

スーパーコンピュータと計算科学

担当 佐藤

1

もくじ

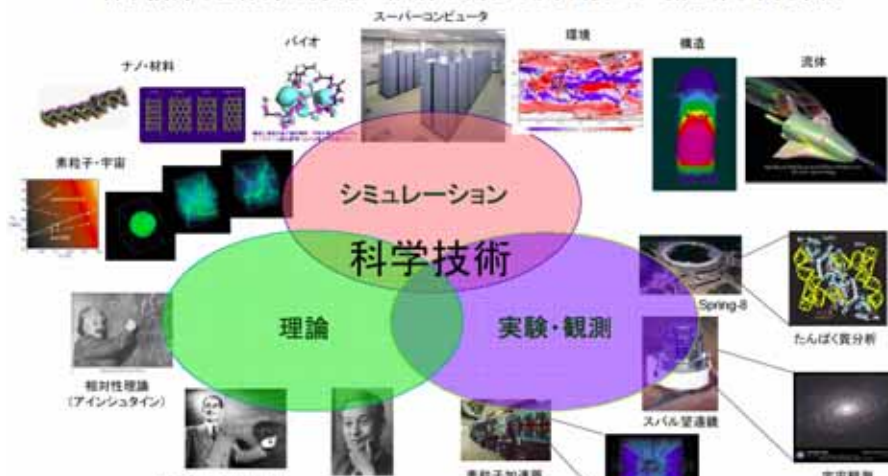
- ◆ 計算科学とは
- ◆ コンピュータはどのくらい早くなったのか
- ◆ スーパーコンピュータと並列処理
- ◆ 計算科学研究センター、PACS-CSの紹介
- ◆ 何ができるのか、何ができたのか
- ◆ 課題、最近の話題

- ◆ グリッドコンピューティングへ

2

科学の三本柱としての計算科学

- 超高速計算機（スーパーコンピュータ）を用いた大規模シミュレーションを中心とした研究
- 科学技術の全分野で、実験・観測、理論と並ぶ、重要且つ最先端の研究手段



計算科学：第三の科学

◆ 第三の科学

－ 実験できない領域

- 素粒子科学 Quantum chromodynamics (量子色力学)
 - クォークの質量
- 宇宙物理学
 - 宇宙の成り立ち
- 地球温暖化 予測
 - このまま進むとどうなるか

－ 第一原理的手法を使用すれば、実験不可能なことでも、シミュレーションによって解明される、であろうことが明らかになりつつある。

- バイオ、ナノテクノロジー
- 現在の計算機リソースでは不可能なものも多い：例えば、AMD Opteron × 1024 クラスターの10,000倍規模の計算機必要など

－ 実験、観測の検証。それでも実験より手軽に実施可能である。

- 流体シミュレーション、構造設計
 - 自動車、ビル設計、飛行機・ロケットの設計

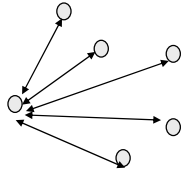
4

代表的な計算科学の手法

◆ 粒子計算

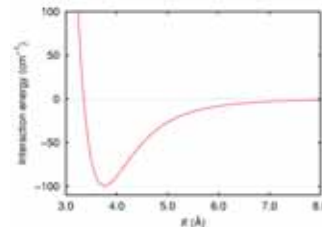
- 個々の粒子の相互作用を計算する
- 宇宙：重力
 - $F_{ij} = G m_i m_j / r^2$
- 分子シミュレーション：Van der Waals forces
 - 分子動力学
 - 2体（或いはそれ以上）の原子間ポテンシャルの下に、古典力学におけるニュートン方程式を解いて、系の静的、動的安定構造や、動的過程（ダイナミクス）を解析する手法。

$$F_i = \sum F_{ij}$$



$$\begin{aligned} F &= m a \\ v &= v + a \Delta t \\ p &= p + v \Delta t \end{aligned}$$

各ステップで
運動方程式
を解く



代表的な計算科学の手法

◆ 連続体シミュレーション

- 偏微分方程式を解く
- 離散化
 - 本来連続体である対象を要素の集合体によって近似する。
- 有限要素法
 - 計算対象の構造に外力が加わって変形する場合等を解析する際、対象の構造をメッシュで区切り（有限数の小さな要素で区切る）、各々の要素内で成り立つ連立一次方程式を作成する（未知数は変位、速度、圧力など）。次に、各要素における方程式を全解析領域分足し合わせることで大きな連立一次方程式（マトリクス方程式）を作成し、解を求める。
- 境界要素法
 - 解くべき対象である偏微分方程式を、境界上の積分方程式の問題に置き換えて解く手法。
- 流体計算
 - ナビエーストークス方程式

代表的な計算科学の手法

◆ 第一原理計算

- 量子力学の基本方程式であるシュレディンガー方程式を解いて、原子レベルのシミュレーションを行う方法
 - 「既存の実験結果（事実）を含めて経験的パラメーター等を一切用いない」という強いものから、「実験結果に依らない」とする比較的緩い解釈まである。
- 量子化学：分子軌道法（MO）
- 密度汎関数法（DFT計算）
- 原理的にはどのような現象でも解けるはずであるが、膨大な計算量を必要とする

計算機はどのくらい早くなったか

◆ 能力はどうやって計るのか？

- 1秒あたりの演算可能回数
- Top500

◆ マイクロプロセッサの発展

- クロックスピードにほぼ比例して早くなる

◆ スーパーコンピュータは並列処理の時代へ

- 単一プロセッサでは限界！
- スーパーコンピュータは並列処理により早くなっている

いろいろなマイクロプロセッサ(1)

- ◆ マイコン(4ビットマイコン)
 - 4004(世界初、1971年、750KHz)
- ◆ 8ビットマイコン
 - 8008(1972年、500KHz、インテル)
 - 8080(1974年、2MHz、インテル)
 - Z80(1976年、10MHz、ザイログ)
 - MC6800(1974年、1MHz、モトローラ)
 - MC6809
- ◆ 16ビットマイコン
 - 8086(1978年、インテル)
 - IBM PC/MS-DOS
 - 80286(1982年、インテル)
 - MC68000(1979年、モトローラ)
 - UNIX
- ◆ 8ビット、16ビットとは、バスの幅、メモリ空間のビット幅のこと。

いろいろなマイクロプロセッサ(2)

- ◆ 32ビットプロセッサ
 - 80386(1985年)、80486(1989年、40MHz~)
 - MC68020(1984年)、MC68030(1987年)
 - 仮想記憶
 - Pentium(1995年、100MHz~200MHz)
 - Pentium II(1998年、300MHz~)
 - SSE/MMX
 - Pentium III(1997年、900MHz~)
 - 1GHzを超える
 - Pentium 4(2000年、~3.2GHz)
 - AMD K9, AMD Athlon
- ◆ 64ビットプロセッサ
 - Itanium(2000)、Itanium II(2001)
 - AMD Opteron(2003)
- ◆ 30年間で、1MHzから1GHz、1000倍の進歩

高速化とは

- ◆ コンピュータの高速化
 - デバイス
 - 計算機アーキテクチャ
- ◆ 計算機アーキテクチャの高速化の本質は、いろいろな処理を同時にやること
 - CPUの中
 - チップの中
 - チップ間
 - コンピュータ間

パイプライン、
スーパースカラ

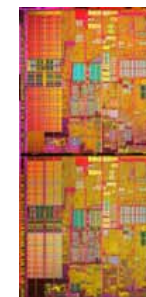
マルチ・コア

共有メモリ
並列コンピュータ

分散メモリ並列コンピュータ
グリッド

プロセッサ研究開発の動向

- ◆ さらなる高速化、高性能化へ
 - クロックの高速化、製造プロセスの微細化
 - いまでは3GHz、数年のうちに10GHzか
 - インテルの戦略の転換 マルチコア
 - プロセスは90nm 65nm, 将来的には45nm
 - 量子的な限界?
 - アーキテクチャの改良
 - スーパーパイプライン、スーパースカラ、VLIW...
 - キャッシュの多段化、マイクロプロセッサでもL3キャッシュ
 - マルチスレッド化、Intel Hyperthreading、複数のプログラムを同時に処理
 - マルチコア: 1つのチップに複数のCPU



インテル® Pentium® プロセッサ
エクストリーム・エディションのダイ

AMD's Next Generation Processor Technology

Native quad core die

Optimized for 65nm SOI and beyond

Expandable shared L3 cache

IPC enhanced CPU cores

Enhanced Direct Connect Architecture and Northbridge

- ▶ 32B instruction fetch
- ▶ Improved branch prediction
- ▶ Out-of-order load execution
- ▶ Up to 4 DP FLOPS/cycle
- ▶ Dual 128-bit SSE dataflow
- ▶ Dual 128-bit loads per cycle
- ▶ Bit Manipulation extensions (LZCNT/POPCNT)
- ▶ SSE extensions (EXTRQ/INSERTQ, MOVNTSD/MOVNTSS)

- ▶ HyperTransport™ links (5.2GT/sec)
- ▶ Enhanced crossbar
- ▶ DDR2 with migration path to DDR3
- ▶ FBDIMM when appropriate
- ▶ Enhanced power management and RAS

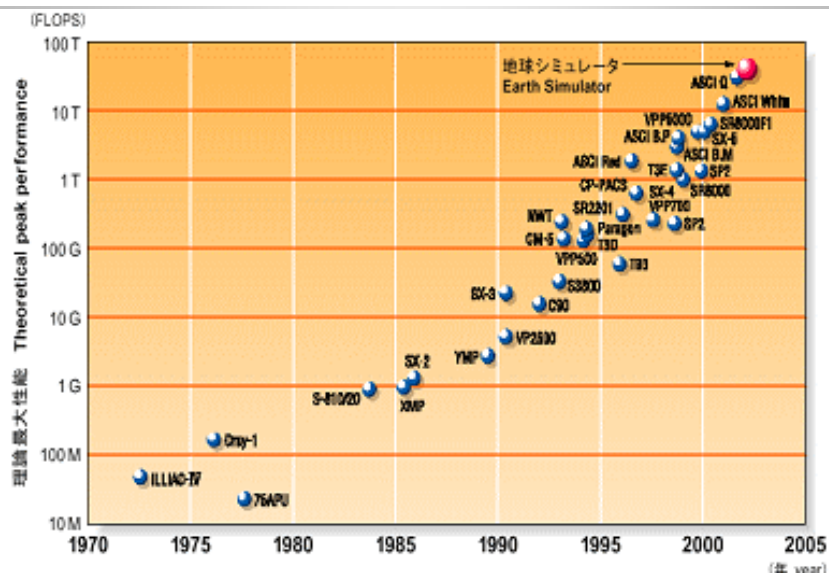
AMD

30 05/16/06 C. Moore - SPF 2006

計算機ハードウェアの歴史

- ◆ 1.5年に2倍の割合で処理速度が増加している（Moore's Law）
- ◆ 1983年：1 GFLOPS，1996年：1 TFLOPS，2002年：36 GFLOPS
 - MFLOPS: Millions of Floating Point OperationS.（1秒間に 10^6 回の浮動小数点処理）
 - GFLOPS： 10^9 回，TFLOPS： 10^{12} 回，PFLOPS： 10^{15} 回
 - 2010年頃にはPFLOPS（Peta FLOPS）マシンが登場するといわれている。
- ◆ 「地球シミュレータ」はピーク性能 40 TFLOPS
- ◆ 2005年にはIBM BlueGene/L（367 TFLOPS）が完成

計算機ハードウェア発達の歴史



計算機の進歩

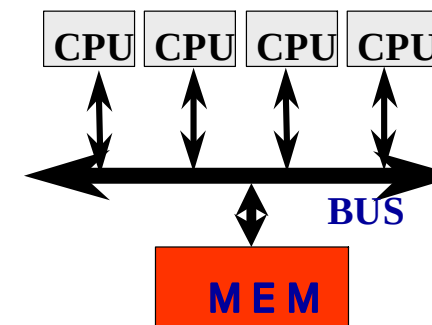
- ◆ **個々のプロセッサ**
 - ベクトルプロセッサ 10年前はこのタイプが多かった
 - 一つのプロセッサで行列演算を効率的にできる
 - スカラープロセッサ：Pentium, Power, Alpha, Itanium
- ◆ **並列計算機のアーキテクチャの種類**
 - 分散メモリ型並列計算機
 - SMP (Symmetrical Multi Processor)：共有メモリ型並列計算機
 - SMPクラスタ型並列計算機（「Constellation」とも言う）
- ◆ **並列計算機の別の分類**
 - PCクラスタ
 - 専用並列計算機
 - 専用機・・・GRAPE等：ブラックホールのシミュレーション

並列処理と分散処理

- ◆ 並列処理(parallel processing)とは、複数のプロセッサを用いて、処理を高速化する技術
 - HPC(High Performance Computing)
 - 数値シミュレーションなど
 - HTC (High Throughput Computing)
 - 大量のデータ処理
- ◆ 分散処理(distributed processing)は、もちろん、複数のプロセッサを用いるため処理を高速化することもあるが、本質的にはいろいろな場所で行う様々な処理、あるいは機能を結合し、機能分担させることが目的であり、必ずしも高速化だけが目的ではない
 - 分散オブジェクト技術
 - RMI , J2EE, Jini...

17

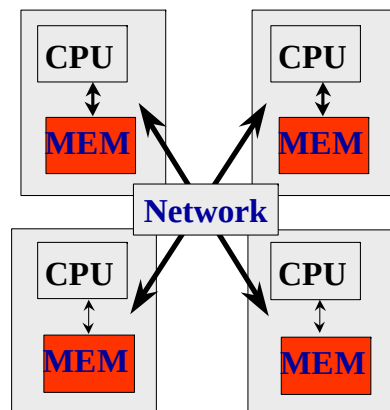
共有メモリ型計算機



- ◆複数のCPUが一つのメモリにアクセスするシステム。
- ◆それぞれのCPUで実行されているプログラム(スレッド)は、メモリ上のデータお互いにアクセスすることで、データを交換し、動作する。
- ◆大規模サーバ

18

分散メモリ型計算機



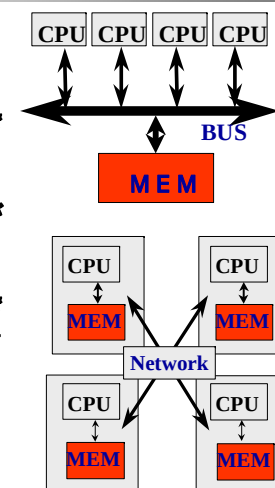
- ◆CPUとメモリという一つの計算機システムが、ネットワークで結合されているシステム
- ◆それぞれの計算機で実行されているプログラムはネットワークを通じて、データ(メッセージ)を交換し、動作する
- ◆超並列 (MPP : Massively Parallel Processing) コンピュータ
- ◆クラスタ計算機

19

並列処理の利点

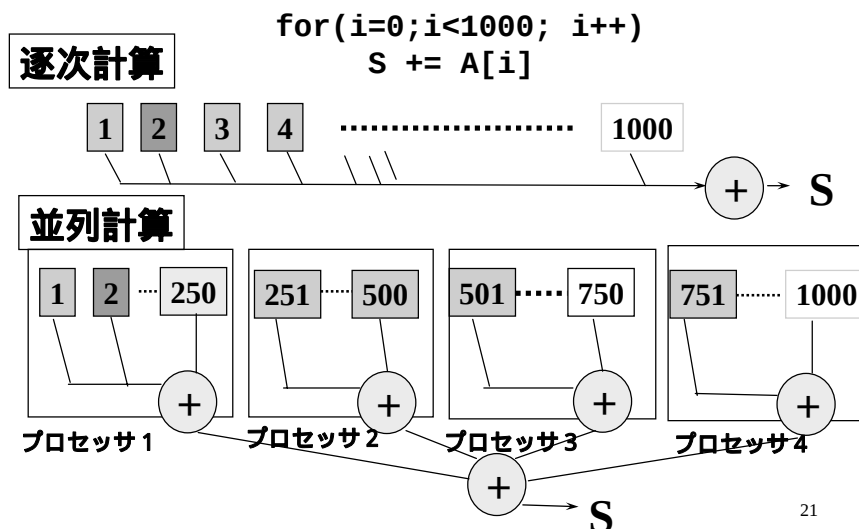
- ◆ 計算能力が増える。
 - 1つのCPUよりも多数のCPU。
- ◆ メモリの読み出し能力(バンド幅)が増える。
 - それぞれのCPUがこのメモリを読み出すことができる。
- ◆ ディスク等、入出力のバンド幅が増える。
 - それぞれのCPUが並列にディスクを読み出すことができる。
- ◆ キャッシュメモリが効果的に利用できる。
 - 単一のプロセッサではキャッシュに載らないデータでも、処理単位が小さくなることによって、キャッシュを効果的に使うことができる。
- ◆ 低コスト
 - マイクロプロセッサをつかえば。

→ クラスタ技術



20

並列処理の簡単な例



21

スーパーコンピュータの速さはどうやって測るのか

◆ 演算速度

- 個々プロセッサの演算速度×台数

◆ Top500 List

- 世界のコンピュータの速度を決める基準になっている

22

TOP 500 List

<http://www.top500.org/>

◆ LINPACKと言われるベンチマークテストを実施する。

- 密行列を係数とする連立一次方程式を解く
- ベクトル機でもスカラー機でも性能が出やすい
- 例えば地球シミュレータはピーク性能40TFLOPS（設計値）に対して35 TFLOPS以上の性能が出ている。

◆ 実際のアプリケーションではこれほどの性能は出ない

- 差分法, スペクトル法系の手法：ピーク性能の60%程度
 - AFES on the Earth Simulator：26 TFLOPS（ピーク性能の65%）
- 有限要素法
 - GeoFEM on the Earth Simulator（512ノード）：10 TFLOPS（30%）
 - スカラー機ではこれほど出ない：5%～10%

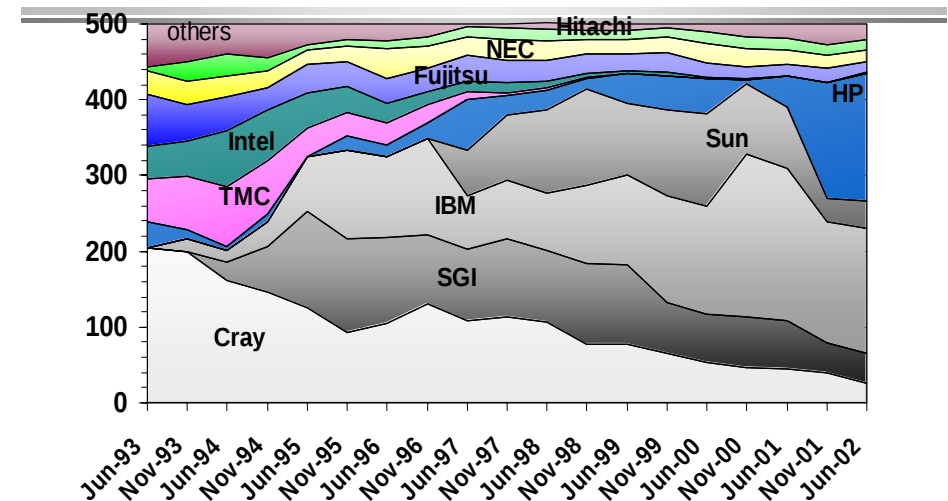
23

Rank	Site	Computer	Processors	Year	R _{max}	R _{peak}
1	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM	131072	2005	280600	367000
2	IBM Thomas J. Watson Research Center United States	BGW - eServer Blue Gene Solution IBM	40960	2005	91290	114688
3	DOE/NNSA/LLNL United States	ASC Purple - eServer pSeries p5 575 1.9 GHz IBM	12208	2006	75760	92781
4	NASA/Ames Research Center/NAS United States	Columbia - SGI Altix 1.5 GHz, Voltaire Infiniband SGI	10160	2004	51870	60960
5	Commissariat a l'Energie Atomique (CEA) France	Tera-10 - NovaScale 5160, Itanium2 1.6 GHz, Quadrics Bull SA	8704	2006	42900	55705.6
6	Sandia National Laboratories United States	Thunderbird - PowerEdge 1850, 3.6 GHz, Infiniband Dell	9024	2006	38270	64972.8
7	GSIC Center, Tokyo Institute of Technology Japan	TSUBAME Grid Cluster - Sun Fire X64 Cluster, Opteron 2.4/2.6 GHz, Infiniband NEC/Sun	10368	2006	38180	49868.8
8	Forschungszentrum Juelich (FZJ) Germany	JUBL - eServer Blue Gene Solution IBM	16384	2006	37330	45875
9	Sandia National Laboratories United States	Red Storm Cray XT3, 2.0 GHz Cray Inc.	10880	2005	36190	43520
10	The Earth Simulator Center Japan	Earth-Simulator NEC	5120	2002	35860	40960
11	Barcelona Supercomputer Center Spain	MareNostrum - JS20 Cluster, PPC 970, 2.2 GHz, Myrinet IBM	4800	2005	27910	42144
12	ASTRON/University Groningen Netherlands	Stella - eServer Blue Gene Solution IBM	12288	2005	27450	34406.4



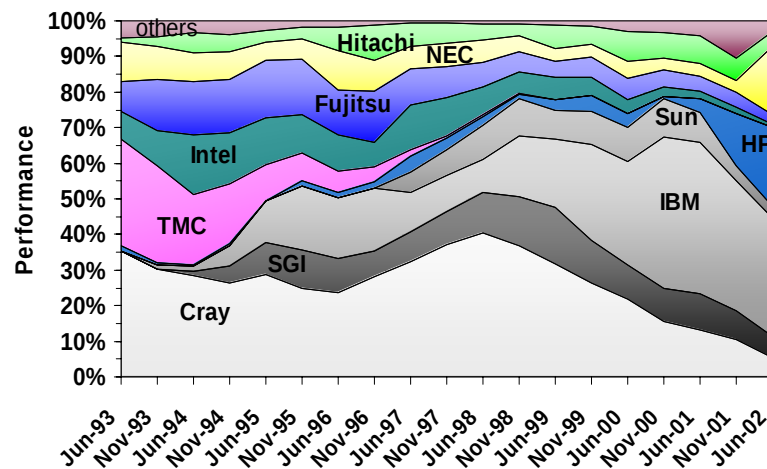
<http://www.top500.org/>

Manufacturers



<http://www.top500.org/>

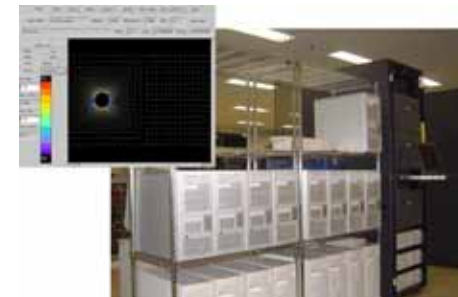
Manufacturers



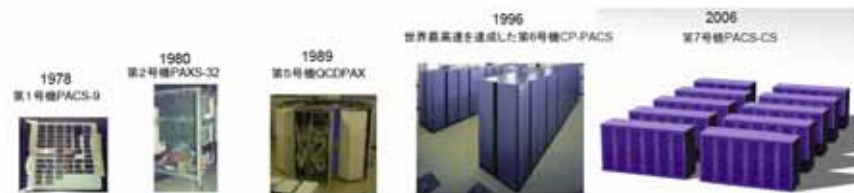
<http://www.top500.org/>

クラスタコンピューティング

- ◆ PCやネットワークなど、汎用（コモデティ）の部品を組み合わせ、並列システムを構成する技術
 - － PCが急激に進歩した！
 - － ネットワークも早くなっている！
 - － 安価に、個人レベルでも並列システムを作れる！



筑波大学における超並列計算機の開発の歴史



完成年	名称	計算速度
1978年	PACS-9	7千回/秒
1980年	PACS-32	50万回/秒
1983年	PACS-128	4百万回/秒
1984年	PACS-32J	3百万回/秒
1989年	QCDPAX	14億回/秒
1996年	CP-PACS	614億回/秒
2006年	PACS-CS	14兆3千億回/秒

- 1977年(星野・川合)に始まる超並列計算機の研究・開発のバイオニア
- 1996年完成のCP-PACS(岩崎・中澤)は世界最高速を達成
- 計算科学者と計算機工学者の協力による、科学の明確な目標を持った計算機開発
- 開発における計算機メーカーとの産学連携

29

PACS-CS

Parallel Array Computer System for Computational Sciences

◆ 計算科学計算センターで7月から稼動

◆ 国内で開発されたスパコンとしては、2位

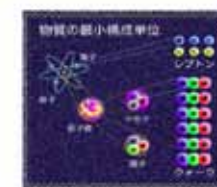
Peak 14.34TF
Linpack 10.35TF

34位(2006/6)

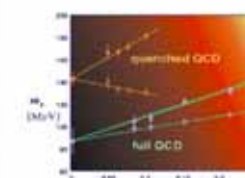


Theoretical peak speed	14.3TFLOPS(*) (64bit double precision, 2560 nodes)	
Total main memory	5.0TByte	
Total distributed disk	0.41PByte	
Node	processor	Intel Low Voltage Xeon 2.8GHz
	memory	2GByte DDR2/400
	FSB	800MHz
	Local disk	160GByte x 2 (RAID-1 mirror)
Network	Topology	3-dimensional Hyper Crossbar (16x16x10)
	Data transfer bandwidth	250MByte/sec (per direction) 750MByte/sec(node aggregate on 3-D)
	Method	Specially developed driver
Shared disk storage	10 TByte RAID-5 disk	
External connection	Trunked Gigabit Ethernet	Theoretical peak bandwidth 250Mbyte/sec
System size	59 racks/100m ² 2 foot print	
Power dissipation	545 kW maximum	

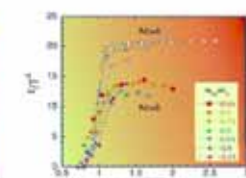
PACS-CSによる素粒子と宇宙の研究



物質の階層構造と基本粒子クォーク



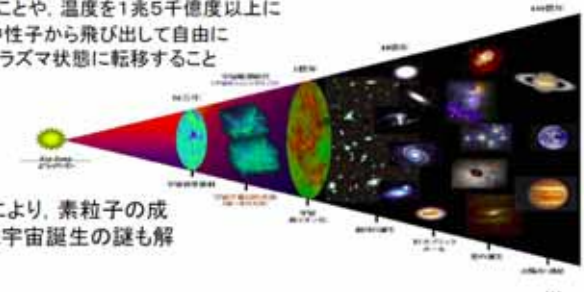
クォークの質量(重さ)の決定



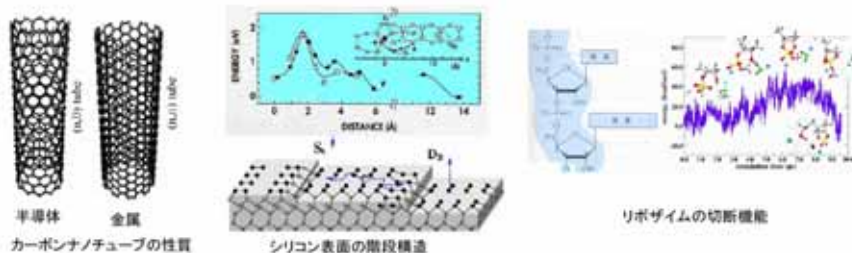
宇宙初期にあったと考えられている、クォーク・グルオン・プラズマへの相転移の研究

現在までの研究により、物質の最小の構成単位クォークは予想されていたより約50%も軽いことや、温度を1兆5千億度以上に上げると、クォークが陽子や中性子から飛び出して自由に飛び回るクォーク・グルオン・プラズマ状態に転移することなどが判明しつつある。

これらの研究を進めることにより、素粒子の成り立ちが解明され、さらには宇宙誕生の謎も解明されることが期待される。

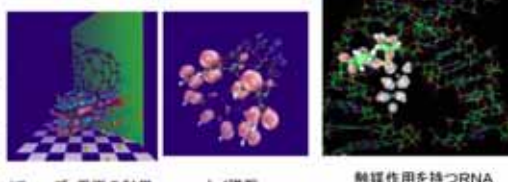


PACS-CSによる物質と生命の研究

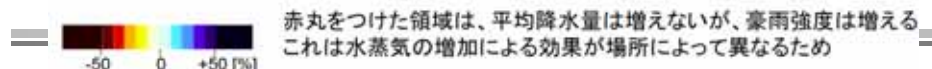


現在までの研究により、カーボンナノチューブや、蛋白質など、ナノバイオ物質の性質を解明するには、空間サイズ10ナノメートル以上、原子数1万個以上の量子力学シミュレーションが重要なことが判明しつつある。

このようなシミュレーションを実現することにより、ナノバイオ物質が、どのようにして様々な機能を持つかを理解し、さらに、新機能を持つナノバイオ新物質の設計・制御への道が開拓されることが期待される。

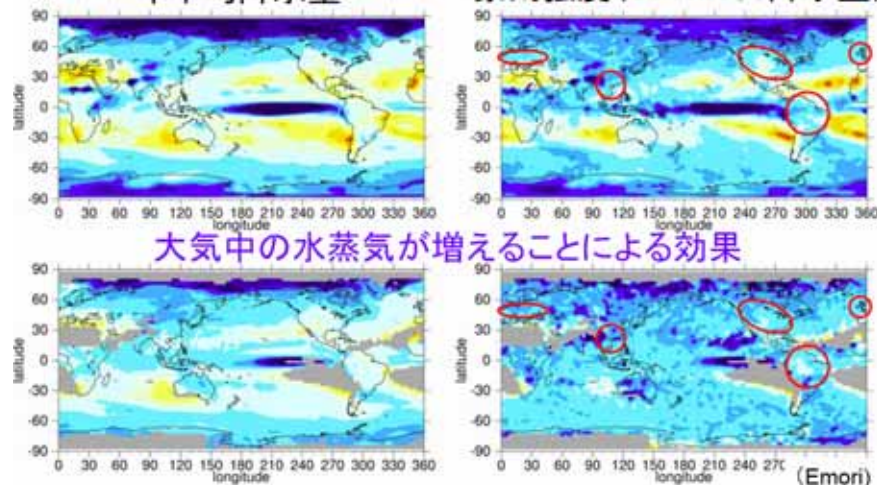


温暖化による平均降水量および豪雨強度の変化率



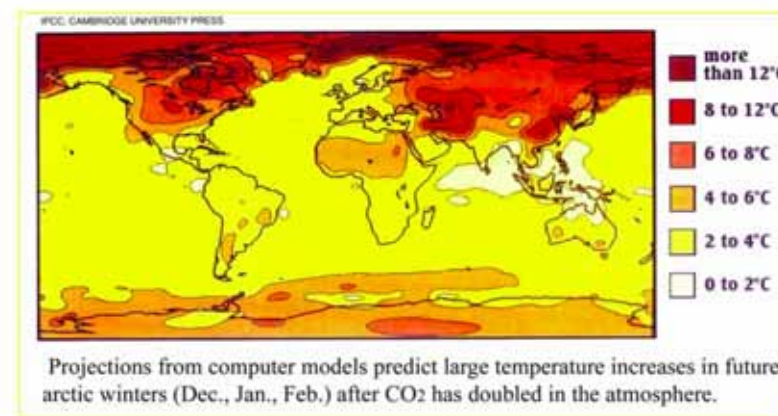
年平均降水量

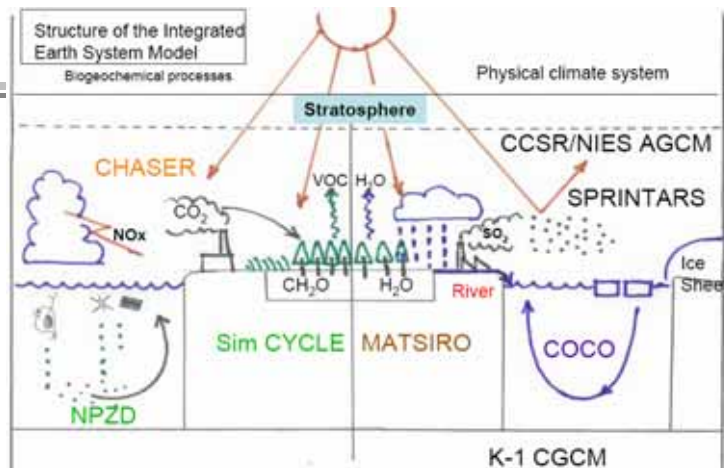
豪雨強度 (99%ile日降水量)



CO2倍増実験

大気中の二酸化炭素増加を想定したコンピューターの気候予測は特に冬期の北極圏の気温上昇を予測している。





これをシミュレーションするには、複雑なモデル、膨大な計算能力が必要

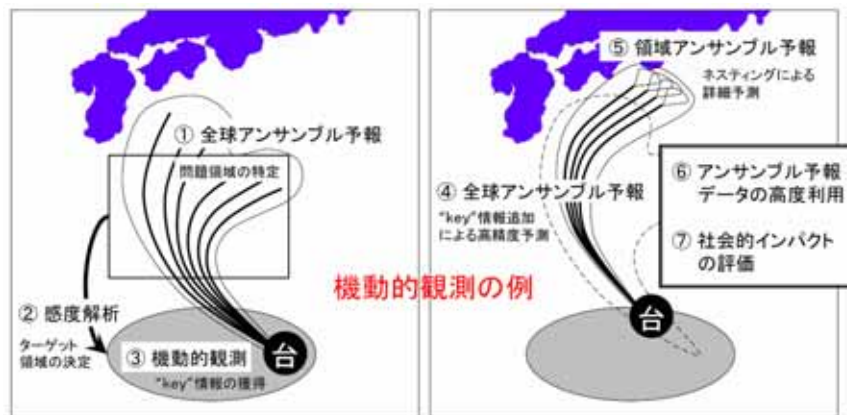
37

これからの計算科学の課題

- ◆ 計算を細かくやっても限界があることがある
 - 天気予報はあたるか？ カオスの問題
 - バタフライエフェクト：ある場所で蝶がはばいたただけで地球の反対で台風が起こる！？
 - アンサンブル計算
- ◆ 計算精度の問題
 - 現在は64ビットの浮動小数点という形式を使っている
 - $-1.79769 \times 10^{308} \sim 1.79769 \times 10^{308}$
- ◆ まだまだ性能が足りない
 - 「京速計算機プロジェクト」
 - 10PFlopsのスパコンの開発を目指す

38

アンサンブル予報



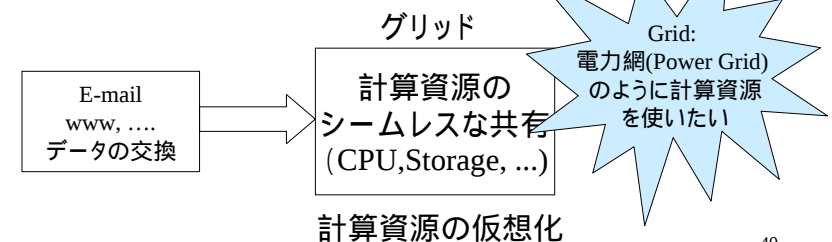
台風予報を例にとった、本計画のねらいの概念図

39

グリッドコンピューティングとは

- ◆ グリッド技術とは広域の高速ネットワーク上において、「安全に」大量のデータ、計算資源、貴重な装置等を共有し、協調作業、資源の有効活用するネットワーク基盤技術（ソフトウェア、ネットワーク、ハードウェア）と、これを活用する応用技術

従来のインターネット技術



計算資源の仮想化

40