

研究開発課題の事後評価結果

【次世代生命体統合シミュレーションソフトウェアの研究開発】

平成25年3月

次世代生命体統合シミュレーションソフトウェアの研究開発

事後評価委員会

「次世代生命体統合シミュレーションソフトウェアの研究開発」
事後評価委員会 委員名簿

青柳 睦	国立大学法人九州大学情報基盤研究開発センター センター長
甘利 俊一	独立行政法人理化学研究所脳科学総合研究センター 特別顧問
小原 雄治	大学共同利用機関法人情報・システム研究機構 国立遺伝学研究所 特任教員
佐久間 一郎	国立大学法人東京大学大学院工学系研究科 教授
清水 謙多郎	国立大学法人東京大学大学院農学生命科学研究科 教授
富田 勝	学校法人慶應義塾大学先端生命科学研究所 所長 環境情報学部 教授
○長洲 毅志	エーザイ株式会社 理事

○ 主査

計 7 名（敬称略 50音順）

「次世代生命体統合シミュレーションソフトウェアの研究開発」の概要

1. 課題実施期間及び評価実施時期

平成18年度～平成24年度

中間評価 (第1回) 平成20年10月, (第2回) 平成22年12月,

事後評価 平成25年3月

2. 研究開発概要・目的

基礎方程式に基づく解析的アプローチと大量の実験データから未知の経路と法則に迫る実験的アプローチの二つを同時に進めることで、ペタスケールという桁違いの性能を持つスーパーコンピュータの性能をフルに発揮できるソフトウェアを開発し、従来の計算機性能では不可能であった規模で計算することで、生体现象の深い理解と新たな発見を目指すと同時に、医薬品や医療機器、診断や手術方法の開発につなげることを目的とする。

具体的には、以下の六つの研究開発チーム・ワーキンググループを設置し、グランドチャレンジアプリケーションの開発等を行う。

(1) 分子スケール研究開発

量子化学、全原子分子動力学、粗視化モデルの三つの階層のソフトウェアとそれらの階層を接続するソフトウェアを完成させ、これらのソフトウェアを「京」上で大規模に動作するように高度化することによって、タンパク質系のマルチスケールシミュレーションを実現する。

具体的には、以下のソフトウェア開発を行う。

- (i) 密度汎関数法に基づくタンパク質全電子波動関数計算 (ProteinDF)
- (ii) 量子化学計算 (Platypus-QM), 量子化学計算/分子動力学計算 (Platypus-QM/MM)
- (iii) ハイブリッドQM/MM反応自由エネルギー計算 (Platypus-QM/MM-FE)
- (iv) 全原子分子動力学計算 (MARBLE)
- (v) レプリカ交換分子動力学計算インターフェイス (Platypus-REIN)
- (vi) マルチコピー・マルチスケール分子シミュレーション法開発の基盤となるクラスライブラリ (Platypus-MM/CG)
- (vii) 粗視化モデル計算 (CafeMol)

(2) 細胞スケール研究開発

これまでは、細胞内を均一な場と仮定した解析がなされていた細胞シミュレーション研究に対し、細胞内に存在するオルガネラなどの不均一な場を有する細胞モデルを構築し、そのモデル中において代謝や物質の移動等の現象を統一的に記述し、様々なシミュレーション機能を搭載して、細胞内の時空間シミュレーションを実現することを目標とする。

具体的には、以下のソフトウェア開発を行う。

- (i) 肝細胞シミュレータの開発
- (ii) 肝小葉シミュレータの開発と病態予測解析への応用
- (iii) 細胞シミュレーション統合プラットフォーム (RICS) の開発

(3) 臓器全身スケール研究開発

病態の予測や治療法の検討のための生体力学シミュレーションを行い、シミュレーションによる次世代型治療器の設計支援や、より低侵襲な治療法の開発支援を目標とする。

具体的には、以下のソフトウェア開発を行う。

- (i) 血栓シミュレータ (ZZ-THROM (=ZZ-EFSI + Kinetic MC+MD))
- (ii) 超音波治療シミュレータ (ZZ-HIFU)
- (iii) 心臓シミュレータ (UT-Heart)
- (iv) 全身血管網シミュレータ (ZZ-VASC)
- (v) 重粒子線シミュレータ (ZZ-DOSE)

(4) データ解析融合研究開発

ゲノムと疾患の間にあるシステムと薬剤応答を解析するソフトウェアを、ゲノムデータ解析、遺伝子ネットワーク解析、タンパク質ネットワーク解析、及びデータ同化によるシミュレーションの観点から開発し、最適な数十遺伝子規模のネットワークからゲノムワイドネットワークまで、動的・静的・パーソナルネットワークの推定することで疾患の理解と新たな知見を見いだすことを目標とする。

具体的には、以下のソフトウェア開発を行う。

- (i) 大規模遺伝子ネットワーク推定プログラム (SiGN)
- (ii) 網羅的タンパク質ドッキング解析プログラム (MEGADOCK)
- (iii) 大規模ゲノム関連解析 (ParaHaplo, EXRat, NGS analyzer)
- (iv) 生命体データ同化プログラム (LiSDAS)
- (v) データ解析融合プラットフォーム (SBiP: Systems Biology integrative Pipeline)

(5) 脳神経系研究開発

神経細胞、局所回路レベルでの情報処理、また、脳全体レベルでの入出力（刺激－運動）変換の処理及びその処理の学習を対象としたソフトウェアの研究開発を実施、具体的には下記の項目を目標とする。

具体的には、以下のソフトウェア開発・シミュレーションを行う。

- (i) 神経細胞シミュレーション (NeuroMorphoKit)
- (ii) 局所回路シミュレーション (NEST, CMDN)
- (iii) 視覚系シミュレーション (NEST)
- (iv) 昆虫嗅覚系シミュレーション (IOSSM)

(6) 生命体基盤ソフトウェア開発・高度化

本プロジェクトが完了するまでの期間、特に「京」の性能実証、及び他チームのアプリケーション高速化への貢献に注力し、以下の目標のため、ソフトウェア開発等を行う。

- (i) 2012年のゴードンベル賞を目標に、本プロジェクトで開発するアプリケーションの中でも

- 第一優先度で「京」上での走行を計画している第一走者アプリケーションの高度化を進め、「京」での科学的成果創出・性能デモンストレーションに貢献する。
- (ii) プロジェクトで開発されたアプリケーションの「京」への移植支援、性能チューニングを実施する。
- (iii) 10 億化合物規模の大規模仮想化合物ライブラリの開発を行う。また、その公開・実利用を推進する。

3. 研究開発の必要性等

【必要性】

過去においては、生命科学分野では他の分野（物理、工学、気象など）に比べて計算科学技術、とりわけスーパーコンピュータの活用が行われてこなかった。生命科学では他分野と異なり、その複雑さゆえに基礎方程式から出発して計算で到達できる範囲で今まで生命現象を計測できなかった、若しくは基礎方程式が明らかになっていなかった。しかし、次世代スーパーコンピュータの開発や計測技術の進歩によって、これらの問題が解決されるのではないかという期待が高まっていた。

解析的アプローチと大量の実験データから未知の経路と法則に迫る実験的アプローチの二つを同時に進めることで、ペタスケールという桁違いの性能を持つスーパーコンピュータの性能をフルに発揮できるソフトウェアを開発し、従来の計算機性能では不可能であった規模で計算することで、生体現象の深い理解と新たな発見を目指すと同時に、医薬品や医療機器、診断や手術方法の開発につなげることができるかと期待される。

【有効性】

生命科学に関連した分野において分子・細胞・臓器全身・データ解析・脳神経系などの計算科学は、このプロジェクト以前はそれぞれ別々の研究領域であった。次世代スーパーコンピュータ「京」のソフトウェア開発のために、これらの研究領域を統合して一つのプロジェクトとして実施することにより、複雑な現象が多階層で関連する生命体の本質に迫っていく上で、大変意義がある。世界的に見てもこれだけの領域が一つにまとまったプロジェクトは過去に類がなく、我が国が「京」によって開発された計算技術によってリードしていく上でも、生命科学を統合的にシミュレーションするためにも、貢献するものである。

【効率性】

そのような背景の中で、本プロジェクトでは、集中的に「京」に向けて、生命科学分野の計算科学技術の研究開発を行い、他の科学技術分野を凌駕するソフトウェア群を開発する。これはスーパーコンピュータの専門であるソフトウェアの高度化チームが各研究者と密接に連携し、始めてなし得ることである。このように開発されたソフトウェア群は、高速なスーパーコンピュータを生命科学分野で活用する上で重要な資産となっていくことと考える。

また開発ソフトウェアについては、ホームページ等での公開を行うとともに、代表機関である独立行政法人理化学研究所において HPCI 計算生命科学推進プログラム及び情報基盤センターが協力して、今後も引き続き利用者への提供、利用支援を行っていくことで、社会還元、計算生命科学分野に貢献する体制を構築することとしている。

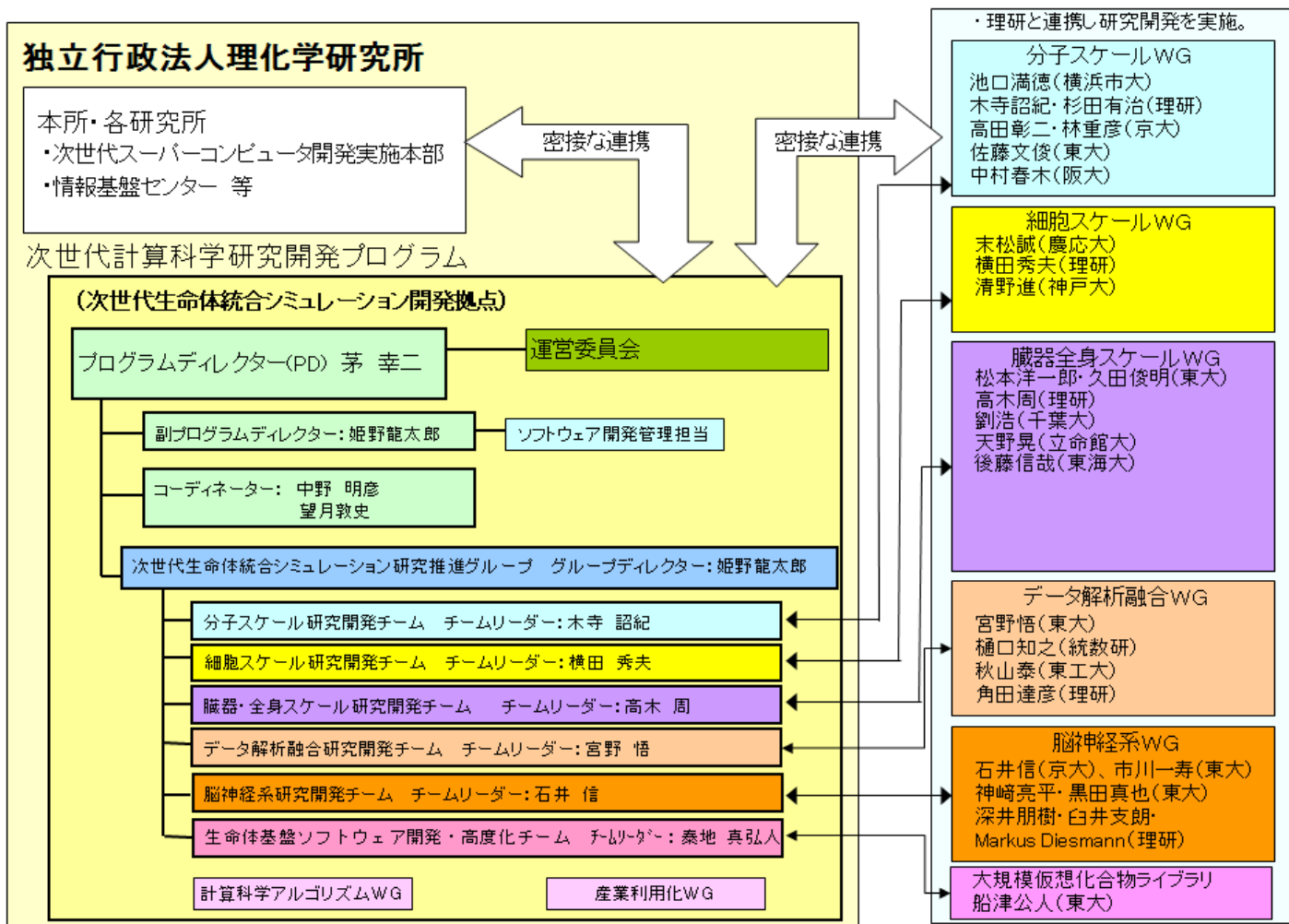
4. 予算の変遷

年度	18年度	19年度	20年度	21年度	22年度	23年度	24年度
予算額(百万円)	166	1,600	1,500	1,423	1,095	675	552
実施機関	理化学研究所	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 東海大学 統計数理研究所 東京工業大学 北陸先端科学技術大学院大学 東北大学 広島大学	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 東海大学 統計数理研究所 東京工業大学 北陸先端科学技術大学院大学 東北大学 広島大学	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 立命館大学 東海大学 統計数理研究所 東京工業大学 北陸先端科学技術大学院大学 東北大学 広島大学	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 立命館大学 東海大学 統計数理研究所 東京工業大学 北陸先端科学技術大学院大学 東北大学 広島大学	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 立命館大学 東海大学 統計数理研究所 東京工業大学	理化学研究所 横浜市立大学 京都大学 東京大学 大阪大学 慶應義塾大学 神戸大学 千葉大学 立命館大学 東海大学 統計数理研究所 東京工業大学

5. 課題実施機関・体制

研究代表者 茅 幸二 (独立行政法人理化学研究所
次世代計算科学研究開発プログラム プログラムディレクター)

【平成 22 年度 12 月 第 2 回中間評価後の体制】



事後評価票

(平成25年3月現在)

1. 課題名 次世代生命体統合シミュレーションソフトウェアの研究開発

2. 評価結果

(1) 課題の達成状況

ライフサイエンス分野及びその融合領域を対象としたグランドチャレンジアプリケーションの開発において、以下のとおり、31本のアプリケーションが開発され、細胞や全身での統合的なシミュレーション、データの高度な解析が可能となり、ライフサイエンス分野における計算科学技術を大きく発展させ、医用工学等への計算科学技術を駆使した新技術へ道筋を明らかにした。

具体的に6つの研究開発チームを設置し、以下のアプリケーションを開発した。

- ① 分子スケールでは密度汎関数法に基づくタンパク質全電子波動関数計算 (ProteinDF-K) をはじめとする8本のアプリケーション
- ② 細胞スケールでは、統合解析プラットフォーム (RICS-K) の1本の統合アプリケーション
- ③ 臓器全身スケールでは、心臓シミュレータ (UT-Heart) をはじめとする4本のアプリケーション
- ④ データ解析融合では、ハプロタイプ関連解析における統計検定ソフト (ParaHaplo-K) をはじめとする9本のアプリケーション
- ⑤ 脳神経系では、神経細胞シミュレーションツール (NEST) をはじめとする5本のアプリケーション
- ⑥ 生命体基盤ソフトウェアとして大規模並列用 MD コアプログラム (cppmd-K) をはじめとする4本のアプリケーション

産業利用化ワーキンググループを設置し、計31本中25本のアプリケーションについて、一般公開・ダウンロード可能としており、利用マニュアル等の利用者へ提供、利用支援を行っていることは高く評価できる。またチームリーダーを中心にマネジメントがよく機能しており、チーム内の効果的な連携を行っている。また並列化を指標とした進捗管理を導入し、進捗に応じた技術支援を行う等プロジェクトとしてのマネジメントが発揮されている。これは中間評価結果への対応の努力が見られるところである。

技術支援チームを設置したことは、プロジェクト全体の成果に大きく貢献したと言える。このようなグループは必要であり、プロジェクト開始後速やかに設置したことは評価されるべきである。また一部アプリケーションについては、固有のテーマを扱うため研究開発チーム内での孤立が見られるため、チームの枠を越えた解析に対する連携、マルチスケール解析等による統合的な解析まで発展させ、有機的な連携を更に進めてほしかった。

中間評価結果を受け階層固有のテーマを絞った結果、個別アプリケーションの完成度が飛躍的に高まり、「京」のライフサイエンス分野への応用のための先駆けとして十分機能するものとなった。この意味で、公募時の課題は十分達成していると言える。

(2) 成果

「京」で使用するアプリケーションの開発としては、並列化が難しい問題に対し、並列度を上げることに努力し、一定の成果を上げている。臨床という重要な応用に貢献する成果があり、更に成果が産業応用されている点など高く評価できる。

ライフサイエンス分野のモデリング、シミュレーション手法についてはまだ開発途上の技術でありもう少し長いスパンでの研究が必要である。継続的なプロジェクトなどのもので、さらなる展開に期待したい。

特にビッグデータ時代を迎えるバイオインフォマティクスにとってはライフサイエンスのための新しいアルゴリズムは非常に重要である。本プロジェクトにおいては、既存の確固としたアルゴリズムが存在する中での超並列化がほとんどであり、例えばベイズ解析と非線形解析との結合を行って新規アルゴリズムへのチャレンジが見られた点などは評価できるが、本格的なビッグデータの時代に向けてさらなる努力が必要である。

全体として、応用を意識したシミュレーションと出口にこだわりつつも、基礎的検証が必要な部分については地についた深掘りした研究姿勢が随所に見られ、プロジェクトとして今後の広がり期待が持てる。適切なシミュレーションモデルを設定し、並列計算の効果を有効に活用して、生命現象の一端の再現・解析が可能となっており、高く評価できる。

開発したアプリケーションによるシミュレーションについて、十分な計算時間を確保することで、ライフサイエンス全体に影響を与えるような計算結果を出すようなプロジェクトが推進されることを期待する。今後の、アプリケーションプログラムのさらなる改善・発展と社会への還元をこれから期待したい。

本プロジェクトで雇用した研究員69名は、プロジェクト終了後、アカデミックへ46名（うち9名が研究所PI職、24名が大学教員）、民間企業へ5名、残りが他プロジェクトの研究員として輩出しており、多くが関連分野での研究者として研究を継続することとなっており、人材育成は良好であったと言える。

(3) 今後の展望

アプリケーション開発については、HPCI 戦略プログラム分野1で一部課題の継続が決まっている。また HPCI 戦略プログラムの課題からはずれたテーマでも今後も着実に進めていく展望を述べており、今後の継続・発展に期待できる。

今後、更に発展するには現状の延長線だけでなく、各シミュレータの基本的な構想から高度化・発展させる必要がある。現時点でできることは達成したと考えるが、さらなる発展の可能性を秘めているため、新たなモデルを構築する等の取組を行い、ライフサイエンスとしての予測を超えたメカニズムの理解や原理の発見まで期待したい。しかし、HPCI や新しいバイオインフォマティクスに裏打ちされたライフサイエンス分野の研究は基礎だけでなく直接ヒトの健康医療に結びつく知見を生み出すものである。したがって、社会への還元についても企業との共同研究によるシミュレーション結果の応用を継続するよう期待したい。

HPCI 戦略プログラムでの継続課題については、産業応用として、創薬・医療へと更に結びつくことを期待したい。また、スーパーコンピュータの利活用は今後も継続・発展させていく必要があるため、今

回実施した並列化支援の技術，ノウハウをどのように生かし発展させていくかも重要な課題である。開発した「京」及び「京」以外のスーパーコンピュータに対するアプリケーションの公開・利用者支援については，代表機関である独立行政法人理化学研究所で行うことが決まっているが，アカデミアばかりでなく応用分野を担う産業界の利用がカギとなるので情報発信等の利用促進のさらなる充実も重要である。

「HPCI の整備」 事後評価結果

HPCI 計画推進委員会

平成 25 年 4 月 12 日現在

氏名	所属・職名
小柳 義夫	神戸大学大学院システム情報学研究科特命教授
笠原 博徳	早稲田大学理工学術院教授
関口 和一	株式会社日本経済新聞社論説委員兼産業部編集委員
鷹野 景子	お茶の水女子大学副学長兼附属図書館長/学術・情報機構長
所 眞理雄	株式会社ソニーコンピュータサイエンス研究所代表取締役会長
主査 土居 範久	慶応義塾大学名誉教授
土井 美和子	株式会社東芝研究開発センター首席技監
根元 義章	独立行政法人情報通信研究機構 耐災害 ICT 研究センター センター長
村上 和彰	九州大学大学院システム情報科学研究院教授
矢川 元基	東京大学名誉教授

合計10名

(50音順)

「HPCIの整備」の概要

1. 課題実施期間及び評価実施時期

平成22年度～平成24年度
事後評価 平成25年3月

2. 研究開発概要・目的

HPCIの整備・運営に関するプロジェクト目標は下記のとおり。

ユーザ等からなるコンソーシアムを形成し、この主導により、平成24年11月を目途に次世代スーパーコンピュータ、国内の主要スーパーコンピュータ、ストレージを用いた高度なコンピューティング環境を実現するインフラ(HPCI)を構築し、運用を開始する。

このうち、HPCIの整備に関する各項目の研究開発目標は次のとおり。

1)利用者支援システム整備(東京大学)

ヘルプデスク及び情報共有コンテンツマネジメントシステム(CMS)の整備を行う。実運用環境に合わせたワークフロー、グループ構成、アクセス権、運用手順等の調整とそれに対応するマニュアルの整備、障害・インシデント対応マニュアルの整備を図る。

2) 運用ツール群整備(大阪大学)

2-1) 認証基盤サブシステム整備として、HPCI 共通運用システムにおける認証基盤の運用に際し、大学共同利用機関法人情報・システム研究機構と連携協力して、ID 連携によるシングルサインオン機能を実現するための認証基盤サブシステムの整備を行う。

2-2) 運用事務局ソフトウェア整備として、HPCI の利用促進における課題選定及び共通窓口の運用に際し、登録機関と連携協力して、HPCI 運用事務局による課題公募や申請受付、利用課題選定、成果報告を支援するシステムの整備を行う。

2-3) 資源管理・検索システム及びアカウント集計ツール整備として、HPCI 運用事務局と HPCI システム構成機関が円滑な情報共有を行うための支援システムを整備する。

3) 共用ストレージソフトウェア整備(筑波大学)

共用ストレージシステムのソフトウェア整備を行う。整備に当たり、アクセス性能及び信頼性の向上のためのソフトウェア改良、性能モニタリングシステムの整備、負荷試験を実施する。ソフトウェア整備に伴う、ドキュメントの整備を行う。

4) 認証基盤整備(NII)

認証基盤運用に必要となるドキュメントの整備及び認証局ソフトウェアを開発する。

5) 先端ソフトウェア運用基盤整備(東工大)

仮想計算機システムの可用性向上、及び、利用者支援機能を向上する。

3. 研究開発の必要性等

HPCI は、世界トップクラスのスーパーコンピュータ「京」やその他の計算資源をユーザが容易に利用できる環境を実現するものであり、我が国の科学の進展、産業競争力の強化に資するとともに、グリーンイノベーションやライフイノベーション等のイノベーション創出の基盤となることが期待される。

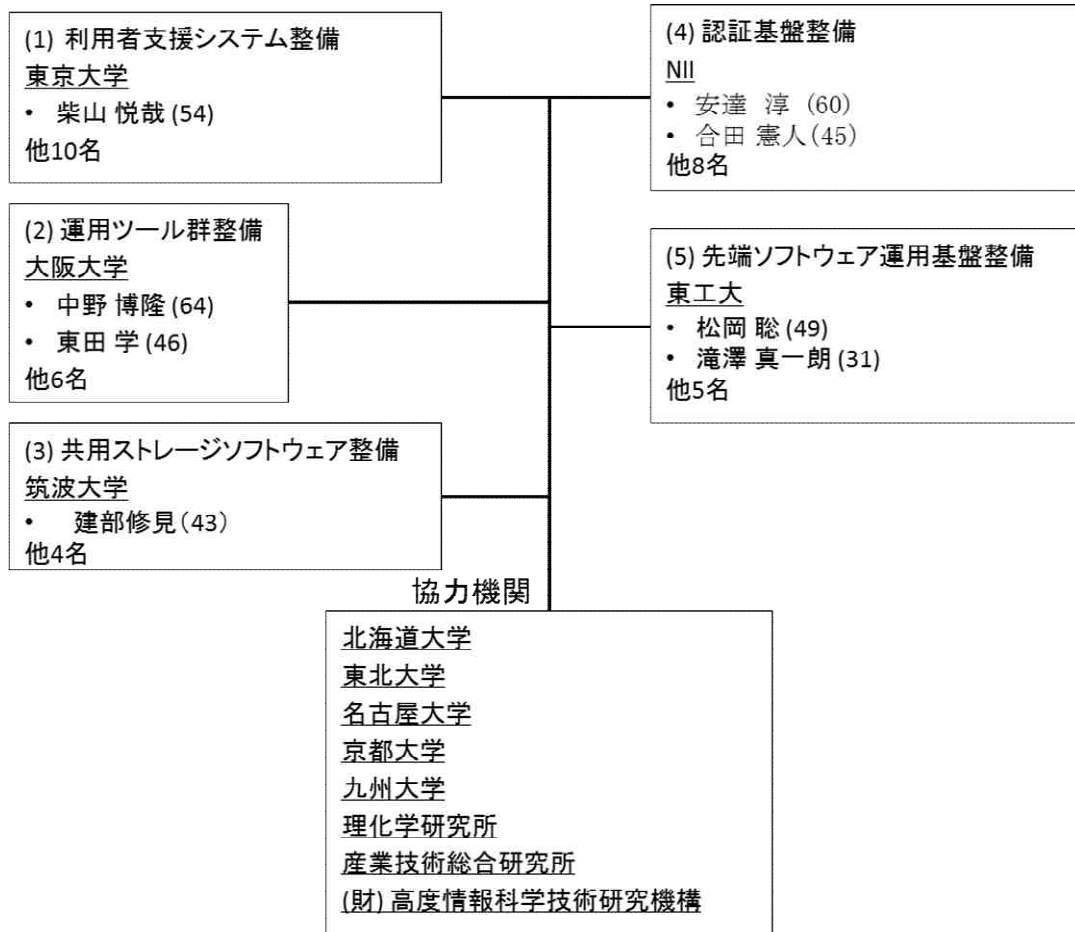
また、このような基盤の整備は、人材育成や利用者の拡大にも貢献すると考えられる。

4. 予算(執行額)額の変遷

年度	H22(初年度)	H23	H24	総額
予算額	0.5億	1.8億	2.4億	4.7億

5. 課題実施機関・体制

研究代表者 国立大学法人東京大学 情報基盤センター長 石川 裕
 主管研究機関 国立大学法人東京大学
 共同研究機関 国立大学法人大阪大学、国立大学法人筑波大学、国立大学法人東京工業大学、
 大学共同利用機関法人情報・システム研究機構



事後評価票

(平成25年4月現在)

1. 課題名 HPCIの整備
2. 評価結果
(1)課題の達成状況
○研究開発目標 利用者支援システム、運用ツール群、共用ストレージソフトウェア、認証基盤及び先端ソフトウェア運用基盤の整備が着実に実施され、世界トップクラスのスーパーコンピュータやその他の計算資源をユーザが容易に利用できる環境が構築され、平成24年9月の共用開始が実現していることから高く評価できる。
○研究開発体制 東京大学(利用者支援システム)、大阪大学(運用ツール群)、筑波大学(共用ストレージソフトウェア)、大学共同利用機関法人情報・システム研究機構(認証基盤)及び東京工業大学(先端ソフトウェア運用基盤)のそれぞれの機関において、密な連携体制の下で研究開発を遂行した。その結果、統一的な視点を持って遅滞なく全体の開発が進められており高く評価できる。
(2)成果
○研究開発成果 HPCI 準備段階コンソーシアムがとりまとめた「HPCIとその構築を主導するコンソーシアムの具体化に向けて(最終報告)」(平成24年1月30日)に基づき、平成22年度「HPCIの基本仕様に関する調査検討」、平成23年度「HPCIの詳細仕様に関する調査検討」を踏まえて平成24年度にHPCIシステムの整備が行われている。 システム全体として、課題申請からアカウント発行、利用支援までの一括管理システムを実現したことは他にない大きな成果であり、特に外国における類似のプロジェクトと比較しても機能面で充足している点は評価できる。 共用ストレージソフト Gfarm は新しいコンセプトによるものであり、システムの再構築なしにストレージの増強が可能である。これをペタバイト級のファイルシステム上で実現したことは大きな成果である。 認証基盤については、NAREGI-CA を改良した実運用システムを可能にした。 先端ソフトウェアについてはサービス群を仮想計算機で実行する基盤を整備した。これは他にない独創的な環境である。
○研究開発成果の利活用 本事業によって各システム(利用者支援システム、運用ツール群、共用ストレージソフトウェ

ア、認証基盤及び先端ソフトウェア運用基盤)が整備されたことにより、平成 24 年 9 月より HPCI システムの共用開始が実現している。

これにより、355 件の課題が申請されるとともに、1706 件の HPCI-ID、1018 件の HPCI アカウントが発行されるなど、着実な利用実績を重ねている。

さらに、認証基盤整備の中で開発された認証局ソフトウェア(NAREGI-GA)について、将来的にオープンソースとして公開することは評価に値する。

○人材育成

HPCI 運用基盤ソフトウェア群の整備・運用について、HPCI システムに参画する各大学の情報基盤センター系職員により実施されており、従来の外部委託による実施ではなく、基盤センター内において運用技術者の育成が行われたことは高く評価できる。

(3) 今後の展望

このような実運用を目的としたシステムが大学を中心とした研究開発グループで開発されたことの意義は大きい。本事業によって整備された「京」を中核とする HPCI システムの利用により、HPCI の運用が促進され、多くの科学技術の発展に寄与できると考えられる。

本事業で整備されたシステムは、今後「HPCI の運営」の実施機関に引き継がれることとなっており、HPCI システムの円滑な運営を目的として HPCI システムを構成する機関からなる委員会が設置されている。今後、この仕組みを活用して、利用者からのフィードバックを受け、システムの問題点を継続的に改善できる体制を確立する必要がある。さらに、後継人材の育成にも取り組まれない。

HPCIの枠組み

参考

- 「京」を中核とする国内のスパコンやストレージを高速ネットワークでつなぎ、ユーザー窓口の一元化などにより、利便性の高い利用環境を構築。
- 「HPCIの整備・運営」として、各機関への委託事業により実施。

