

次世代スーパーコンピュータプロジェクトの概要と 進捗状況

平成20年3月
理化学研究所
次世代スーパーコンピュータ開発実施本部
プロジェクトリーダー 渡 辺 貞

プロジェクトの概要

「次世代スーパーコンピュータ」プロジェクト

平成20年度予算案	:	145	億円
平成19年度補正予算	:	42	億円
平成19年度予算額	:	77	億円
平成18～24年度（総事業費）	:	1,154	億円

1. 目的 世界最先端・最高性能の次世代スーパーコンピュータの開発・整備及び利用技術の開発・普及

2. 概要

理論、実験と並び、現代の科学技術の方法として確固たる地位を築きつつある計算科学技術をさらに発展させるため、長期的な国家戦略を持って取り組むべき重要技術（国家基幹技術）である「次世代スーパーコンピュータ」を平成22年度の稼働（平成24年の完成）を目指して開発する。

今後とも我が国が科学技術・学術研究、産業、医・薬など広汎な分野で世界をリードし続けるべく、

（1）世界最先端・最高性能の「次世代スーパーコンピュータ（注）」の開発・整備

（注）10ペタFLOPS級

（1ペタFLOPS：1秒間に1千兆回の計算）

（2）次世代スーパーコンピュータを最大限利活用するためのソフトウェアの開発・普及

（3）上記（1）を中核とする世界最高水準のスーパーコンピューティング研究教育拠点（COE）の形成を文部科学省のイニシアティブにより、開発主体（理化学研究所）を中心に産学官の密接な連携の下、一体的に推進する。

3. 体制

（1）開発主体である独立行政法人理化学研究所を中心とした産学連携体制を構築。

（2）特定先端大型研究施設の共用の促進に関する法律を整備し（平成18年7月施行）、産学官の研究者等に幅広く開かれた共用施設として位置付け。

全体スケジュール

		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)
						稼動▲	完成▲	
システム	演算部 (スカラー部、ベクトル部)	概念設計 詳細設計 試作・評価 製造・据付調整						
	制御フロントエンド (トータルシステム ソフトウェア)		基本設計	詳細設計	製作・評価 性能チューニング・高度化			
	共有ファイル		基本設計	詳細設計	製造・据付調整			
ソフトウェア シミュレーション	次世代ナノ統合 シミュレーション	開発・製作・評価						実証
	次世代生命体統合 シミュレーション	開発・製作・評価						実証
施設	計算機棟		設計	建設				
	研究棟		設計	設計微修正等	建設			

立地地点の選定

理化学研究所において、次世代スーパーコンピュータ施設の立地地点を客観的・科学的な観点から検討するため、外部有識者からなる立地検討部会(部会長：黒川 清 内閣特別顧問)を設置し、平成18年7月より、15の候補地について評価を実施。

理化学研究所は立地検討部会の評価報告書に基づき、有力とされた神戸及び仙台について総合的に評価・検討を行い、平成19年3月に、神戸(ポートアイランド第2期内)を立地地点とすることを決定。

所在地	兵庫県神戸市中央区港島7丁目(ポートアイランド第2期内) ・ポートアイランド南駅より徒歩約1分(JR新神戸駅から25分)
-----	---



土地所有者：神戸市



4

次世代スーパーコンピュータ施設の整備

施設の建設

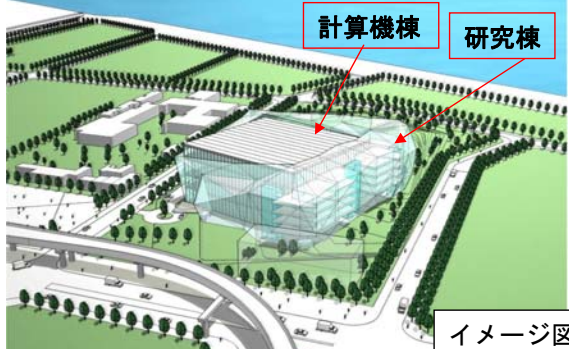
- ・現在、理化学研究所において建屋(計算機棟、研究棟等)の設計を実施中。
- ・平成19年度内に着工、平成22年度に完工予定。

整備の基本方針

- (1) 次世代スーパーコンピュータの性能を最大限引き出す設備・能力の確保
- (2) 世界最高水準のスーパーコンピュータティング研究教育拠点(COE)として相応しい研究・教育環境の整備
- (3) ランニングコストと環境負荷の低減化

施設の特徴

- (1) 計算機の性能を常時保証できる床耐荷重及び免震構造とするとともに、必要な電源設備及び冷却設備を整備
- (2) 共用施設としての運用上の利便性を高めるとともに、研究交流や多様な知識の融合を促進するため、計算機棟と研究棟を整備
- (3) 廃熱利用の推進や排水処理への配慮などによりランニングコストと環境負荷の低減を実現



イメージ図

【計算機棟】

- 延床面積 約17,500㎡
- 建築面積 約6,300㎡
- 構造 鉄骨造り地上3階地下1階

【研究棟】

- 延床面積 約8,500㎡
- 建築面積 約1,900㎡
- 構造 地下1階、地上6階

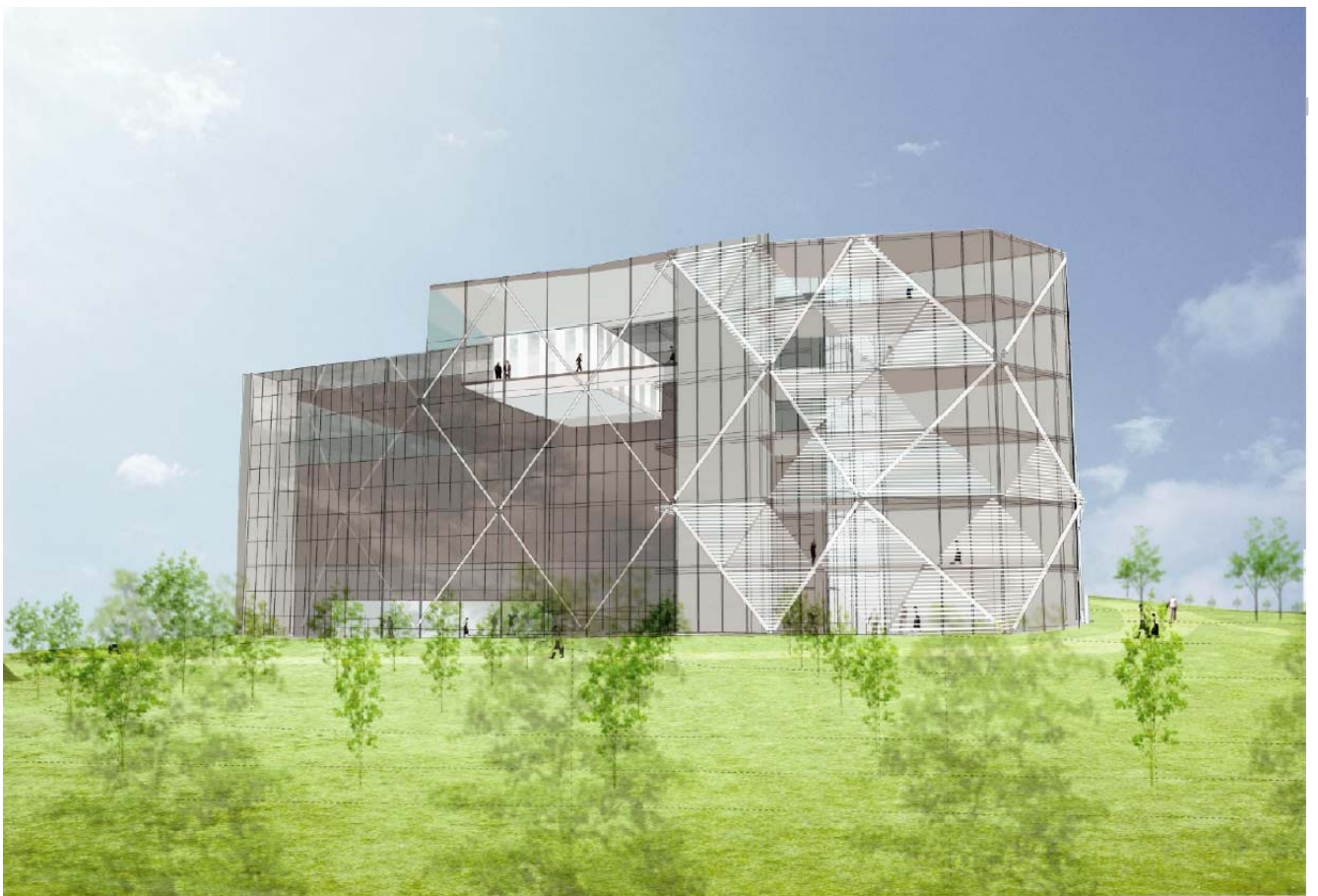
その他、電源を供給する特高受変電設備、計算機棟の空調機を冷却する冷却設備、及び環境負荷低減のためのC.G.S(自家発電)設備等を設置

5



北西から研究室側を見る

6



北からエントランスロビー側を見る

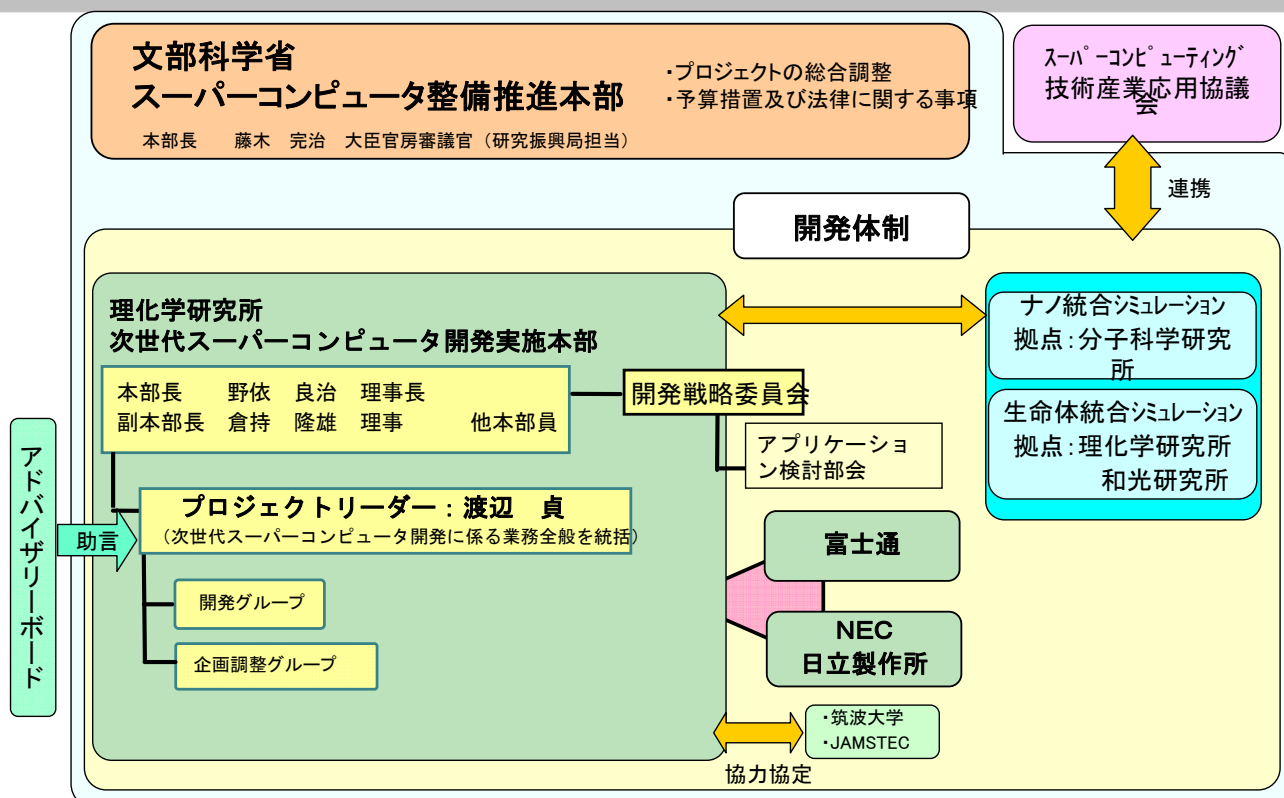
7

次世代スーパーコンピュータを最大限活用するためのソフトウェアの研究開発
(グランドチャレンジアプリケーション)

	ナノテクノロジー分野 〔次世代ナノ統合シミュレーション ソフトウェアの研究開発〕	ライフサイエンス分野 〔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〕
概要	ナノ電子デバイスの設計やバイオ燃料生成用の酵素設計等に役立つシミュレーションソフトウェアを研究開発	タンパク質分子の反応や、細胞・臓器の働きの詳細な解析により、製薬・医療に役立つシミュレーションソフトウェアを研究開発
応用例	現状 10万原子 2千原子程度(デバイスの一部)の計算が可能 ↓ 次世代スパコン 10万原子(デバイス全体)の計算が可能 ↓ アウトカム 高速、低消費電力のナノ電子デバイスの実現を加速	現状 様々な薬剤候補物質 150年の計算時間 ↓ 新薬設計 タンパク質の活性部位 ↓ 次世代スパコン 6ヶ月(更なる短縮を検討) ↓ アウトカム 新薬候補物質の最適化を効率化し、新薬開発の期間短縮及びコスト削減を実現
体制	分子科学研究所を中核に、東京大学物性研究所、東北大学金属材料研究所、産業技術総合研究所等と連携した研究開発体制を構築	理化学研究所を中核に、東京大学医科学研究所、慶應義塾大学等、13機関と連携した研究開発体制を構築

8

プロジェクト推進体制



9

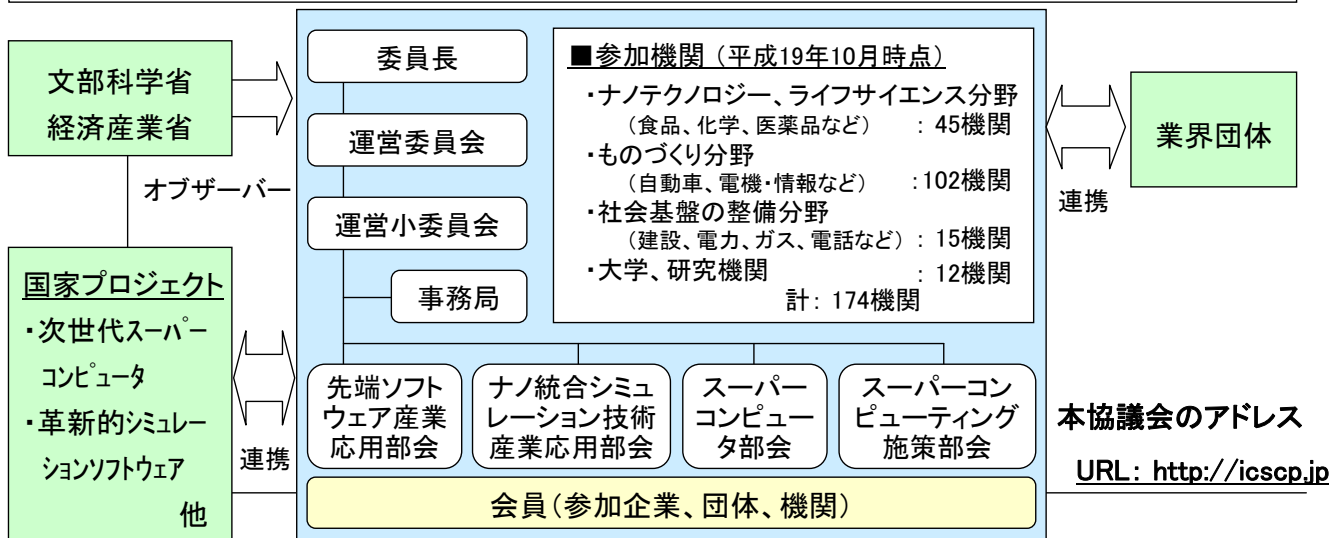
スーパーコンピューティング技術産業応用協議会

[設立趣旨]

産業界におけるスーパーコンピューティングの産業活用を推進するため、スーパーコンピュータ用の先進的応用ソフトウェアの評価、利用技術の開発、普及や超高速コンピュータ網を活用したスーパーコンピュータ利用技術の開発、普及などを推進する目的で平成17年12月に設立

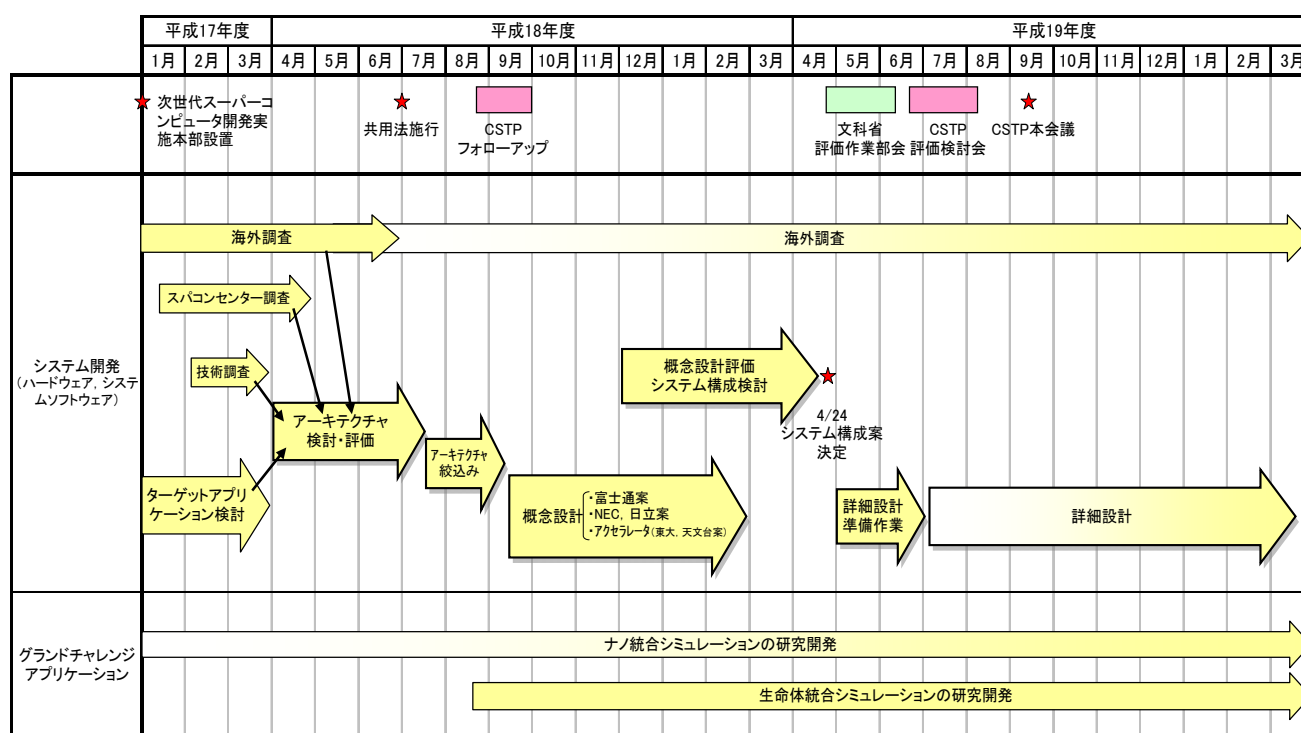
[活動内容]

4つの部会を設置し、スーパーコンピュータを活用したシミュレーションソフトウェアの産業界での利用技術の普及、次世代スーパーコンピュータ開発に関する産業界からの要望事項の検討・取りまとめなどを実施



システム開発

システム開発の経緯と進捗状況



12

システム構成について(概念設計の実施)

【世界最速のシステム】⇒ 1秒間に1京(ケイ=10の16乗)回の計算性能

(現最速計算機の約36倍)

【汎用システム】⇒ 科学技術・産業で用いられる多様なアプリケーションやこれまで不可能だった複雑かつ大規模なシミュレーションが実行可能

【革新的なシステム】⇒ 先端技術の積極的導入により、画期的な省電力、省スペースを実現

理化学研究所とメーカー3社(NEC、日立、富士通)による共同開発・共同負担により、日本の技術力の総力を結集して開発

[検討経緯および今後の予定]

平成18年4月
平成19年4月

理化学研究所において、システム構成案の検討を開始
NEC・日立及び富士通の提案を基礎に、理化学研究所においてシステム構成案を作成。

平成19年3～6月

文部科学省において、概念設計評価作業部会を設置し、理化学研究所のシステム構成案について評価を実施

平成19年6～9月
平成19年9月

総合科学技術会議において評価を実施

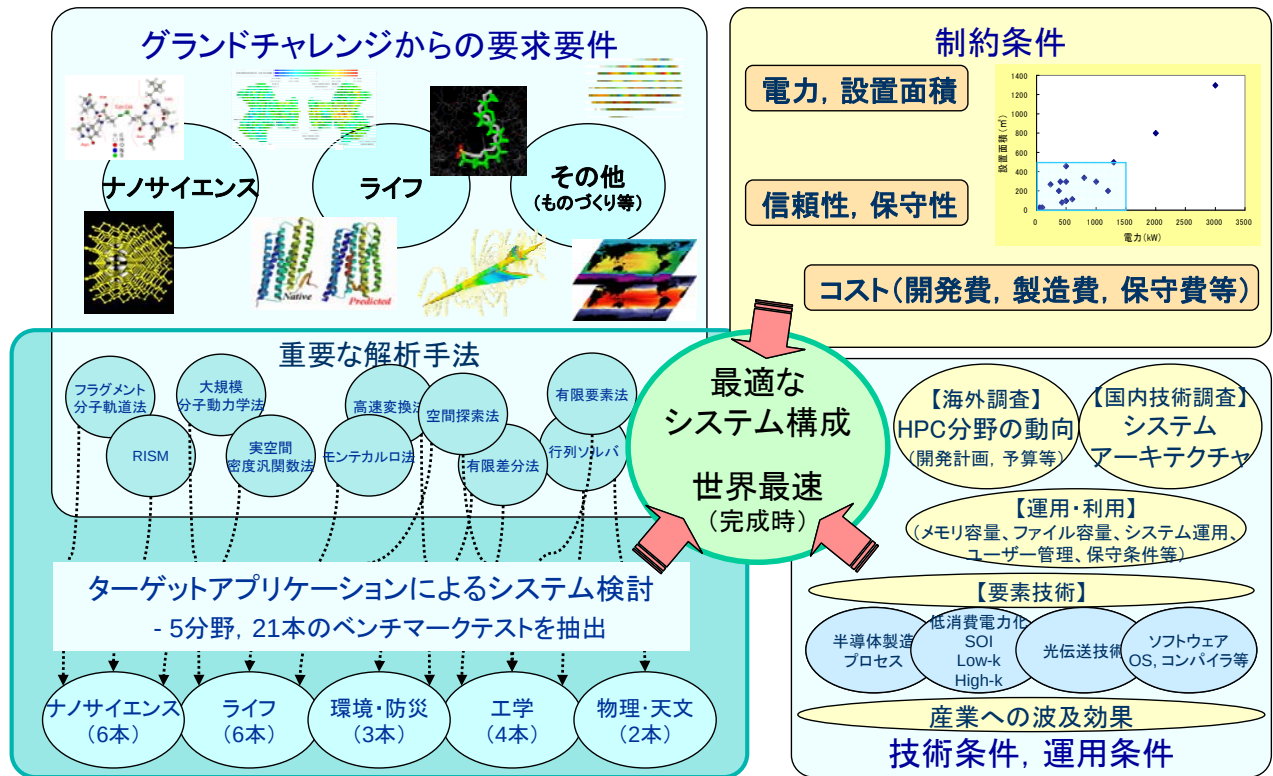
文部科学省及び総合科学技術会議の評価を踏まえ、理化学研究所においてシステム構成を正式決定

平成19～20年度

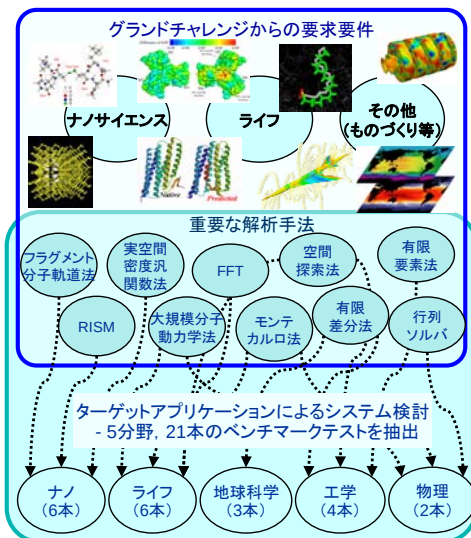
詳細設計

13

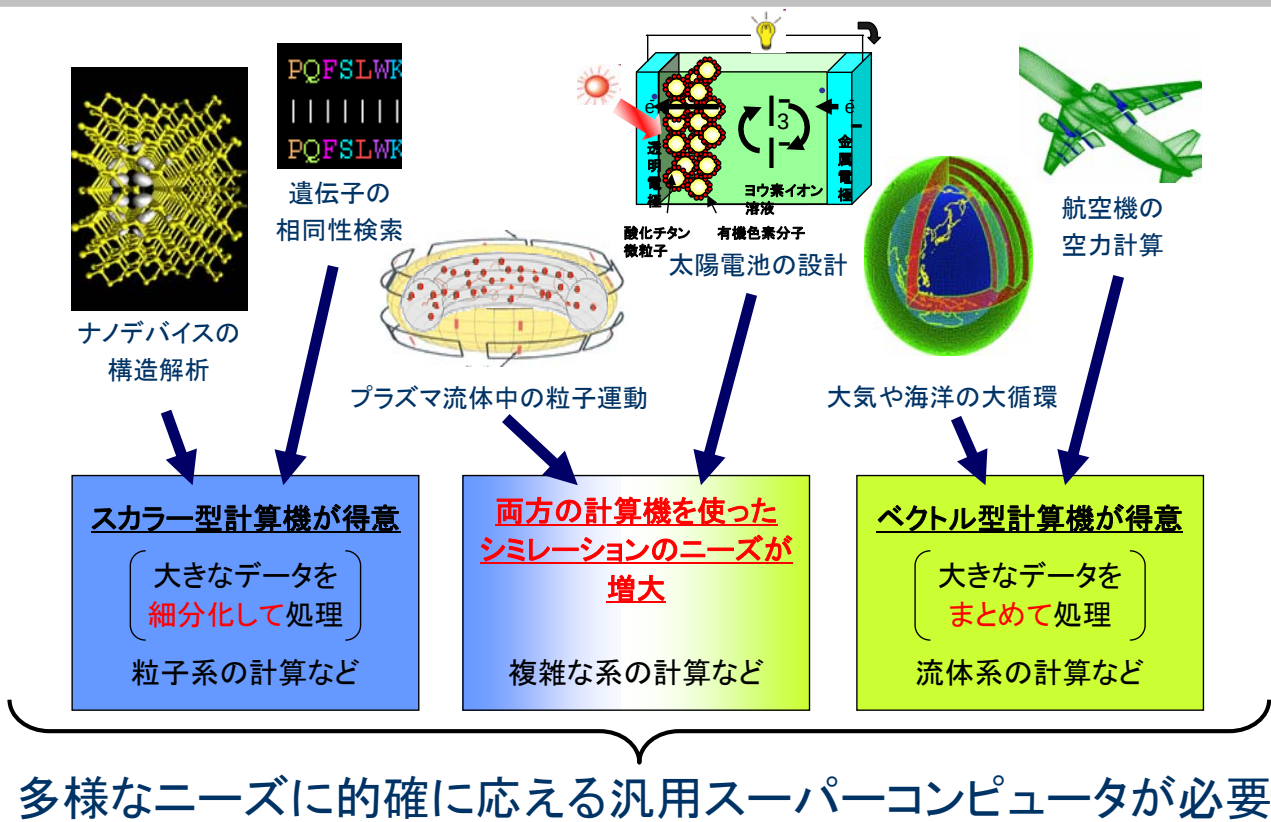
システム最適化の考え方



評価したターゲット・アプリケーション



ベンチマーク	分野	アプリケーション
SimFold	ライフ/ナノ	ライフタンパク質立体構造の予測
GAMESS/FMO		分子軌道法計算
Modyas		高並列汎用分子動力学計算
RSDFIT		実空間第一原理分子動力学計算
NICAM	地球科学	全球雲解像度大気大循環モデル
LatticeQCD	物理	格子QCDシミュレーション
LANS	工学	圧縮性流体計算



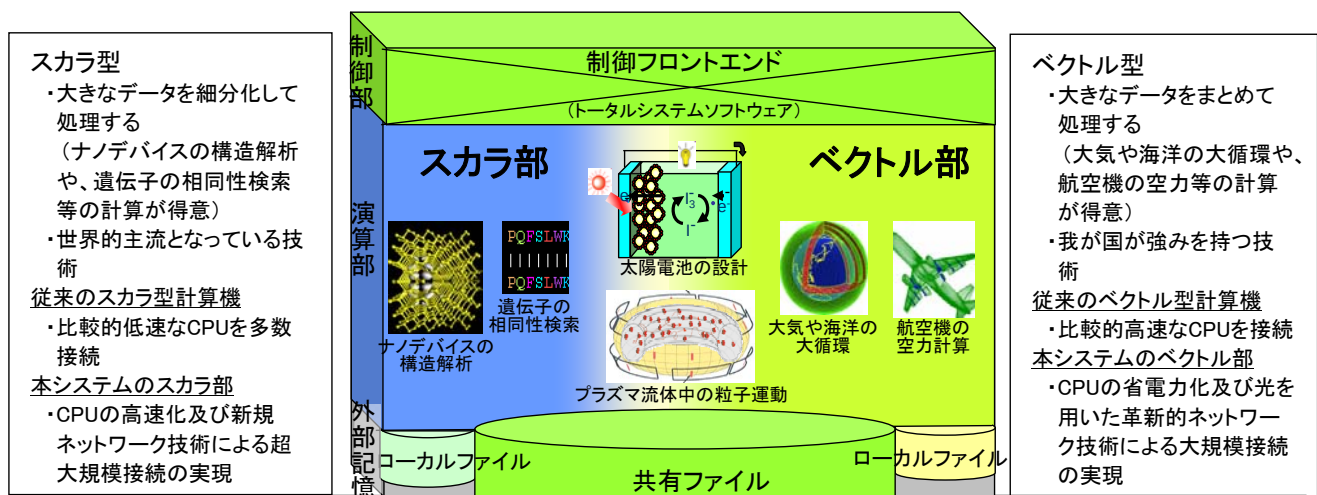
16

10

システムの特徴

スカラー部とベクトル部から構成される複合汎用システム

- 異なる2つの演算部の特性を活かし、あらゆるシミュレーションに対応可能な高い汎用性
- 計算能力に関する高い拡張性
- 大学や研究機関向けの高性能な計算機への展開性
- スカラー及びベクトルの両技術の確保による、我が国の技術力の強化と、国際競争力の向上



次世代スーパーコンピュータシステムの概要

【世界最速を達成する最先端システム】

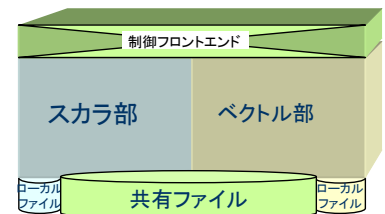
- Linpack性能10ペタFLOPSの達成のみならず、アプリケーションの実行においても世界最高性能
- 先端技術(45nm半導体プロセス,光インターコネクトなど)により画期的な省電力、省スペースを実現
- 理研とメーカー3社が共同で開発

【科学技術・産業の競争力を発展させる将来型システム】

- 高機能スカラ部と革新的ベクトル部から構成される複合汎用システム
- 複雑系問題、多階層問題などシミュレーションの革新を先導する計算環境を提供
- 次々世代以降の開発と利用を見据え、我が国の国際競争力を牽引

【我が国の科学技術基盤となる複合汎用システム】

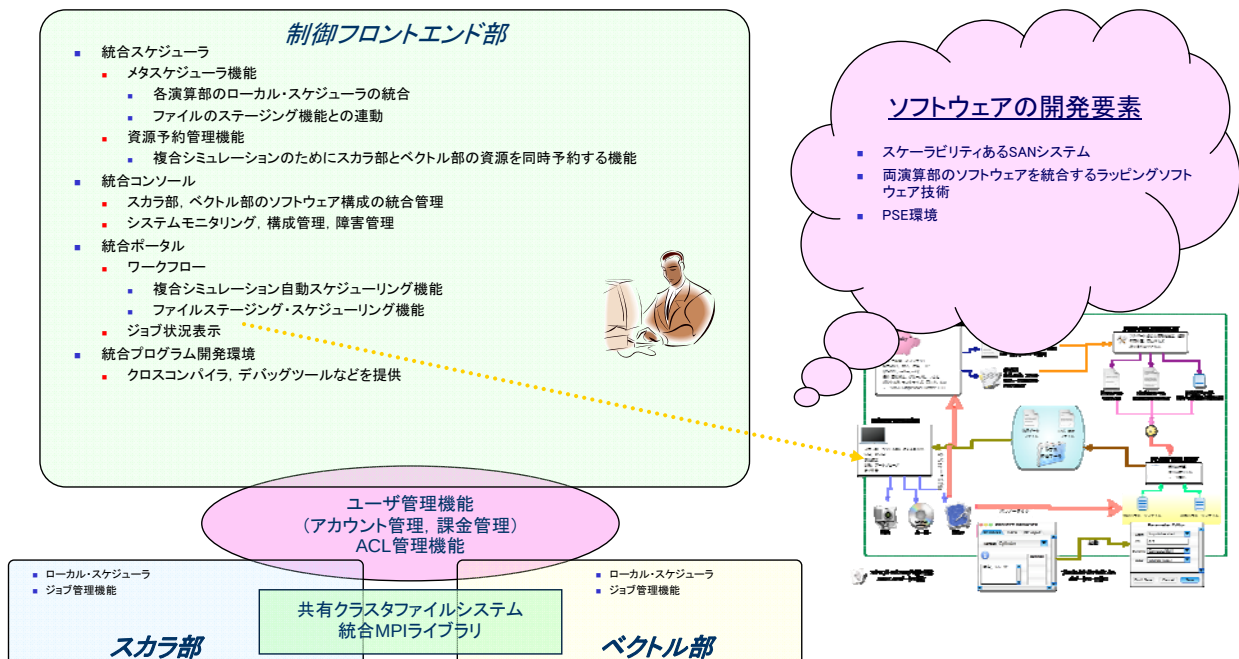
- 様々なアプリケーションを効率よく実行する複合汎用システム
(⇒より多くのユーザーが利用可能)
- 全国の産学の研究者等の共用施設
- アプリケーション資産を最大限活用
- 大学等のスパコンセンターへの展開性大
(⇒我が国の計算機環境を質的にも量的にも革新)



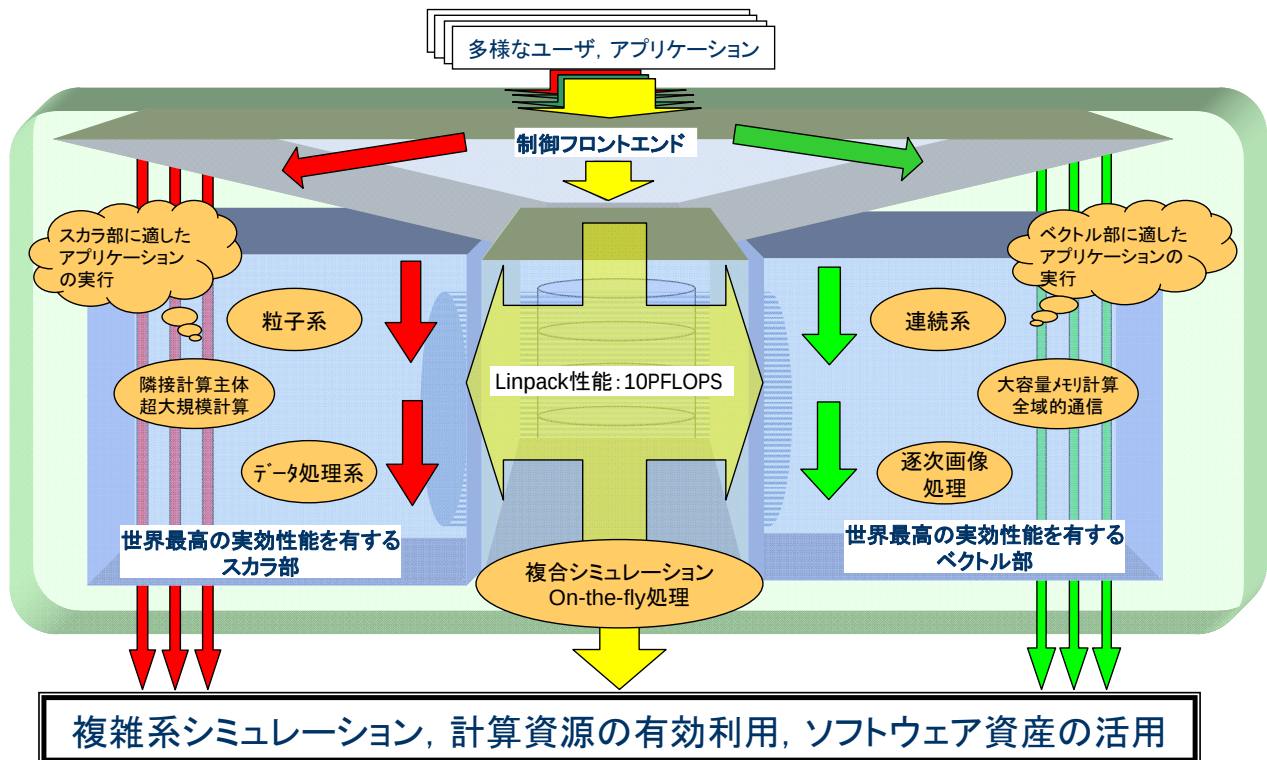
18

複合システムとしての機能

- 複合汎用スーパーコンピュータシステムの効率良い運用のために、以下のトータルシステムソフトウェア機能の開発を予定.



＜次世代汎用スーパーコンピュータシステム概念図＞
 =多様な計算ニーズに対して最適な計算環境を提供する複合型システム

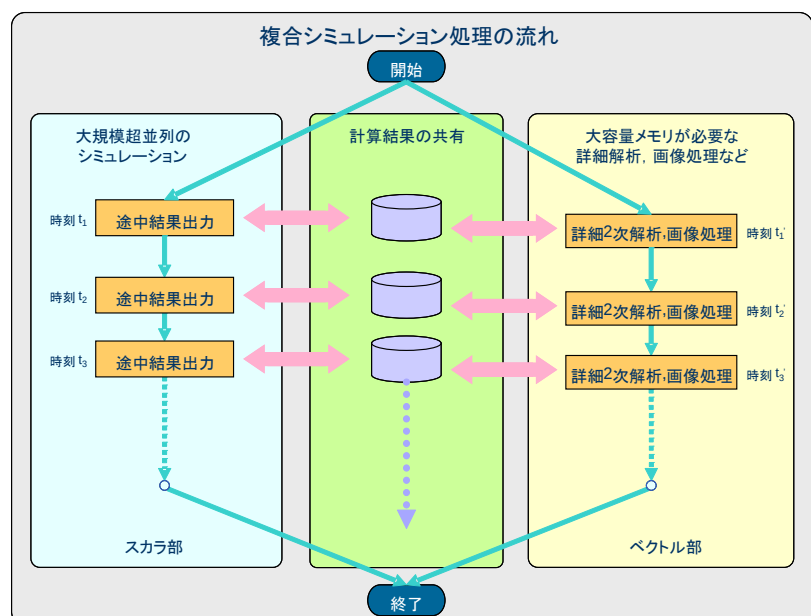


20

On-the-fly複合シミュレーション

- 複合シミュレーションにおいて、ある時刻ごとに出力される途中結果の流れ作業的なデータ解析、及び画像処理に最適なシステム。

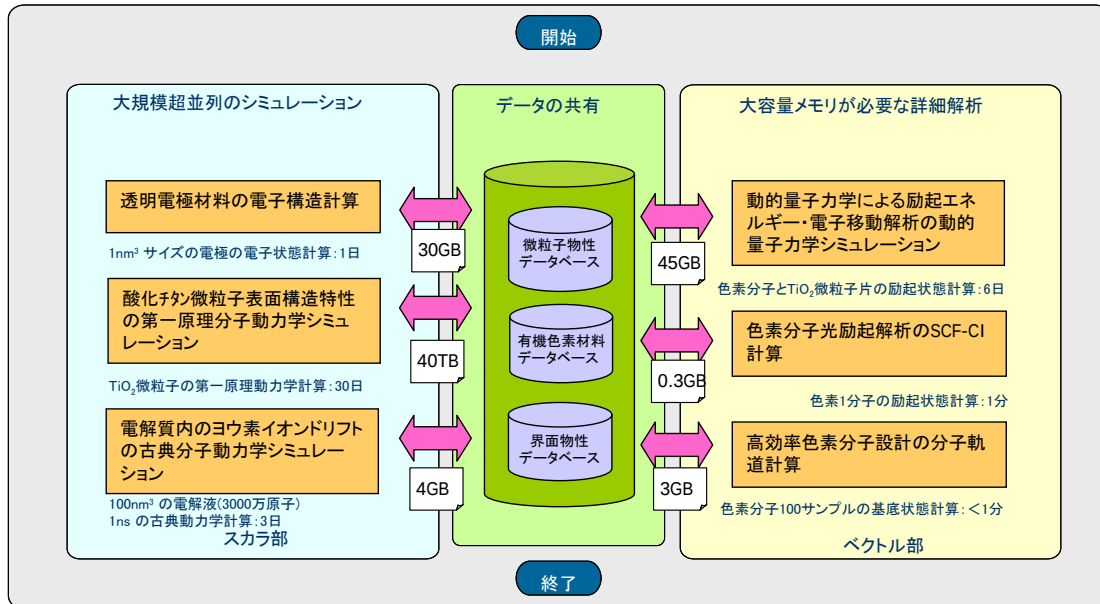
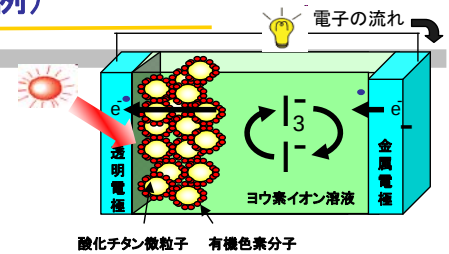
- 各演算部の連携により、一連のデータ処理の短縮化が可能。
- 大規模かつ長時間シミュレーションの途中結果のモニタリングによる計算資源の有効活用が可能。



21

複合シミュレーション(太陽電池設計の例)

- 色素増感型太陽電池の構成要素のシミュレーション
- 構成要素毎に異なるシミュレーション技術が必要
- スカラ部とベクトル部の複合計算でデバイス設計が可能

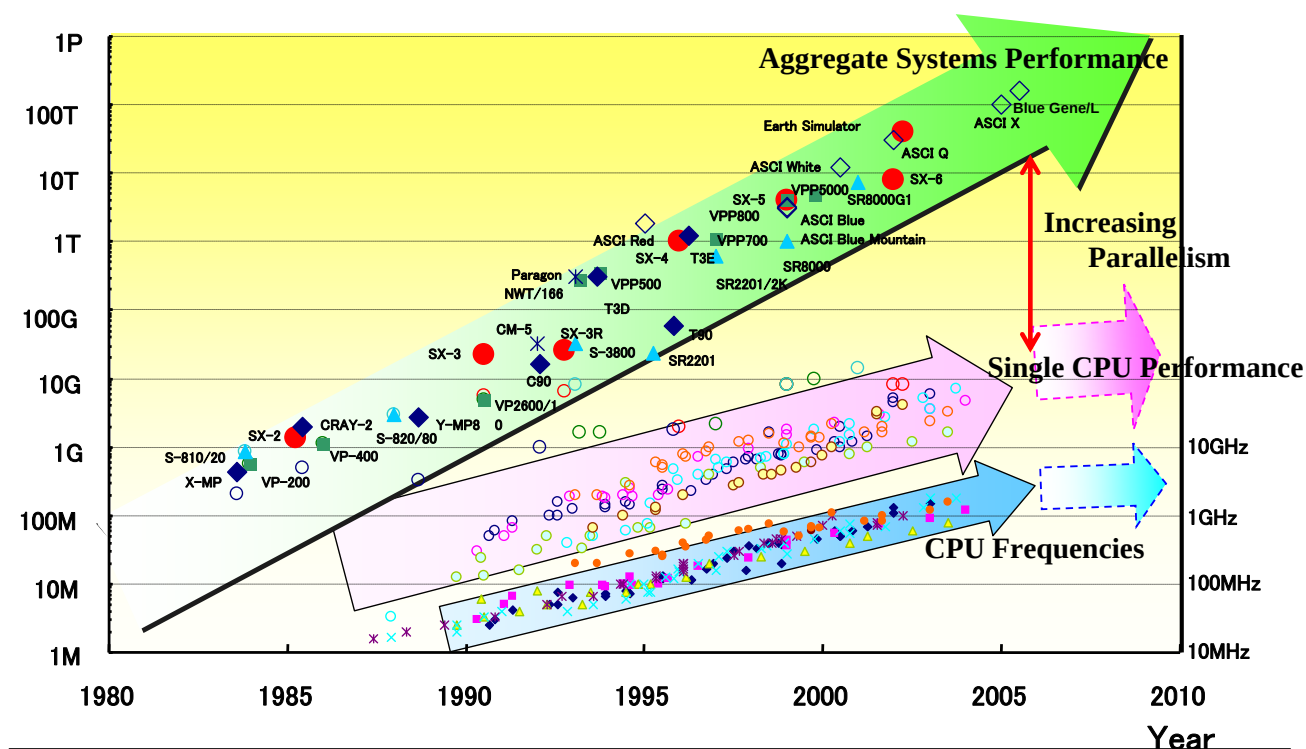


22

次世代スーパーコンピュータの 技術課題

23

スーパーコンピュータの性能推移



24

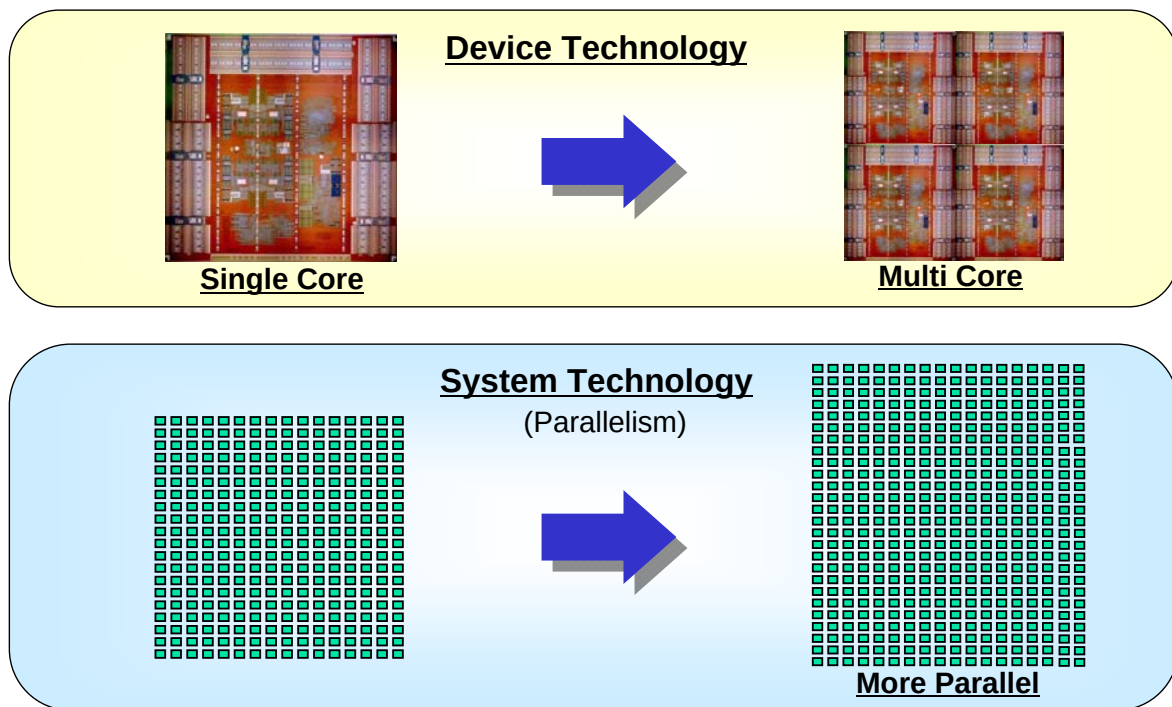
HPCアーキテクチャの変遷

	単一CPU型		共有メモリ型	(共有)分散メモリ型		
システム構成イメージ						
CPU数	1CPU	1CPU	～数十CPU	～数百CPU	～数千CPU	>数万CPU
CPU種類	スカラー	ベクトル	スカラー/ ベクトル	スカラー/ ベクトル	←	←
言語	Fortran	ベクトル化Fortran	並列(ベクトル) Fortran	並列(ベクトル) Fortran/C, MPI	←	←
チューニング	コンパイラ レベル	ベクトル チューニング	並列/ベクトル チューニング	並列 チューニング	←	←

P: Processor、P (S):スカラー、P (V):ベクトル
M: Memory

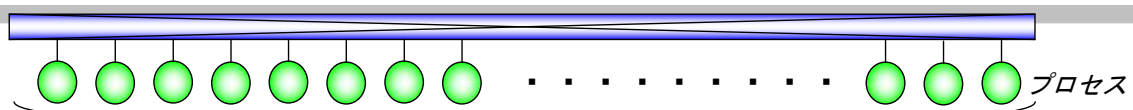
25

Increasing Parallelism



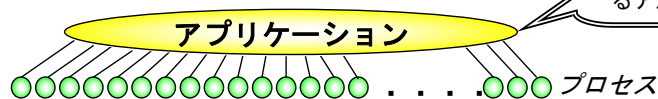
26

並列化チューニング

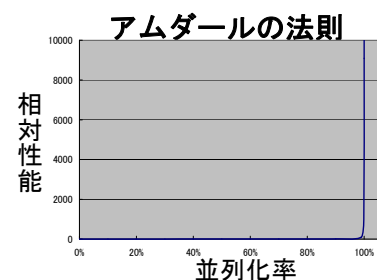


＞ 数万プロセス

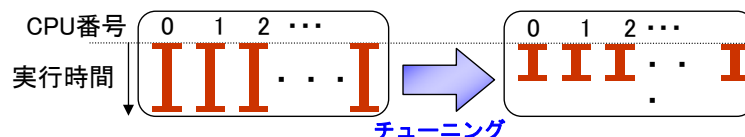
① 並列度の増大 → 並列化率 : 99.999%以上



高並列化率を達成するアルゴリズム



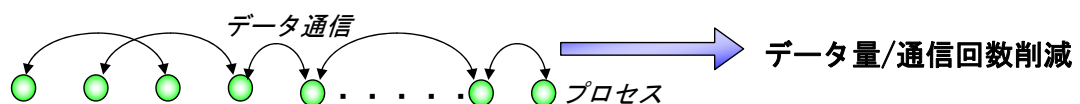
② CPU (プロセス) チューニング



■ チューニング手法

- キャッシュ/ソフトウェアパイプライン
- ベクトルチューニング
- ...

③ 通信量削減

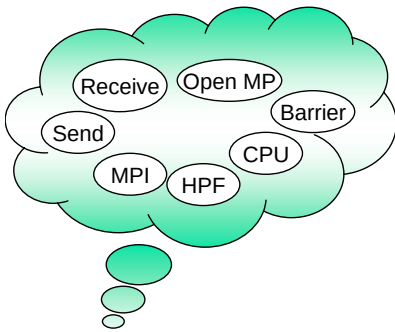


④ ロードバランス



27

Increasing Complexity of Programming



Serial Code

```
real*4 X(400)
do 10 i=1,400
  X(i)=i
10 continue
S=0
do 30 i=1,400
  S=S+X(i)
30 continue
write(6,*) '1+...+400=',S
stop
end
```

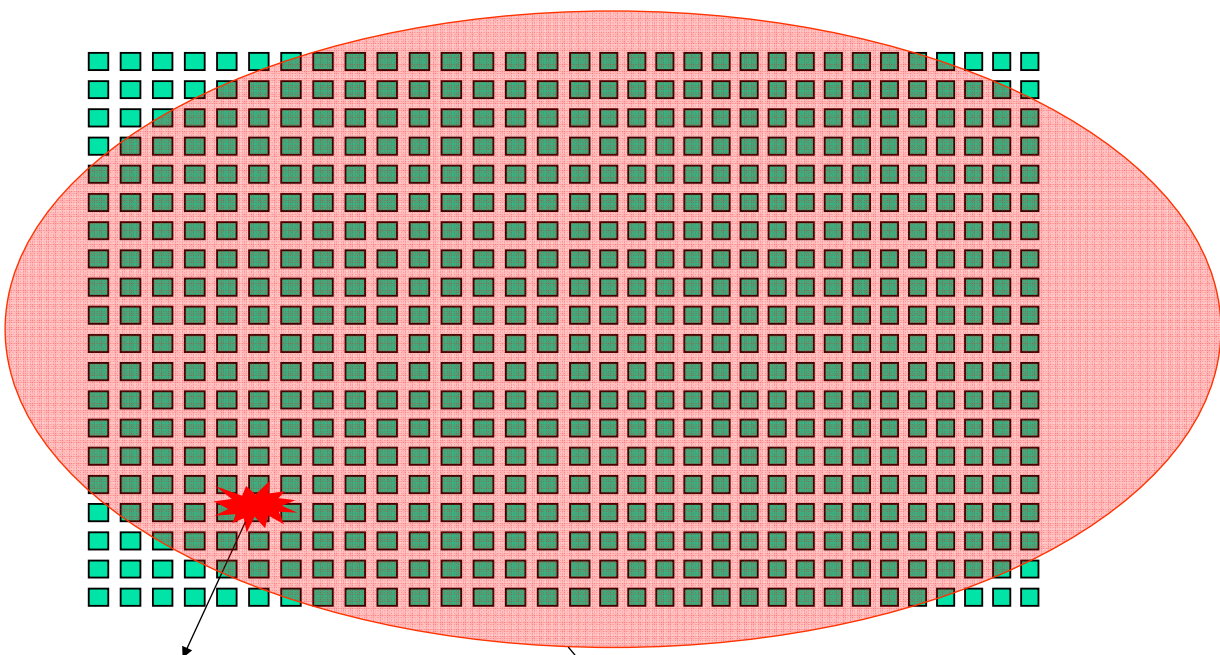
MPI

```
parameter (n=400, np=4)
parameter(masterpid=0)
real*4 X(400)
integer to_p, from_p, tag, mypid, pnum
call MPI_init(4)
call MPI_comm_rank(MPI_COMM_WORLD, mypid)
call MPI_comm_size(MPI_COMM_WORLD, pnum)
if (mypid.eq.masterpid) then
  do 10 i=1,400
    X(i)=i
10 continue
  do 20 to_p=1,3
    tag=0
    call MPI_send(X(100*to_p+1), 100, MPI_REAL, to_p, tag, MPI_COMM_WORLD)
20 continue
  else
    from_p=0
    tag=0
    call MPI_recv(X(1), 100, MPI_real, from_p, tag, MPI_COMM_WORLD, idummy)
  endif
  S=0
  do 30 i=1,100
    S=S+X(i)
30 continue
  if (mypid.ne.masterpid) then
    to_p=0
    tag=1
    call MPI_send(S, 1, MPI_REAL, to_p, tag, MPI_COMM_WORLD)
  else
    do 40 from_p=1,3
      tag=1
      call MPI_recv(SS, 1, MPI_REAL, from_p, tag, MPI_COMM_WORLD, idummy)
      S=S+SS
40 continue
    write(6,*) '1+...+400=',S
  endif
  call MPI_barrier(MPI_COMM_WORLD)
  call MPI_finalize
stop
end
```

28

System Reliability

- To Minimize Total System Down
- To Maintain Data Integrity (Error Checking)

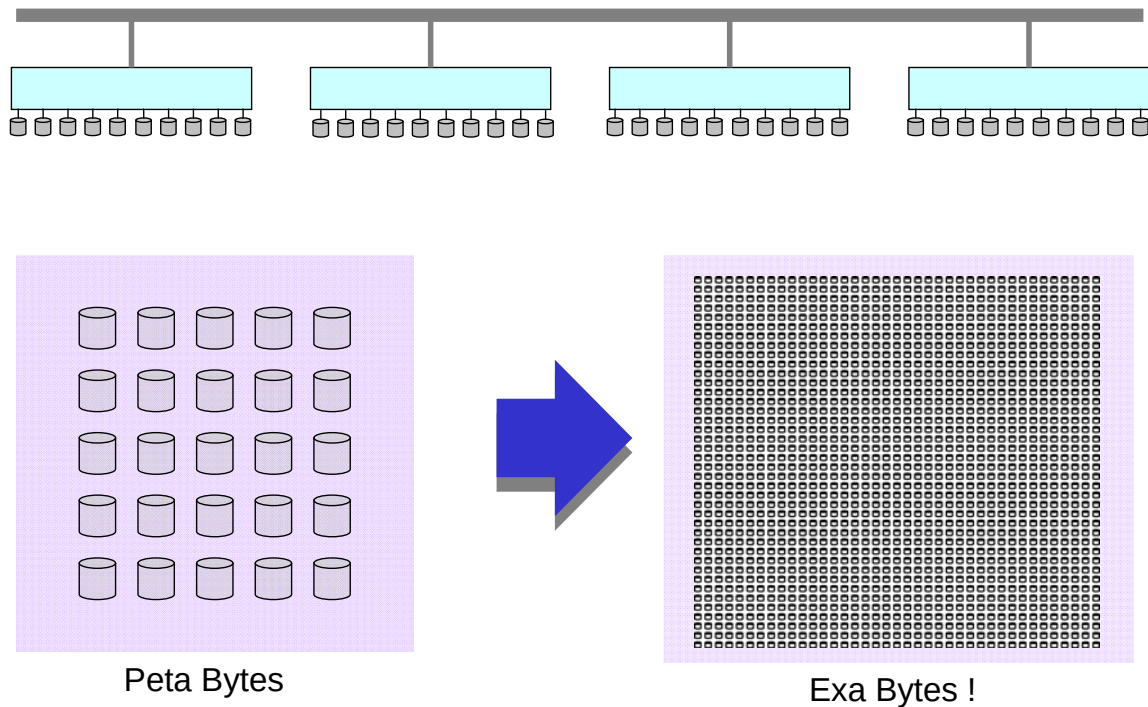


A Single Point of Failure Must not Cause Total System Crash.

Solid and Intermittent Errors of Data Paths, Control, Arithmetic Units and Memory Must be Strictly Checked.

29

Increasing Storage (Data Explosion)



30

Conclusion

- Nine Lessons Learned in the Design of CDC6600 by N. R. Lincoln (1977) –
It's Really not as much Fun Building Supercomputer as it is Simply Inventing One.
(Proc. of the Symposium on High Speed Computer and Algorithm Organization, 1977)

Lesson 9

The success or failure of any supercomputer development is finally going to rest on the ability and willingness of users to adapt to the strange world of parallel processing, and the consequent need to restructure algorithm, if not total processes.

END