



全国共同利用施設

東京大学情報基盤センター

Information Technology Center, The University of Tokyo



東京大学
THE UNIVERSITY OF TOKYO

学際大規模情報基盤共同利用・共同研究拠点
Joint Usage / Research Center for Interdisciplinary Large-Scale Information Infrastructures



T2K(東大)の現状と 東大情報基盤センターの活動

中島研吾

東京大学情報基盤センター
スーパーコンピューティング研究部門

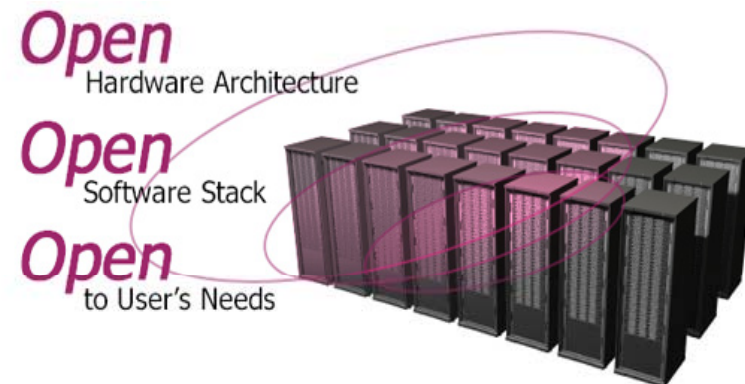
2010年7月12日, T2Kシンポジウム2010@東大

Overview

- T2K(東大)の概要
 - T2K(東大)がもたらしたものの
- T2K(東大)の現状
 - 稼働状況
 - 東大情報基盤センターの活動
- 今後の動向

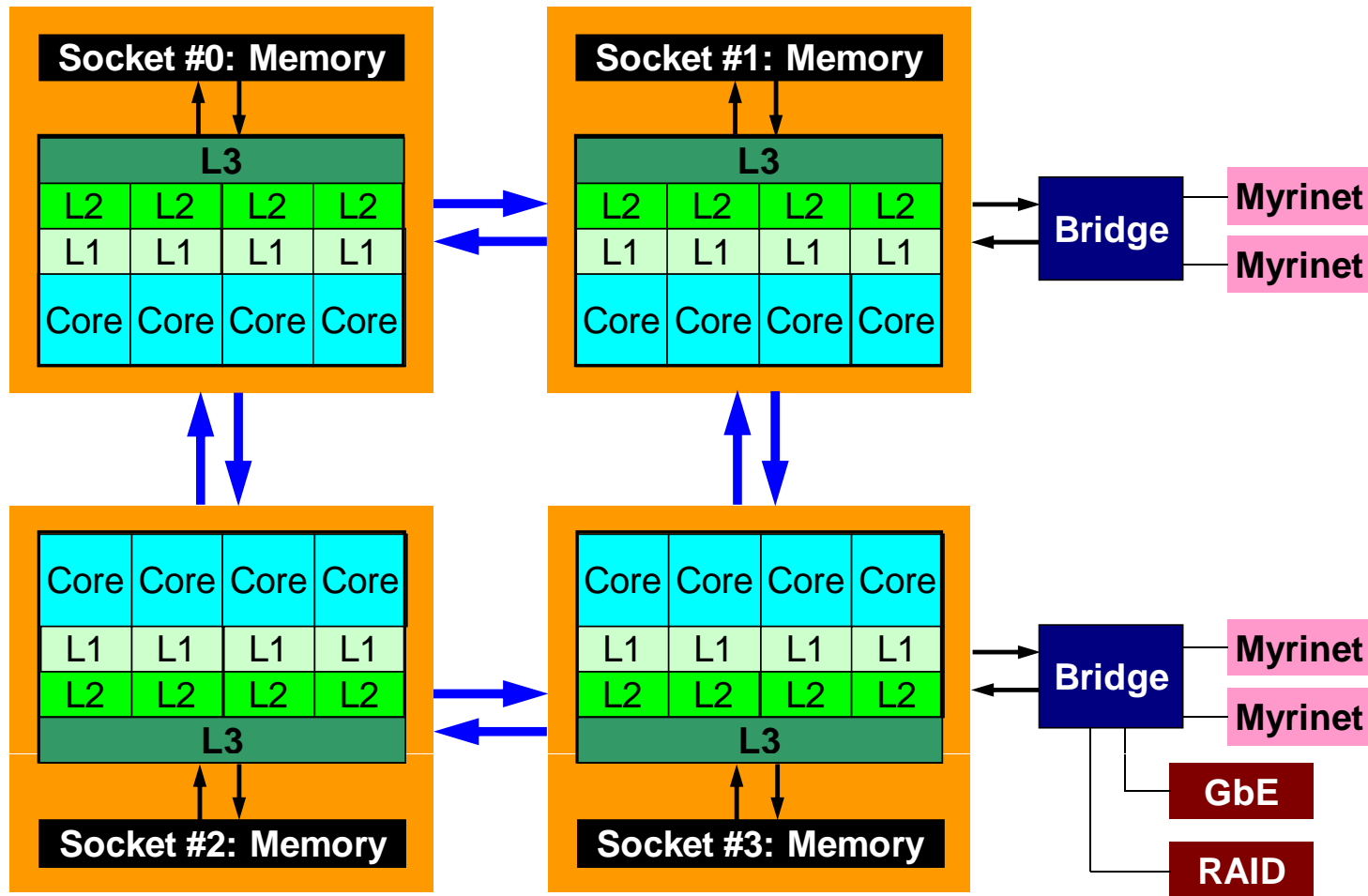
T2Kオーpensパソコン(東大), T2K(東大)

- 2008年6月稼動開始
- T2Kオーpensパソコン仕様
- 日立製作所製
 - 952ノード, 15,232コア
 - ピーク性能140.1 TFLOPS
 - HA8000クラスシステム



T2K(東大): Hitachi HA80000

- AMD Quad Core Opteron 2.3GHz (9.2GFLOPS)
- Quad Core × 4 ⇒ 1node(16cores)
- cc-NUMA

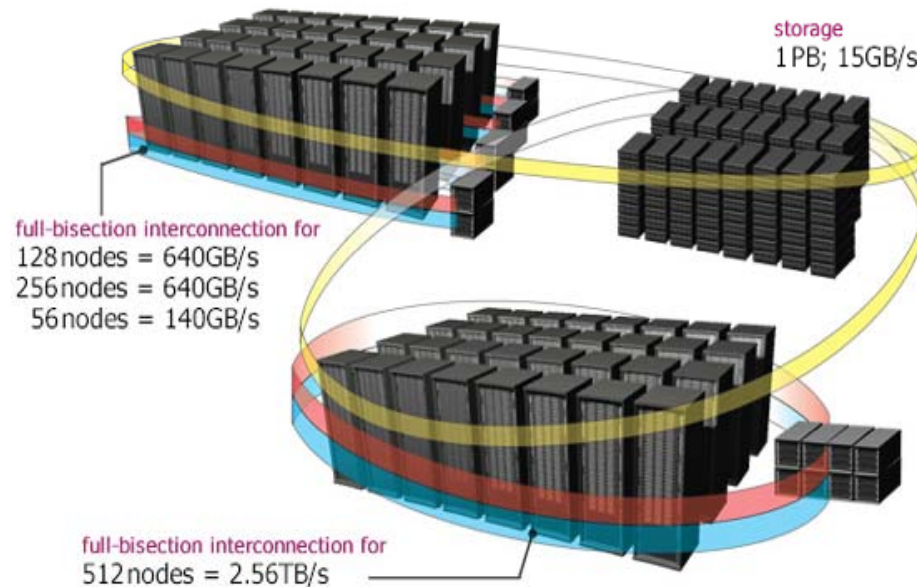


T2K Open Supercomputer Alliance

<http://www.open-supercomputer.org/>

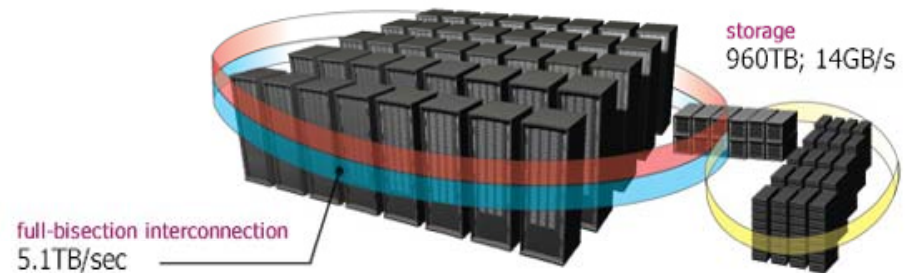
University of Tokyo

nodes = 952 Rpeak = 140.1TFlops Memory = 31TB



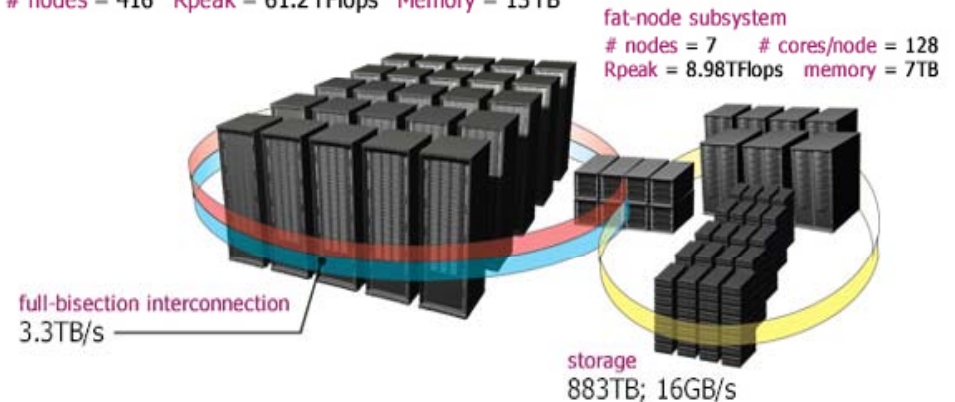
University of Tsukuba

nodes = 648 Rpeak = 95.4TFlops Memory = 20TB



Kyoto University

nodes = 416 Rpeak = 61.2TFlops Memory = 13TB



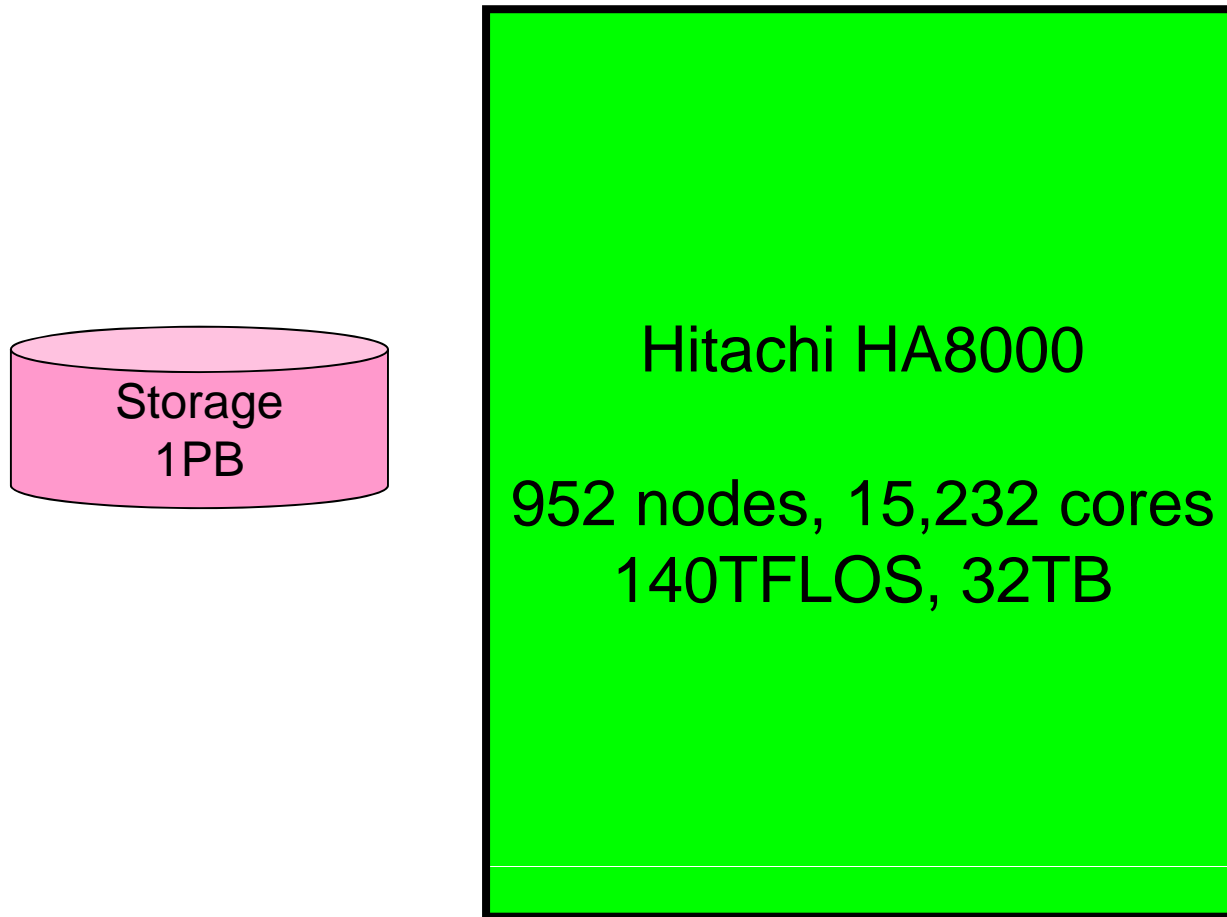
T2Kオープンスパコン達 @SC08



T2Kオープンスパコン達 @SC08

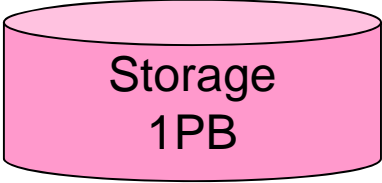


T2K Open Supercomputer (Todai Combined Cluster)

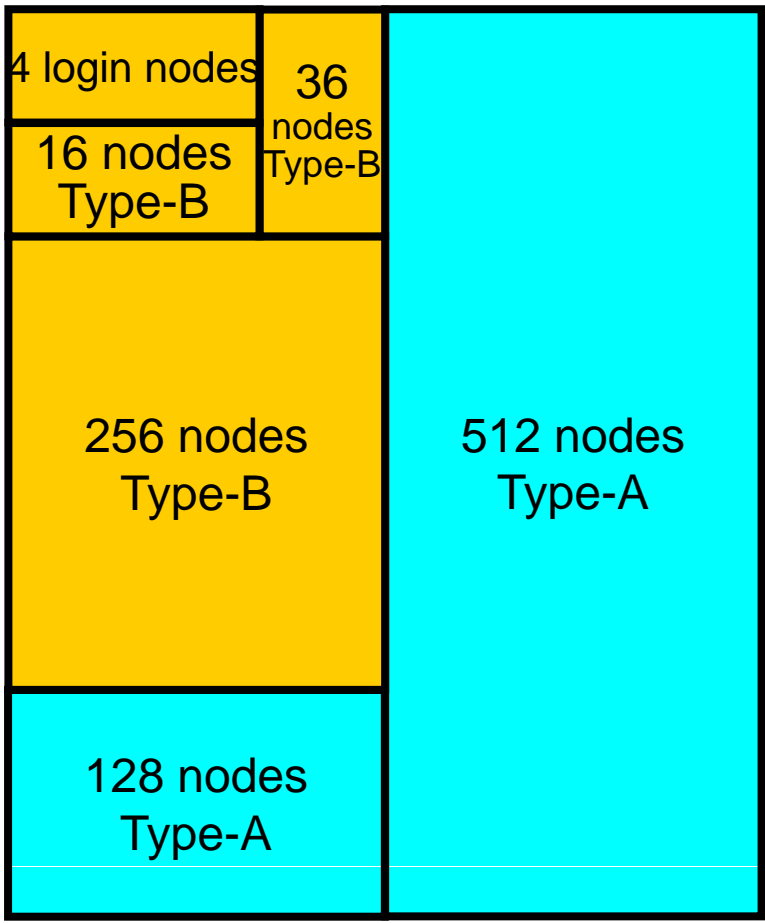


T2K Open Supercomputer (Todai Combined Cluster)

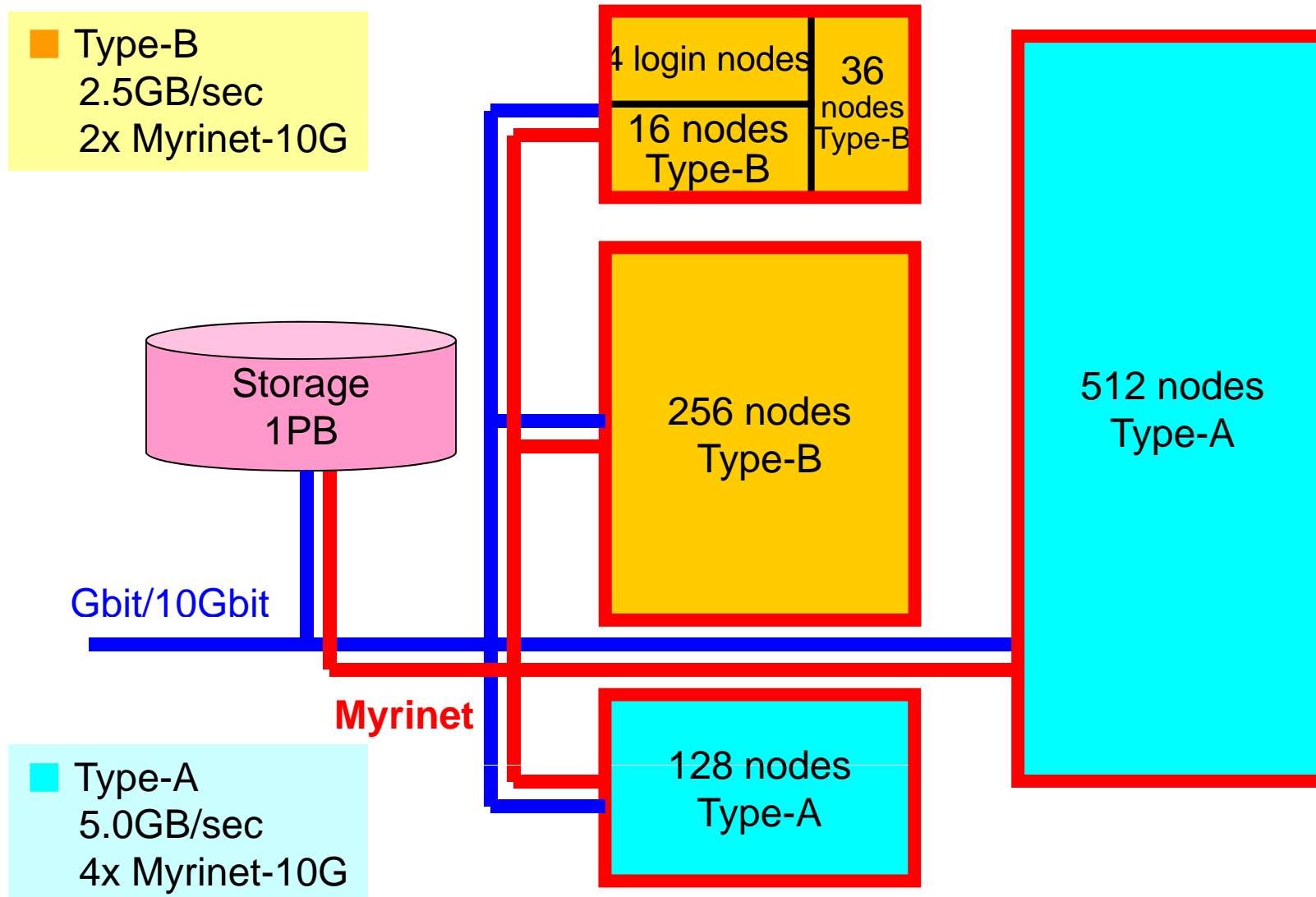
■ Type-B
2.5GB/sec
2x Myrinet-10G



■ Type-A
5.0GB/sec
4x Myrinet-10G



T2K Open Supercomputer (Todai Combined Cluster)



Results of LINPACK

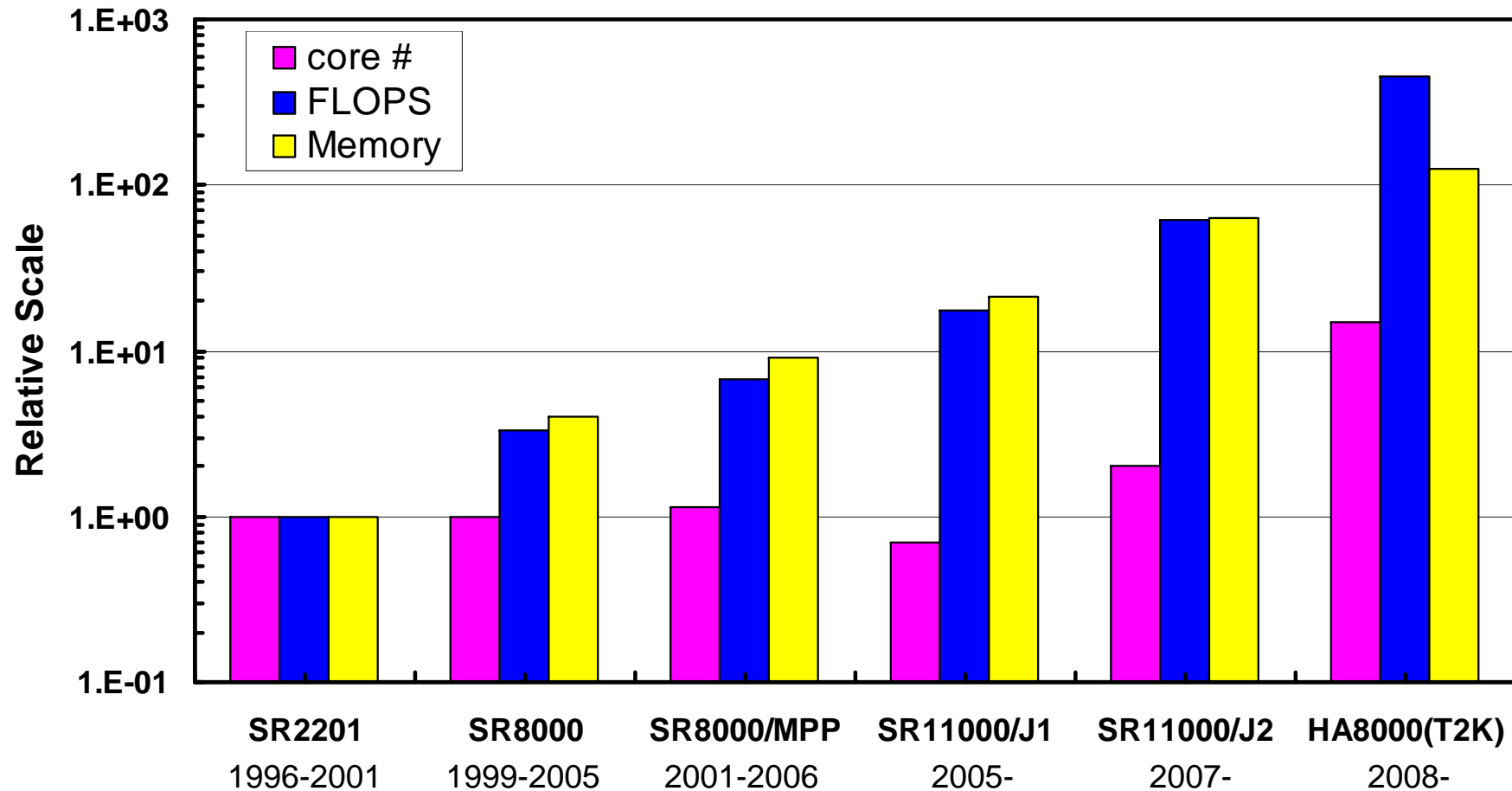
944 nodes, 139 TFLOPS peak

TOP500: 45th at Nov.09, 53rd at Jun.10

T/V	N	NB	P	Q	Time	Gflops
WC00L2R4	1740800	160	32	118	34566.94	1.017e+05
Ax-b _oo / (eps * A _1 * N) =					0.0004115 PASSED
Ax-b _oo / (eps * A _1 * x _1) =					0.0007797 PASSED
Ax-b _oo / (eps * A _oo * x _oo) =					0.0000009 PASSED

東大センタースパコンの歴史

SR2201 (ピーク性能300GFLOPS)を1



東京大学情報基盤センター スパコン

1システム～6年, 3年周期でリプレース

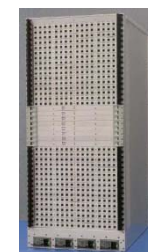
HITACHI SR11000 model J2

Total Peak performance	: 18.8 TFLOPS
Total number of nodes	: 128
Total memory	: 16384 GB
Peak performance per node	: 147.2 GFLOPS
Main memory per node	: 128 GB
Disk capacity	: 94.2 TB
IBM POWER5+ 2.3GHz	

T2K(東大)(HA8000クラスシステム)

Total Peak performance	: 140 TFLOPS
Total number of nodes	: 952
Total memory	: 32000 GB
Peak performance per node	: 147.2 GFLOPS
Main memory per node	: 32 GB, 128 GB
Disk capacity	: 1 PB
AMD Quad Core Opteron 2.3GHz	

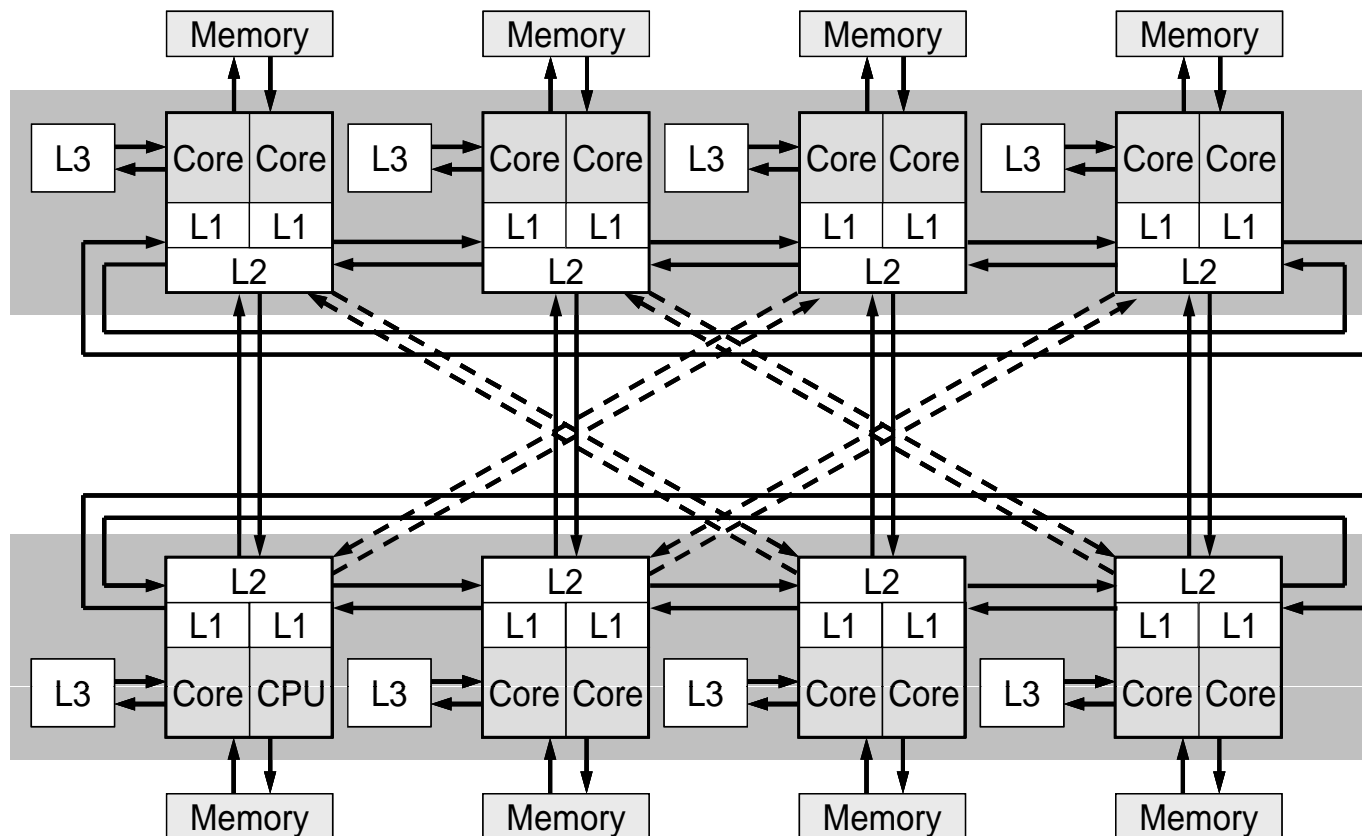
ノード製品名 : HITACHI HA8000-tc/RS425



Hitachi SR11000 model J2

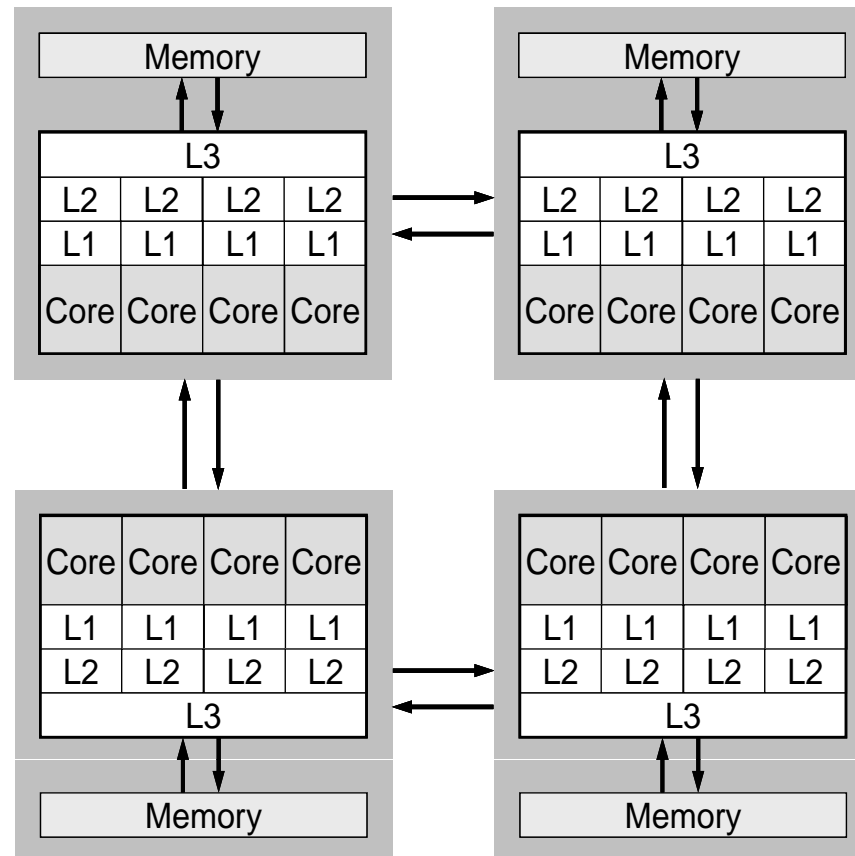
- IBM POWER5+ 2.3GHz (9.2GFLOPS)
- Dual Core × 4 ⇒ 「MCM (Multi Core Module)」 × 2 ⇒ 1node (16cores)
- based on NUMA architectures, but small latency of memory, huge cache

青木, 中村, 助川, 齋藤, 深川, 中川, 五百木 (2005)
スーパーテクニカルサーバSR11000モデルJ1のノードアーキテクチャと性能評価
情報処理学会論文誌: コンピューティングシステム 45-SIG12(ACS11), 27-36 より作成



T2K(東大): Hitachi HA80000

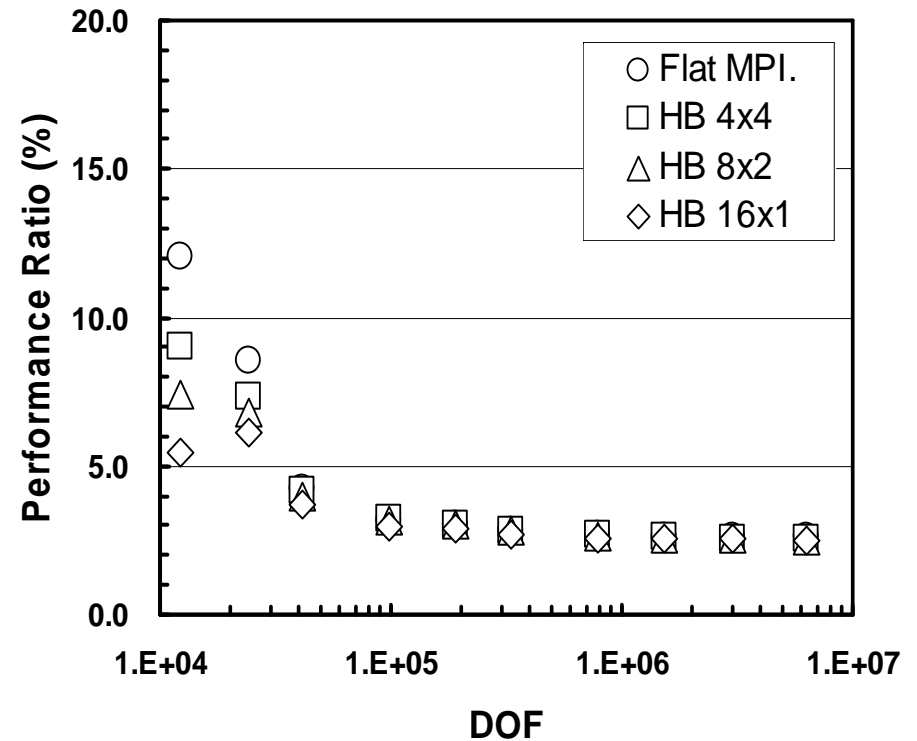
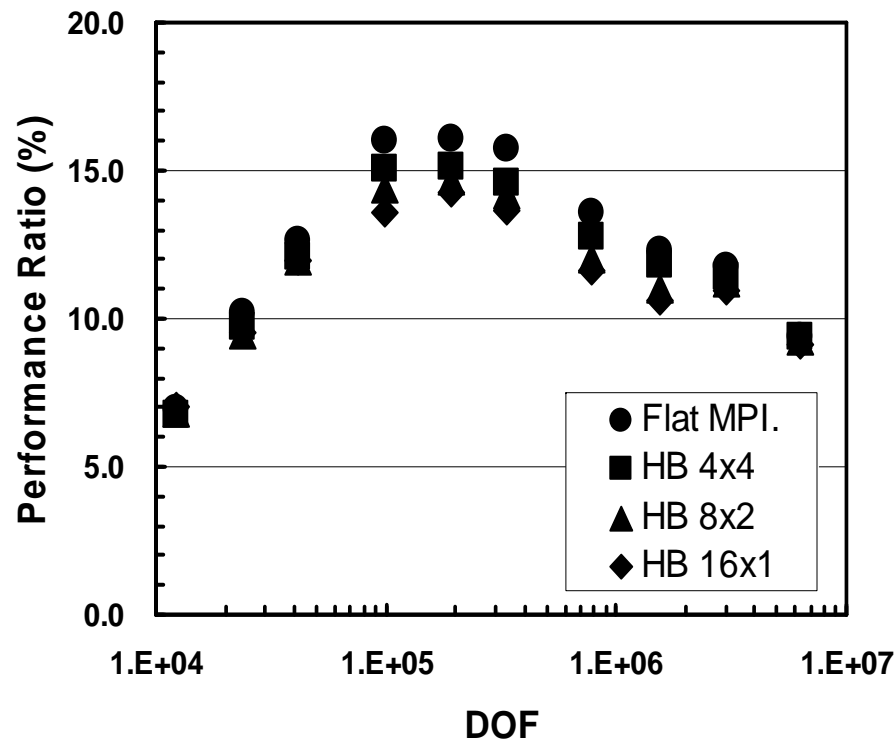
- AMD Quad Core Opteron 2.3GHz (9.2GFLOPS)
- Quad Core × 4 ⇒ 1node(16cores)
- ccNUMA
- Careful control of core-memory-data configuration required: numactl



GeoFEM Benchmark 1-node with 16-cores: ICCG

Hitachi SR11000/J2
Power 5+ 2.3GHz x 16
147.2 GFLOPS/node
100 GB/s for STREAM/Triadd

T2K/Tokyo
Opteron 2.3GHz x 16
147.2 GFLOPS/node
20 GB/s for STREAM/Triadd

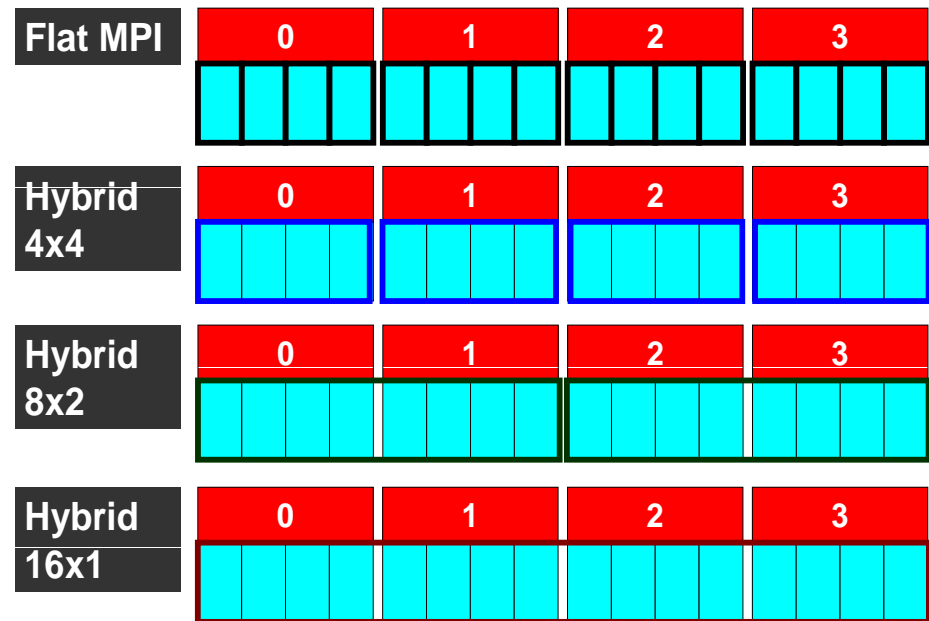
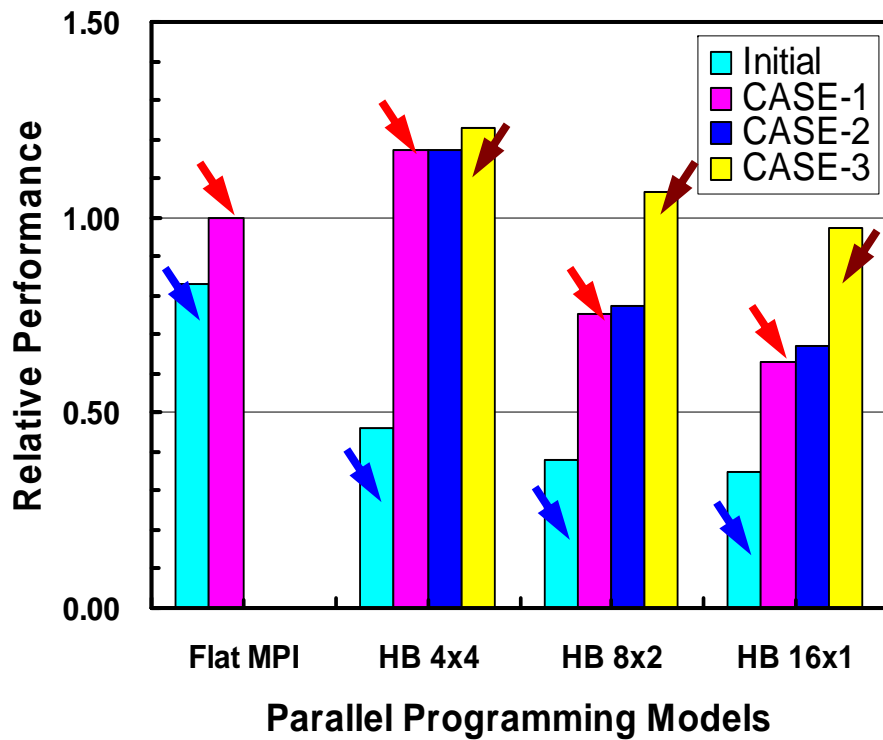


パーソナルコース(後述)で, 同じ料金で使える資源量はT2KはSR11000の4倍に設定 ⇒ 利用促進

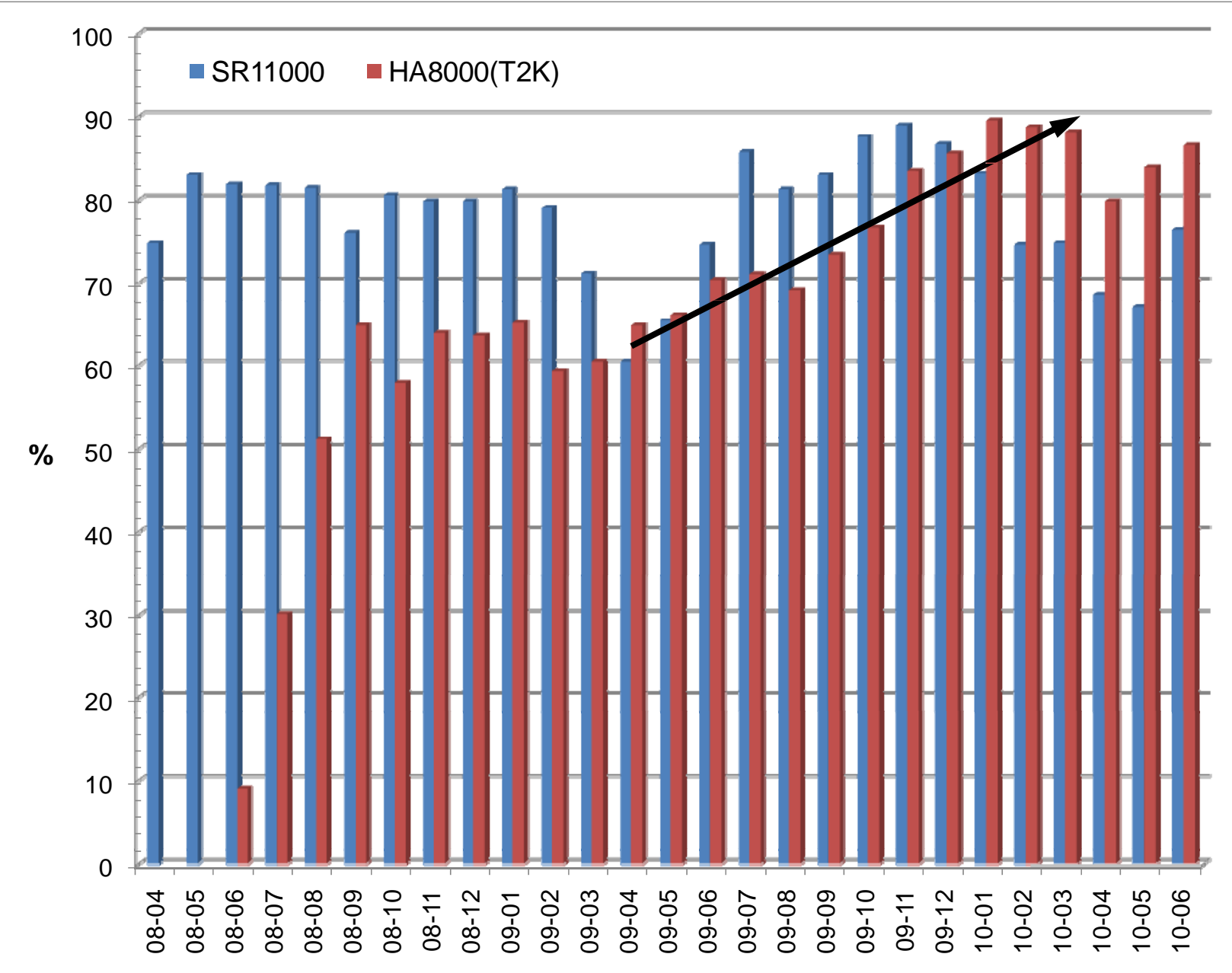
OpenMP/MPI Hybrid

FEM Applications
32 nodes, 512 cores

CASE-1: NUMA control
 CASE-2: + F.T.
 CASE-3: + Further Reordering



稼働率変遷(ノード固定除く)



T2K(東大)がもたらしたもの

- 利用者に対して
 - 「超並列」「マルチコア」時代の到来というメッセージ
 - 単体性能は低いが(低いため), 並列性能は高い
 - 「そろそろまじめに考えた方がいいですか?」というSR11000利用者
- 東大情報基盤センターに対して
 - 2006年から始まった「T2K」導入に向けての取り組みは, 「センター史上初」の様々なことを開始する契機となった
 - 他大学への営業活動
 - 公募型共同研究プロジェクト
 - お試しアカウント付き講習会
 - 利用者による発表成果登録, 公表
 - 可視化ソフト導入(AVS/Express PST) などなど
 - 「世界史」, 「日本史」と同じように「東京大学情報基盤センター史」というものがあるとすれば, 「癸丑甲寅以来」のごとく「T2K以前」, 「T2K以降」に明確に区分されることであろう

- T2K(東大)の概要
 - T2K(東大)がもたらしたもの
- T2K(東大)の現状
 - 稼働状況
 - 東大情報基盤センターの活動
- 今後の動向

情報基盤センター(スパコン部門)活動

<http://www.cc.u-tokyo.ac.jp/>

- サービス
 - 学術・民生利用
 - システム整備
- 普及・人材育成
- 共同研究
- 広報活動
 - スーパーコンピューティングニュース(年6回+特集号)

T2K(東大)の利用形態

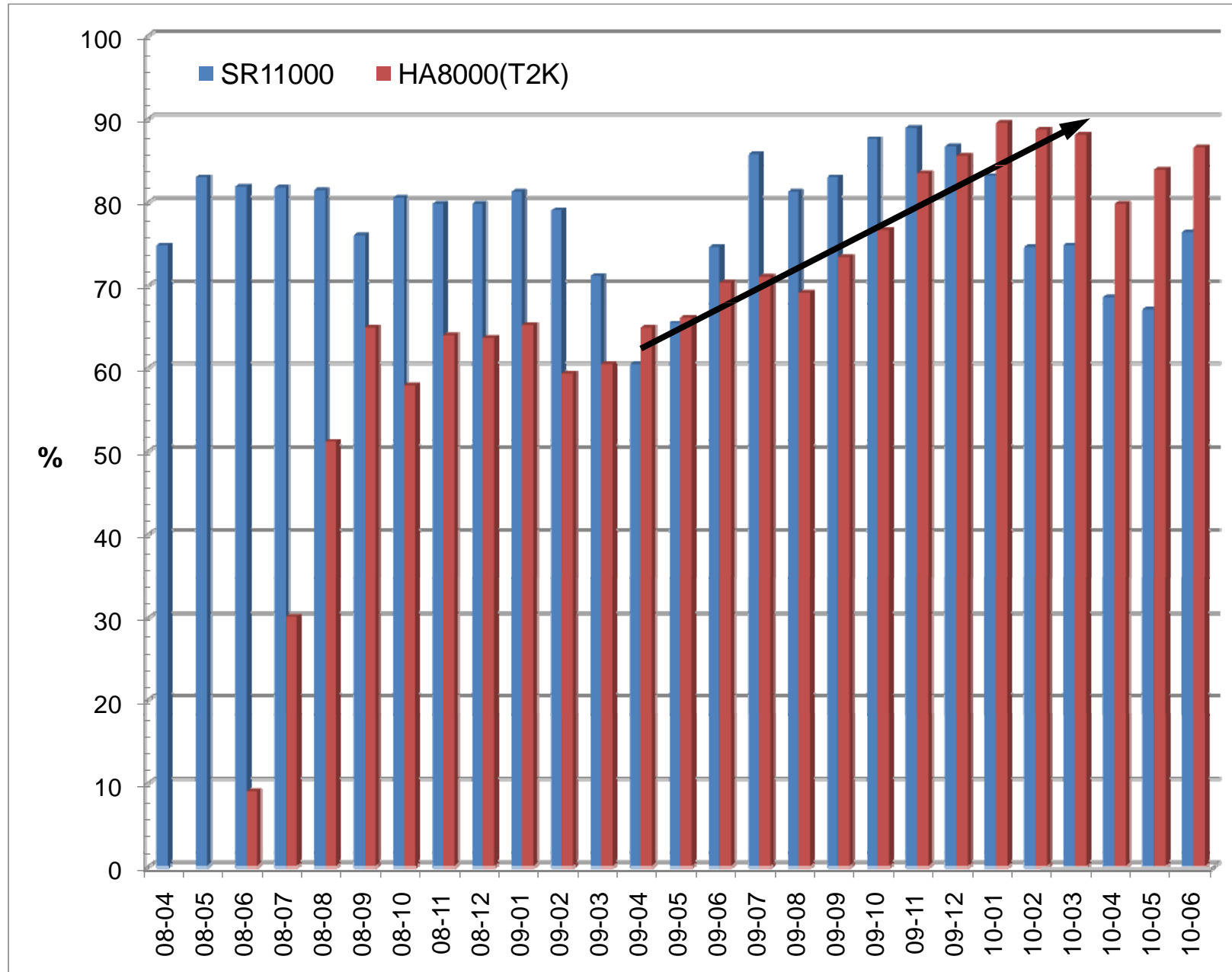
<http://www.cc.u-tokyo.ac.jp/service/ha8000/>

- パーソナルコース
 - 5コース, コース5:最大64ノード, 440千円/年(大学等)
 - 同じ料金でSR11000の4倍の計算資源を利用可能
- 専用キュー(グループ利用)
 - 1,000千円/8ノード/年(大学等), 最大48時間
- ノード固定
 - グループ利用, 商用コード利用・ノードカスタマイズ要
- 先端大規模計算利用サービス(共用イノベーション), 企業利用:別途審査あり
 - 全体の10%以下
- 共同研究プロジェクト, 若手利用
- 教育利用, 講習会等, その他(特別プロジェクト)

T2K(東大)利用者数等(2009年度末)

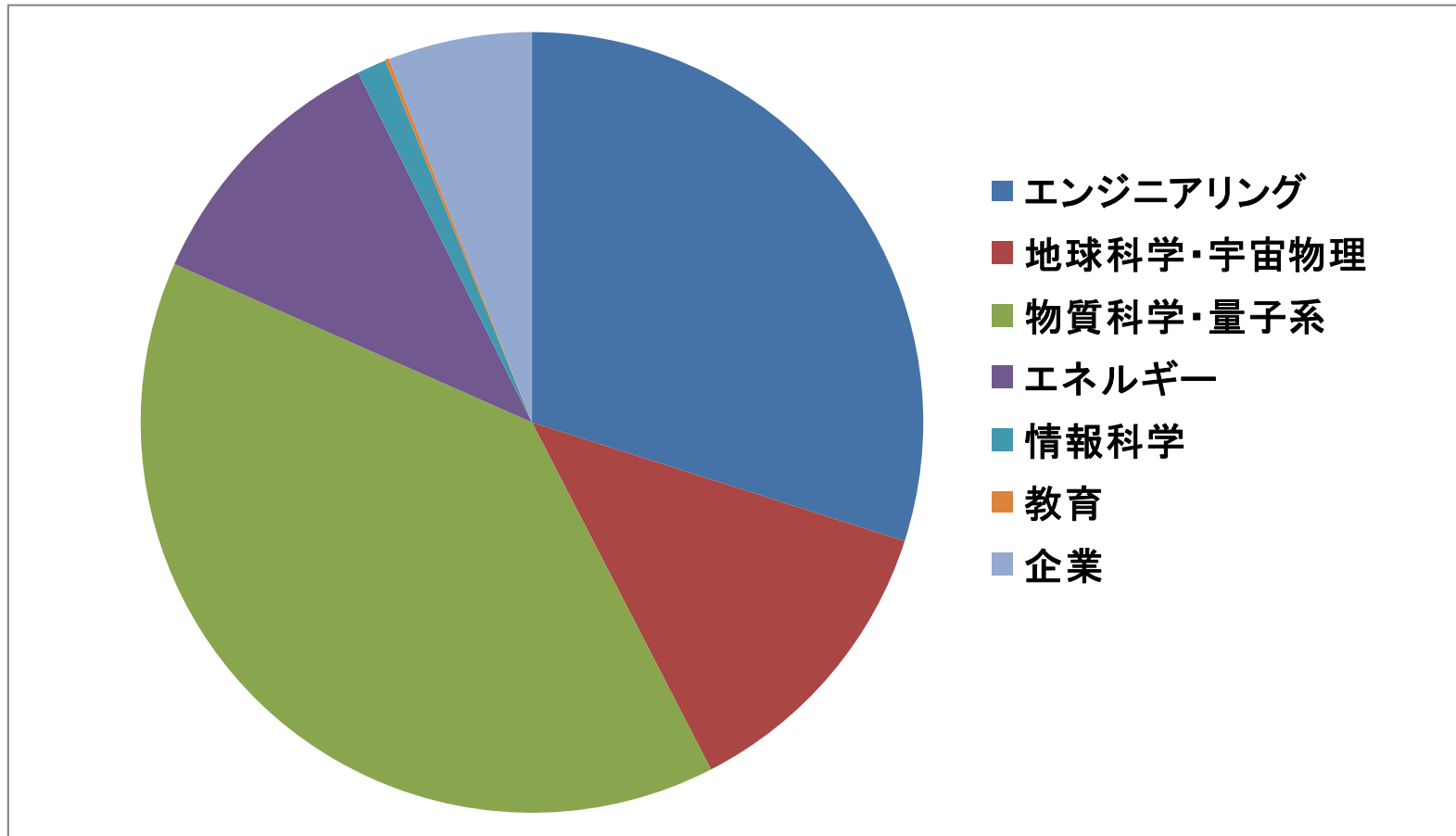
- 登録利用者数
 - 938名(うちパーソナルコース171名, 教育利用283名)
 - SR11000:488名
- 専用キュー
 - 20グループ, 612ノード
- ノード固定
 - 3グループ, 160ノード
 - e-Science, 大規模データ解析(ゲノム, 衛星観測データ)
- 共同研究プロジェクト
 - 10グループ, 64ノード(共用)
- 企業利用・イノベーション
 - 4グループ, 54ノード
 - ASP事業, エンジニアリングコンサルタント

稼働率変遷(ノード固定除く)



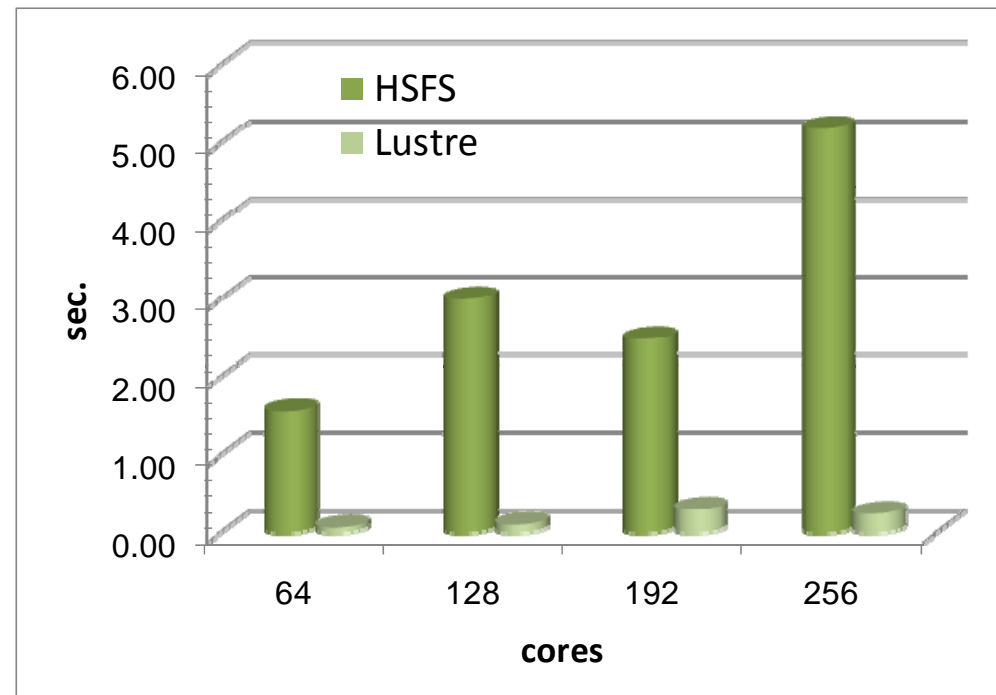
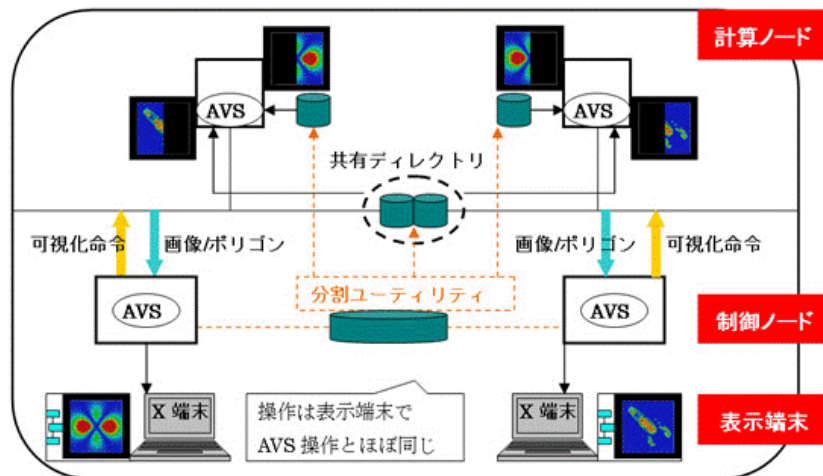
利用ノード時間積による利用分野

専用キュー＋教育＋企業・イノベーション(全体の約40%)



2009年度に開始したサービス

- ファイルシステム増強
 - 大容量NFS
 - 並列ファイルシステム(Lustre)
 - 例:FEM通信テーブル作成:最大数KB単位の入出力繰り返す
- 並列可視化システムAVS/Express PST導入
- 研究成果登録(2009年)
 - SR11000 44件
 - T2K 163件



情報基盤センター(スパコン部門)活動

<http://www.cc.u-tokyo.ac.jp/>

- サービス
- 普及・人材育成
 - 学際計算科学・工学 人材育成プログラム
 - 全学的なHPC教育
 - お試しアカウント付き講習会(T2K)
 - 1週間有効な「お試しアカウント」を利用した実習付き
 - MPI並列化(基礎, 応用), OpenMP・マルチコア, ライブラリ利用
 - 次回7月29日・30日(MPI基礎)
 - 2009年9月より企業からの受講, アカウント発行可能
 - 3分の2程度が企業からの受講者
- 共同研究
- 広報活動
 - スーパーコンピューティングニュース(年6回+特集号)

講習会(2009年度)

名称	開催日	利用計算機	申込者数	備考
第5回お試しアカウント付き並列プログラミング講習会(試行)	2009年7月1日～2日	HA8000クラスタシステム	26人	情報基盤センター主催
第6回お試しアカウント付き並列プログラミング講習会(試行)	2009年9月7日～8日		12人	
第7回お試しアカウント付き並列プログラミング講習会(試行):科学技術計算のためのマルチコアプログラミング入門:Hybrid並列プログラミングモデルへの道	2009年9月14日～15日		38人	情報基盤センター主催
第8回お試しアカウント付き並列プログラミング講習会(試行)	2009年3月17日		17人	情報基盤センター主催
CCS HPCサマーセミナー2009	2009年7月23日～24日	—	7人	筑波大学主催
京都大学 並列プログラミング入門講習会	2009年9月3日～4日	—	5人	京都大学主催

お試しアカウント付き講習会(2010年度)

http://www.cc.u-tokyo.ac.jp/publication/kosyu/schedule_kosyu.html

名称	期間	時期(予定)	内容
MPI基礎	1日半 ~2日	2010年7月29・30日 2010年9月2・3日 2011年3月17・18日	<ul style="list-style-type: none"> • MPIによる並列プログラミングの基礎に関する講習, 実習 並列化の基礎知識 • MPIのAPI説明 • 行列積の並列化実習 • makeを使った分割コンパイルと並列処理 • T2K(東大)による実習
MPI応用	1日半	2010年5月13・14日 2010年12月(未定)	<ul style="list-style-type: none"> • MPIを使用した並列アプリケーション開発手法に関する講習, 実習 有限体積法によるポアソン方程式ソルバーの概要 • 並列データ構造の考え方 • 領域分割手法 • 並列化手法 • T2K(東大)による実習
OpenMP (基礎+応用)	1日半 ~2日	2010年9月(未定)	<ul style="list-style-type: none"> • OpenMPによるマルチコアプロセッサ向け並列プログラミング, 最適化手法に関する, 実アプリケーションに基づく講習, 実習 有限体積法によるポアソン方程式ソルバー, ICCG法の概要 • OpenMPの基礎 • リオーダーリングによる並列化, 最適化 • T2K(東大)による実習
ライブラリ利用	2日	2010年9月27・28日 2011年2月2・3日	<ul style="list-style-type: none"> • 密行列ライブラリBLAS, LAPACK, ScaLAPACK、および、疎行列ライブラリPETsc, Lisの利用法に関する講習, 実習 数値解法の原理と特徴の説明 • 数理的モデリング, 離散化, データ格納 • ブロック化, データ分散の考え方 • T2K(東大)による実習

情報基盤センター(スパコン部門)活動

<http://www.cc.u-tokyo.ac.jp/>

- サービス
- 普及・人材育成
- 共同研究
 - 公募型プロジェクト
 - 企業との共同研究
- 広報活動
 - スーパーコンピューティングニュース(年6回＋特集号)

公募型プロジェクト

- T2K東大(HA8000クラスタシステム)のみ
 - 「超並列」指向
- T2Kオープンスパコン(東大)共同研究プロジェクト
- スーパーコンピューター若手利用者推薦(試行)
- HA8000 クラスタシステム 512 ノードサービス
 - 2010年度より開始, 既存ユーザーに限定しない
 - 512ノード, 8,192コアを24時間占有可能
 - 1月1回実施, 次回締切は9月10日(金)
 - http://www.cc.u-tokyo.ac.jp/use_info/512node/

T2K 共同研究プロジェクト

- 様々なシミュレーションのアルゴリズムの開発, プログラムの高速化に関する研究を本センターの教員と共同で実施
 - 成果は一般ユーザーにも還元
- 共同研究プロジェクト期間中は64ノード専用キュー(1,024コア)を無料で使用(他グループと共用)
- H20年度より開始(T2K), H21年度は8件中, 7件採択
- H20報告書は受付にあります

陰山 聡	神戸大学大学院 システム情報学研究科	地球ダイナモの新しいシミュレーションコード開発とその応用
坂本 雄三	東京大学大学院 工学系研究科	市街地における風・温熱・光・音環境総合数値予測データベースの開発
羽角 博康	東京大学大気海洋研究所	海洋大循環のマルチスケール連結階層モデリング
久田 俊明	東京大学大学院 新領域創成科学研究科	超並列計算によるマルチスケール・マルチフィジックス心臓シミュレーション
深沢圭一郎	九州大学宙空環境研究センター	電磁流体コードによる惑星磁気圏シミュレーション
古村 孝志	東京大学大学院情報学環/ 地震研究所	津波発生伝播の大規模 3次元シミュレーション
松元 亮治	千葉大学大学院 理学研究科	天体活動現象の輻射磁気流体シミュレーション

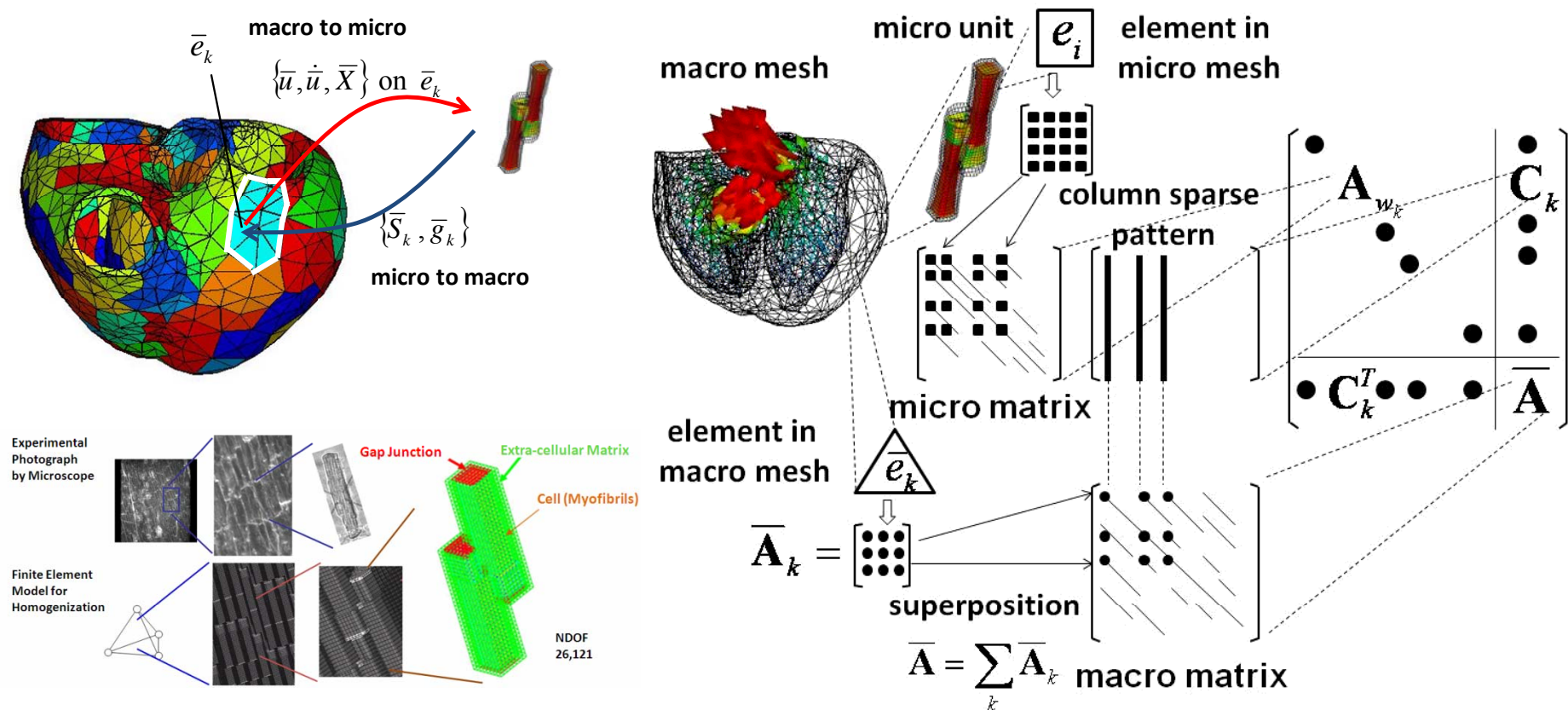
T2K 共同研究プロジェクト

- 様々なシミュレーションのアルゴリズムの開発, プログラムの高速化に関する研究を本センターの教員と共同で実施
 - 成果は一般ユーザーにも還元
- 共同研究プロジェクト期間中は64ノード専用キュー(1,024コア)を無料で使用(他グループと共用)
- H20年度より開始(T2K), H21年度は8件中, 7件採択
- H20報告書は受付にあります

陰山 聡	神戸大学大学院 システム情報学研究科	地球ダイナモの新しいシミュレーションコード開発とその応用
坂本 雄三	東京大学大学院 工学系研究科	市街地における風・温熱・光・音環境総合数値予測データベースの開発
羽角 博康	東京大学大気海洋研究所	海洋大循環のマルチスケール連結階層モデリング
久田 俊明	東京大学大学院 新領域創成科学研究科	超並列計算によるマルチスケール・マルチフィジクス心臓シミュレーション
深沢圭一郎	九州大学宙空環境研究センター	電磁流体コードによる惑星磁気圏シミュレーション
古村 孝志	東京大学大学院情報学環/ 地震研究所	津波発生伝播の大規模 3次元シミュレーション
松元 亮治	千葉大学大学院 理学研究科	天体活動現象の輻射磁気流体シミュレーション

A Multi-Scale Heart Simulation on Massively Parallel Computers

A.Hosoi, T.Washio, Y.Kadooka, K.Nakajima & T.Hisada
 accepted as a Technical Paper at SC10 (Nov.,2010)



スーパーコンピュータ 若手利用(試行)

- H19年度～
- (概ね)35歳未満を対象
- 利用負担金(パーソナルコース)の全額を負担(T2K:64ノード, 月末256ノード)
- 人材育成の一環
- 半年ごとに募集, 最大1年

Rothkopf, Alexander	東京大学大学院 理学系研究科	Study of QCD and effective models in the context of equilibrium and out of equilibrium quantum field theory with application to Heavy Ion Collisions, the Quark Gluon Plasma and Early Universe dynamics
坂下 達哉	電気通信大学大学院 情報システム学研究科	量子i.i.d.状態のシミュレーションとその理論的考察
田村 純一	埼玉大学大学院 理工学研究科	実対称固有値問題に対する多分割の分割統治法の分散メモリ型並列計算機への実装
高木 洋平	静岡大学工学部 物質工学科	直接メタノール形燃料電池内の流動現象シミュレーション

HA8000 クラスタシステム 512 ノードサービス

- 2010年度より開始
- 既存ユーザーに限定しない
- 512ノード, 8,192コアを24時間占有可能
 - 国内のオープンな計算資源としては恐らく最大
 - 大規模計算で見えてくるもの:物理だけではない
 - 試行期間限定HPC特別プロジェクト(2008年)
- 外部委員も含めた審査
- 1月1回実施
- 次回締切は9月10日(金)
- http://www.cc.u-tokyo.ac.jp/use_info/512node/

HA8000 クラスタシステム

512 ノードサービス

2010年度第一期採択課題

課題名	代表者名	代表者所属
流体解析ソフトウェアFrontFlow/blue ver.6.0の大規模ベンチマークテスト	加藤 千幸	東京大学 生産技術研究所
電磁流体コードによる宇宙天気シミュレーション性能測定	深沢 圭一郎	九州大学 宙空環境研究センター
密行列固有値解析の大規模並列化のための通信性能評価	大島 聡史	東京大学 情報基盤センター
超並列計算によるマルチスケール・マルチフィジックス心臓シミュレーション	久田 俊明	東京大学 新領域創成科学研究科
酸化チタンと金属材料の界面構造の電子状態	大野 隆央	東京大学 生産技術研究所

試行期間限定 HPC特別プロジェクト

2008年7月～9月，受付に報告書

- 512ノード(8,192コア)利用を目的とした特別プロジェクト(無料)
(特集号)
 - 応募:29件, 採択:10件(一次:4件, 二次:6件)
 - 64ノード専用キュー(10グループで共有)
 - 月末:256ノード(4,096コア)利用可能
 - 2008年7月～9月の週末:各グループ1日(24時間程度)512ノード占有

①=R:次世代ターゲットアプリ
①=G:Gordon Bell Finalists
②:Flat MPI
③:Hybrid並列

	課題名・代表者(所属)	①	②	③
A	高性能直接法N体計算ベンチマーク, 似鳥啓吾(東大・天文学専攻)	R	○	
B	海洋循環形成プロセスの高解像度シミュレーション, 羽角博康(東大・気候センター)	R	○	○
C	3次元不均質場での地震波伝播の大規模シミュレーション, 古村孝志(東大・情報学環)	R	○	
D	密度行列繰り込み群法と行列対角化による強相関量子系のシミュレーション, 町田昌彦(原子力機構)	G	○	○
E	プロセッサアフィニティ制御を組み込んだフレームワークによる実用大規模並列シミュレータの性能評価, 小野謙二(理化学研究所)		○	○
F	T2Kオープンスパコンへのインヤン地球ダイナモコードの移植, 陰山聡(海洋研究開発機構)	G	○	
G	革新的シミュレーションソフトウェア, 加藤千幸(東大・生産研)	R	○	○
H	GXPシステムとそれを用いた大規模テキスト処理の実行, 黒橋 禎夫(京大・情報学研究科)		-	-
I	全球雲解像正20面体格子非静力学大気モデル(NICAM)の開発, 佐藤正樹(東京大・気候センター)	R	○	
J	超並列計算によるマルチスケール・マルチフィジックス心臓シミュレーション, 久田俊明(東大・新領域)		○	

学際大規模情報基盤 共同利用・共同研究拠点 JHPCN

<http://jhpcn-kyoten.itc.u-tokyo.ac.jp/>

- 北大, 東北大, 東大, 東工大, 名大, 京大, 阪大, 九大の8大学情報基盤センター(中核機関:東大)
- 平成22年度より正式に活動開始(6年ごとに見直し)
 - 2010年6月より正式共同研究開始
- 分野
 - 超大規模数値計算系応用分野, 超大規模データ処理系応用分野, 超大容量ネットワーク技術分野
 - 融合分野(大規模情報システム関連研究分野)
- グリッド的な運用も視野
- スパコン: 資源の10%程度を供出(東大はT2K 72ノード)

公募型プロジェクトの今後

- T2Kオープンスパコン(東大)共同研究プロジェクト
 - 2010年3月を以て終了
 - 学際大規模情報基盤共同利用・共同研究拠点公募型研究へと発展的解消
- スーパーコンピューター若手利用者推薦(試行)
 - 継続, 本年度前期採択3課題
 - 教育利用と関連して制度化を目指す
 - 若手研究者のencouragingの場
 - 継続的なサポート, 共同研究へのパス:今はAward
- HA8000 クラスタシステム 512 ノードサービス
 - とりあえず究極の超並列

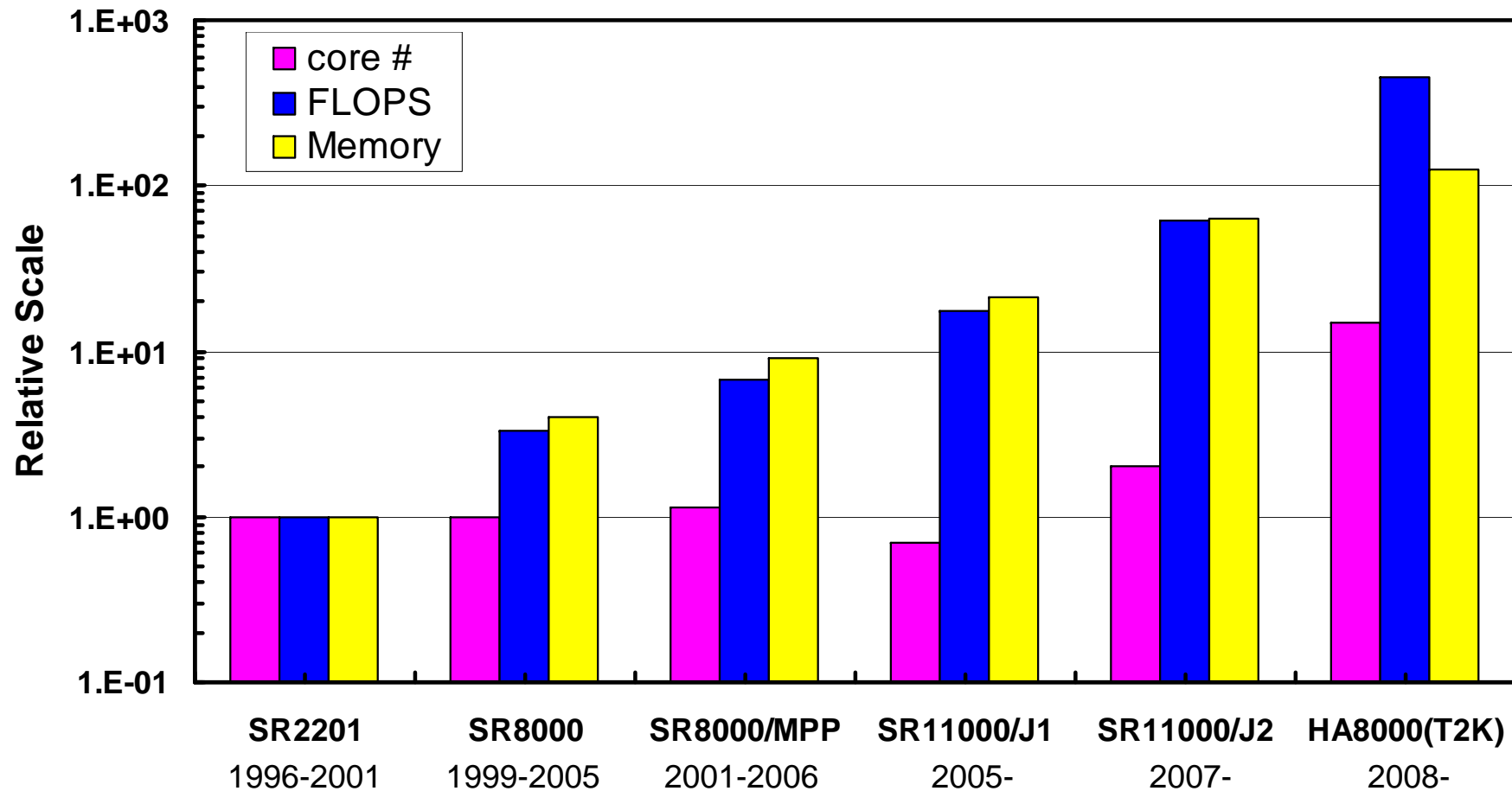
何を目指すのか?: 基本は「超並列」

- T2K利用環境の改善
- 「次世代スーパーコンピュータ」への準備
 - 様々な技術的な問題の洗い出し, 対処
 - アプリケーション
 - ライブラリ, ミドルウェア
 - システム
- 新しいユーザーの開拓
 - 次世代を担う人材の育成: 科研費等の獲得
 - 新規分野
 - 「科学技術シミュレーション」の枠にとらわれる必要は無い
 - **オープンソースアプリケーション: RSS21/RISS, OpenFOAM**
- 次期システムへの準備
 - 様々な要望

- T2K(東大)の概要
 - T2K(東大)がもたらしたものの
- T2K(東大)の現状
 - 稼働状況
 - 東大情報基盤センターの活動
- 今後の動向

東大センタースパコンの歴史

SR2201 (ピーク性能300GFLOPS)を1



東大情報基盤センター スパコンの今後

- SR11000後継機種
 - 2011年秋, 柏に設置予定
 - SMP:ピーク性能50TFLOPS(予定), こういうのはこれが最後
 - MPP:ピーク性能1PFLOPS(予定), アクセラレータ無し
- ポストT2K
 - 2014年4月以降
 - 詳しくは「パネルディスカッション」で
- 留意点
 - 大容量・高速で安定なファイルシステム
 - オープンソースアプリケーション, ライブラリの充実
 - 可視化