

今後のHPCIシステムの構築と その利用に関する 基本的な考え方について

一般社団法人 HPCIコンソーシアム
理事長 加藤千幸

目次

第1章 前書き

第2章 HPCIの達成目標とHPCIの推進について

- 2.1 HPCIが達成しようとしているビジョン
- 2.2 HPCIの推進方法

第3章 HPCIのこれまでの成果と顕在化した課題

- 3.1 フラッグシップ計算機「京」の開発と運用
- 3.2 第二階層計算資源の運用
- 3.3 利用者拡大・利用支援
- 3.4 アプリケーションの開発・利用・普及
- 3.5 コミュニティ形成と人材育成

第4章 HPCIで利用されるハードウェアとソフトウェアの開発・整備・運用に関する基本的な考え方

- 4.1 HPCIシステムが提供するハードウェア
 - (ア)フラッグシップ計算機
 - (イ)第二階層計算資源
 - (ウ)共有ストレージ
 - (エ)高速ネットワーク
 - (オ)その他の計算資源

目次

4.2 HPCIシステムで利用されるソフトウェア

(ア)システムソフトウェア

(イ)ソフトウェアツール

(ウ)国プロで開発されるアプリケーション・ソフトウェア

(エ)商用アプリケーション・ソフトウェア

(オ)オープン・ソースのアプリケーション・ソフトウェア

(カ)利用者(利用する機関)自らが開発したアプリケーション・ソフトウェア

第5章 HPCI利用課題の実施の在り方

5.1 利用課題に関する基本的な考え方

5.2 フラッグシップ計算機における実施課題

5.3 第二階層計算資源における実施課題

5.4 産業利用に関する基本的な考え方

5.5 HPCIの利用料金に関する基本的な考え方

5.6 利用者支援について

第6章 あとがき

各章の要旨・要点と報告書原文(抜粋)

第1章 前書き

- 我が国の革新的ハイパフォーマンス・コンピューティング・インフラストラクチャー（以下、HPCI）はフラッグシップ計算機「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、すでに7年あまりの年月が経過している。「京」は令和元年（2019年）8月16日を以って、ほぼ7年間に及んだ共用が終了された。一方、アプリケーションで「京」の最大100倍の性能を発揮することが既に確認されている、次期フラッグシップ計算機「富岳」の共用が早ければ令和3年（2021年）4月に開始される見通しである。
- HPCIの運用が開始されてから7年間以上が経過し、その間に、人工知能（AI）に代表されるデータ科学の台頭など、高速計算技術（HPC）を取り巻く環境にも大きな変化が生じている。また、フラッグシップ計算機停止期間における、計算科学の持続的振興の在り方など、幾つかの課題も顕在化している。次期フラッグシップ計算機「富岳」の運用開始が早ければ1年後に迫っている中、HPCIがこれまで以上に大きな成果を創出し、我が国のHPCIが長期にわたり持続的に発展を続けるためには、これまでのHPCIの運用実績、ならびに、それによる成果を検証するとともに、今後のHPCIの構築と運用に関する基本的な考え方を改めて整理し、明確にしておく必要がある。
- 本報告書は上記の目的を達成するために、HPCIの運営主体等からのヒアリング結果などに基づき、HPCIコンソーシアムが主体的に検討し、ユーザコミュニティの意見も集約した上で、取り纏めたものである。

第1章 報告書原文(抜粋)

【第1段落(1ページ)】

「HPCIはフラッグシップ計算機を中核として、全国の大学情報基盤センター等のスーパーコンピュータ(以下、スパコン)を高速ネットワークで接続するとともに、10 PB(ペタバイト)級の大容量ストレージを全国に2カ所配置し、コミュニティにおけるデータの共有環境も含めて、利用分野を特定しない、延べ13,000人以上の多くのユーザに計算資源を提供してきた。」

【第3段落(1ページ)】

「HPCIの運用が開始されてから7年間以上が経過し、その間に、人工知能(AI)に代表されるデータ科学の台頭など、高速計算技術(HPC)を取り巻く環境にも大きな変化が生じている。また、フラッグシップ計算機停止期間における、計算科学の持続的振興の在り方など、幾つかの課題も顕在化している。」

【第3段落(1ページ)】

「HPCIがこれまで以上に大きな成果を創出し、我が国のHPCIが長期にわたり持続的に発展を続けるためには、これまでのHPCIの運用実績、ならびに、それによる成果を検証するとともに、今後のHPCIの構築と運用に関する基本的な考え方を改めて整理し、明確にしておく必要がある。」

第2章 HPCIの達成目標とHPCIの推進について

- 多様なユーザニーズに応えることができる最先端の計算機環境を提供することにより、画期的な研究成果を創出し、科学技術の発展や産業競争力強化に資するとともに、人材の育成やスパコン利用の裾野の拡大に貢献する。
- AIやデータ科学がHPCと融合することにより飛躍的に発展していくためには、大規模なデータを多様なユーザが共有し、新しい方法論の提案や知見の発見のために活用することが重要であり、高速ネットワークで相互接続された大容量ストレージが具備されたHPCI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。
- HPCに関連する機関だけではなく、AIやデータ科学に関連する、さらに多くの機関がHPCIコンソーシアムの活動に参加することにより、より幅広いユーザ層から集約された意見に基づいて、HPCIコンソーシアムがHPCIの運営を主導し、フラッグシップ計算機の開発・運用主体である理化学研究所、第二階層計算資源を提供している全国の大学情報基盤センター等、フラッグシップ計算機および第二階層計算資源の利用者選定や利用支援を行う登録施設利用促進機関等の三者が一体となって、HPCIを運営していくべきである。

第2章 報告書原文(抜粋)

【2.1節第2段落(2ページ)】

「HPCIがその構築と運用を通じて達成しようとしているビジョンは、多様なユーザーニーズに応えることができる最先端の計算機環境を提供することにより、画期的な研究成果を創出し、科学技術の発展や産業競争力強化に資するとともに、人材の育成やスパコン利用の裾野の拡大に貢献することである。」

【2.1節第2段落(2ページ)】

「高速ネットワークで相互接続された大容量ストレージが具備されたHPCI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。」

【2.2節第2段落(3ページ)】

「AIやデータ科学に関連する、さらに多くの機関がHPCIコンソーシアムの活動に参加することにより、より幅広いユーザ層から集約された意見に基づいて、HPCIコンソーシアムがHPCIの運営を主導していくことが望まれる。そのうえで、HPCIシステムの中核を成すフラッグシップ計算機の開発・運用主体である国立研究開発法人理化学研究所(以下、理化学研究所)、第二階層計算資源を提供している全国の大学情報基盤センター等、フラッグシップ計算機および第二階層計算資源の利用者選定や利用支援を行う登録施設利用促進機関(以下、登録機関)等の三者が一体となって、HPCIを運営していくべきである。」

第3章 HPCIのこれまでの成果と顕在化した課題

- 「京」は超並列計算の実行環境を提供し、汎用性、実効性能ともに優れたスパコンとして、大きな成果を創出した。また、コンソーシアム型「京」産業利用を設定するなどの工夫もあり、利用者は拡大し、「京」は延べ11,000人以上に利用された。一方、ISAの問題もあり、商用アプリケーションの移植はあまり進まず、「京」の商用機の普及も限定的であった。
- 第二階層計算資源は、フラッグシップ計算機とともに、HPCIシステムの中核的な計算資源を成し、多様な計算機アーキテクチャにおけるベンチマーク環境を提供し、「京」とともに、多くの成果を創出した。一方、第二階層計算資源を提供している大学情報基盤センター等が、より積極的にHPCIの運営に参画できる仕組みを構築する必要がある。
- アプリケーションの高度化支援などにより、延べ13,500人あまり、年間平均2,000人近くの研究者等がHPCIを利用したが、新規利用者数はあまり拡大しなかった。
- 「HPCI戦略プログラム」(総額100億円)により、2年連続してゴードンベル賞を受賞するなど、卓越する数々の成果が創出された。また、アプリケーション・ソフトウェアの普及、アウトリーチ活動などを通して、強固なユーザコミュニティが形成され、現在のユーザコミュニティの原型が形成された。

第3章 報告書原文(抜粋)

【3.1節第2段落(4ページ)】

「1,000以上の計算ノードを利用したジョブの実行が全体の6割から7割を占めるなど、「京」がその開発・運用の目標とした「超並列計算機環境の提供により、従来できなかった規模の計算の実行を可能にする。」というミッションは十分に果たすことができた。」

【3.1節第7段落(7ページ)】

「超並列計算機環境を提供し、従来は実行不可能であった大規模な計算を実現することにより、産業利用も含めて大きな科学的・社会的な成果を創出するという目標は達成されたが、フラッグシップ計算機を独自に開発し、それに関連したハードウェア技術も普及させるという波及効果は期待どおりではなかった。」

【3.1節第8段落(7ページ)】

「フラッグシップ計算機「京」には、成果の公開が義務付けられないが、利用料金を徴収する「個別利用課題」が設定され、利用料収入は「京」の運用経費の一部に充てられた。しかし、「京」の運用経費に較べて利用料収入はごくわずかなものに留まった。(中略)開発・整備に加えて、多額の運用経費が必要となるフラッグシップ計算機を中核として、長期にわたり持続可能なものとしてHPCI計画を推進していくためには、運用経費の一部を利用料金により充当するという考え方もあり得る。」

【3.2節第2段落(9ページ)】

「HPCの本格的な実用化を検討している企業に対して、第二階層計算資源がHPCの利用効果を定量的に予測し得る、多様な計算機アーキテクチャにおけるベンチマーク環境を提供したという意義は大きい。」

第3章 報告書原文(抜粋、つづき)

【3.2節第6段落(11ページ)】

「第二階層計算資源を提供している各情報基盤センター等が、HPCI計画の推進に対して、より主体的・積極的に参画できる体制を構築していくことが、長期的に見て、我が国のHPCI計画が強い求心力を維持しながら推進されるために必須であると考えられる。」

【3.3節第6段落(13ページ)】

「利用者が固定化する、すなわち、同一の研究者等が異なる利用研究課題のためにHPCIシステムを長期にわたり利用するという傾向が見られ、新規利用者数の増加という観点では必ずしも利用者が拡大したとは言えない状況にある。」

【3.4節第1段落(13ページ)】

「平成23年(2011年)4月から平成28年(2016年)3月の5年間にわたり、フラッグシップ計算機「京」の実証研究事業である「HPCI戦略プログラム」が総額100億円以上の研究経費を掛けて実施され、フラッグシップ計算機「京」に対応した高度なアプリケーション・ソフトウェアにより、卓越する数々の成果が創出された。(中略)本事業のもう一つの柱であった、「計算科学技術推進体制の構築」により、人材育成、開発されたアプリケーション・ソフトウェアの普及、アウトリーチ活動などが実施され、各戦略分野において、強固なユーザコミュニティが形成され、現在のユーザコミュニティの原型が形成された。」

【3.5節第3段落(17ページ)】

「理化学研究所や高度情報科学研究機構を中心とした上記のような人材育成事業の成果により、コミュニティを繋ぐネットワークが形成され、これまでは各大学や大学情報基盤センター等で個別に実施されていた計算科学人材の育成事業が全国的に展開されるようになったことは特筆に値する。」

第4章 HPCIで利用されるハードウェアとソフトウェアの 開発・整備・運用に関する基本的な考え方

- フラッグシップ計算機は重要な研究基盤として、今後とも開発・整備すべきであるが、画期的な成果を創出するためには、集約的に開発・整備すべきである。また、調達は不可能なため、独自開発が前提となるが、アプリケーションとの協調開発が必須という意味からも、独自に開発する必要がある。ただし、ハードウェア技術や商用機の普及・商用アプリケーションの移植性にも十分に配慮して開発を進めるべきである。
- 第二階層計算資源は、単に、フラッグシップ計算機を補間するためではなく、多様なユーザーニーズに応えることができる、最新鋭のスパコン環境として、今後も整備していくべきである。また、先導的な第二階層計算資源による研究成果をフラッグシップ計算機の開発計画に反映させていくべきである。
- 国プロによるアプリケーション・ソフトウェアの開発にあたっては、開発・実証・維持・普及も含めたプロジェクト設計をすべきである。
- フラッグシップ計算機や第二階層計算資源に、商用のアプリケーションの整備が進むような連携体制を構築すべきであり、また、ユーザーニーズが高いオープンソースソフトウェアの整備も進めるべきである。

第4章 報告書原文(抜粋)

【4.1節(ア)第1段落(19ページ)】

「我が国は今後も科学技術発展のための重要な研究基盤としてフラッグシップ計算機の設備整備を続けていくことが妥当である。」

【4.1節(ア)第2段落(19ページ)】

「実行可能な最大計算規模の観点から、フラッグシップ計算機は中規模なスパコンを複数台開発・整備するのではなく、最大規模のものを開発し、設備整備すべきである。」

【4.1節(ア)第5段落(20ページ)】

「今後のフラッグシップ計算機の開発整備にあたっては、半導体技術等の進歩速度にも配慮しながら、できる限りフラッグシップ計算機の端境期が生じないような開発・整備方法を検討すべきである。」

【4.1節(ア)第6段落(20ページ)】

「アプリケーション開発者がスパコンの設計段階からハードウェア開発に積極的に関与できる環境がなければ、フラッグシップ計算機の性能を十分に引き出すことは不可能である。ハードウェア(スパコン)本体とアプリケーションの協調的开发を可能にするために、フラッグシップ計算機を独自に開発することが必須となる。」

【4.1節(ア)第8段落(21ページ)】

「フラッグシップ計算機を独自に開発する場合、それを利用した成果の創出だけでなく、その波及効果、すなわち、CPUやインターコネクト技術、あるいはフラッグシップ計算機の商用機が広く普及することが、我が国が今後もフラッグシップ計算機の独自開発を続け、世界をリードしていくためには重要となる。」

第4章 報告書原文(抜粋、続き)

【4.1節(ア)第10段落(22ページ)】

「文部科学大臣が告示する基本方針に沿って、登録機関と連携しながら「富岳」の利用環境の構築・整備・運用ルール具体化を行い、具体化されたルールに従って「富岳」を運用すべきである。」

【4.1節(イ)第3段落(23ページ)】

「第二階層計算資源は情報基盤センター等が独自のポリシーに基づいて整備・運用をしている計算資源の一部がHPCIの計算資源として提供されたものであるため、その最大の特長は提供されている計算資源の多様性にある。特に、情報科学分野やAI分野への展開も含めて、HPCIの利用分野の拡大や利用者の裾野拡大に対して大きく貢献していくことが期待される。」

【4.1節(イ)第3段落(24ページ)】

「第二階層計算資源全体で見た場合、フラッグシップ計算機よりもはるかに短いインターバルで最先端のスパコンが設備整備されていることになり、第二階層計算資源は多様なユーザーニーズに応えることができる、最新鋭のスパコンの利用環境を提供しているということも重要な意義を持っている。以上のような観点から、第二階層計算資源は今後もHPCIの重要な計算資源として整備・運用をしていくべきである。」

【4.1節(エ)第1段落(25ページ)】

「AIやデータ科学の時代を迎え、ネットワーク基盤の重要性は従来にも増して高くなることが予想される中、共用ストレージも含めて、HPCI利用者の具体的な利用形態を早期に検討し、適切な予算措置を実施すべきである。」

第4章 報告書原文(抜粋、続き)

【4.2節(ウ)第1段落(27ページ)】

「「富岳」の共用開始よりも1年以上前の令和2年(2020年)3月現在、「京」の100倍以上の性能を発揮できるアプリケーションが既に複数本開発されるなど、コデザインにより大きな成果が得られ、「富岳」の共用開始と同時にその本格利用により卓越した成果が創出されることが期待されている。」

【4.2節(ウ)第1段落(27ページ)】

「フラッグシップ計算機の開発は今後もコデザインにより実施すべきあり、そのためにもフラッグシップ計算機を独自に開発することは必須の条件となる。」

【4.2節(ウ)第2段落(28ページ)】

「実証後の当該アプリケーションの維持・普及の在り方も含めた、長期的な観点からプロジェクトを設計して当該事業を実施すべきである。」

【4.2節(エ)第1段落(28ページ)】

「フラッグシップ計算機や大学情報基盤センター等のスパコンに対する商用アプリケーション・ソフトウェアの移植作業はソフトウェアの開発者(あるいは、ソフトウェアベンダー等)が行うことになるが、それを加速するための取り組みを、国からの適切な支援策の下、フラッグシップ計算機や第二階層計算資源を提供している大学情報基盤センター等が実施することが望まれる。」

第5章 HPCI利用課題の実施の在り方

- 画期的な研究成果の創出し、科学技術の発展や産業競争力強化のために優れた研究課題をフラッグシップ計算機においては実施すべきである。産業利用に関しても、優れた「研究開発」のために利用すべきである。一方、フラッグシップ計算機の利用は多くのユーザに対して開かれたものであるべきであり、比較的小規模な計算を実施すること目的とした課題などを、情報基盤センター等の利用料金も参考にした上で、適切な利用料金が設定された、有償の課題として実施すべきである。
- 第二階層計算資源における課題の実施にあたっては、多様性・運用の機動性を活かした実施を検討すべきである。また、HPCIの新規利用分野の開拓や利用者の裾野拡大に対する貢献度が評価される課題を積極的に選定すべきである。
- フラッグシップ計算機や第二階層計算資源を利用した、きわめて優れた課題に関しては、それによる成果の創出が加速されるように補助金等の交付により無償の課題として実施されるべきである。
- ユーザ支援事業やアプリケーションの整備事業は、HPCIの利用拡大のために必須であるため、今後も積極的にこれを実施すべきである。その際、各分野におけるコミュニティの特性にも配慮し、また、将来的には長期的な事業としてこれが実施されるようにすべきである。

第5章 報告書原文(抜粋)

【5.2節第4段落(31ページ)】

「フラッグシップ計算機は基本的には他の計算機では実施できない研究開発に利用され、フラッグシップ計算機上では、科学技術の発展や産業競争力の強化などに対して大きな寄与が期待される課題が実施されるべきである。」

【5.2節第6段落(32ページ)】

「一方、国費により開発・整備された研究基盤施設はその利用に関心を持つ、多くのユーザに対して開かれた利用を促進すべきであり(中略)、より多くの利用者に対してその利用を拡大できるような課題の審査と選定を実施すべきである。(中略)たとえば、フラッグシップ計算機の計算資源の内の一定量に関しては、課題の応募を随時受け付け、1週間から2週間程度の短期間の簡略化された審査により、上記のような利用課題を選定することも検討すべきである。」

【5.3節第1段落(32ページ)】

「第二階層計算資源の最大の特徴は提供されている資源の多様性とその運用の機動性にあることを踏まえ、実施される利用課題に関しては、HPCIの新規利用分野の開拓や利用者の裾野拡大に対する貢献度も重要な観点として、課題審査と審査結果に基づく課題選定を実施すべきである。」

【5.4節第2段落(33ページ)】

「2.1節で言及したHPCIの達成すべきビジョンを踏まえ、フラッグシップ計算機を含めた今後のHPCI計算資源の産業利用に関しても、優れた「研究開発」のために利用されるべきである。」

第5章 報告書原文(抜粋、続き)

【5.5節第1段落(34ページ)】

「この目的を達成するために、HPCIの利用課題の内、産業利用課題も含めて、特に優れた課題の利用料金は補助金等の国費により充当されるべきである。」

【5.5節第2段落(34ページ)】

「フラッグシップ計算機の利用は多くのユーザに対して開かれたものであるべきであり、5.2節に言及したような、比較的小規模な計算を実施すること目的とした課題などは、情報基盤センター等の利用料金も参考にした上で、適切な利用料金が設定された、有償の課題として実施されるべきである。」

【5.5節第2段落(34ページ)】

「産業利用の拡大を図るために、成果の公開を義務付けなかったり、一定の期間、成果の公開を延期することを認めたりする有償の課題も設定すべきである。さらに、これらの有償利用課題に関しては、フラッグシップ計算機の筐体毎の占有利用を認めたり、計算機の繁忙期においても短時間以内に計算が実行されることを保証したりする課題なども検討すべきである。」

【5.5節第3段落(34ページ)】

「上記のような有償の課題利用が拡大することは、より幅広いユーザ層に対する利用の拡大に加えて、フラッグシップ計算機を含めたHPCI計画の安定的・持続的な推進にも繋がる
18
ことが期待される。」

第6章 後書き

- 本報告書はHPCI計画推進委員会における決定を受けて、HPCIの運営事業の実施機関の一つである、高度情報科学技術研究機構の下に設置された、「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、HPCIコンソーシアムが中心となり、調査・検討した結果を取り纏めたものである。
- 本報告書を取り纏めるにあたり、フラッグシップ計算機の開発・運用機関である理化学研究所やHPCIの運営事業の実施機関である高度情報科学技術研究機構などの関係機関にヒアリングしたり、「京」の中間評価結果を改めて精査したりすることにより、これまでの7年間あまりのHPCIの運用実績やそれにより得られた成果、および顕在化した課題を検証した。
- 本報告書の中でも言及したように、AIに代表されるデータ科学の台頭という、新たな潮流が現れている中で、HPCIを具体的にどのように整備・運用していくべきか、あるいは、「富岳」の次のフラッグシップ計算機の具体的な開発の在り方などに関しては、残念ながら検証できるデータと検討する時間が十分ではなかったため、本報告書では言及していない。今後も世の中の変化やHPCIの運用実績や成果などを定期的に自己検証し、HPCIの持続的発展に繋げていく必要がある。