

KMATHLIB: 京での数値計算ライブラリ整備 の試みについて

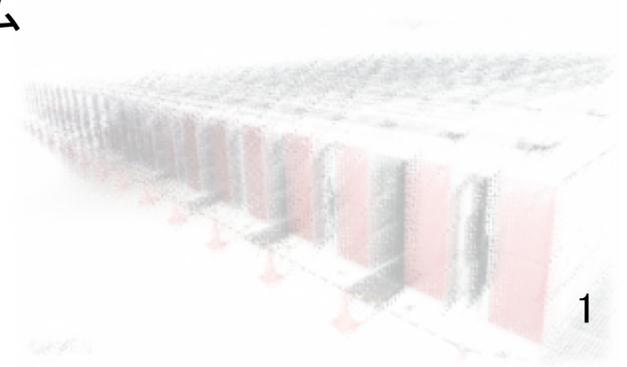
今村俊幸¹³

共同研究者: 山田進²³, 町田昌彦²³
櫻井鉄也³, 山本有作³, 張紹良³, 蔵増嘉伸³, 星健夫³

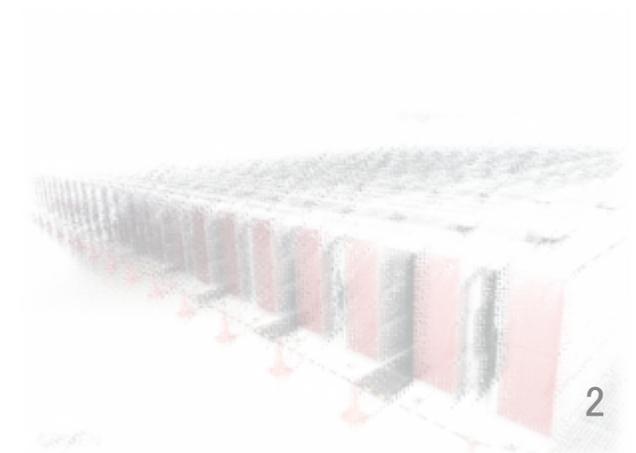
1: 理化学研究所 計算科学研究機構
大規模並列数値計算技術研究チーム

2: 日本原子力研究開発機構

3: JST CREST



1. はじめに
2. 数学ソフトウェア
3. KMATHLIB開発計画
4. 戦略分野との連携
5. まとめ



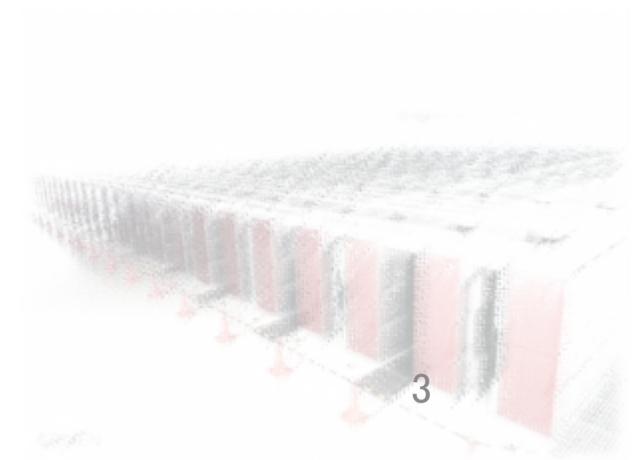


§1

はじめに

2013/01/16

第5回 HPCI戦略プログラム合同研究交流会



• Top500, November 2012

Rank	Site	System	Cores	Rmax (TFlops)	Rpeak (TFlops)	Power (KW)
1	DOE/SC/Oak Ridge National Laboratory, United States	Titan - Cray XK7, Opteron 6274 16C 2.20GHz, Cray Gemini Interconnect, NVIDIA K20x, Cray Inc.	560640	17590.0	27112.5	8209
2	DOE/NNSA/LLNL, United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom, IBM	1572064	16324.0	29132.7	7990
3	RIKEN Advanced Institute for Computational Science (AICS), Japan	K computer, SPARC64 VIIIx, Fujitsu	705024	10510.0	11280.4	12660
4	DOE/SC/Argonne National Laboratory, United States	Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom, IBM	786432	8162.4	10066.3	3945
5	Forschungszentrum Juelich (FZJ), Germany	JUQUEEN - BlueGene/Q, Power BQC 16C 1.60GHz, Custom Interconnect, IBM	393216	4141.2	5033.2	1970
6	LabInfo-Rechenzentrum, Germany	SuperMUC - DataPlex DX3RIM4, Xeon E5-2680 8C 2.70GHz, InfiniBand FDR, IBM	147456	2097.0	3105.1	3423
7	Texas Advanced Computing Center/Univ. of	Stampede - PowerEdge C8220, Xeon E5-2680 8C	204900	2660.3	3959.0	



1st TITAN <http://www.olcf.ornl.gov/titan/>

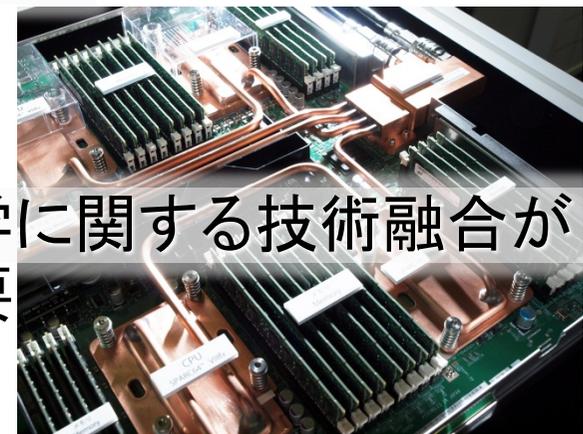
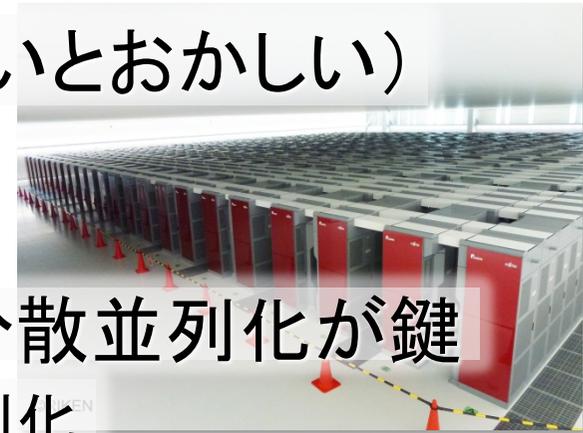


2nd Sequoia <https://asc.llnl.gov/>

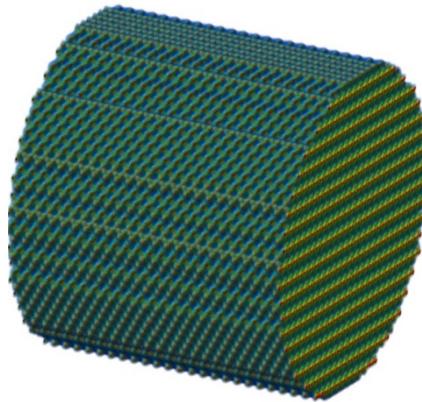


3rd K-computer <http://www.aics.riken.jp/>

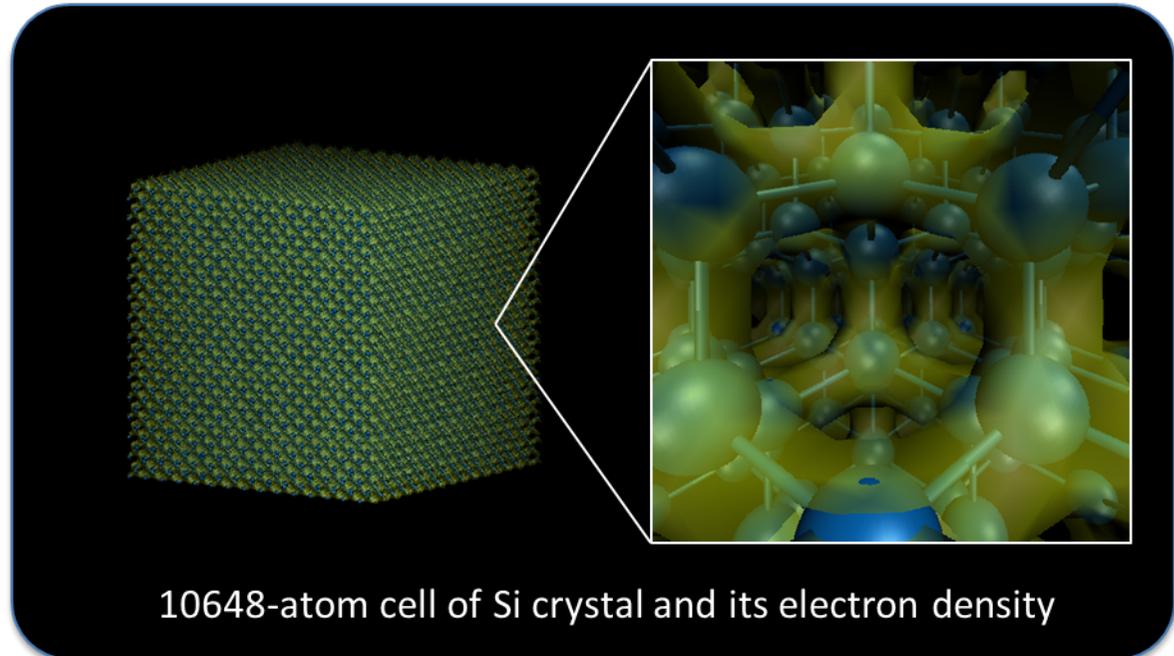
- 10PFLOPS級システムの衝撃
 - 100T~1PFLOPS が一般ユーザの通常利用範囲に
 - Kでは1000ノード以上が簡単に流せる
 - 科学シミュレーションは変わる(ないとおかしい)
 - 数桁以上のスケールアップ
 - 高精度計算結果への期待
 - パラメタサーベイの使い方でも、分散並列化が鍵
 - 単一ノードでもメモリ不足→分散並列化
 - 分散並列版数学ソフトウェア
 - 数学ソフトウェアの重要性
 - 高度に洗練された数学と計算機科学に関する技術融合がペタスケールの計算機環境では必要



- For grand challenge applications like a nano-material simulation, they demand a high-performance eigensolver strongly.



Silicon nanowire



10648-atom cell of Si crystal and its electron density

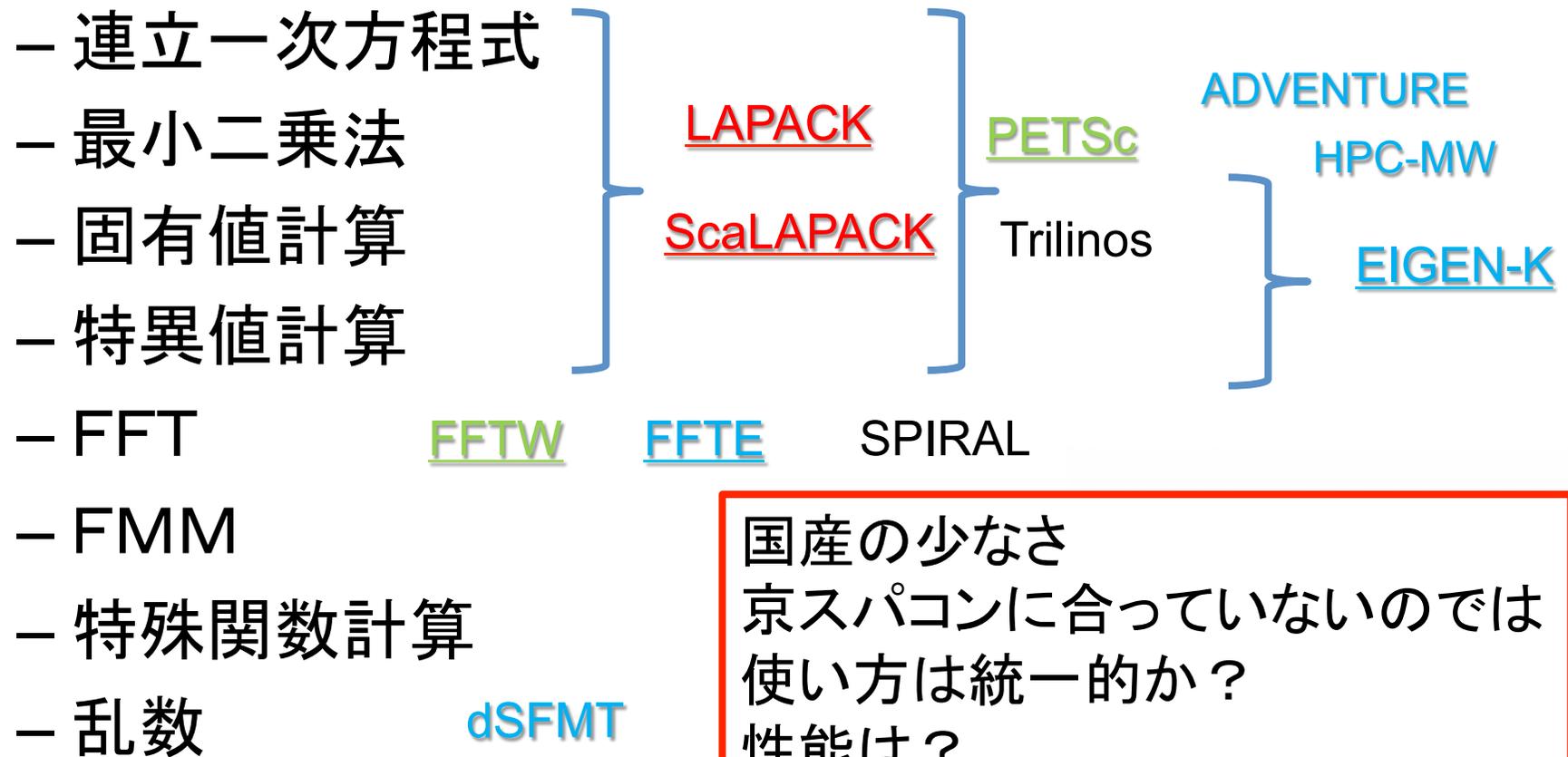
Y.Hasegawa, et al @AICS.RIKEN, GBP winner, SC2011

§2.

PFLOPS級ソフトウェアに欠かせないもの ＝「並列化された」数学ソフトウェア

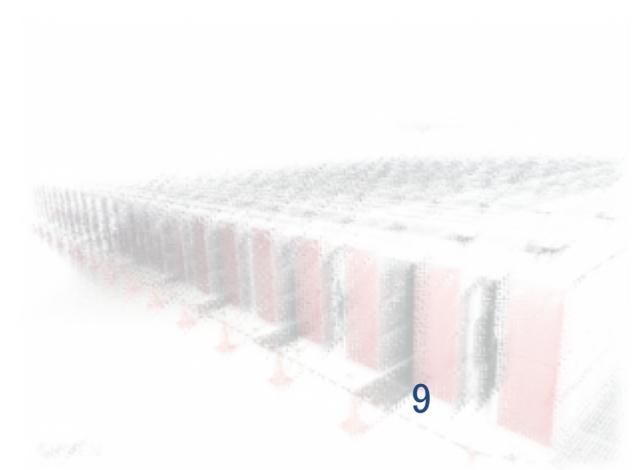
- シミュレーションに必要な数値計算法と数学ソフトウェア（分散並列で実績のあるものなど）

下線付：京もしくはFX10で対応済



§3.

KMATHLIB開発計画



京の高度利用化のための研究部門

- 計算機科学(5チーム), 共通基盤(3チーム), 計算科学(8チーム +3ユニット)

大規模並列数値計算技術研究チーム

- * 京における数値計算ライブラリの研究開発
- * 統一的なAPI整備とデータ形式サポートが目標



大規模並列数値計算アルゴリズムおよびライブラリの研究開発 新報

- 京コンピュータにおける数千万コアを超える超並列システムにおいては、並列性を十分に引き出すための数値計算ライブラリの充実が急務。
 - 一般的には現在、米国を中心とした数値計算ライブラリ(OpenMP, PETSc, など)が主に使われている。
 - 京においては基盤的にハイパーカー(専ら通信提供の数値計算ライブラリ)に依存しているが、数億コア並列規模に必要となる、カスタムライブラリをはじめ、いっさいのアプリにおいては大学・研究機関等で開発した独自の数値計算ライブラリを有用しており、標準化されていないのが現状。
 - 計算科学アプリの開発者が、京を使いこなすには並列性を十分に引き出すための数値計算ライブラリの充実が必要
- 京を使いこなし、さらに効果的に使えるようにするためには、以下のような課題・制約の克服・解消が必要
 - 京の一斉に障害が発生した場合、通常はアプリケーションの計算が停止する。障害が発生した場合においても障害箇所を回避して計算を継続するよう数値計算アルゴリズムが必要。
 - 数億コアに分散された計算結果を統合する際、計算結果が大規模になるにつれて計算結果に誤差が生じるため、大規模システムにおいては、大規模問題の計算結果と精度を確保する、誤差解析・結果検閲手段が必要。
 - 並列の数値計算ライブラリは、高並列のアプリケーションに導入できない、並列列に対して高次元数値計算の研究開発及びライブラリの提供が必要。
- このために以下のような研究を行い、新たな、従来にない先進的な並列数値計算ライブラリを開発し、ユーザーの利用に供する幅広い分野の多くのユーザーの利用の進展につなげる

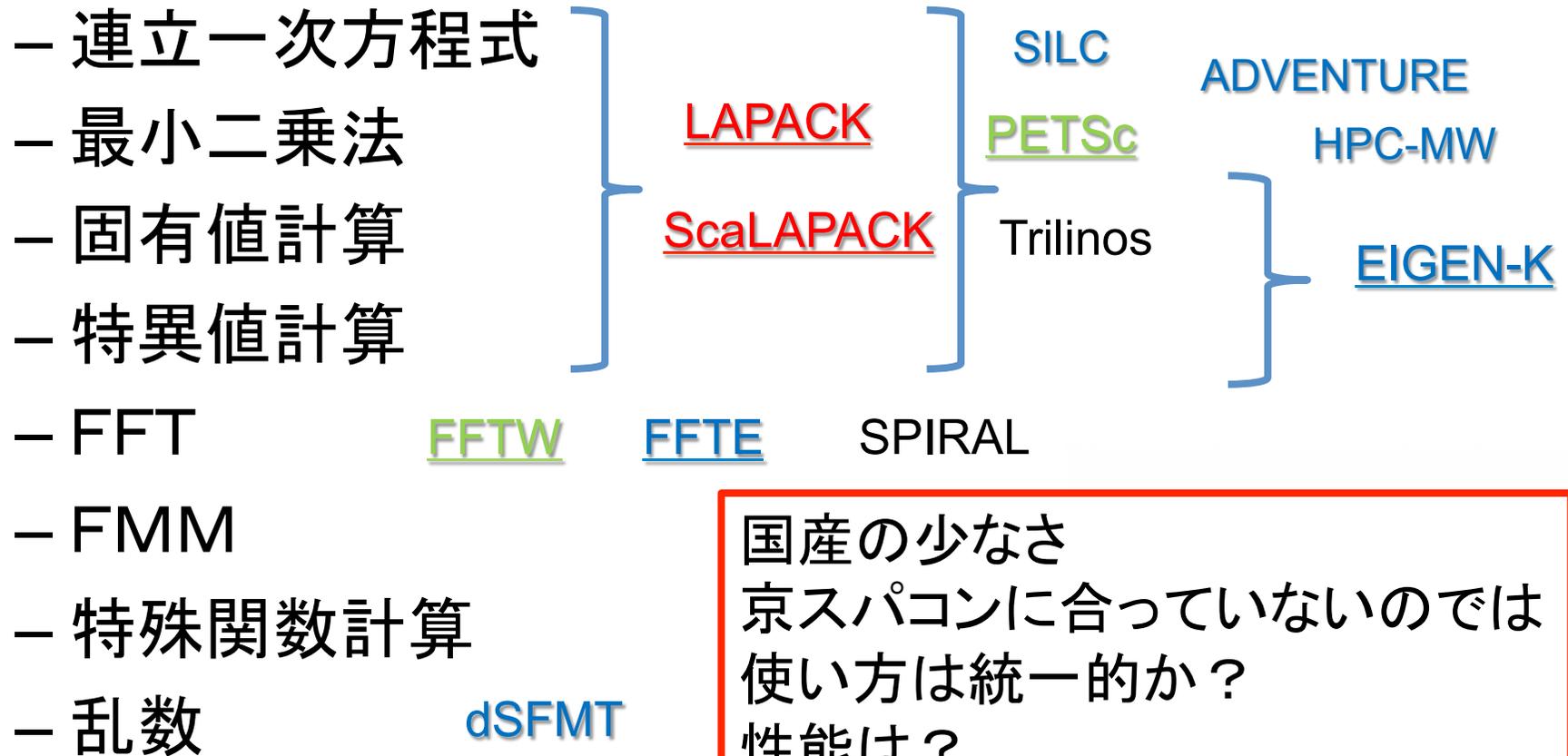
(研究開発目標)

 - 京の一部に障害が発生しても、計算を継続できる、耐障害機能を持つ数値計算アルゴリズムおよびソフトウェアの開発
 - 大規模になった時に問題になる高次元計算および精度検閲手段の開発
 - 京の高性能プロセッサを活用し、数億並列に対応する疎行列ソルバー、固有値ソルバー、2次元FFTなどの並列数値計算アルゴリズムおよびライブラリを開発
- 本研究は計算科学者が新たに設計開発するのではなく、計算科学分野の研究をベースとして行うことが効率的
- 戦略機関との連携

戦略機関には計算科学分野がないため、本研究開発は連携を求められている。

- シミュレーションに必要な数値計算法と数学ソフトウェア（分散並列で実績のあるものなど）

下線付：京もしくはFX10で対応済



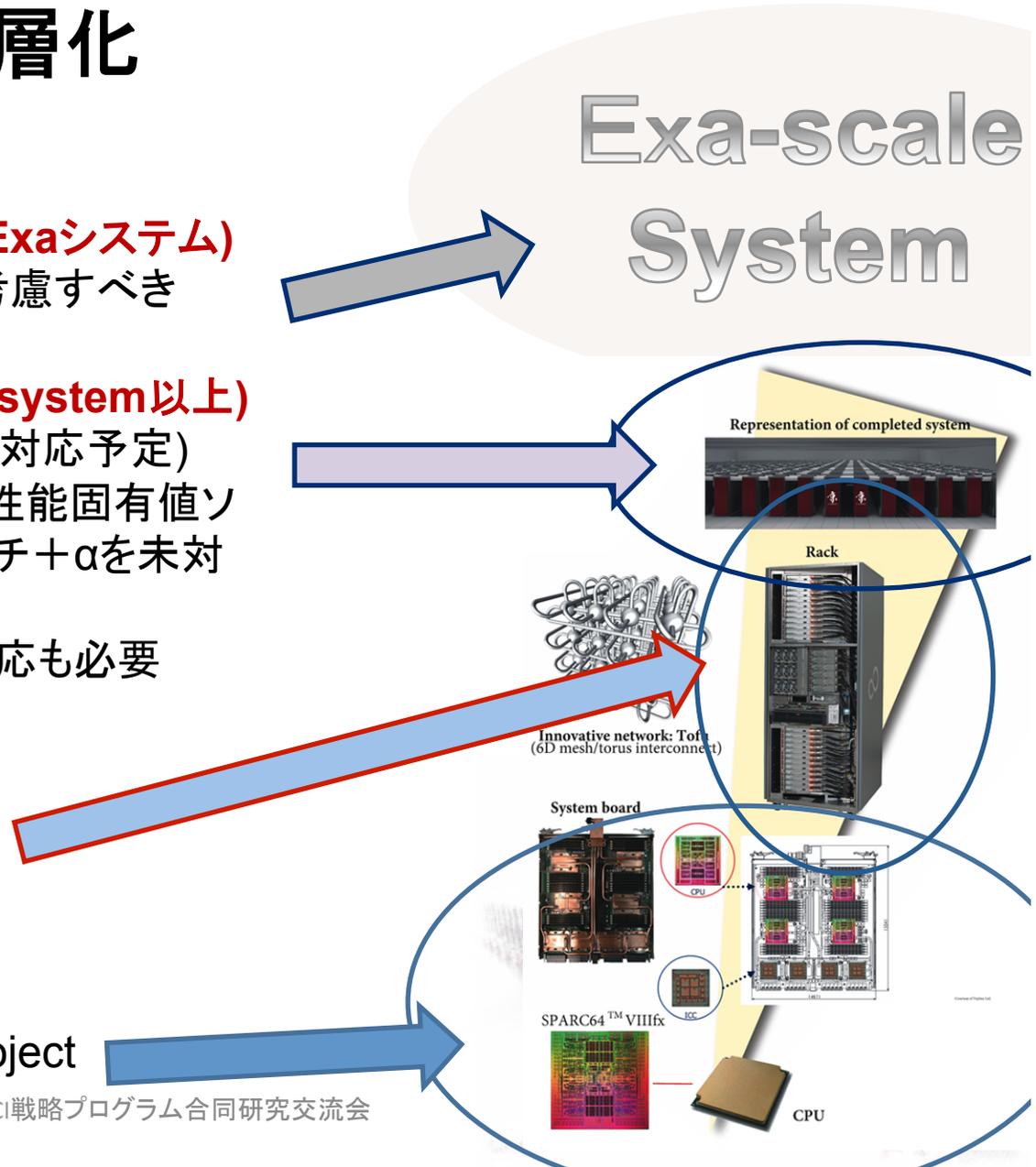
• 並列度に応じた階層化

□ **O(100K)ノード、O(M)~コア (Exaシステム)**
 超難: Hetero・階層的な並列性を考慮すべき

□ **O(K)~O(100K)ノード (京full system以上)**
 米澤先生CREST(H24年度で一部対応予定)
 FX10&T2KでO(K)ノードまでの高性能固有値ソルバは開発済み、同様のアプローチ+αを未対応ソルバに適用
 (後述)Tofuなどネットワークへの対応も必要

□ **~O(K)ノード程度**
 従来数値ライブラリで対応可能

□ **逐次版もしくはノード内並列**
 階層マルチコア対応
 参考: 米国PLASMA&MAGMA project



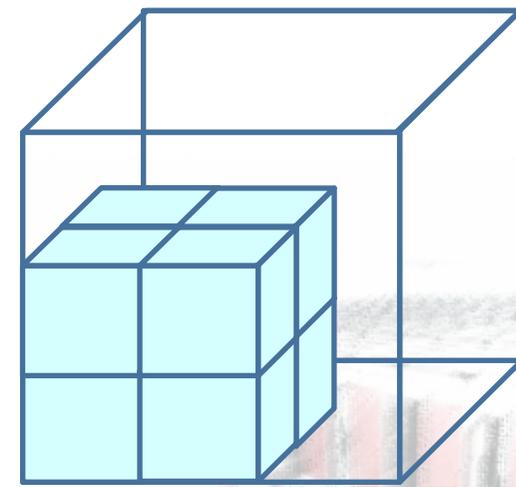
- 各階層に合ったソルバの整備
 - $O(\sim K)$, $O(\sim M)$, $O(M\sim)$ の3階層で、メタドライバAPIにより内部ソルバを自動選択, → 規模・ソルバの差異を意識させない様に
 - H24~25: 既存OSS (性能+メタドライバAPI&データ変換)調査
 - H25から利用頻度が高く性能が出ていないものを独自開発

	$\sim O(K)$ cores	$\sim O(100K)$	$\sim O(M)$	$O(M)\sim$
SSL/II	??	??	??	??
ScaLAPACK	OK, so so	??	NG	??
PETSc	??	??	??	??
Trilinos	??	??	??	??
FFTW	??	??	??	??
FFTE	OK	??	??	??
dSFMT	??	??	??	??

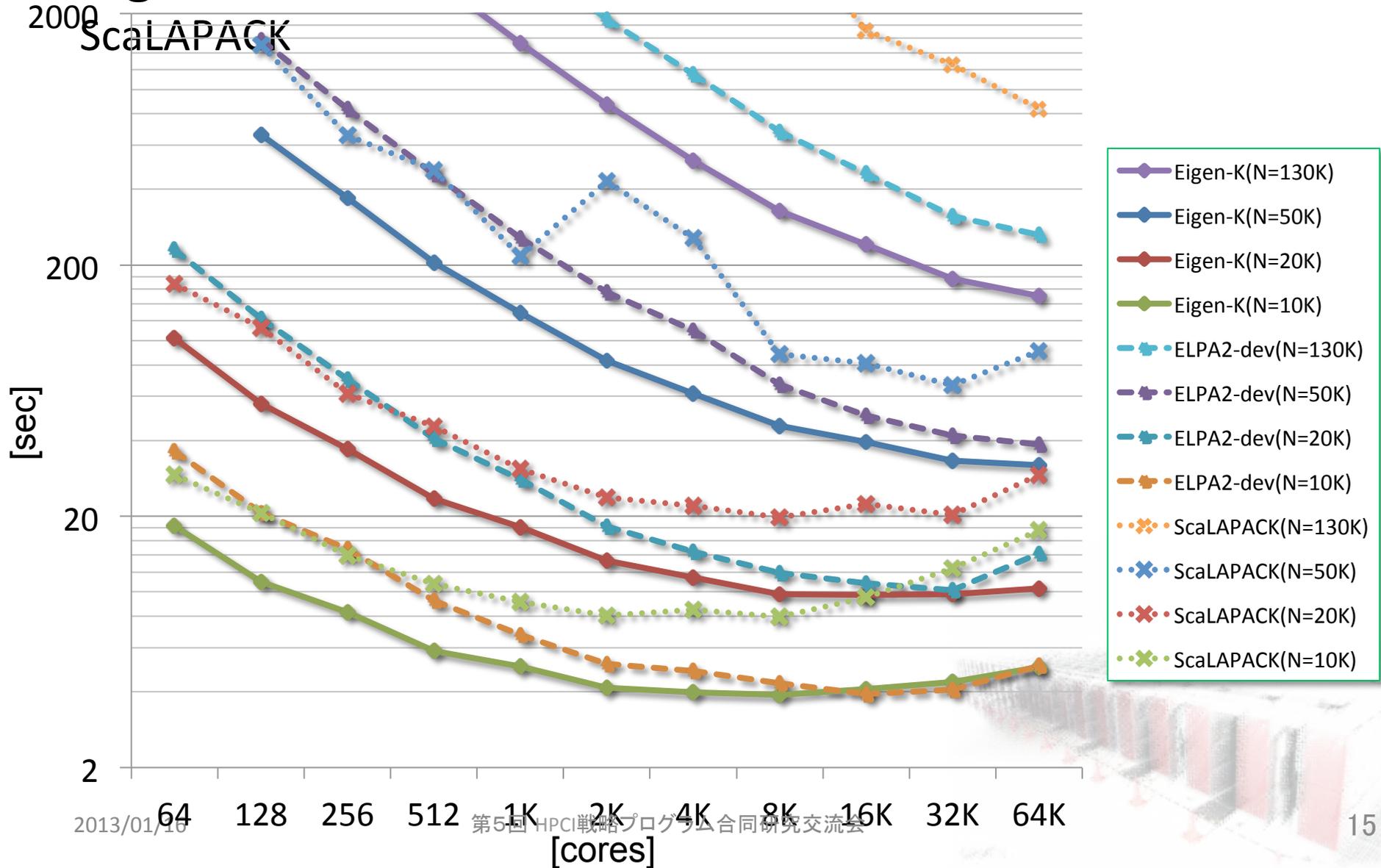
- 固有値計算
 - 標準固有値問題 (Eigen-K/Exa by CREST&JAEA)
 - 実対称 JAEAより公開中
 - エルミート JAEAより公開中
 - **一般化固有値問題**
 - 実 開発中
 - エルミート H25開発予定
 - 非対称固有値問題
 - 既存ソルバーの性能調査と高並列アルゴリズムを選定中
 - 疎行列版は櫻井先生(@筑波大)開発の(z-Pares)を取り込む予定
- FFT
 - **三軸分割(cube分割)FFT**
 - 既存の2軸分割(pencil分割)より高並列に
- 乱数
 - **分散プロセス乱数生成器**
 - メルセンヌツイスター乱数生成器をエンジンに
 - 分散環境下で乱数系列の一貫性を確保

$$Ax = \lambda x$$

$$Ax = \lambda Bx$$

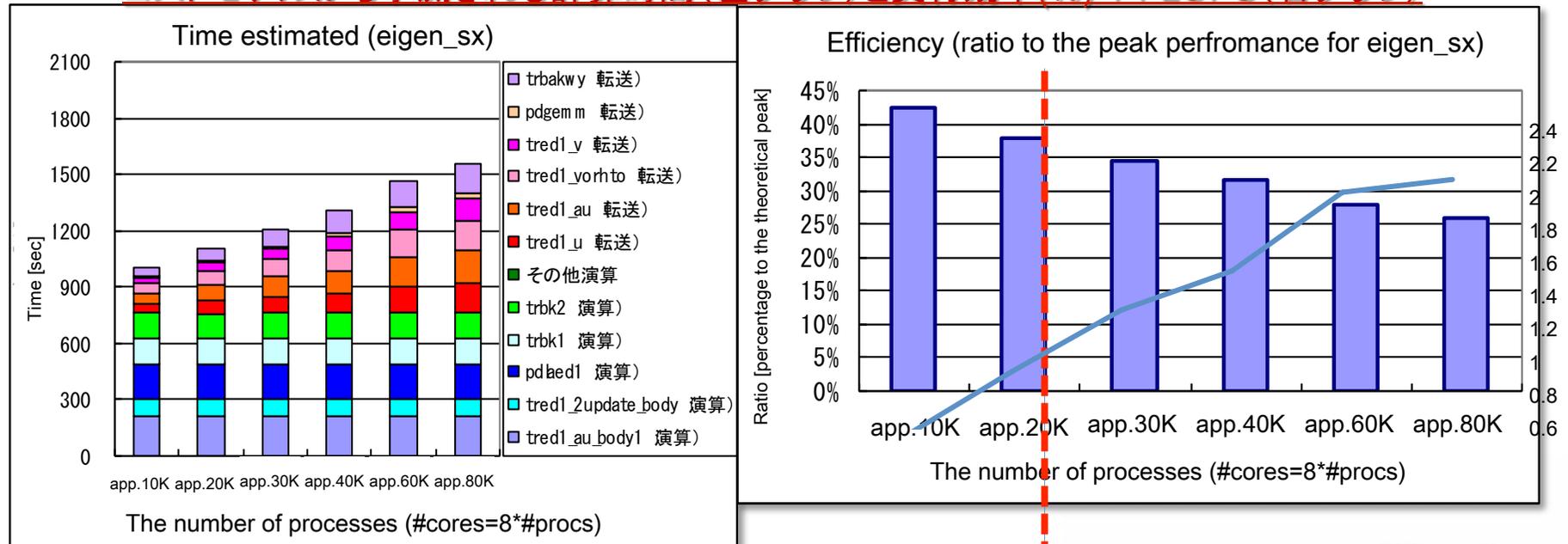


- Eigen-K (2012.12, N=10K, 20K, 50K, 130K), ELPA2-devel, and ScaLAPACK



- 数値ライブラリは大規模であるほど効果大
 – 小規模問題も最適化されています

コストモデルから予測される計算時間(左グラフ)と実行効率(%)・PFLOPS(右グラフ)



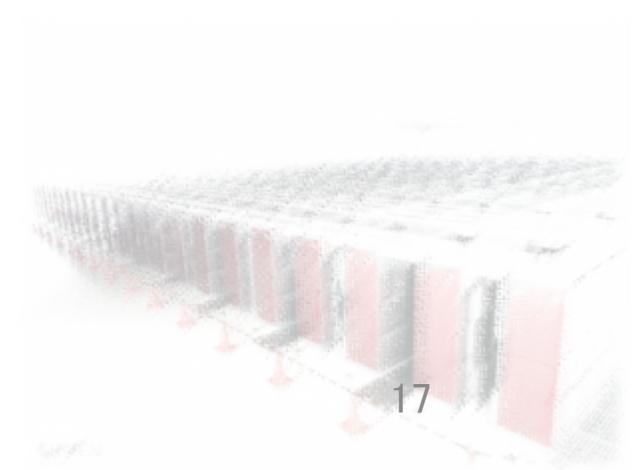
N	#procs
437,149	App.10K
549,908	20K
693,452	30K
794,145	40K
908,238	60K
1,000,000	80K

→ PFLOPS

*This estimation is **optimistic**, so it might reduce down to less than 20% in the worst case.*

§4.

戦略分野との連携



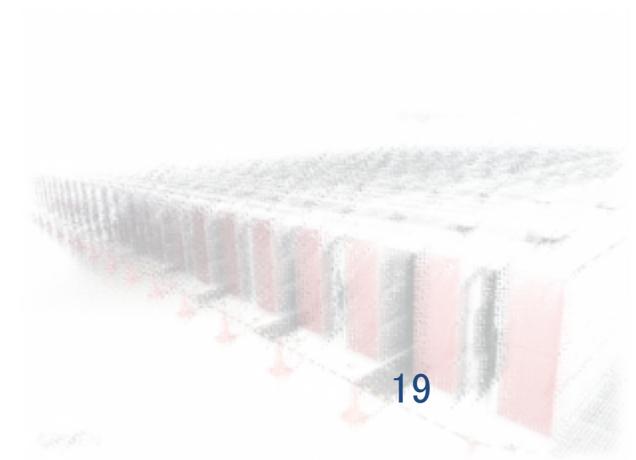
- 京ユーザーへのヒアリング・アンケートを実施しています。
 - 戦略分野2:「新物質・エネルギー創成」
 - 固有値計算、FFT
 - 産業界との接点::分野1や分野4など
 - 連立一次方程式、固有値計算、FFT?

- 数値計算部分は独自に開発をされる前に、並列化された数値ライブラリを使っていただけだと思います。

- また、京などの分散並列環境の利用促進に資する目的で
 - 教育ツール: チュートリアル作成
 - 使い勝手: 抽象度を上げたドライバ関数
 - 問題規模: 1000x1000? or $10^6 \times 10^6$?
 - 精度: 倍精度浮動小数点よりも低or高精度なデータ型
といった部分も視野にいれています

§5.

その他＋まとめ



- 京コンピュータを中心に
 - 数値計算ライブラリK-MATHLIB整備
 - 連立一次方程式(PETSc)、最小二乗法、固有値計算(Eigen-K)、特異値計算、FFT(FFTW&FFTE)、FMM、特殊関数計算、乱数(dSFMT)
- 京ユーザーだけでなく、皆様からの数値計算ライブラリへのご意見をお待ちしております。
 - 一般的なご意見、使い勝手、性能、精度など

