階層計算機の多様性を確保するためには、どのようなアーキテクチャでどのようなスペックのシステムをどのようなタイミングで導入するかについて、各システム構成機関の事情・利益も考慮しつつ、議論・検討する場が必要と考えられる。また、次世代計算基盤の運営体制については、ユーザや運用者の問題意識を利用制度の改善に迅速に反映させる体制の整備が必要等の改善提案が運用技術 FS 等でなされているが、こうした運営体制の改善についても、引き続き関係機関で議論・検討していくことが期待される。これらの議論・検討についても、ユーザビリティ向上のための取組の強化と同様に、次期フラッグシップ開発プロジェクトの中で、あるいは運用技術 FS を継続して、取り扱うことが望まれる。

なお、フラッグシップシステムは、本来のその卓越した計算性能から、入れ替わりに伴う端境期が生じると、第二階層の計算資源で代替ができるものではない。このため、次世代計算基盤として、

どうすれば端境期が生じないようなシステムを将来的に実現できるかについて、現在の FS 及びそれに続く次期フラッグシップ開発プロジェクトの中で、様々な観点から議論・検討されることが望まれる。

2.4 産業利用の強化

産業利用は、HPCI 設立当初から取り組みがあり、発展してきた実績がある。「京」での産業利用枠は、高並列計算機の産業課題への適用性検証が主な目的だった(HPCI 準備段階コンソーシアム最終報告書第3章、平成24年1月30日)が、「富岳」では産業上の成果をより求めるようになった。産業界が「富岳」のような公的計算資源を使うメリットは、アカデミアの研究力と最先端計算環境、基礎的研究段階における数値的検証、アカデミアも含めた協調領域での連携の場としての計算環境が得られるためである。一方、産業界ではクラウド活用の進展と、AI・データサイエンスの急速な展開は産業構造そのものを変革しつつある。こうした、産業界を取り巻く状況の変化を受けた制度や仕組みの見直しが必要と考えられる。

「京」や「富岳」で開発されたアプリケーションに関して、次世代計算基盤でも効率的に使えるのみならず、ソフトウェアの将来にわたる継続的な維持・発展の担保、各社の戦略に沿った商用クラウド等の計算資源へのスムーズな展開を可能とする技術・新たな産業利用制度を整備する必要がある。産業競争力の革新を目指す戦略的利用制度の対象としては、産学連携利用・コンソーシアム利用の推進、スタートアップ・ベンチャーの育成・助成、データ創出・資源としての活用推進等が挙げられる。また、世界的な産業競争力の観点では、業界のデファクトスタンダードな OSS や商用アプリケーションの利用環境向上が重要であり、少なくともコンパイラやライブラリ等により性能向上が図れることが期待される。さらに、性能向上に果たすコンパイラ等が長期間に亘り、継続的に整備される支援が必要である。

実装すべき仕組みとしては、官民システムのシームレス化がある。フラッグシップシステムや第二階層計算機のような公的設備と商用クラウドや自社計算機のような民間設備の間の認証連携を実現し、いずれの計算機にもコンテナ型仮想化環境が用意されていて、それらのいずれにもポータルサイトからユーザがアクセスして計算・解析ができるといったことが理想である。また、AI for Industry のためのデータ資源と活用システムの仕組み、すなわち、学習・推論のための計算基盤のみならずデータ基盤とのやり取りがシームレスにできる環境も実装すべきである。公的計算環境での共通基盤的大規模言語モデル・汎用的機械学習モデルやものづくり分野における大規模シミュレーション結果等を用いて生成されたデータと、民間計算環境での企業内データとを組み合わせるのであれば、個社の差別化のためのモデル構築が必要となる。ここで、企業が保有するデータの取り扱いが必要になることや、大規模シミュレーション結果をデータとして学習することが必要になってくること等から、計算基盤の視点のみならずデータ基盤の視点を取り入れることが要求される。そこでは、経済安全保障を念頭においたセキュリティの強化も必要となる場合がある。HPCI では共用開始当初より、国の安全保障輸出管理制度の下で、外為法に基づき利用や支援に必要な審査や手続きを実施してきた。内外の計算資源やデータ基盤がシームレスにつながっていく中

で、セキュリティを確保することはますます重要になっている。また、学習に用いる大規模データとしては、フラッグシップシステムでの超大規模計算や多数回行った計算のアウトプット、外部にある大規模実験施設や IoT 技術を通して得られる研究室レベルでのデータも対象とできることが望まれる。

その他に考慮すべきこととしては、テストベッド機能の充実、すなわち、産学連携、ベンチャー支援を戦略的に実施する一方で、個社の方針に沿った試行を希望するユーザの利便性を向上させることが望まれる。また、他省庁との連携、特に経産省等の設備・事業との接続を滑らかにすることで、企業としては産業競争力強化に繋がりやすい面がある。

これらはいずれも、2.2 節で取り上げたユーザビリティ向上やシームレス化として対応が可能なものである。

2.5 分野拡大(AI・データサイエンス、量子コンピュータ関係、その他応用分野)

ユーザビリティ向上のための取組の強化として、クラウドサービスや web ベースでのユーザビリティ向上に向けた開発を進めるための試行機会の提供と結果のフィードバックを、AI・データサイエンスや量子コンピュータ関連、産業利用を含めた様々な分野を対象に実施すれば、分野拡大の促進につながると期待される。

AI・データサイエンス分野の専門家及びユーザがより使いたくなるシステム、つまり、普段使っている商用クラウド等におけるユーザビリティの高さを満たしつつ、GPU クラスタ等を駆使しているユーザが、大規模シミュレーション結果や外部からのリアルタイムデータ等の様々なデータをシームレスに使う大規模解析できるということでの商用クラウド等との差別化が必要である。さらに、これまで HPC を活用してきた人たちが、大規模シミュレーション結果の機械学習、物理法則を満たす機械学習、AI を活用したシミュレーション中の収束計算の高速化等のような AI for Science を実践するために活用しやすいシステムであることも必要である。

これらのいずれにおいても、現在 GPU をはじめとした様々な演算加速機構を駆使して行われているAI・機械学習の大規模データ解析(例えば、<https://www.nvidia.com/ja-jp/ai-data-science/>)を、さらに大規模・高速化することが必要である。これは、量子コンピュータ関連技術における疑似量子アニーラや量子コンピュータのシミュレータに対しても該当する。この大規模・高速化は、まさに次期フラッグシップが担うべきところである。そこで動かすコード群は、コデザインの重要な対象と考える必要がある。一方、これまで HPC コミュニティから遠かった AI・データサイエンス分野の専門家が、大規模計算を手軽にできるようにするためには、OSS の移植・継続的更新やライブラリの並列処理対応など、システム開発側での研究開発とその継続が不可欠である。なお、これまで HPC を活用してきたユーザの利用については、コードの演算加速機構等への対応という意味で、3 章で取り上げる。

量子コンピュータは、次世代計算基盤の計算資源における多様性の一つである。将来的には、複数の量子コンピュータと次期フラッグシップシステムや第二階層計算機群が、直結や SINET 経由等の様々なレベルでネットワークに接続されたハイブリッド環境の構築が望まれる。そのため、次

世代計算基盤のネットワーク開発においては、量子コンピュータ等との接続も視野に入れる必要がある。ここでは、量子コンピュータやそのシミュレータは、従来の HPC が苦手とする特定の課題解決のための道具、つまり一種の演算加速機構、としての活用が期待される。また、量子コンピュータのシミュレータ、最適化アルゴリズム、量子機械学習等の活用は、クラウドからの利用が標準的であり、2.2 節で述べたユーザビリティ向上の取り組みによって、量子コンピュータ関係のユーザが次世代計算基盤を活用する道も拓けると期待される。

分野拡大に向けて、エンドユーザのニーズを取り込むことが運用技術 FS の中で課題とされている。ユーザのニーズをどう調べるかは、非常に重要かつ難しい問題である。ユーザニーズ志向であることは重要だが、ユーザのニーズは今実現しているところからしか出てこないで、システム側でシーズを提供した上で、新たなニーズも開拓しながら進めていく必要がある。その意味で、次世代計算基盤に向けての実証試験的な取り組みである mdx とスパコンの連携や商用クラウドとの連携等を積極的に活用して、幅広いユーザのニーズを引き出すことが効果的と考えられる。その際、「富岳」や第二階層の計算資源を mdx から利用することを通して、HPC を十分に活用するために必要なハードウェア・ミドルウェア・アプリケーションに求められる機能やプロトタイプの開発・試行・改善を行い、次期フラッグシップシステムやその周辺環境の設計・開発に反映することも必要となる。

mdx を活用するためには、大学等での認証連携を実現している学術認証フェデレーション(学認)に現時点で直接加わっていない HPCI 関係の研究機関や企業のユーザ、AI・データサイエンス分野や量子コンピュータ関連分野、高エネルギー・天文等の大規模実験・観測分野、IoT を活用する応用分野や産業界等のユーザも、mdx をトライできる環境を整備・宣伝する必要がある。その上で、活用事例を増やしつつ、活用のために必要なニーズに応えるための開発を、ユーザビリティ向上の取組を強化として実施することが望まれる。また、産学あるいは国際間で柔軟な認証連携を可能とするための次世代認証連携への対応や、様々なサイエンス分野や企業のデータを扱う上で、データマネジメントプランに従った研究データ管理を実践することも同様に取り組むべき課題である。

幅広いニーズを引き出す上で、mdx、次世代認証連携 (https://www.nii.ac.jp/openforum/2023/day2_auth3.html)、研究データ基盤など HPCI の外の活動との連携、高エネルギー、天文等の大規模実験・観測分野、IoT を活用する応用分野や産業界とのコミュニケーションも重要である。例えば、WLCG(<https://wlcg.web.cern.ch/>)は、CERN が運用している大型ハドロン衝突型加速器(LHC)が出力する巨大なデータを世界の研究者が解析しなければならないという明確な目的に応えるための機能が情報系から提供され、うまくかみ合っている。このような応用分野との連携をしっかりとりながら進めていくことも重要であろう。このためには、ユーザビリティ向上の取組の強化の中で、ニーズ把握や反映のための試行の実施機関として、こうした機関にも加わってもらうなど、より積極的な外部との連携が有効と思われる。

2.6 人材育成、分野振興

人材育成や分野振興における課題は、前回の提言(令和 4 年 6 月 21 日)にも指摘されているよ

うに、現フラッグシップである「富岳」の開発期には、「HPCI 戦略プログラムやポスト「京」重点課題を通じて若手研究者の育成が行われ、育成された若手研究者等による成果創出への貢献が多く見られた」が、「富岳」成果創出加速プログラムにおいては大幅な予算削減により、折角育った若手研究者の雇用を大幅に減らさざるを得ず、「富岳」が稼働して研究が本格化した段階で貴重な人材を失った課題も」あった。また、「現在のようにフラッグシップ計算機開発の端境期(運用期)において、人材育成を主眼としない成果創出加速プログラムだけが実施されている状況では、人材育成のための予算が非常に少なく継続的な人材育成が困難」ということであった。今後は、次期フラッグシップのための開発期に入ることになる。したがって、その開発プロジェクトの中で若手研究者の人材育成をどのように進めるのが効果的か、さらには、運用期において、それらの育成され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。

次期フラッグシップを含む次世代計算基盤の一体的な開発を進めるためには、演算機構・メモリ・計算機内外のネットワーク・ストレージといったハードウェア開発、OS、コンパイラ、ライブラリ等のソフトウェア開発、アプリケーション開発、さらにはユーザビリティ向上のためのクラウド環境やユーザインターフェース開発など、あらゆる開発が必要になる。これらの開発のための人材育成を、開発主体の企業だけでなく、大学・国研・産業界等のユーザ、システム構成機関でも実施できるように、次期フラッグシップ開発プロジェクトの中で課題を設定し、人材育成のための予算を配分することが求められる。

その際に、開発期の段階から、開発期間終了後の運用期での継続性を考慮した制度設計も併せて行う必要がある。その意味で、計算物質科学人材育成コンソーシアム(Professional development Consortium for Computational Materials Scientists: PCoMS)は、若手人材の育成から大学・企業等へのキャリアアップまでがつながった重要な参考事例と言える。当該コンソーシアムは、「文部科学省 科学技術人材育成補助事業 国立研究開発法人科学技術振興機構『科学技術人材育成のコンソーシアムの構築事業(次世代研究者プログラム)』」の採択を受けて 2015 年 8 月に設立し、当該分野の科学的知見と HPC 技術を身につけた若手研究者や長期インターンシップ等による経験をしたポスドク・博士課程大学院生を育成し、大学・国立研究機関や企業への雇用を実現することで、非常に高い事後評価を受けている。開発期間後のフラッグシップを含めた次世代計算基盤の運用期間においても、これらの開発・活用人材がキャリアアップして、幅広い分野におけるコーデザインの推進・成果創出のための開発を担う人材として継続的な活躍が求められる。そのためにも上記のような人材育成プログラムの拡大が急務である。

なお、育った後にキャリアアップできるかどうかは、HPCI コミュニティだけで閉じた問題ではなく、Society5.0 などの大きな社会的仕組みの中でキャリアアップできていくことが必要と言える。生成 AI で雇用の問題が取り沙汰されるが、一方で、生成 AI は新たな雇用を生むものでもある。次世代計算基盤の構築によって、関係する人材にとっての雇用が拡大するような仕掛けが必要である。例えば、産業利用の強化に関わる開発で育成された若手研究者が、産業利用を行う企業側で雇用され、活躍するといったことが考えられる。また、AI の進展とシミュレーションの融合で新たなサイエンスを生み出すことで、新たな雇用を生み出せることにも期待したい。

また、計算科学の分野振興についても、「京」時代の HPCI 戦略プログラムやその後のポスト「京」研究開発枠重点課題の後は分野振興を含むプロジェクトがなく、一部の分野を除けば、分野別のコミュニティ活動が停滞している。人材育成でも述べた通り、今後は次期フラッグシップ開発期に入るので、その開発プロジェクトの中で、現在の「富岳」成果創出加速プログラムのようなフラッグシップでの新たな研究プロジェクトによる課題に加えて、計算科学の様々な分野振興や分野間交流を促すためのコード開発・活用を推進するための課題も設定されることが望まれる。その際には、従来の計算科学分野だけでなく、AI・データサイエンス分野や量子コンピュータ関連分野、高エネルギー・天文等の大規模実験・観測分野、IoT を活用する応用分野や産業界等のユーザなども対象として、新規分野の積極的な開拓をしていく必要があることは言うまでもない。

3. 次期フラッグシップシステム検討にあたって

3.1 アーキテクチャの方向性

前章では、次世代計算基盤全体の開発・整備・運用について取り上げたが、ここでは次世代計算基盤のうちの次期フラッグシップシステムについて取り上げる。フラッグシップシステムは、その卓越した計算性能を用いることで、サイエンスにおいてこれまで解けなかった問題を初めて解けるということに加えて、Society5.0 実現に貢献することが求められるため、データも計算も大規模・高速に対応できる必要がある。そのため、大規模シミュレーション向けと AI・データサイエンス向けが高いレベルで両立することが次世代計算基盤に求められる重要な方向性である。

その実現のため、2つのシステム FS では、電力制約のもとでデータ移動を効率よく行える、相対的に高いメモリバンド幅のアーキテクチャを目指している。サイエンスのブレークスルーを実現する道具としてのフラッグシップマシンが、AI for Science を加速できる性能を備えるためには、演算加速機構をフルに活用するような AI・データサイエンス向けの主要なアプリケーションにおいて、現状の AI・データサイエンス向けアーキテクチャでの実効性能と比較しても十分な性能向上が期待できることが望ましい。また、いずれのシステム FS でも、より高い実効性能を実現するには、ムーアの法則の終焉による演算性能の頭打ちや満たすべき電力性能等を踏まえると、演算加速機構が不可欠とされている。そのことを踏まえて、様々な分野を含むアプリケーション調査研究グループと協力して、ベンチマークプログラムの整備も進められている。なお、アプリケーション開発側としては、演算加速機構が加わることを含めて、どのようなアーキテクチャになるとしても、OS やプログラミング言語として標準的なものが採用されることを期待している。

ここで重要なことは、アーキテクチャの持つ性能だけでは実効性能は決まらず、アプリケーション開発側との協調設計、すなわちコデザインがあつて初めて期待されるような実効性能が得られるということである。これまでのフラッグシップ「京」、「富岳」は汎用プロセッサのみであったが、演算加速機構の性能を十分に引き出すためには、アプリケーションの中には大幅なアルゴリズム改変を含む開発が必要になるものが出てくる。開発にあたっては、アプリケーション開発側の自助努力だけでなく、アルゴリズムの工夫や実際のコーディングまで含めたフォローを、FS や開発期からコデザインとして実施することが不可欠である。これが HPC アプリケーション開発側にとってのユーザビリティの向上であり、次期フラッグシップ利用の成果最大化にとって必要不可欠な要素である。アプリケーション開発側も、新しいプログラミングモデルを取り入れる等、コデザインに積極的に参画して柔軟に取り組むことが必要である。

3.2 コデザイン

「富岳」におけるコデザインの振り返りから見出された課題とそれにもとづく提言をここでは取り上げる。「富岳」でのコデザインにおけるシステム開発上の課題としては、設計開始時点ですでにハードウェア設計上の自由度がある程度限られていたことが挙げられる。アプリケーションにとって使いやすいハードウェア機構を加えたいと要望があつた際に、コストの問題や電力制限の問題から、ベンダーとの協議で不採用となるケースもあつた。また、実際のハードウェアがない状態での設計

選択の問題もある。例えば、物理レジスタを含む Out of Order 実行のリソースを増やすことでアプリケーション性能が出しやすいことは定量的に示されていたが、電力性能が実際にどうなるかがまだわからない段階でどこまで増やすかを決めなければならず、より安全側、すなわち、電力消費を抑える側、アプリケーションから見れば性能が出にくい側の選択をすることになった。最終的な電力性能から考えれば、より増やすこともできた可能性はあるが、決定段階では難しい選択だったと言える。もしプロトタイプチップがあり、カーネルだけでなく通信等も含むアプリケーションを使って、より現実的な性能評価までできれば、より適切な選択ができる可能性がある。したがって、設計の方向性の大枠が決まるよりも前の段階、つまり FS における概念設計の段階で、アプリケーション性能を向上するために満たすべき項目を整理して示し、それがハードウェア的に実現可能であるかについて検討を進めて、実際の設計の方向性に反映させるようにすることが望まれる。また、電力効率などの制約が厳しい中での適切な性能評価のためには、プロトタイプチップを製造して、その上で、通信等も含んだ様々なアプリケーションでのベンチマークテストを行えるかどうかを鍵を握ると考えられる。プロトタイプチップの製造は予算的に大きな負担になり、時間もかかることが予想されるが、実機での電力効率とアプリケーション性能や性能を出す効率に大きく影響すること、すなわち運用時点の成果に対するコストパフォーマンスに直結することを考えれば、検討すべき課題の一つと考えられる。あるいは、性能予測のためのシミュレータで十分に高速かつ高信頼性のものを開発することがもしできれば、プロトタイプチップのコストを削減できる可能性はある。いずれにせよ、コデザインにおける設計評価をより精度良く実施できる方策を用意することで、より最適なシステムが開発されることが望まれる。

アプリケーション開発については、ターゲット・アプリケーションが限定されていたことが問題として指摘されている。選定をよりオープンにして、様々なアプリケーションを対象とすることが期待される。ターゲットアプリケーションとして様々なタイプを調べる上で、現在のシステム FS のアプリケーションチームに加えて、AI for Science のロードマップで取り上げられるアプリケーションや新計算原理 FS で取り上げられている疑似アニーリングや量子シミュレータのアプリケーション、さらには、HPCI や JHPCN の課題のアプリケーションで、FS のアプリケーションチームに含まれない分野やアルゴリズムも検討対象とすることが望まれる。

アプリケーション開発にとってのもう一つの課題は、システムソフトウェアやコンパイラ等の開発の継続性である。コデザインをしている段階まではベンダー協力のもとでコンパイラやライブラリ等の開発・性能向上ができていた。一方、共用開始後は、登録施設利用促進機関 RIST による高度化支援によるチューニング等の支援は行われたものの、コデザイン時と同様の支援という訳にはいかず、すべてのユーザのアプリケーション性能向上にはある程度限界があり、開発期間も必要となる場合があった。また、コデザインをしている中で、ハードウェアの制約の下ではあったがアプリケーションの性能向上のための知見が様々に得られていたが、共用開始後すぐに、アプリケーション開発者にそれらの知見が広く提供され、十分活用されたとは残念ながら言えない。なお、コデザインの成功とされるターゲットアプリケーションの性能目標達成は、コンパイラの開発やライブラリの最適化等によって、アプリケーションの性能がより出しやすくなったことはあるものの、各アプリケーション

の開発グループが、ベンダーや理研のソフトウェア開発チーム等と密接に連携して、アーキテクチャの特性に合わせたチューニング、場合によってはアルゴリズムの大幅な変更を含むコード改変を行なったことによることも明記しておく必要があると思われる。以上のことから、アプリケーション性能を引き出すためのコンパイラの開発・改良や、アーキテクチャの特性に合わせたアプリケーション開発に対する支援が、可能な限り様々なタイプのターゲットアプリケーションに対して行われるとともに、コデザイン段階からはもちろんのこと、運用開始後にも継続して行われることが不可欠である。

HPCの世界における性能向上率は依然として高く、次期フラッグシップシステムが、その計算性能で世界と勝負できるのは最初の3-4年程度までであると思われる。したがって、稼働開始当初からシステムの性能を十分に発揮させることが、次期フラッグシップシステムを利用した成果の最大化に直結するといえる。その意味で、コンパイラを含めたシステムソフトウェアの整備が、システムの稼働開始までに高い水準でなされることが不可欠である。

3.3 アプリケーション・ソフトウェアの開発について

アーキテクチャの検討は、予算、電力、技術、スケジュール等のハードウェアの設計を行う上での制約条件のもとでのトレードオフの中で進められるが、アプリケーション開発側からは、そのトレードオフの幅が、FSの段階である現時点でも、狭く限定されているように見える。例えば、総容量だけでなく、ノード単位やコア単位でのメモリ容量の制約等である。FSの概念設計段階では、まず、演算性能、メモリ性能等の各ハードウェアの特性に対して、どのような制約条件からどのようなトレードオフがあるのかを、できるだけ広い範囲のパラメータに対して示されることが望ましい。それらのトレードオフが、様々な特性をもつアプリケーションに対してその特性を再検討しつつ、どのような性能の出し易さ・出しにくさに関係しているのかを網羅的に整理し、提示されることが必要である。

また、詳細設計でアーキテクチャの特性が取捨選択された後には、トレードオフにおいてどの選択が行われ、何が捨てられたのかを明確に示すとともに、アプリケーション側でそれを補うためにどのような工夫が必要とされることになるかを示す性能向上指針が与えられることが不可欠である。特に、加速機構をはじめとして、一部あるいは多くのアプリケーションにとって、性能を引き出すための大幅なアルゴリズムの改変が求められる選択がなされる場合には、秘密保持契約に抵触しない範囲で、できるだけ早い段階で、定性的にでも情報が開示される必要がある。3.1節や3.2節でも述べた通り、加速機構への対応をはじめとして、大幅なアルゴリズム改変が必要となるアプリケーションについては、アルゴリズムの工夫や実際のコーディングまで含めた支援を、コデザインとして実施するとともに、運用開始後も継続した支援が不可欠である。

4. あとがき

本提言は、次世代計算基盤の調査研究チームの研究代表や「富岳」開発関係の方々、WG 委員等へのヒアリング、HPCI のユーザやシステム構成機関の意見を踏まえて、次世代計算基盤を利用した成果を最大化するために、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。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	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
副主査	小林 広明	東北大学大学院情報科学研究科
副主査	藤堂 眞治	東京大学大学院理学系研究科
委員	合田 憲人	国立情報学研究所
委員	青木 尊之	東京工業大学 学術国際情報センター
委員	伊藤 聡	公益財団法人 計算科学振興財団
委員	上田 修功	理化学研究所 革新知能統合研究センター
委員	塩原 紀行 ¹⁾	一般財団法人 高度情報科学技術研究機構
委員	田浦 健次朗	東京大学 情報基盤センター
委員	高木 亮治	宇宙航空研究開発機構 宇宙科学研究所 学際科学研究系
委員	野澤 剛二郎	清水建設株式会社 技術研究所
委員	森 雅博 ¹⁾	一般財団法人 高度情報科学技術研究機構

※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、富田浩文(HPCI コンソーシアム理事長)、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

1) 令和5年6月1日付けで森委員から塩原委員に交代

検討の記録

- | | |
|---|---------------------|
| 令和4年(2022年)5月10日(火) | HPCI コンソーシアム第72回理事会 |
| ・HPCI コンソーシアム新理事体制の確定 | |
| ・「将来のスーパーコンピュータの在り方の検討」業務体制の継続、及び 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、「調査検討WG」)の設置を決定 | |
| 令和4年(2022年)7月14日(木) | HPCI コンソーシアム第74回理事会 |
| ・令和4年度 HPCI コンソーシアム理事の業務分担の決定 | |
| 令和4年(2022年)10月5日(水) | HPCI コンソーシアム第75回理事会 |
| ・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4年(2022年)11月9日(水) | 第1回調査検討WG |
| ・次世代計算基盤に係る検討状況について文科省からヒアリング | |
| ・過去の検討の成果(HPCI コンソーシアムの提言)の確認 | |
| ・調査検討項目及びスケジュールを決定 | |
| 令和4年(2022年)12月22日(木) | 第2回調査検討WG |
| ・計算科学ロードマップについてのヒアリング | |
| ・「富岳」コデザインの実験について～システムの視点からのヒアリング | |
| ・「富岳」コデザインの実験について～アプリケーションの視点からのヒアリング | |
| 令和5年(2023年)1月12日(水) | HPCI コンソーシアム第76回理事会 |
| ・調査検討WGの進捗を確認 | |
| 令和5年(2023年)1月25日(水) | 第3回調査検討WG |
| ・次世代計算基盤に係る調査研究(近藤チームリーダー)の検討状況のヒアリング | |
| ・次世代計算基盤に係る調査研究(牧野チームリーダー)の検討状況のヒアリング | |
| ・現在のHPCIの課題の洗い出し | |
| 令和5年(2023年)2月17日(金) | 第4回調査検討WG |
| ・意見交換会の計画について | |
| ・中間取りまとめ案について | |
| 令和5年(2023年)3月6日(月) | 意見交換会 |
| ・中間取りまとめについての意見聴取 | |
| 令和5年(2023年)3月13日(月) | HPCI コンソーシアム第77回理事会 |
| ・意見交換会の報告 | |

- | | |
|----------------------|---|
| 令和5年(2023年)3月15日(水) | 第5回調査検討WG |
| | ・中間取りまとめの確認について |
| 令和5年(2023年)4月24日(月) | HPCIコンソーシアム第78回理事会 |
| | ・令和4年度WG報告書(提言に向けた中間取りまとめ)として提出することの承認 |
| | ・令和5年度WG年間計画の承認 |
| 令和5年(2023年)5月8日(月) | 第1回調査検討WG |
| | ・令和4年度WG報告書(提言に向けた中間取りまとめ)の確認 |
| | ・令和5年度WG年間計画の確認 |
| 令和5年(2023年)5月19日(金) | HPCIコンソーシアム総会 |
| | ・令和4年度WG報告書(提言に向けた中間取りまとめ)の報告 |
| 令和5年(2023年)6月13日(火) | 第2回調査検討WG |
| | ・次世代計算基盤に係る調査研究(埜チームリーダー)の検討状況のヒアリング |
| | ・シームレス化(合田委員)に関するヒアリング |
| 令和5年(2023年)7月5日(水) | HPCIコンソーシアム第79回理事会 |
| | ・調査検討WGの進捗を確認 |
| 令和5年(2023年)7月25日(火) | 第3回調査検討WG |
| | ・クラウド(RCCS・庄司運用部門長)に関するヒアリング |
| | ・次世代計算基盤に係る調査研究(天野チームリーダー)の検討状況のヒアリング |
| 令和5年(2023年)8月25日(金) | 第4回調査検討WG |
| | ・AI・データサイエンス(上田委員)に関するヒアリング |
| 令和5年(2023年)9月27日(水) | 第5回調査検討WG |
| | ・産業利用産応協(みずほ池田業務推進役、富士通松本エキスパート)に関するヒアリング |
| | ・産業利用中小企業(伊藤委員)に関するヒアリング |
| | ・産業利用バイオ創薬(バイオグリッド関西志水理事)に関するヒアリング |
| 令和5年(2023年)10月4日(水) | HPCIコンソーシアム第80回理事会 |
| | ・令和5年度WG報告書案(提言案)のまとめ方の確認 |
| 令和5年(2023年)10月17日(火) | 第6回調査検討WG |
| | ・令和5年度WG報告書案(提言案)の検討(1) |
| 令和5年(2023年)11月20日(月) | 第7回調査検討WG |

- ・令和5年度 WG 報告書案(提言案)の検討(2)

令和5年(2023年)12月11日(月) 第8回調査検討WG

- ・令和5年度 WG 報告書案(提言案)の検討(3)

令和5年(2023年)12月20日(水) HPCIコンソーシアム第81回理事会

- ・令和5年度 WG 報告書案(提言案)の確認

令和6年(2024年)1月15日(月) 意見交換会

- ・令和5年度 WG 報告書案(提言案)についての意見聴取

令和6年(2024年)1月18日(木) 第9回調査検討WG

- ・令和5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

令和6年(2024年)2月26日(月) 第10回調査検討WG

- ・令和5年度 WG 報告書案(提言案)についての WG 委員からの意見反映、次年度に向けた検討事項

令和6年(2024年)3月11日(月) HPCIコンソーシアム第82回理事会

- ・令和5年度 WG 報告書案(提言案)の確認