

## ご挨拶

茅 幸二

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



### 発表者紹介

昭和41年 3月 東京大学大学院理学研究科博士課程修了  
昭和45年 9月 東北大学理学部化学科助教授  
（昭和48年 9月～昭和49年 9月 米国ベル電話研究所研究員）  
昭和56年 4月 慶應義塾大学工学部教授  
平成11年 4月 岡崎国立共同研究機構分子科学研究所長  
平成11年 4月 慶應義塾大学名誉教授  
平成16年 4月 独立行政法人理化学研究所和光研究所・中央研究所長  
平成18年 10月 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(兼)  
平成20年 4月 次世代スーパーコンピュータ開発実施本部 副本部長  
平成20年 5月 基幹研究所 特別顧問(兼)  
平成22年 4月 社会知創成事業 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(兼)

### 研究分野

ナノ物質化学

## ご挨拶

茅 幸二

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

理化学研究所が2006年から推進してきた次世代スーパーコンピュータ開発は、2011年10月に目標速度 10 ペタフロップスを達成し、SC11において世界最速を含む多くの賞を獲得した。また、本コンピュータの愛称「京(けい)」が、神戸の研究機構の近隣駅(モノレール)にも採り入れられるなど、神戸市の誇る科学技術の拠点となっている。

「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」はKコンピュータ開発に併せて、生命現象に対する理解を深め、特に創薬・医療に資することを目的に2006年10月に発足し、終了まであと一年となった。本プログラムは、分子、細胞、臓器全身、脳科学などの分野を含む生命科学分野全般に涉って、基礎方程式に基づく解析的アプローチ(シミュレーション)と、大量のデータから生命現象の経路と法則を抉り出す情報学、統計学的アプローチの両面からの定量的解明を推進し、10万並列を超える高度なソフトの手法開発チームを含む、6チームが、たがいに協力しつつ研究開発に励んでいる。

理研を拠点とする14研究機関の研究者の努力の結果、31本のソフトウェアの大半が、Kコンピュータに挑戦するレベルに到達した。それらのなかのいくつかのものはペタスケールへの予備段階において、すでに Nature 誌に掲載されるなどの成果をだし、Kコンピュータを利用した医薬、医療、生命機能解明への大きな寄与が期待されるに至っている。

本報告会は、本プログラムに参加された研究者の方々の献身的な努力の結晶であるソフトウェア開発の成果を纏めたものである。今後一年間にそれらのソフトウェアがKコンピュータ上で花咲き、複雑で混沌とした生命現象の理解の基盤としての計算科学の役割をより発展させていくことを願っている。

# 生命科学のグランドチャレンジ ISLiM の開発状況概要

姫野龍太郎

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  
副プログラムディレクタ



## 発表者紹介

1979 年 3 月、京都大学大学院工学研究科修了  
1979 年 4 月、日産自動車株式会社入社  
1998 年 1 月、理化学研究所入所、情報環境室室長  
2003 年 10 月～現在、情報基盤センター長  
2006 年 1 月～2008 年 8 月、次世代スーパーコンピュータ開発実施本部開発グループ・グループディレクタ  
2006 年 10 月～現在、次世代計算科学プログラム・副プログラムディレクタ  
情報基盤センター

## 研究分野

高性能計算、計算流体力学、計算生体力学

## 生命科学のグランドチャレンジ ISLiM の開発状況概要

姫野龍太郎

次世代計算科学研究開発プログラム  
副プログラムディレクタ

### 1. はじめに

2006年10月に始まったISLiMも5年が経過、来年度末のプロジェクト終結に向けて、いよいよ残り1年ほどになった。今年度は京が実際に使えるようになり、京の性能を十分に引き出す生命科学分野でのソフトウェア開発を進めるといふ点では大きく進展することができた。京はTop500で二期連続して世界最速となり、我々には世界最速の京を使って世界をあっと言わせる研究成果を出すことが期待されている。

これまでこのプロジェクトは、二度中間評価を受け、軌道修正を行いながら進めてきた。これまでの開発状況と今後の計画を報告するために、今回の成果報告会を企画した。ここでは全体の概要を紹介する。

### 2. プロジェクト全体の目標

分子・細胞・臓器全身という3つのスケールに分かれた基礎方程式に基づくアプローチと、大量の実験データに基づくアプローチの二つを元に、脳神経系の解析チームと、ペタを超えるスーパーコンピュータを有効に使う高性能計算のチームの合計6チームが力を合わせ、生命現象の統合的な理解とともに、その応用として医薬品の開発や革新的治療機器の開発に使えるソフトウェア基盤を開発提供してゆくことを目標にしている。

### 3. 開発の状況

各チームで開発しているソフトウェアは全部で34本で、そのうち31本が京で動くことを想定している。本年4月から京で試験利用が可能になり、31本のソフトウェアは順次京での性能チューニングを開始した。我々はこれらの31本で具体的にどんな問題を解くかという問題設定を含めて優先順位をつけ、3つのグループ：第一走者、第二走者、第三走者に分け、効果的で効率的な開発が進められるように配慮している。計算速度に関しては複数のソフトウェアで1PetaFLOPSを超える速度を達成することを目指している。

現在までのところ、31本中11本で京での実行を確認、そのうち3本では京の理論性能のおよそ40%性能を引き出すことができるようになった。3月にはGordon Bell賞に応募するために、特別に用意された京の実行環境(京全体の約1/3のシステム)で、分子動力学計算ソフトcppmdが1.3PetaFLOPSを達成した。また、心臓シミュレータであるUT-Heartと血流解析ソフトウェア(構造流体連成解析ソルバ):ZZ-EFSIも十分な並列性能と高い実効性能を示しており、京の約2000ノードを使った計算で、実効性能で理論性能の約40%の性能を出している。

その他のアプリケーション・ソフトウェアでは、これほどの性能はまだ出せていないが、並列性能は十分あることが分かっている。今後の性能改善が期待される。

京は現在もコンパイラやライブラリ、ファイルシステム、OSなど、日々テストや修正が行われており、開発が続けられている。今回の成果報告会で報告される京の性能数値は、ここで示した数値も含め、その時点のものであり、今後改善される可能性もあることを承知置きたい。

### 4. 今後

これまでのところ、京から十分な性能を引き出すことができたソフトウェアは3つであるが、今後更に増えてゆくことだろう。今後はこれらの性能チューニングから、京を使って研究成果を出すことに力を注ぐことことになる。同時に、開発したソフトウェアの利用を普及させてゆくことも期待されており、今後進めてゆく。

### 5. 謝辞

このプロジェクトは全体として次世代スーパーコンピュータ開発プロジェクトの中にあり、文部科学省の委託費で開発が行われている。また、京の利用にあたって、配慮がなされていることを関係者の皆さんに感謝します。






## 生命科学のグランドチャレンジISLIMの開発状況

### 概要



理化学研究所  
次世代計算科学プログラム  
副プログラムディレクタ  
姫野龍太郎  
himeno@riken.jp









## 京、10Peta達成

- 二期連続一位
- 8.16PetaFLOPs:2011/06 ISC
- 10.5PetaFLOPs:2011/11 SC11

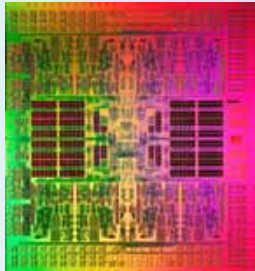




Fujitsu Lunch Party at ISC2011

## 次世代スーパーコンピュータ“京”とは

- 新しく開発したプロセッサCPU128GFLOPS
- 2010年、試作、その時点でプロセッサ単体として世界最速
- 8万個以上集めて10PetaFLOPSを目指す
- 一秒で一京回(一兆の一万分)の計算ができる
- 864櫃体 x 96 CPU/櫃体=82,944 CPU= 663,552cores
- I/O nodes: 5,184
- 現在、ハードは既にすべて搬入済み、テスト中

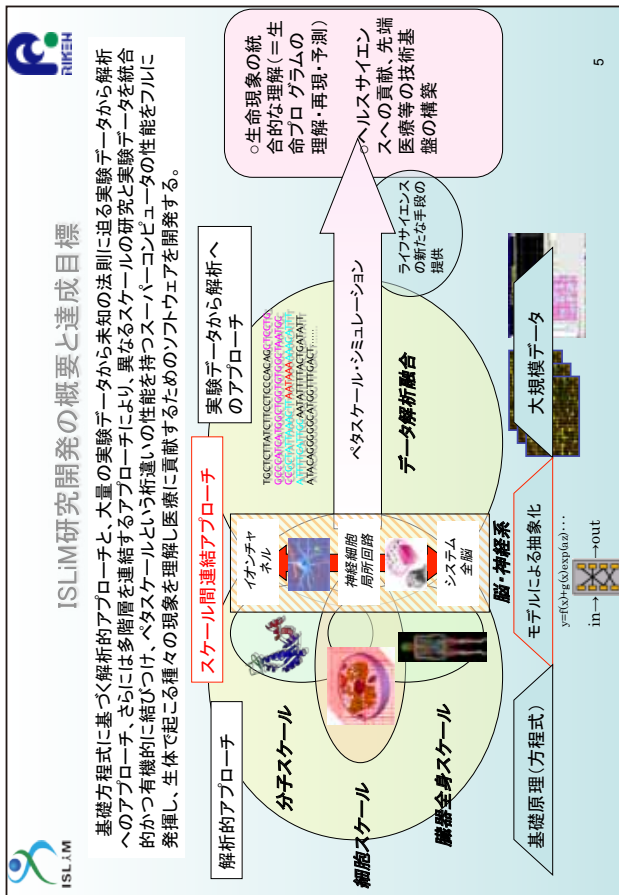
「SPARC64™ VIIIfx」  
富士通提供

計算機室に設置される次世代スーパーコンピュータ(イメージ図)

### 次世代スーパーコンピュータの研究開発プロジェクトの開発計画

	平成18年度 (2006)	平成19年度 (2007)	平成20年度 (2008)	平成21年度 (2009)	平成22年度 (2010)	平成23年度 (2011)	平成24年度 (2012)	
システム	概念設計		詳細設計		試作・評価・製造			性能 チューニング
	次世代ナノ融合 シミュレーション		開発・製作・評価		開発・製作・評価			実証
次世代生命体統合 シミュレーション	次世代生命体統合 シミュレーション		開発・製作・評価		開発・製作・評価			実証
	計算機棟		設計		建設		建設	
機	研究棟		設計		建設		建設	
							4	





### ISLiM 参加機関

13機関、東京大学、大阪大学、京都大学、東北大学、北海道大学、慶応大学、横浜市大、東工大、広島大学、統計数理研究所、理化学研究所  
 約200名の研究者(ポストドク約80名)

参加機関のロゴ: ISLiM, RIKEN, JAIST, 大阪大学, KOBE UNIVERSITY, 広島大学, 東京工業大学, TOKAI, 長岡技術科学大学

### 国際比較

全電子計算では世界をリード (ProteinDF)

組視化モデルによる分子モデル

心臓

- 分子スケール
  - アメリカは超並列計算で先行
  - 組視化モデルによる計算、全電子を入れたタンパク質の量子化学計算では日本がリード
- 細胞スケール
  - システムバイオロジー (E-cell等) は日本発
  - 細胞群から臓器へ、分子から細胞へのアプローチはともこれから
- 臓器・全身スケール
  - 国際的な研究開発体制が構築されつつある
  - 日本は心臓モデル、健康者からの詳細人体モデルでリード
- バイオインフォマティクス
  - 個人の多様性と医療をつなぐところで日本は優位
  - 次世代超高速シーケンサー等による巨大データに直面、世界的な課題
- 脳科学
  - スーパーコンピュータによるシミュレーションではBlueBrainなどが先行
  - しかし、計算論的神経科学では日本が世界をリード、BlueBrainではこの観点で日本は抜けている

### 国際比較 (続き)

ISLiM、数万並列規模の計算能力をもったソフトウェア開発を推進

今後、次世代スパコンに向けた並列性能向上を加速  
 ⇒ 分子・細胞・臓器全身の各スケールと実験データ解析、脳神経系を統合して研究開発  
 ★世界的にも全くない試み  
 同時に行うことにより、手法やソフトを互いに利用し、開発を加速  
 ⇒ 次世代スパコンで世界をリードする

チーム間の連携

提供	受	分子スケール	細胞スケール	臓器全身スケール	脳神経系
分子スケール	細胞スケール	細胞スケール	細胞スケール	細胞スケール	細胞スケール
細胞スケール	臓器全身スケール	臓器全身スケール	臓器全身スケール	臓器全身スケール	臓器全身スケール
臓器全身スケール	脳神経系	脳神経系	脳神経系	脳神経系	脳神経系

全電子計算での並列性能 (ProteinDF) 2500コアまでの並列を確認

脳の局所回路シミュレーションの並列性能 (BlueGeneLで1000並列までの速度向上を確認)

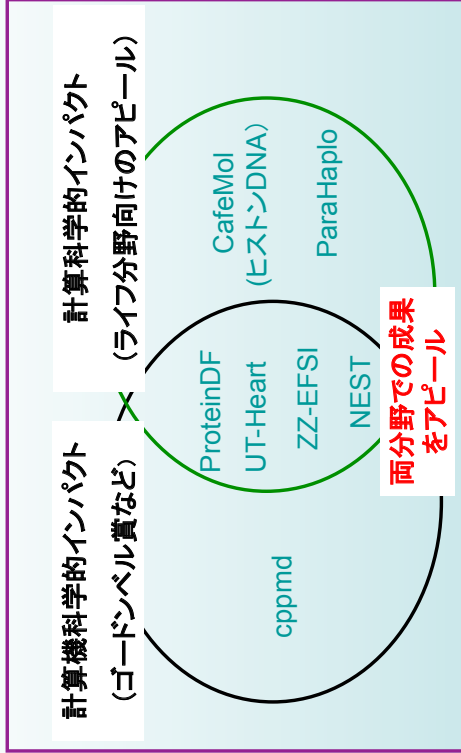
## アプリのリスト(2)

カテゴリ	アプリ名	概要	開発者	実装言語	実行環境
分子	ParHaplo	ハプロタイプ解読解算に於ける統計推定を行うためのソフトウエア	菅野(理研)	C	MPI, OpenMP
	SGN	大規模遺伝子組換えネットワーク推定プログラム	菅野(東大)	CG9	MPI, EP (SGE)
	LION	再結合正常化法による主体内分子の大規模ネットワーク推定プログラム	菅野(東大)	R	EP (SGE)
	SSM	イデオモर्फリによる群集系列データからの遺伝子ネットワーク推定プログラム	菅野(東大)	C, C++, JAVA	MPI, OpenMP
	SBP	データ駆動型ハイブリッド結合プロトフォームの構築	菅野(東大)	C++	プラットフォーム
	MEGADOCK	細胞シグナル伝達ネットワーク解析プログラム	秋山(東大)	C++	MPI
	LISDAS	生命体データ同化プログラム	樋口(東大)	Fortran90, C, C++	MPI, OpenMP
	NCS analyzer	次世代シーエンシー解析プログラム	樋口(東大)	perl, shell	EP
	EnAT	拡張MAT法による25NP結合データの全ゲノム網羅的解析ソフトウェア	樋口(東大)	C++	MPI
脳神経	NEST	Neural Simulation Tool	渡邉(理研)	C++	MPI
	NeuroMicrofit	神経細胞シミュレーションのための共有プラットフォーム	渡邉(理研)	SLI, C++	MPI
	NeuroMicrofitKit	神経細胞シミュレーションのための共有プラットフォーム	石井(東大)	MATLAB, C, C++	渡邉
	VSM	視覚シミュレーションのための共有プラットフォーム	日井(理研)	C++	プラットフォーム
	IOSIM	昆虫視覚系シミュレーション	神崎(東大)	C++	実験系アプリ
多	cppmd	大規模並列分子シミュレーションソフトウェア	桑島(理研)	C++	MPI
	VLSSL	大規模並列化共有物ライブラリー	船津(東大)	C++	EP
多	LSV	分散並列大規模データ可視化システム	小野(理研)	C++, G	MPI
	SPHERE	アプリケーションモデルウェア	小野(理研)	C++, G	MPI

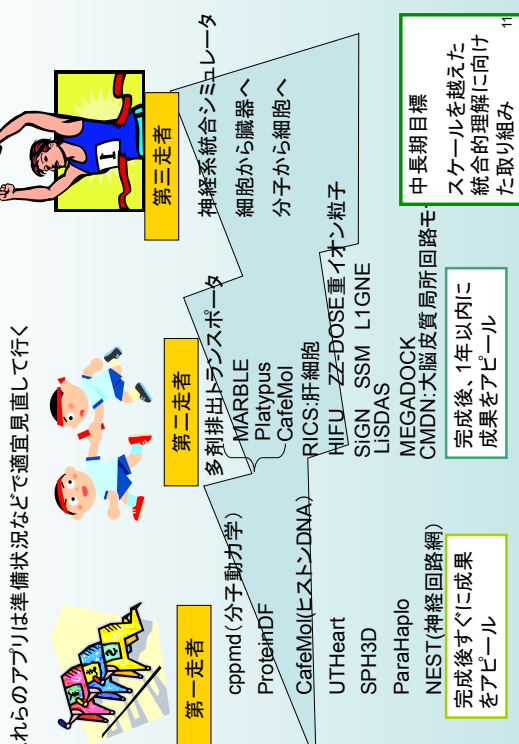
## アプリのリスト(1)

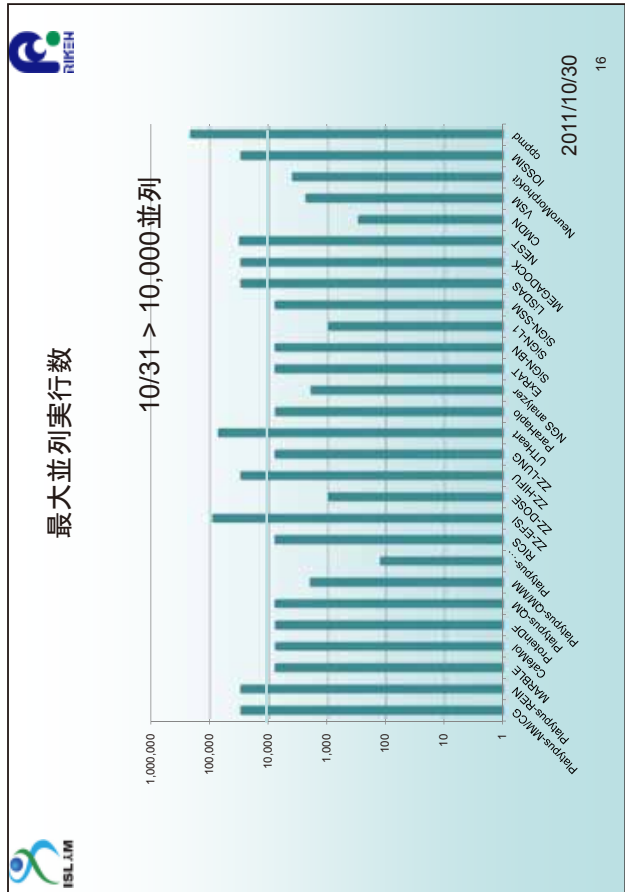
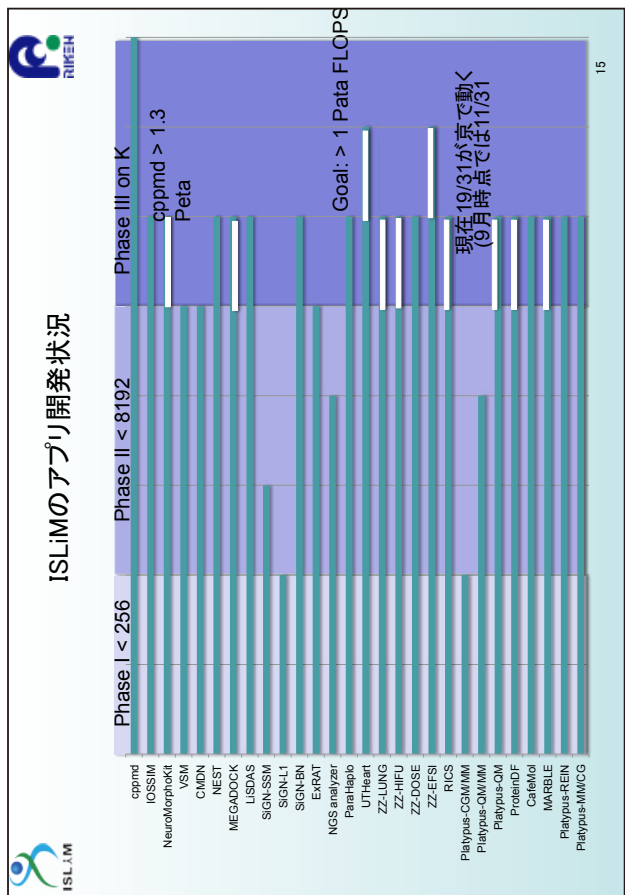
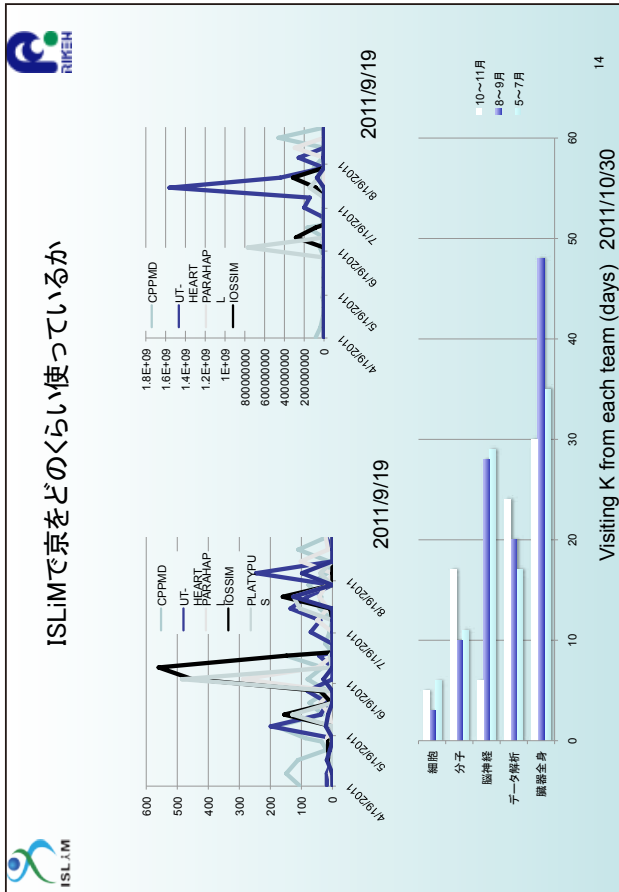
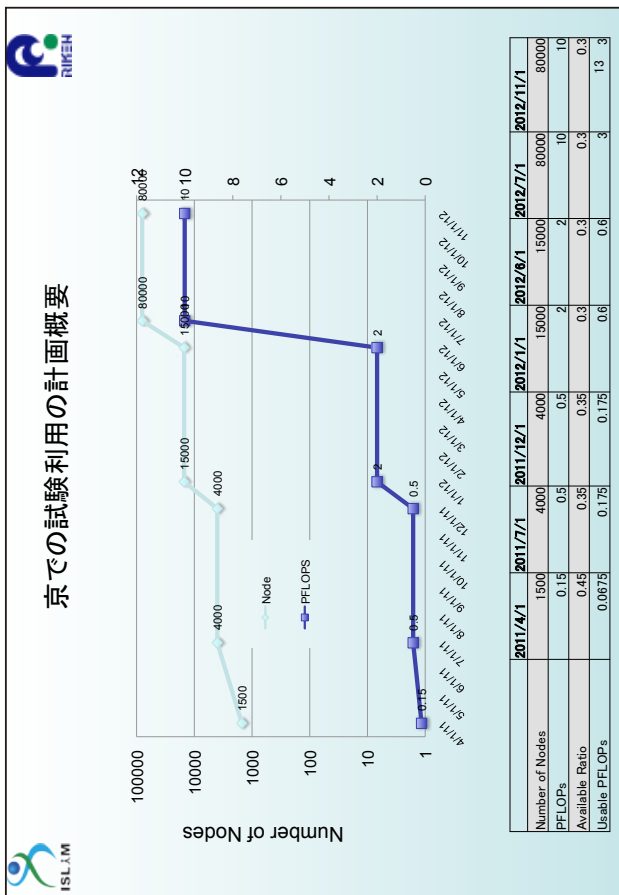
カテゴリ	アプリ名	概要	開発者	実装言語	実装化の方法
分子	ProteinDF	非標準同位体に基づくタンパク質完全電子運動関数計算	佐藤(東大)	C++	MPI, OpenMP
	CafeMol	相対化エネルギー計算	高田(東大)	Fortran90	MPI, OpenMP
	MAHILE	全原子分子軌道計算	池口(横浜国大)	C	MPI
	Platypus-MM/CG	マルチコア・マルチプロセス・マルチプラットフォーム環境での大規模分子シミュレーション	水寺(横浜国大・理研)	C++	MPI
	Platypus-OM/MM/FE	ハイブリッドOM/MM混合自由エネルギー計算	林(東大)	Fortran90	GAMESS/rock50, イソラリ
	Platypus-REN	レプリカ交換分子軌道計算インターフェイス	杉田(理研)	Fortran90	MPI
	Platypus-OM/MM	相対化エネルギー計算/分子軌道計算	中村(東大)	C++	MPI
	Platypus-OM/MM	量子化学計算/分子軌道計算	中村(東大)	Fortran90	MPI
	Platypus-OM	量子化学計算	中村(東大)	Fortran90	MPI
多	RCS	細胞シミュレーションプラットフォーム	横田(理研)	Fortran, C, C++	Sphere(MPI)
	ZZ-EFSI	多スケールシミュレーションプラットフォーム構築と高速計算プログラム	高木(東大・理研)	C++	Sphere(MPI)
多	UTHeart	マルチスケール・マルチプラットフォーム・マルチアーキテクチャの心臓シミュレーション	久田(東大)	Fortran	MP3, OpenMP
	HFU	脳神経シミュレーション	高木・松本(東大)	Fortran90	MP3, OpenMP
	ZZ-RBC	赤血球シミュレーション	高木(東大・理研)	Fortran90	Sphere(MPI)
	ZZ-DOSE	単粒子追加シミュレーション	高木(東大・理研)	Fortran90	MPI
多	ZZ-LUNG	肺呼吸・肺シミュレーション	和田(東大)	Fortran, C, C++	Sphere(MPI)

## 第一走者の分類

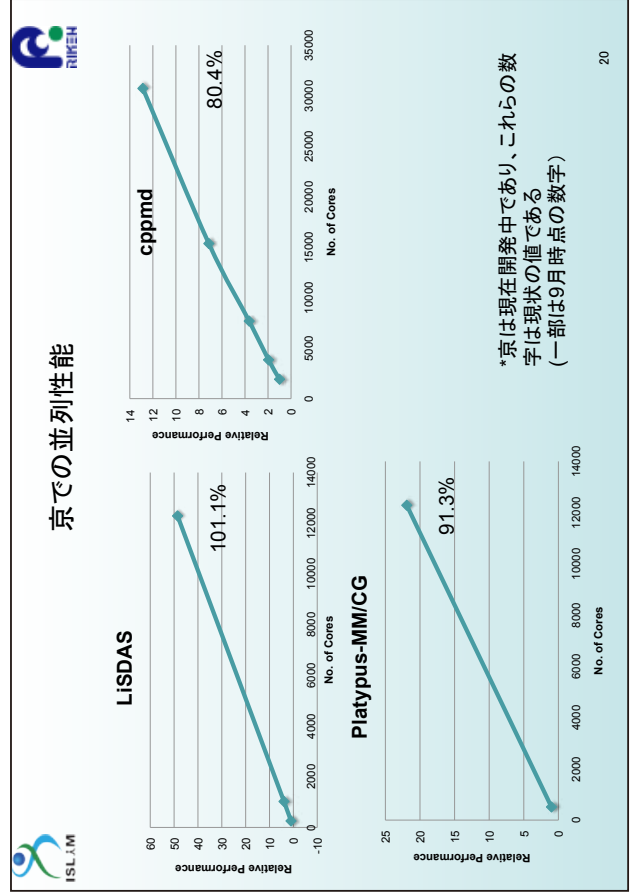
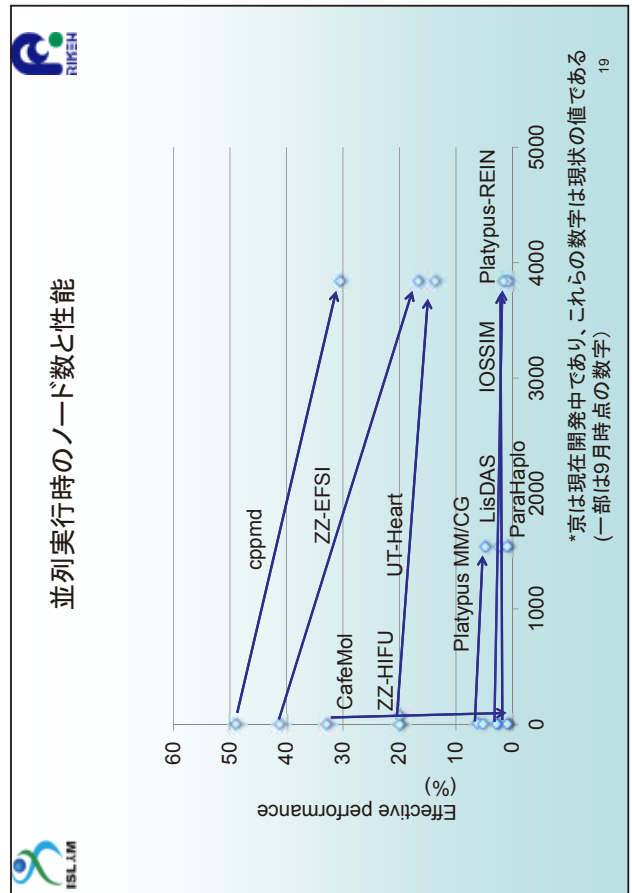
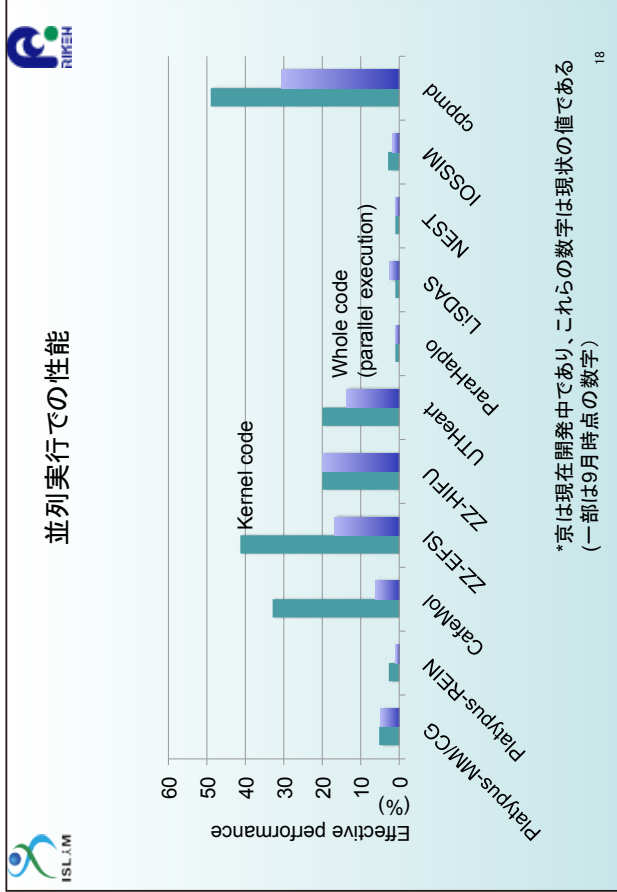
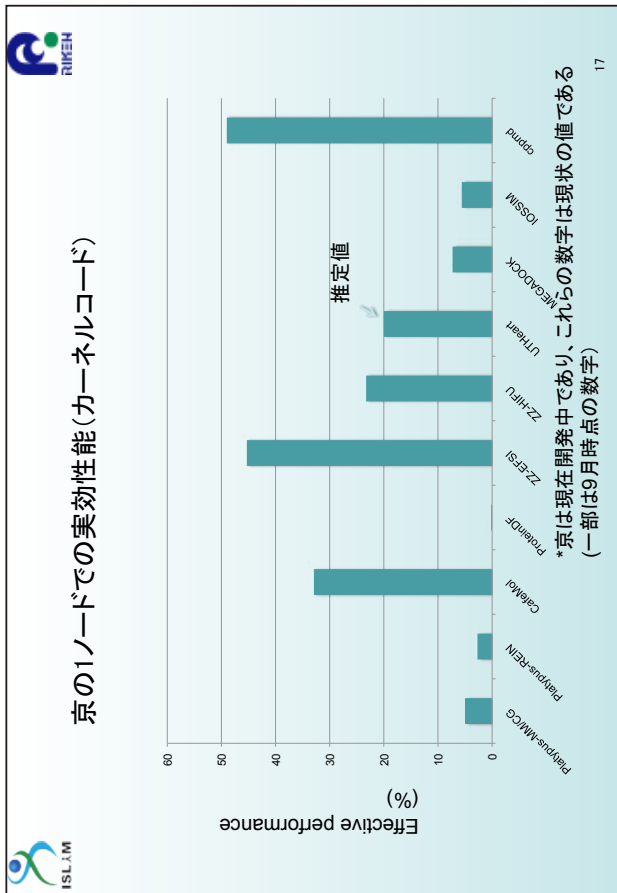


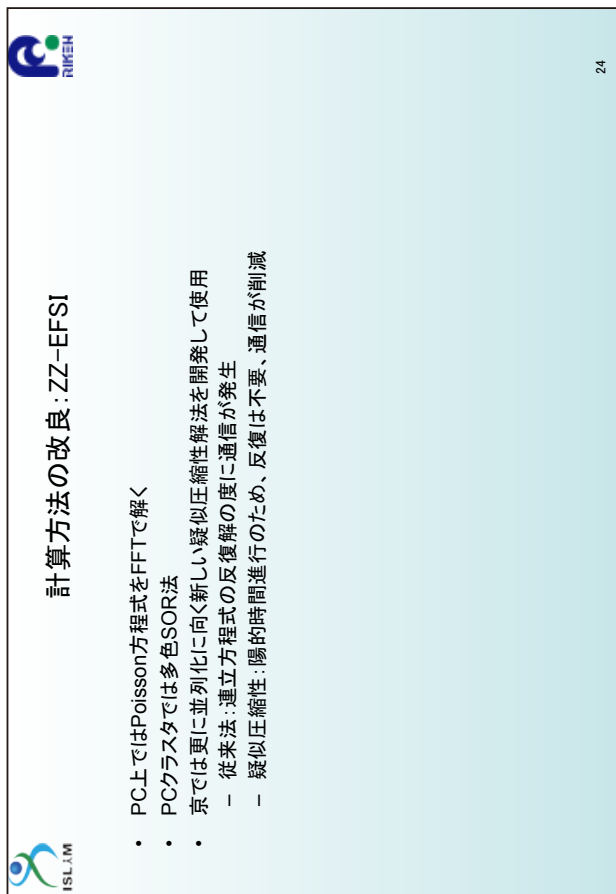
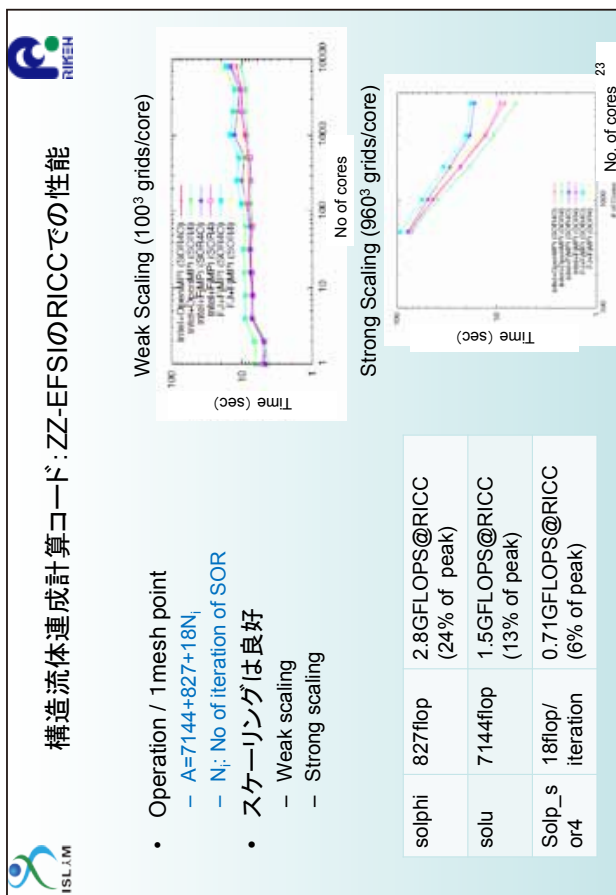
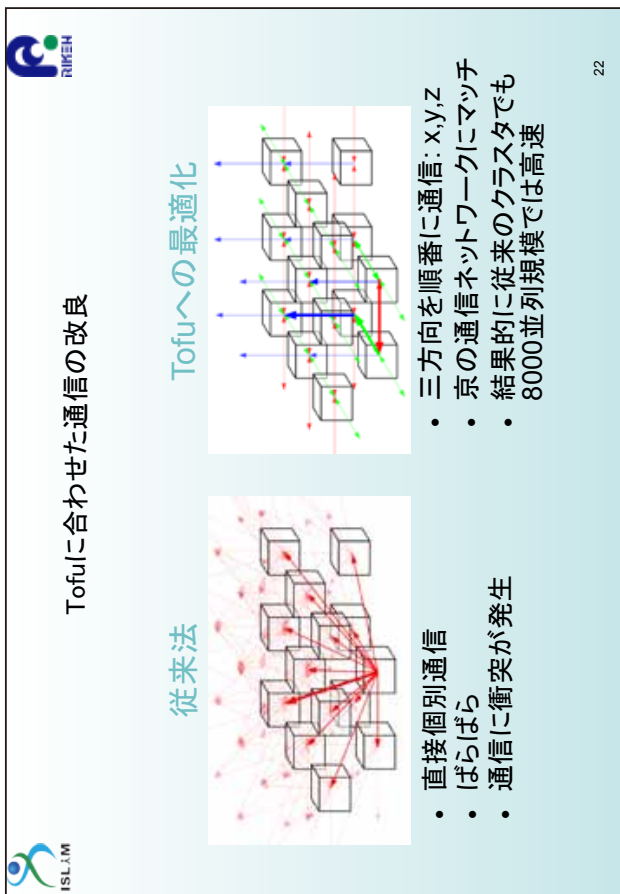
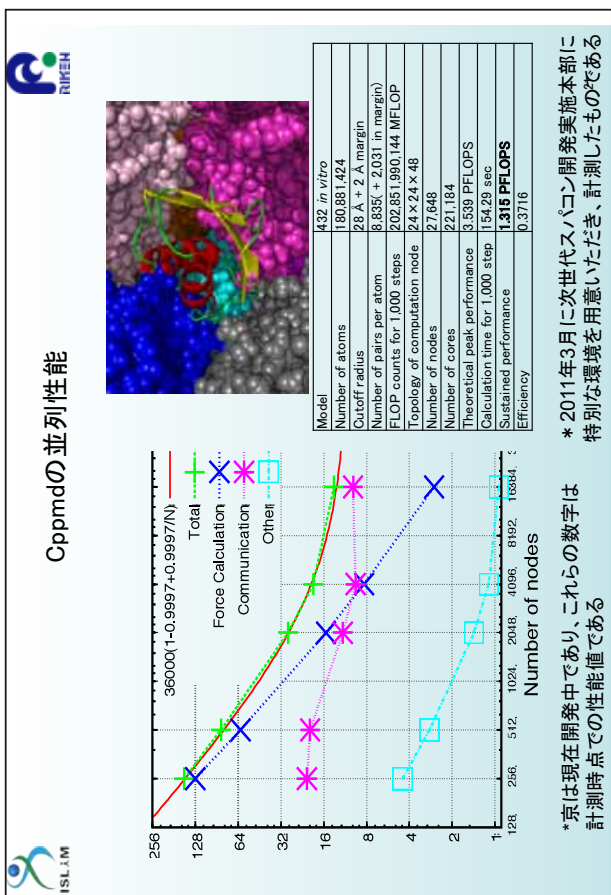
## それぞれの成果アピール時期















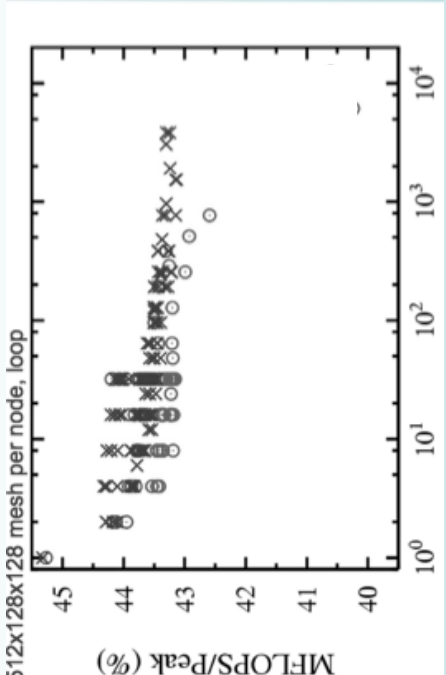
### まとめ

- 京のフル稼働に向けて着々とソフトの最適化を進行中
  - SIMD、hybrid並列に合わせたコード書き換え
  - Tofuネットワークに合わせた通信の最適化
  - 超並列アルゴリズム
- 最適化後のソフトは理論性能の約40%の実効性能
- 今後
  - コードの最適化は継続し、さらなる性能向上に努める
  - 同時に生命科学・医療で成果を上げ、ソフトを普及させる方向に力をシフトする

謝辞：京での計算に関しては京速コンピュータの試験利用、および本年3月での特別運用での結果です。また、PCクラスタでの性能計測に関しては理化学研究所情報基盤センターのRICCを使用しています。

### ZZ-EFSIの京での実効性能



\*京は現在開発中であり、これらの数字は現状の値である

(謝辞)

本資料集に記載されている「京」での計算は、2011年3月の「京」の特別運用およびその後の試験利用によって行われたものです。

また、本資料集に記載されている「京」を使用した測定値は、開発整備中の「京」による、測定時点での数値です。