

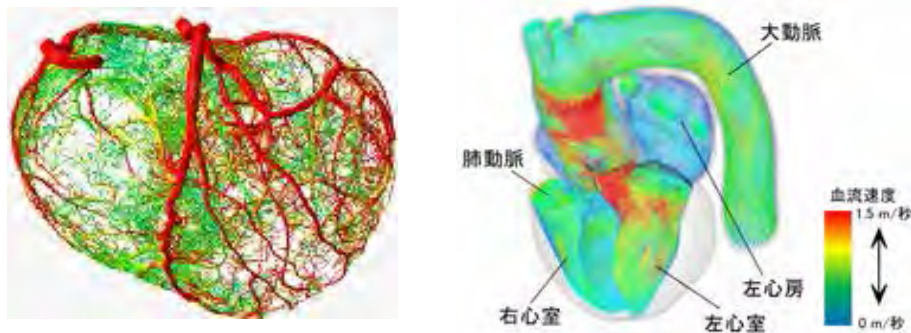
エクサスケール・スーパーコンピュータ 開発プロジェクト(仮称)

平成25年10月10日
文部科学省研究振興局参事官付(情報担当)
計算科学技術推進室

スーパーコンピュータ「京」の主な成果事例

「京」の活用により、心臓全体の挙動を精密に再現することによる心臓病の治療法の検討等がなされ、また、地震・津波の被害の予測の精度が大幅に向上し、地方自治体の災害対策に活用されるなど、画期的な成果をあげつつある。

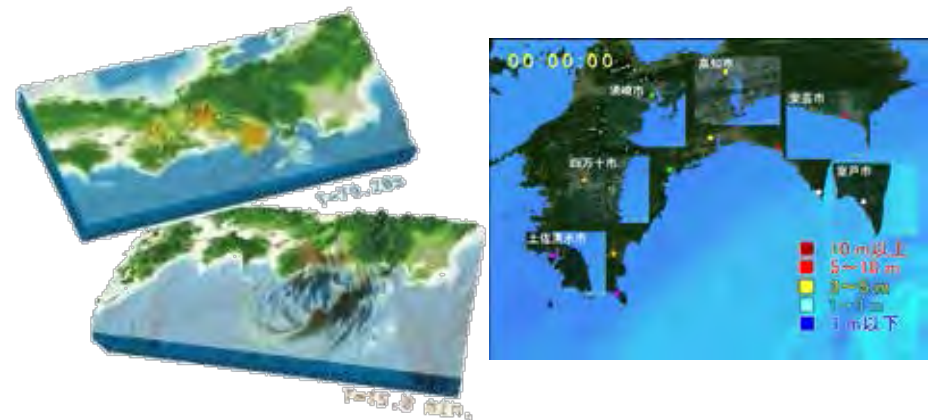
● 心臓シミュレーション (心臓病治療等への応用)



分子レベルからの心臓シミュレーション

分子レベルから心臓全体を精密再現することにより、心臓の難病のひとつである**肥大型心筋症の病態が解明された**。これにより、今後、**治療法の検討や薬の効果の評価**への貢献が期待される。

● 津波シミュレーション (地震・津波の予測への応用)



シミュレーションによる地震・津波の被害予測

高精度な地震・津波シミュレーションの実現により、災害に強いまちづくりや、きめ細かなハザードマップ作成等の防災対策等への貢献が期待される。

これまでの検討経緯

技術的事項の検討

今後のHPC技術の研究開発の検討WG (平成23年4月～7月)

「今後のハイパフォーマンス・コンピューティング技術の研究開発について」(平成23年7月)

アプリケーション作業部会 コンピュータアーキテクチャ・コンパイラ・システムソフトウェア作業部会 (平成23年7月～平成24年2月)

「今後のHPCI技術開発に関する報告書」(平成24年2月)

将来のHPCIシステムのあり方の調査研究 略称：FS (平成24年度から2カ年)

- ・公募によりシステム設計研究チームとアプリケーションソフトウェアチームを選定。
- ・システム設計研究チームでは、技術動向調査、システム設計研究、システムソフトウェアの検討等を行う。
- ・アプリケーションソフトウェアチームでは、サイエンスロードマップの策定、評価用アプリの抽出、それを用いたシステムの評価等を行う。
- ・これらの活動をもとに、**5～10年後の我が国のHPCIシステムに必要な技術的知見を獲得**する。

【選定結果】

分野	主管事業実施機関		
アプリ	理化学研究所		
システム設計	東京大学	筑波大学	東北大学

政策に関する検討

今後のHPCI計画推進のあり方に関する検討WG (平成24年2月～)

- ・研究振興局長の諮問会議であるHPCI計画推進委員会の下に設置
- ・今後10年程度を見据え、国内外の計算科学技術の動向、HPCIシステム構成のあり方、HPCI全体のネットワークや利用体制のあり方、今後の研究開発のあり方等について調査検討を実施。
- ・平成25年6月25日に**中間報告**がとりまとめられ、**2020年頃までにエクサスケールコンピューティングの実現を目指すことが適当とされた。**
- ・平成26年3月頃を目途に**最終報告**をとりまとめる予定。

今後のHPCI計画推進のあり方に関する検討WG システム検討サブWG (平成25年6月～8月)

- ・**中間報告**において、リーディングマシンの研究開発の方向性として、開発主体候補(理研)において、期待される成果、必要なスペック、要素技術等に関するイメージを明らかにした上で、フラッグシップシステムの具体的な方向性等を検討。
- ・開発主体候補からのヒアリングを実施。
- ・平成25年8月30日**報告書要旨公表**

WGの議論に反映



プロジェクトの目的、国として取り組む意義・必要性

目的

第四期科学技術基本計画において、**国家基幹技術として位置づけられている世界最高水準のスーパーコンピュータを国として戦略的に開発・整備**し、科学技術イノベーションに適した環境を創出することで、我が国をとりまく様々な**社会的・科学的課題の解決を先導し、科学技術の振興、産業競争力の強化、安全・安心の国づくり等**に貢献する。

概要

我が国をとりまく様々な社会的・科学的課題の解決を見据え、2020年頃までに、それに要求される**演算性能1エクサフロップス（「京」の約100倍）レベルのスーパーコンピュータ（以下「ポスト「京」という。）を開発・整備**するとともに、**ポスト「京」を最大限活用するためのアプリケーションを開発**する。その際、「京」の開発で培った我が国の技術の継承、システムソフトウェアやアプリケーションの開発、人材育成等の観点で国内開発に利点があることから、**プロセッサ等の中核となる技術については新たに自主開発することを基本方針**とするとともに、迅速な成果創出を図るため**ハードウェアとアプリケーションの協調設計（Co-design）を推進**する。

必要性・国として取り組む意義

最先端のスーパーコンピュータは、演算性能が桁レベルで飛躍することで、理論・実験に並ぶ科学技術の第3の手法である**シミュレーション利活用の全く新しい地平を拓き、人類が経験したことのない先端的成果を実現する研究開発基盤**である。また、ものづくり現場や資源・エネルギー問題、防災・減災等の対策においてもシミュレーションの活用が進みつつある中で、**産業競争力の強化や国家の存立基盤の維持という観点からも重要なツール**である。さらに、ますます複雑化する社会的・科学的課題の解決に向けて、異分野や産学官の間を結びつけ統合させていくツールやビッグデータの処理・解析のツールにもなり得るので、**イノベーション創出の観点からも今後ますますその重要性は高まっていく**。

国が最先端のスーパーコンピュータを開発することは、**国内産業へ様々な波及効果を及ぼす**のみならず、我が国の存立基盤にも関係する重要技術を国内に維持するという点で**国家の総合的安全保障にも影響する**。我が国は「京」プロジェクトで、高性能なプロセッサやネットワーク、優れた省電力機構等の最先端技術を開発したことで、国全体の産業競争力を強化しただけでなく、技術を生み出す人材を国内で育成・確保し**今なおスーパーコンピュータを自主開発できる能力を保っている**。しかし、**国として最先端のスーパーコンピュータの開発から手を引くことは、国際的な開発競争の中での頭脳流出はもとより、将来のスーパーコンピュータ開発を担う人材の枯渇を引き起こすおそれがある**。

国際的にも、国主導の開発・利用が積極的に進められており、米中欧をはじめとして**2020年頃のエクサスケールコンピューティングの実現を見据えて研究開発を進めている**。

以上のように、最先端のスーパーコンピュータは、**国家の競争力の源泉、国力を象徴する技術**であり、さらに2020年頃を目指した開発競争が激化する国際的背景も含め、第四期科学技術基本計画において国家安全保障・基幹技術に位置付けられているところ、**国として本プロジェクトに主導的に取り組む意義・必要性は十分にある**。

ポスト「京」の国家的な位置付けについて

最先端のスーパーコンピュータは、**先端的成果を実現する研究開発基盤**であるとともに、**我が国の存立基盤に影響するもの**である。また、その自主開発は、我が国の存立基盤に影響する重要技術を国内に維持するという点で**国家の総合的な安全保障に密接に関わる**。

よって、**技術的ブレークスルーの原動力**となるとともに、**国益を守ることに重要な貢献を果たすもの**であり、**国家の競争力の源泉、国力を象徴する技術**である。



そのような、国が主導する一貫した推進体制の下で実施され世界をリードする人材育成にも資する長期的かつ大規模なプロジェクトにおいて、**国家の総合的な安全保障の観点も含め経済社会上の効果을最大化するために基本計画期間中に集中的な投資が必要なもの**を、第三期科学技術基本計画においては「**国家基幹技術**」と位置付け、国として協力に推進することとした。



第四期科学技術基本計画においてもその考え方は踏襲されており、**我が国が国際的な優位性を保持し、安全な国民生活を実現していくために、国自らが長期的視点に立って、継続的に、広範囲かつ長期間にわたって研究開発を推進していくべきもの**を「**国家安全保障・基幹技術**」とし、国として、国家存立の基盤に関わる研究開発と位置付けて強力に推進することとした。

ポスト「京」及びそれを最大限活用するアプリケーションは、「**国家安全保障・基幹技術**」に位置付けられている「**革新的ハイパフォーマンスコンピューティング技術**」であり、第四期科学技術基本計画や科学技術イノベーション総合戦略に基づき、**国として責任を持って開発する必要がある**。

社会的・科学的課題解決を先導するHPCIシステム開発

2020年から運用するHPCIシステムの設計指針

• 計算科学ロードマップ（9月20日付公開）に基づくマシン設計

- アプリFSにおいて社会的・科学的課題および必要とする計算資源量について精査を行っている
- ロードマップで見えてくる使い方は
 - 大規模、精密、長時間発展といったcapability computingのニーズ
 - 大規模計算を解く。例えば、京で8万ノード使わないと動かさないプログラム。
 - 複雑な現象を対象とした課題におけるensemble computingのニーズ
 - 気象、気候シミュレーションなど多くの分野では、初期パラメータを変えてシミュレーションし、それらの結果から予測値を得る。例えば、京で80ノード使って動作するプログラムを1000個同時に動かす。このような計算はcapacity computingとも呼ばれる
 - Big data computing、社会科学シミュレーションのニーズ
- これらの要求に応えるシステムの設計が前提
 - Capability/Ensemble computingを両立させるには一か所にcapability computingできるマシンを設置する必要がある

• Sustainable System

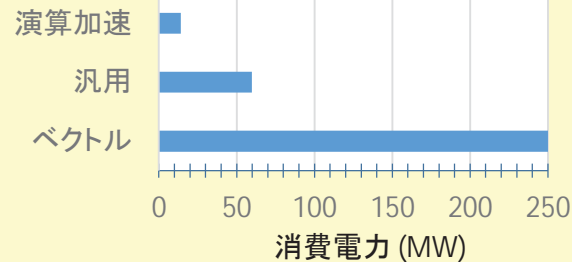
- 京の後継機として京の資産が受け継げる、そして、将来の計算機システム発展動向を見据えたシステム
- 2020年ころの世界の最先端システム、競争力、競合力あるシステム
- 将来の我が国のHPCの発展につながるシステム

• TCO(Total Cost of Ownership)-aware System

- 計算科学ロードマップで必要とされる計算資源量に対して最大限応えられる、低電力、ソフトウェア高移植性、高耐故障性を有するシステム設計

今後の HPCI システムの技術的課題 (1/2)

汎用・ベクトルの限界：性能電力比問題



2020年頃から運用可能な最先端技術を用いて開発した1エクサマシンの電力は演算加速のみを使わない限り50MW以上の電力が必要 (FSの中間結果から推定)。

演算加速部の限界：適用可能アプリの制限

	演算加速	汎用	ベクトル
データ構造	規則構造	不規則・非構造	不規則・非構造
B/F	4	0.2~0.4	1~2
メモリ容量	小 (~1GB/node ~1PB/total)	中 (~64GB/node ~10PB/total)	中 (~64GB/node~ 10PB/total)

演算加速：規則的に配列をアクセスして計算するプログラムでかつ必要とされるメモリ容量が小さい場合に有効 (対象アプリが限定)

汎用&ベクトル：従来の並列アプリケーション

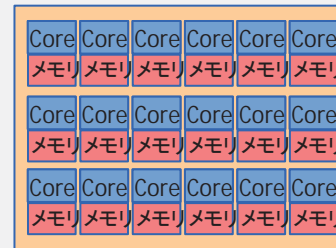
演算加速部 vs 汎用&ベクトル解説

演算加速：単純なコアを1000以上

汎用：汎用コアを数十~100、コア内にキャッシュメモリ

ベクトル：汎用よりもB/F重視

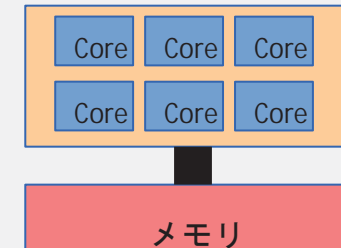
B/Fの解説は次ページ



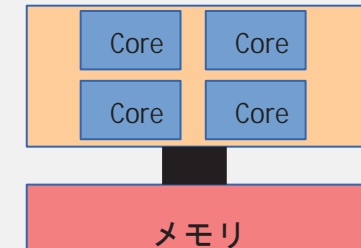
演算加速部

演算加速部はメモリ容量が小さく、また規則的構造には向くが不規則データには向かない

規則データ
DO I=1,1000
A(I)=B(I)+C(I)
ENDDO



汎用



ベクトル

不規則データ
DO I=1,1000
A(IND(I)) = B(IND(I))+C(IND(I))
ENDDO

配列A, B, Cのどの要素を使用するかは別の配列INDに格納されていて、A, B, Cを連続してアクセスするとは限らない

今後の HPCI システムの技術的課題 (2/2)

B/F 問題とコア数およびノード数 & 通信問題

CPU演算性能を上げられても配線数の問題でメモリバンド幅を上げられず、BFが低下

演算加速： チップ内に収まるメモリ容量に限定してB/F値を維持。チップ当たりの性能大。大規模メモリを提供するためチップ数増。高いノード性能およびノード数増大により**低遅延通信機構必須**。

汎用： メモリキャッシュの活用によりデータアクセス局所性のあるアプリケーションに対して高性能。局所性のないアプリは並列実行で全体性能維持。ノード性能向上およびノード数が増えることにより**低遅延通信機構必須**。

ベクトル： 演算性能とメモリバンド幅のバランスを重視。演算加速や汎用に比べてノード数を増やさなくても同等の実効性能だせるが電力ネック。全体の演算資源量は演算加速や汎用に比べて少ないため、多くのアプリ実行資源を提供できるとは限らない。

B/F解説

$B/F = \text{メモリバンド幅(B/s)} / \text{演算性能(Flop/s)}$
 要求B/F: 演算に必要なメモリ転送量/演算数

並列スケールするアプリケーションであればノード単体B/Fが少ない分、ノード数を増やすことで実効性能維持

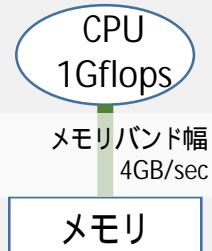
Type C: $B/F = 1$
 総メモリバンド幅: 4GB

プログラム例

```
DO I = 1, 1000
  SUM = SUM + A(I)
ENDDO
```

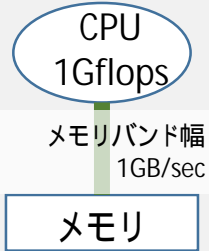
Aは単精度浮動小数点(4byte長)配列とすると、一回の加算をするために4byteのデータがCPUに供給されないと演算できない。このプログラムが要求する要求B/F値は
 要求B/F = 4

Type A: $B/F = 4$

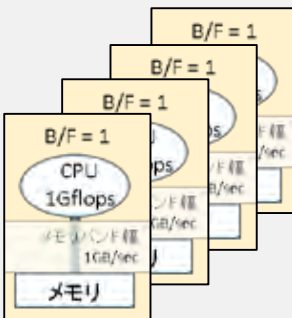


プログラム例実行時
1Gflops性能
 効率100%

Type B: $B/F = 1$



プログラム例実行時
0.25Gflops性能
 効率25%



プログラム例実行時
1Gflops性能(トータル)
 効率25%

通信問題解説

通信によるデータ交換



p ノード性能向上による通信遅延問題

```
DO I = 1, 1000
  ローカル計算 1秒
  データ交換 0.1秒
ENDDO
```

実行時間: 1010秒

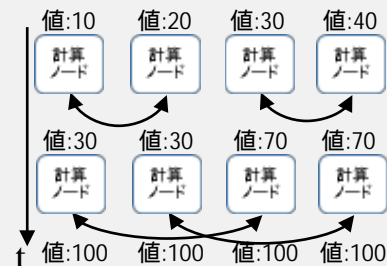
```
DO I = 1, 1000
  ローカル計算 0.01秒
  データ交換 0.1秒
ENDDO
```

実行時間: 110秒

このように全体時間の中で通信時間の割合が増える

p ノード数増大による通信遅延問題

右図は4つの計算ノード上の値の合計値を求める時の通信パターンを示している。この場合、2回の通信フェーズが必要となる。ノード数増大により通信回数が増える。



設計開発基本方針

課題：2020年から運用可能な高い性能電力比と幅広いアプリケーション実行環境を有するエクサスケールマシンの実現

計算機アーキテクチャ基本方針

汎用コアと演算加速コアを有するマシン

- 汎用コアでない性能がだせないアプリ
- 演算加速コアで性能だせるアプリ
- 両コアを使って性能がだせるアプリを棲み分け、全電力時間積削減

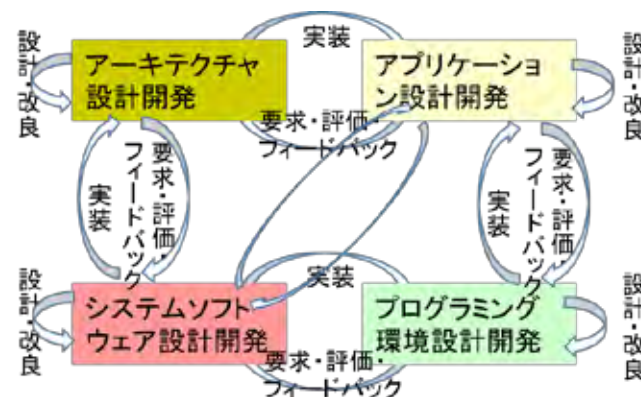
Co-design (協調設計)

- アプリケーションプログラムと計算機アーキテクチャの協調設計：演算性能・並列性能
- アプリケーションプログラムとプログラミング環境の協調設計：記述性・並列性能
- アプリケーションプログラムとシステムソフトウェアの協調設計：ファイルI/O性能・通信性能
- プログラミング環境とシステムソフトウェアの協調設計：ファイルI/O機構API・通信機構API
- 計算機アーキテクチャとシステムソフトウェアの協調設計：ファイルI/O性能・通信性能

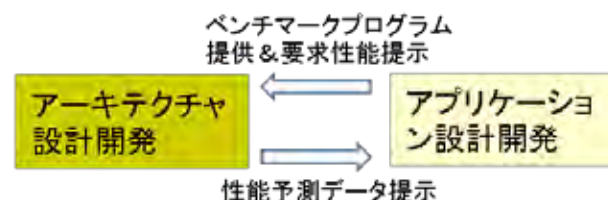
プログラミング環境設計基本方針

- 将来にわたって有効な統一プログラミングモデル、ライブラリ、フレームワークを提供
- 国際連携によるソフトウェア資産の国際的共有化促進

Co-design



従来手法



開発目標

目標性能：京コンピュータの100倍性能を目指す

Linpack性能で100倍ではなくアプリケーション実効性能で100倍

	京	ポスト京
性能	10.51 Pflops (Linpack)	エクサスケール
消費電力	12.7 MW(Linpack)	30 ~ 40MW

開発目標達成検証法（設置時）
社会的・科学的課題を解決するために使用されるアプリケーションの中からベンチマークプログラムとして選定し目標性能を検証する

100倍向上するアプリケーション例

医療・創薬	新薬スクリーニングのための分子動力学シミュレーション
総合防災	被害予測のための地震波伝播シミュレーション
基礎科学	量子色力学から原子核・宇宙初期の解明を目指す格子QCD計算

注1：現在概念設計レベルのため、性能および消費電力は今後精査

注2：評価に用いるベンチマーク群はアプリF Sの意見を反映しながら選定する

10～数十倍向上するアプリケーション例

医療・創薬	病気診断に貢献する心臓シミュレーション
総合防災	実時間ゲリラ豪雨予測
グリーンエネルギー創出と環境	デバイス材料の電子状態計算
基礎科学	太陽活動の長期シミュレーション
ものづくり	車体衝突シミュレーション

注3：平成26年7月までに選定ベンチマークプログラム群を用いて概念設計レベルでの性能推定を実施