

究極の技術への挑戦

—— スーパーコンピュータ ——

NEC / NEC ソリューションズ

渡辺 貞

スーパーコンピュータって何？

“その世代で最も速く、構成上も最大のコンピュータシステム”

“最も高価なコンピュータシステム”

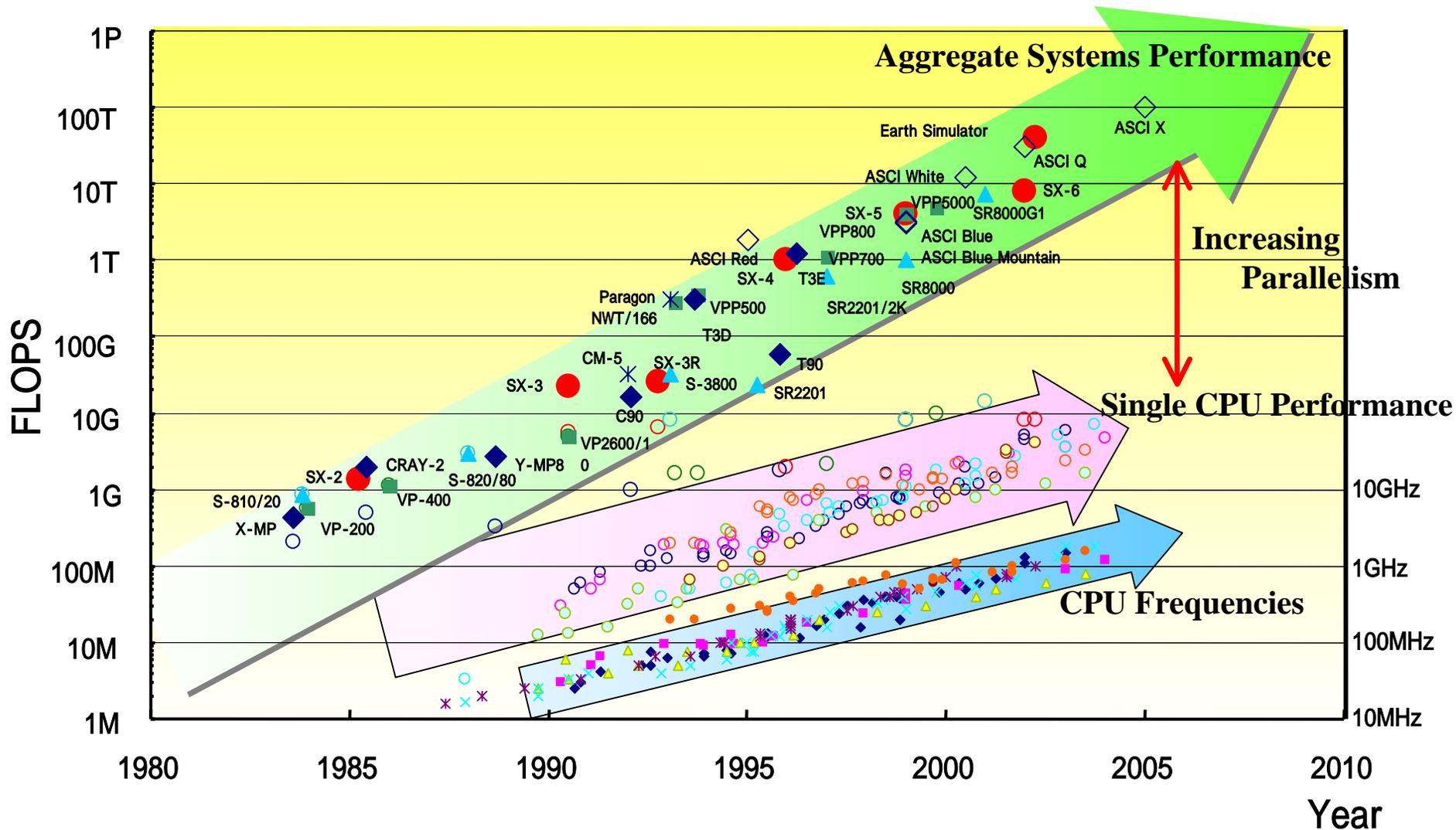
注)性能(速度)の単位:FLOPS

1FLOPS=1演算/秒

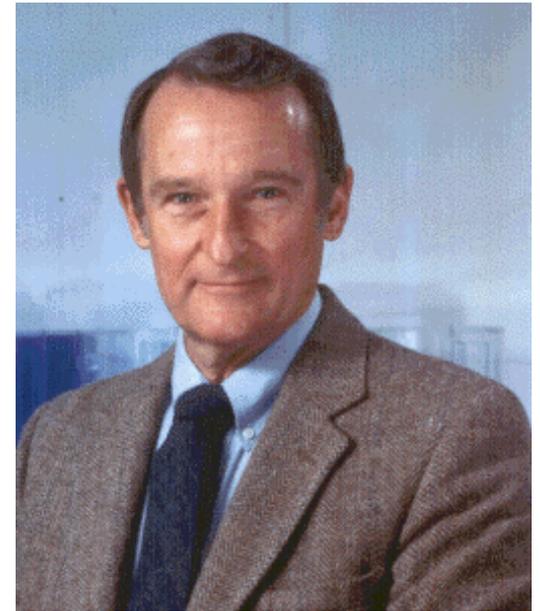
1GF(Giga Flops)=10億演算/秒

1TF(Tera Flops)=1兆演算/秒

スーパーコンピュータの歴史



Cray-1

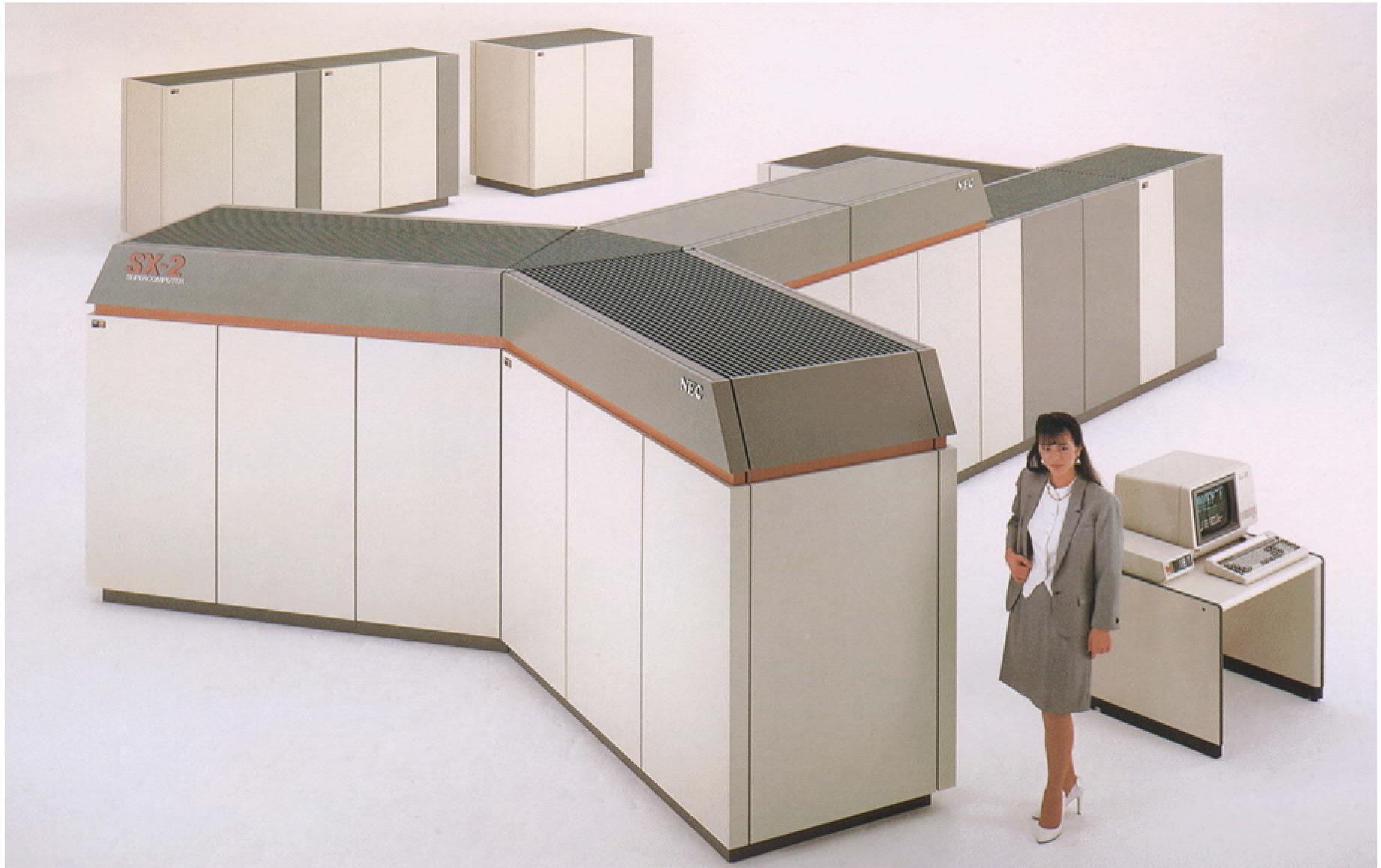


Seymour Cray

Cray-2



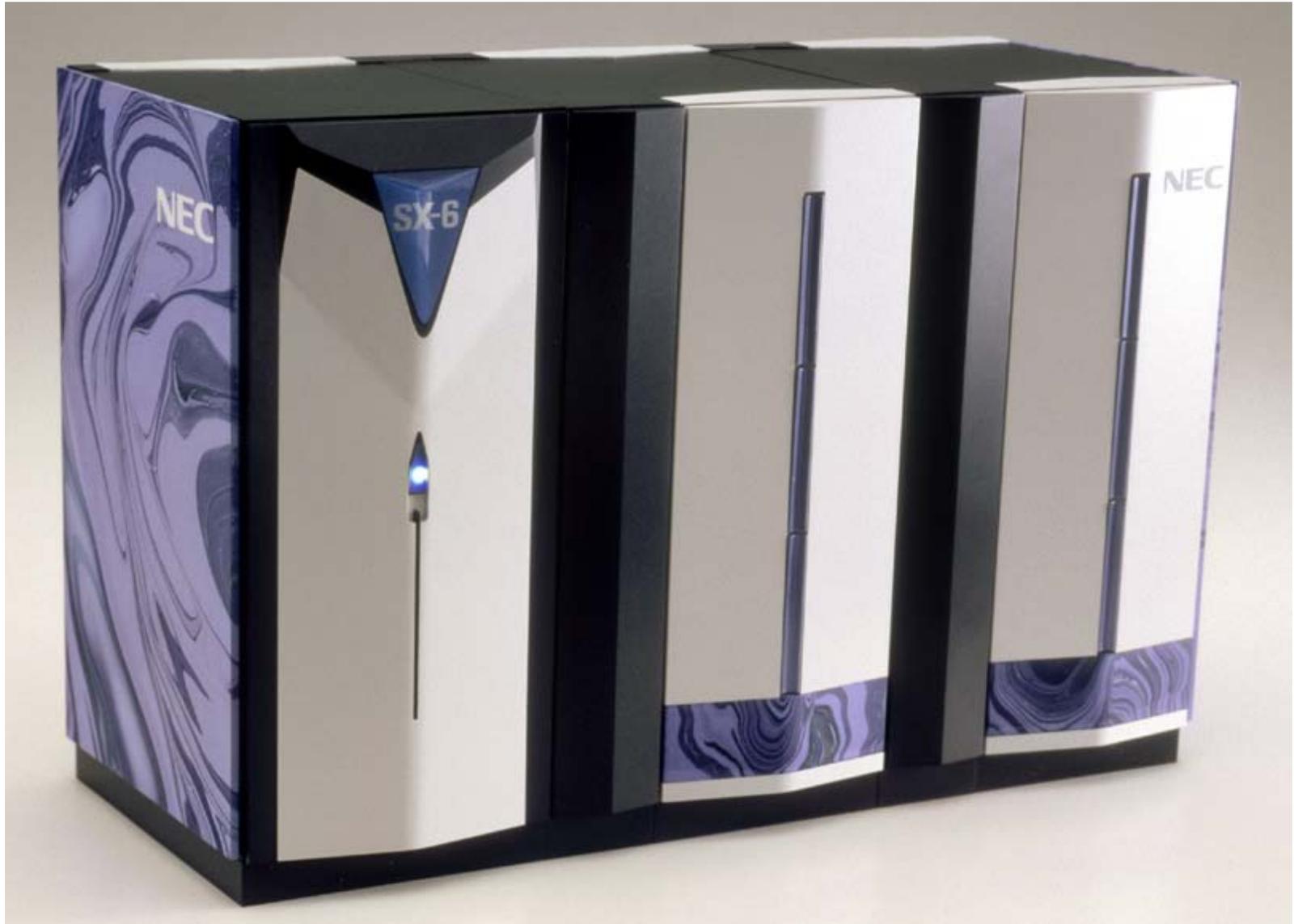
SX-2



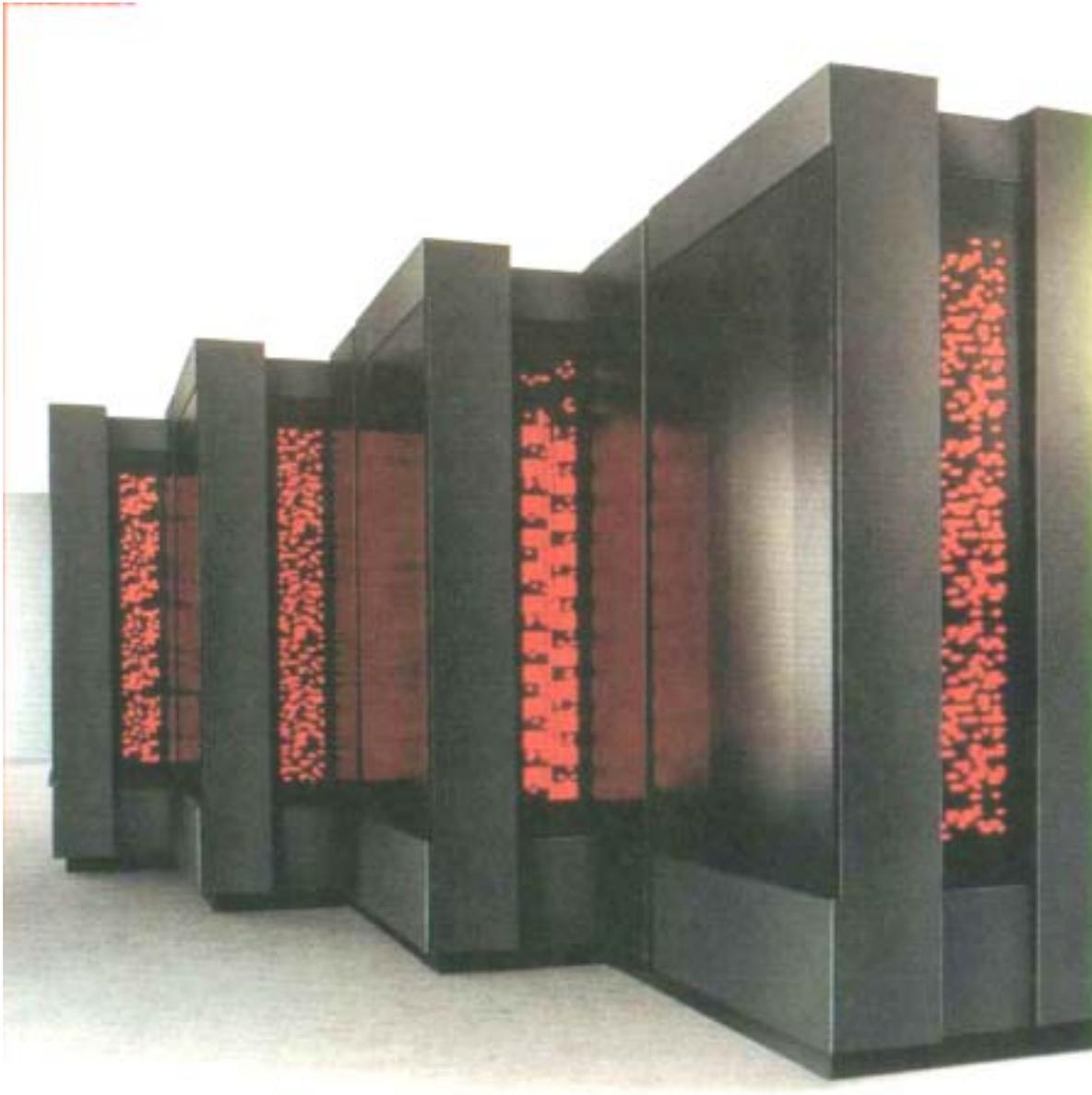
SX-4



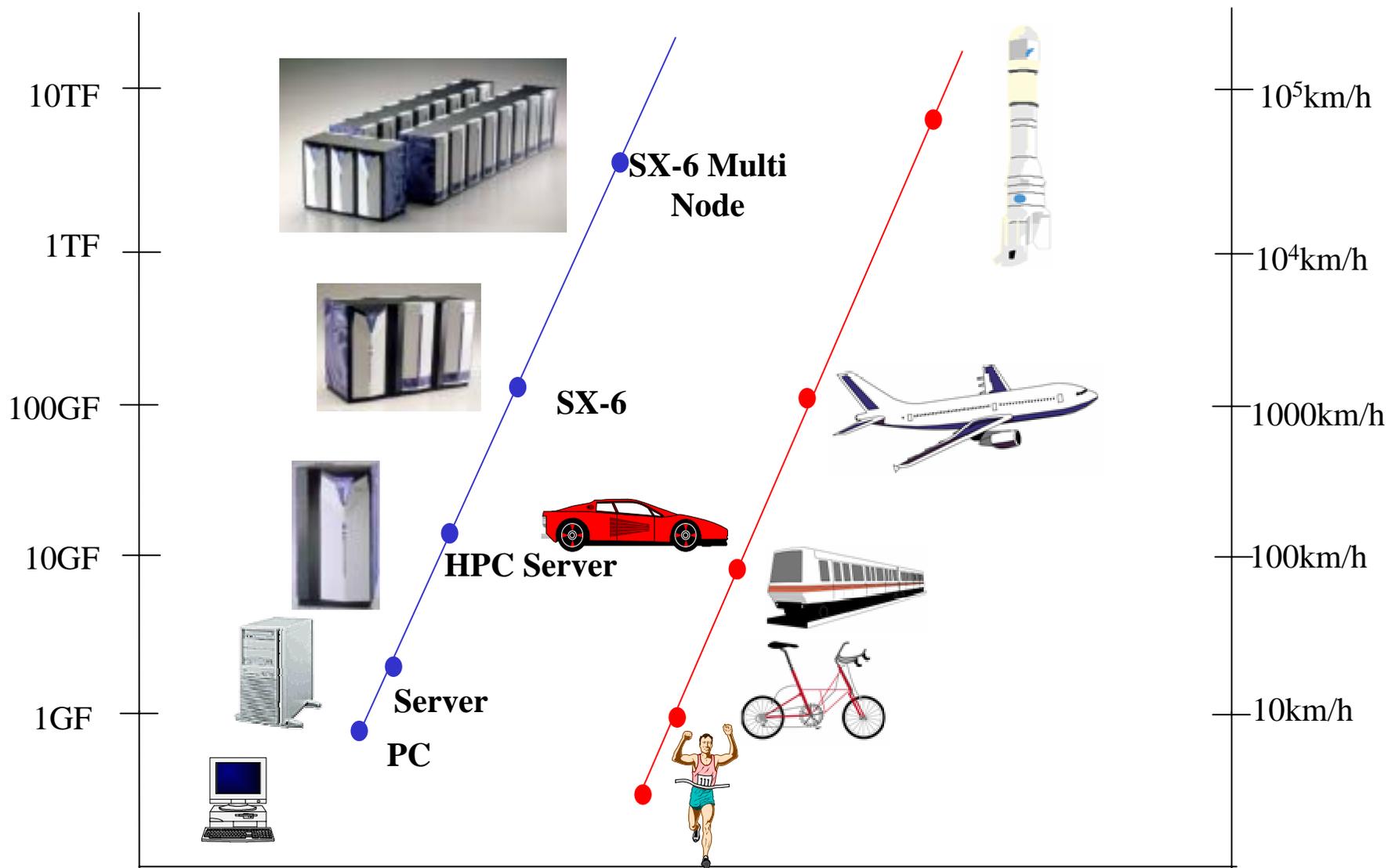
SX-6



CM5



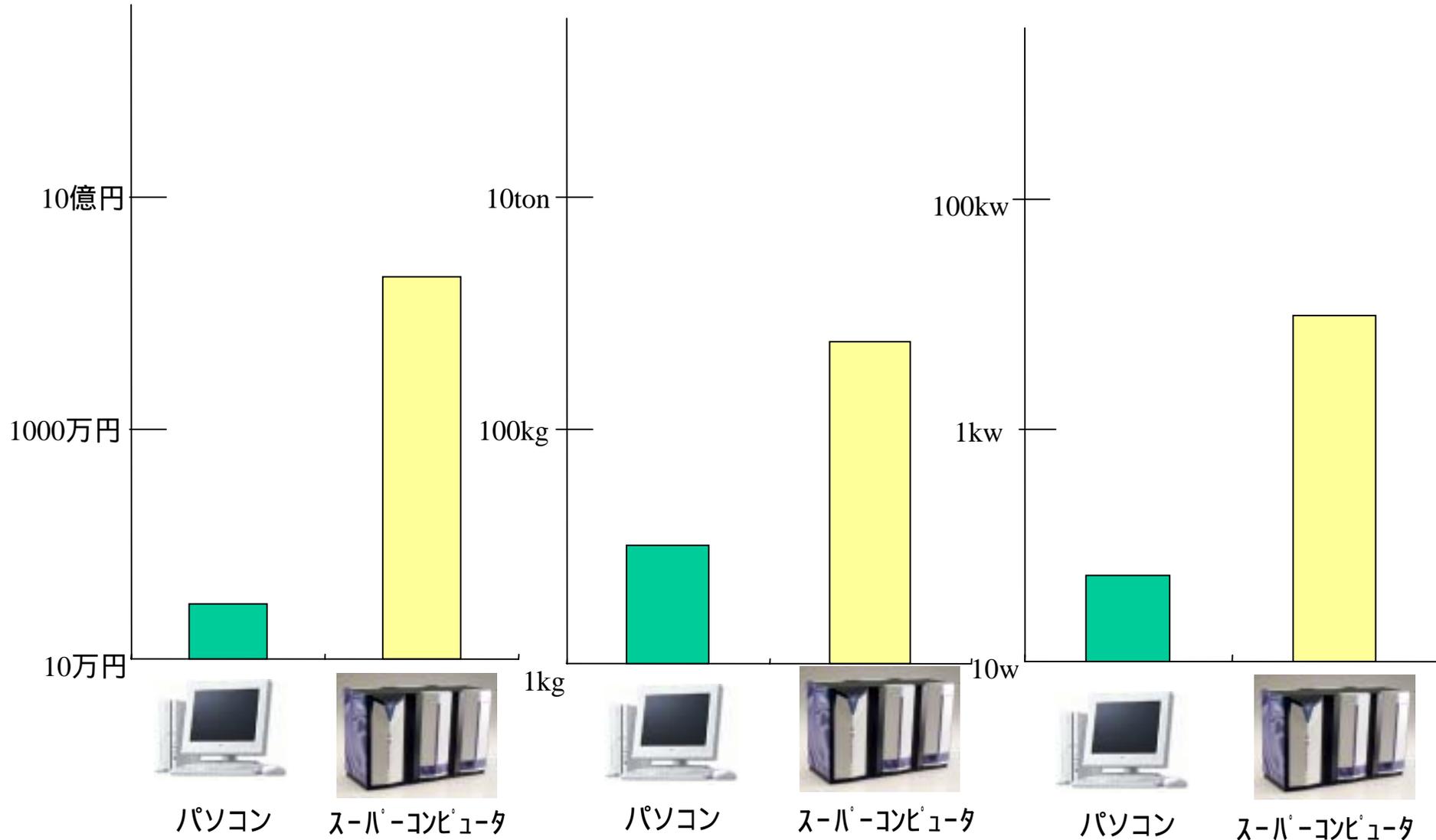
どのくらい速いか？



1GF(Giga Flops)= 10^9 Floating Point Operations per Sec.(10億演算/秒)

1TF(Tera Flops)= 10^{12} Floating Point Operations per Sec.(1兆演算/秒)

価格・重量・消費電力



スーパーコンピュータで何ができるか？

- スーパーコンピュータを使った数値シミュレーションで対象物を**拡大/縮小**あるいは**時間を延長/短縮**することにより、目に見えないもの、予測できないもの、実験不可能なものを**目で見、予測し、実験**を行うことができる。

第3の科学：計算科学

理論

実験

計算
(数値シミュレーション)

超長時間の現象：宇宙，気候，環境

超短時間の現象：核融合，衝突，燃焼

実験不可能：結晶/分子構造，
安全解析，気象

↓
計算機実験=数値シミュレーション

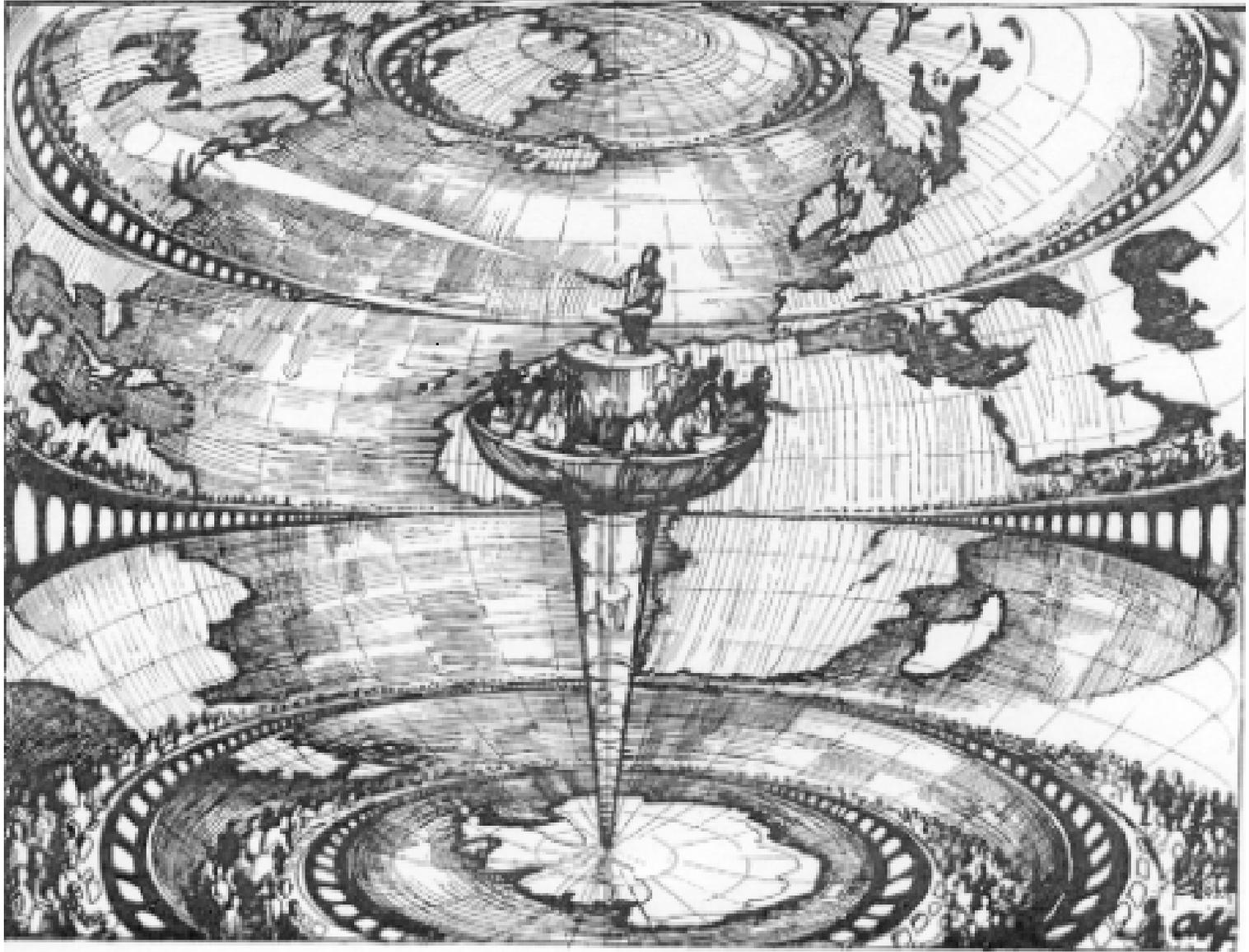
↓
膨大な計算量

↓
超高速コンピュータ(スーパーコンピュータ)

私達の地球



リチャードソンの夢

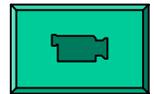


気象予測(地球温暖化)

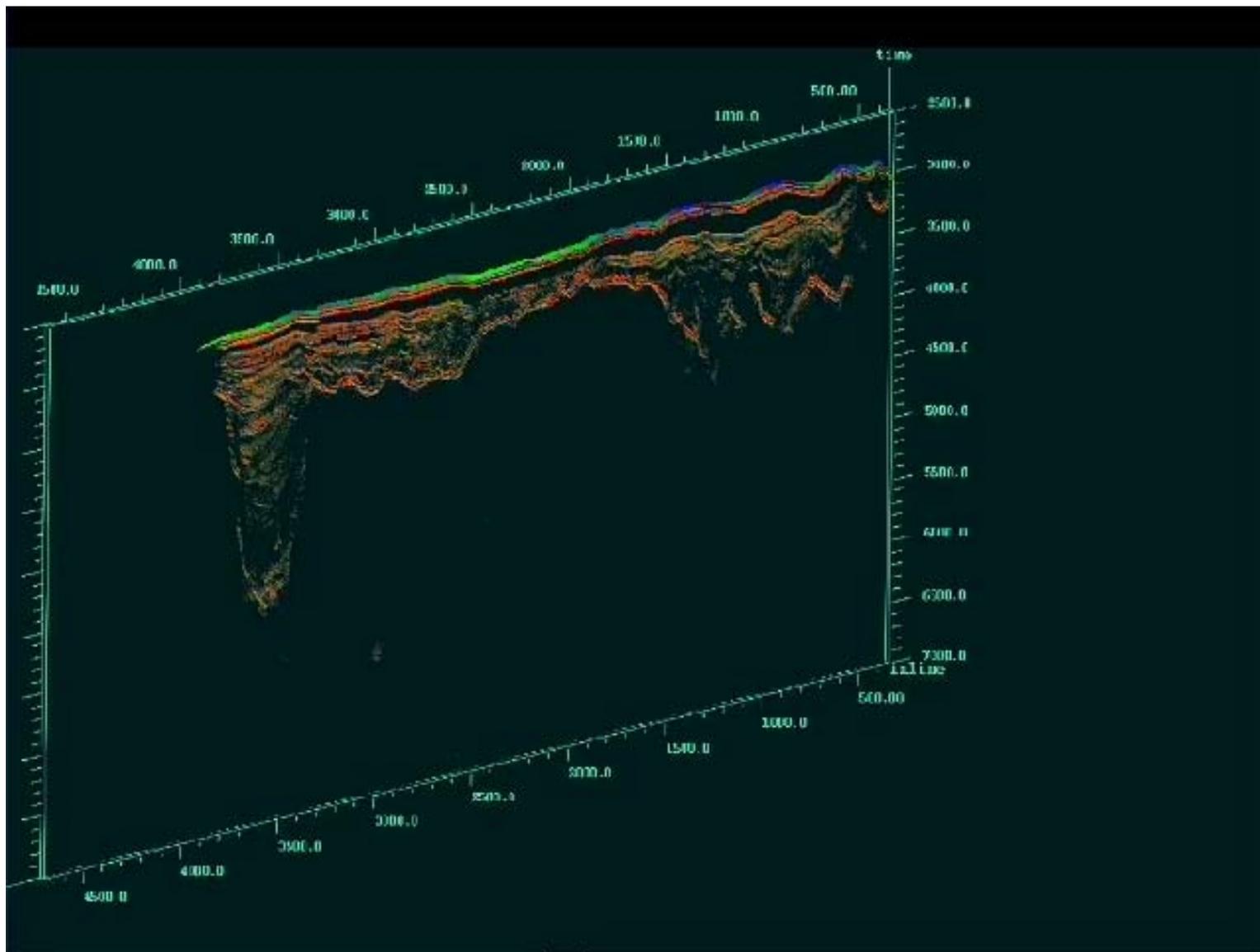
Next animation

*Change of surface temperature
due to increase of CO₂*

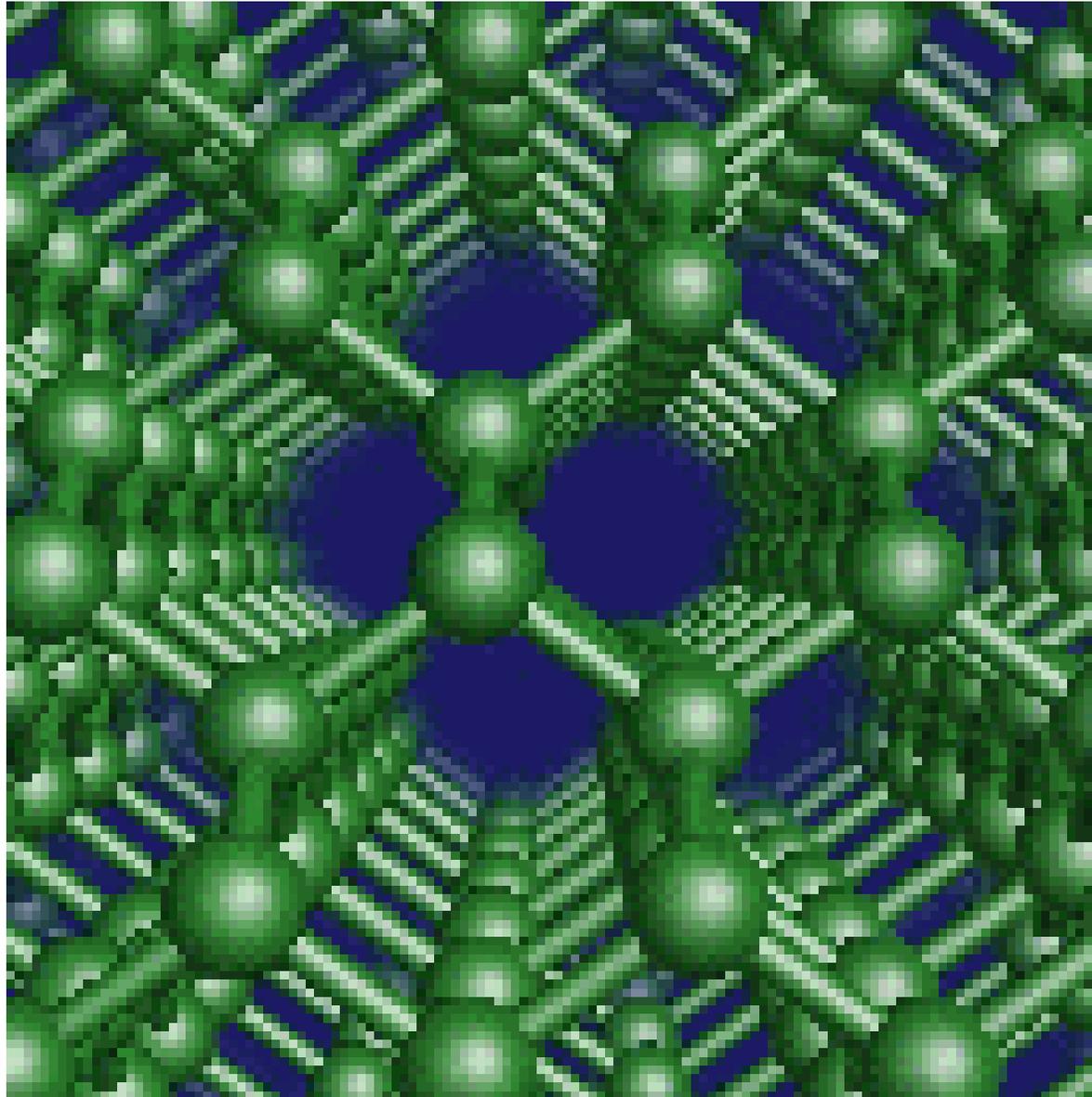
- difference from 1991 level temperatures.*
- every 5years animation*



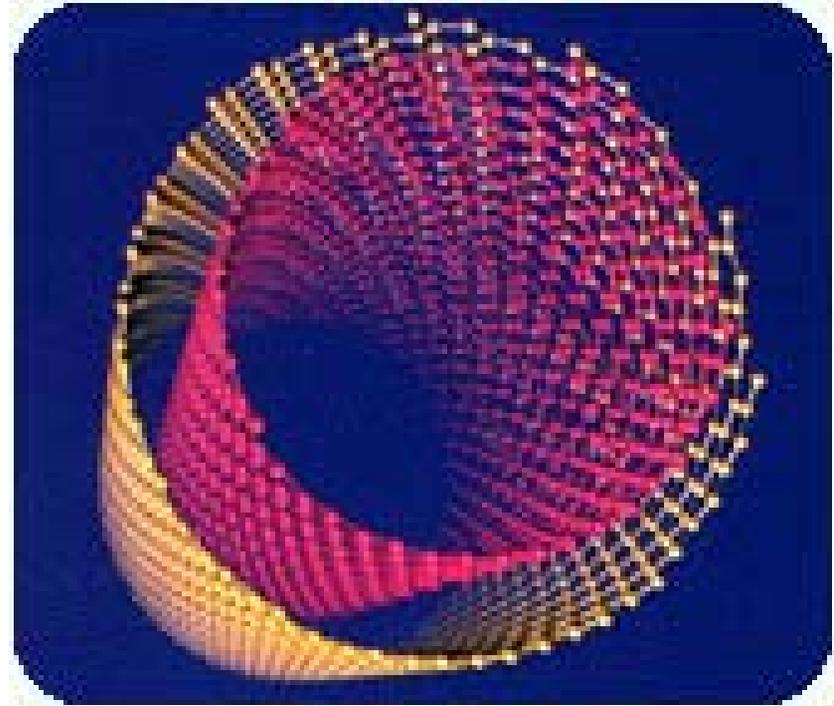
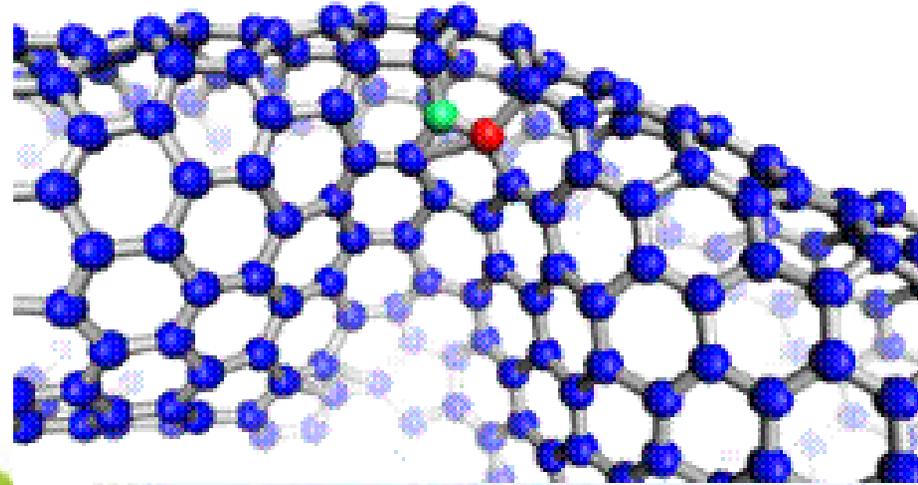
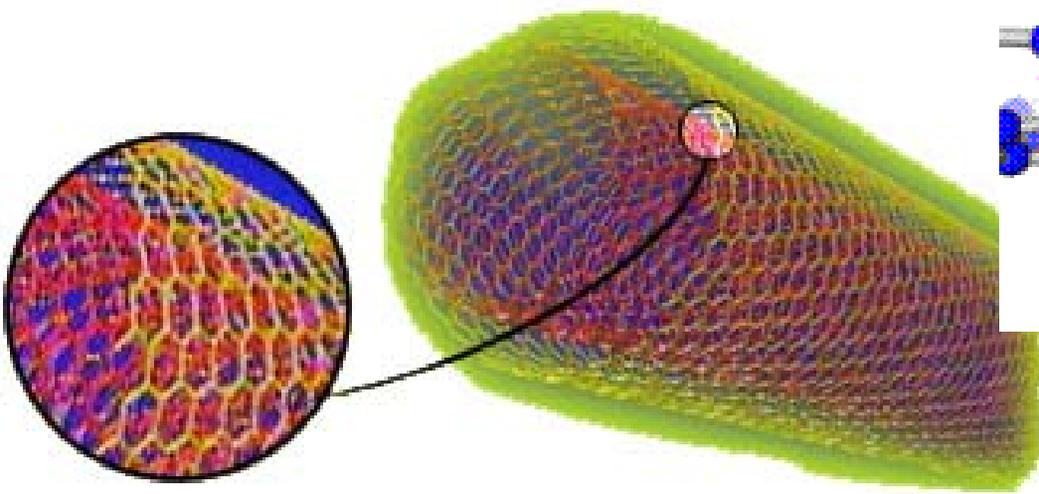
石油探查



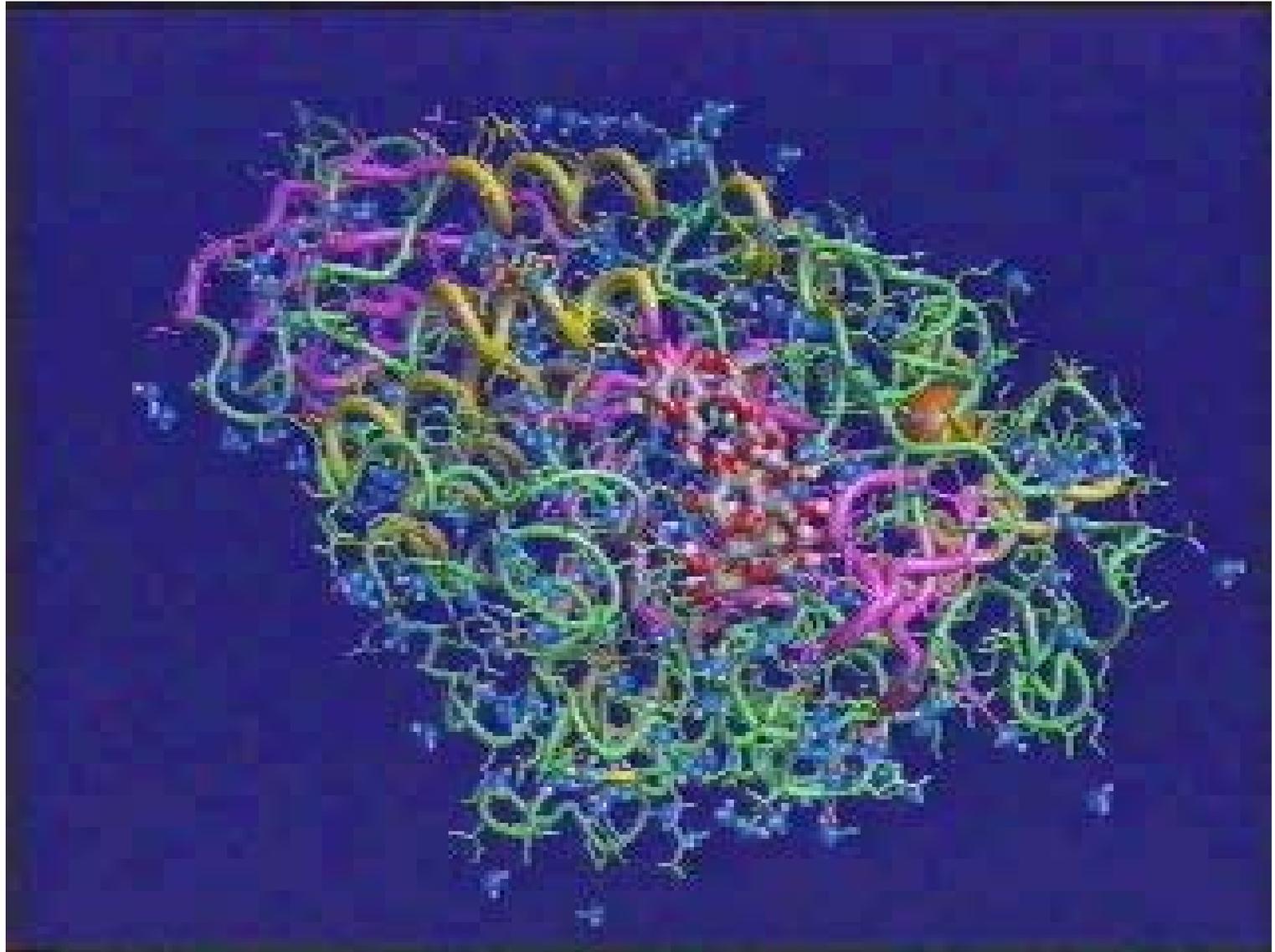
原子と分子の世界



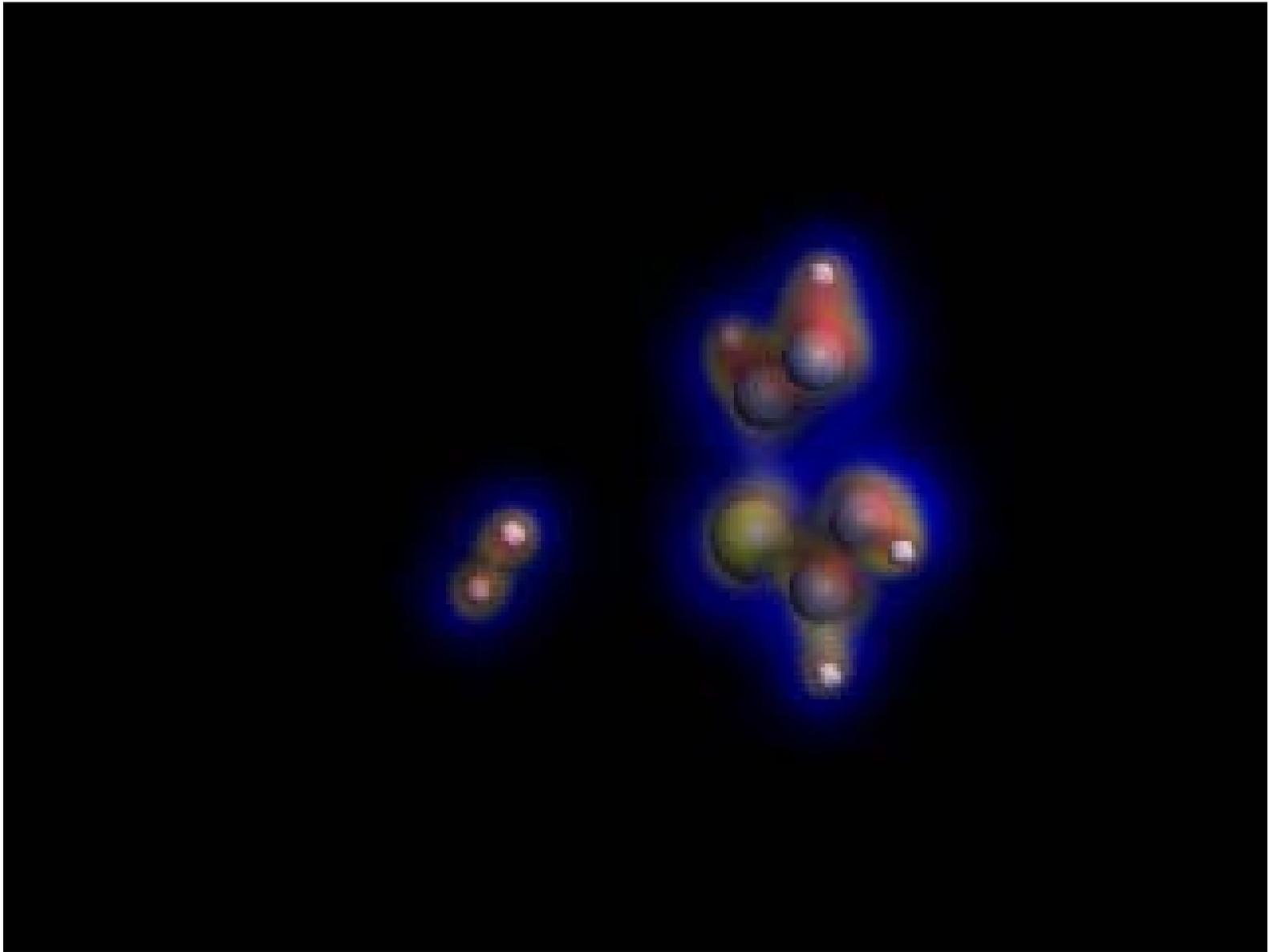
カーボンナノチューブ



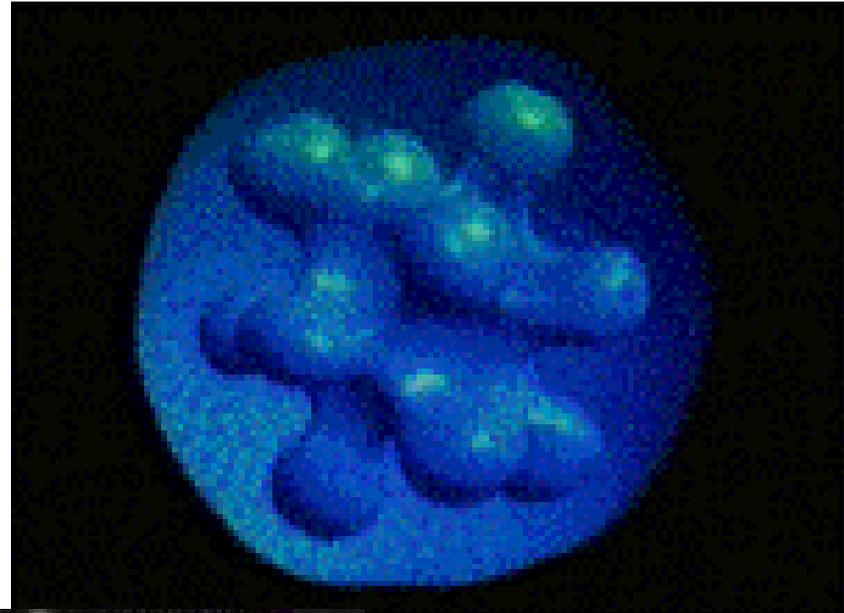
アミノ酸



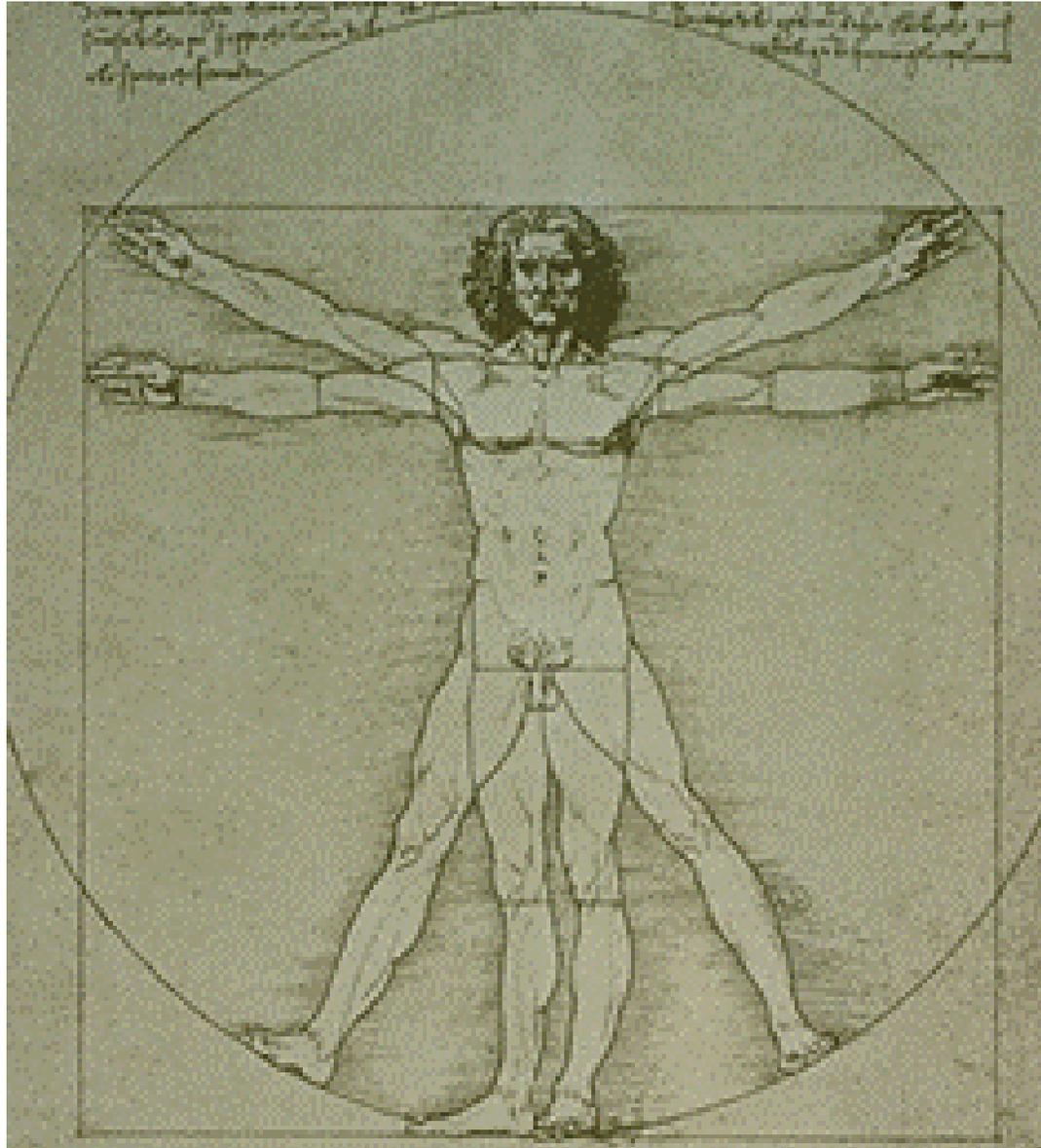
仮想顕微鏡(バーチャルマイクロスコープ) $C_4H_4S+H_2$



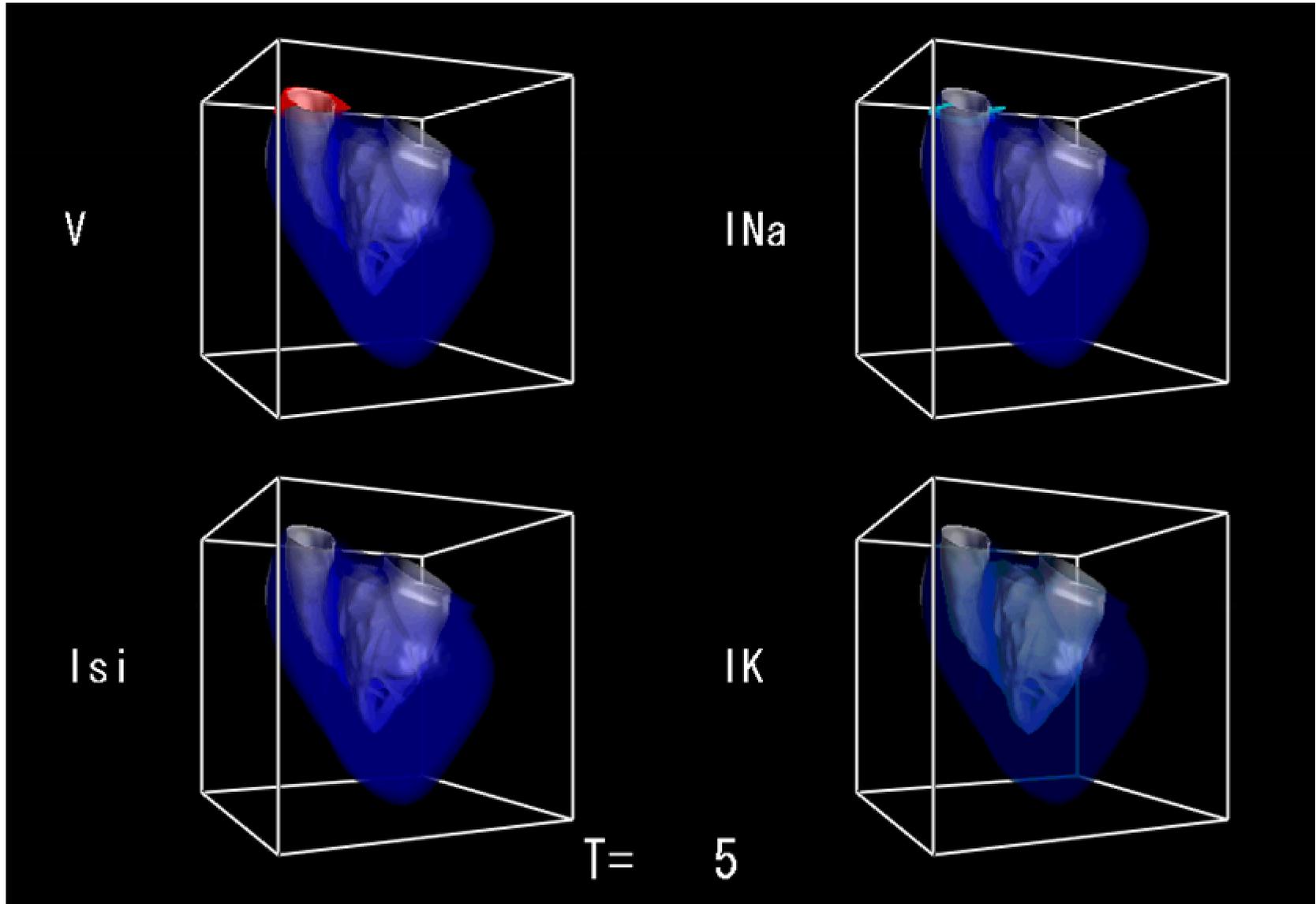
核融合



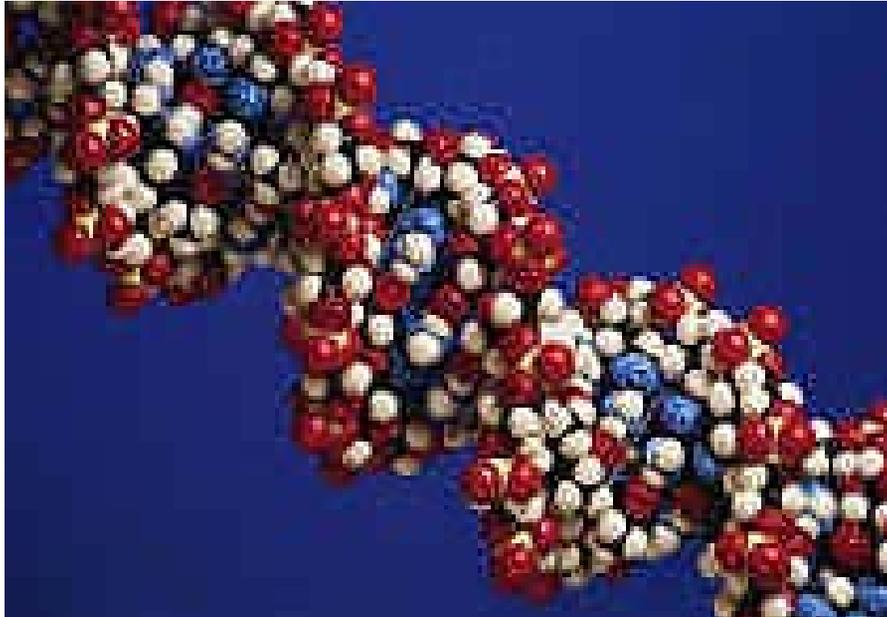
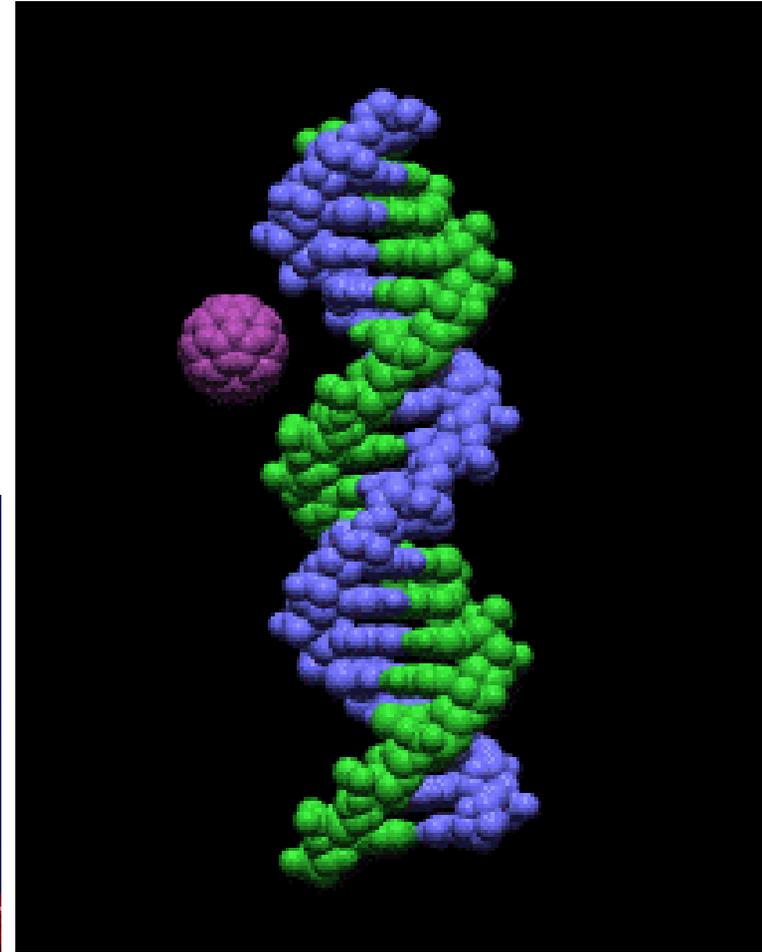
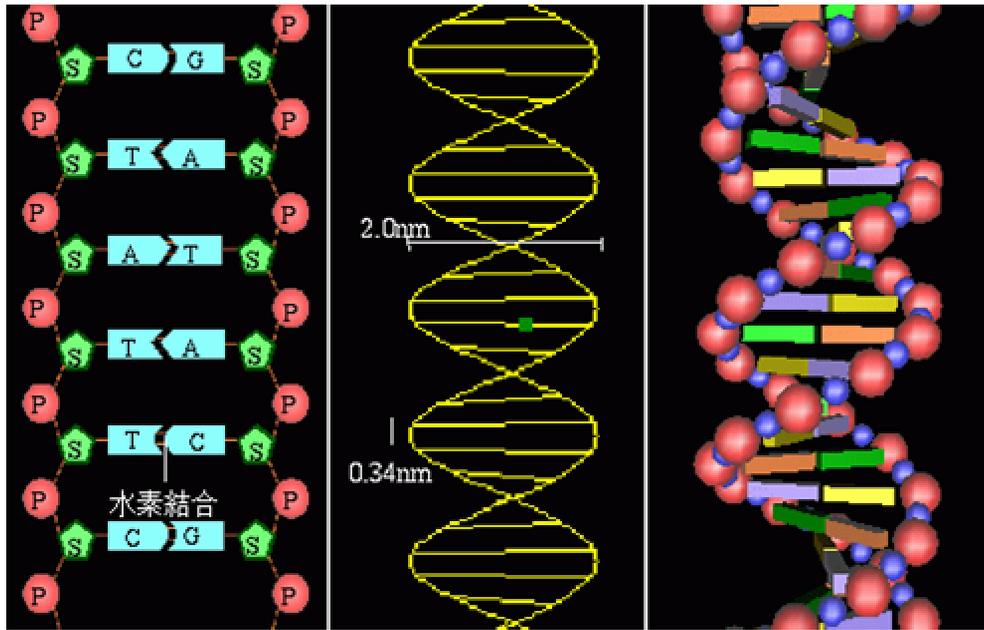
人間の身体 (医療/遺伝子)



血管と血流



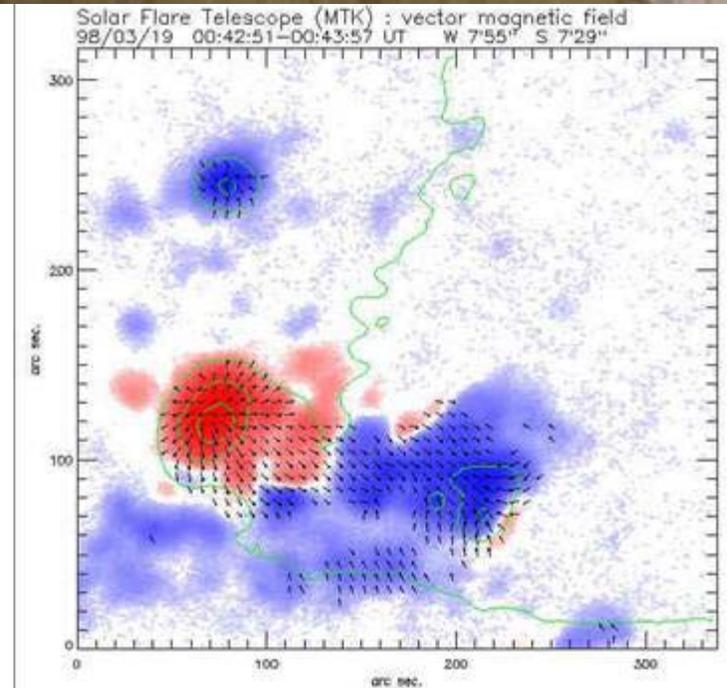
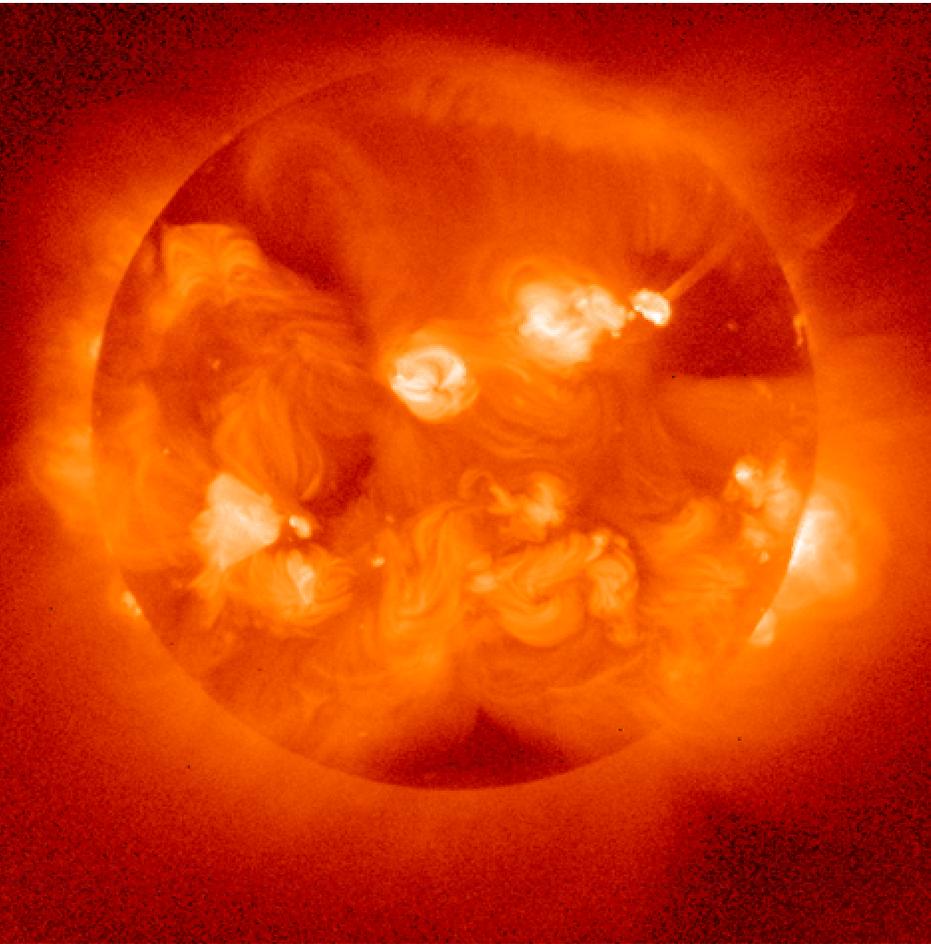
DNA (遺伝子解析)



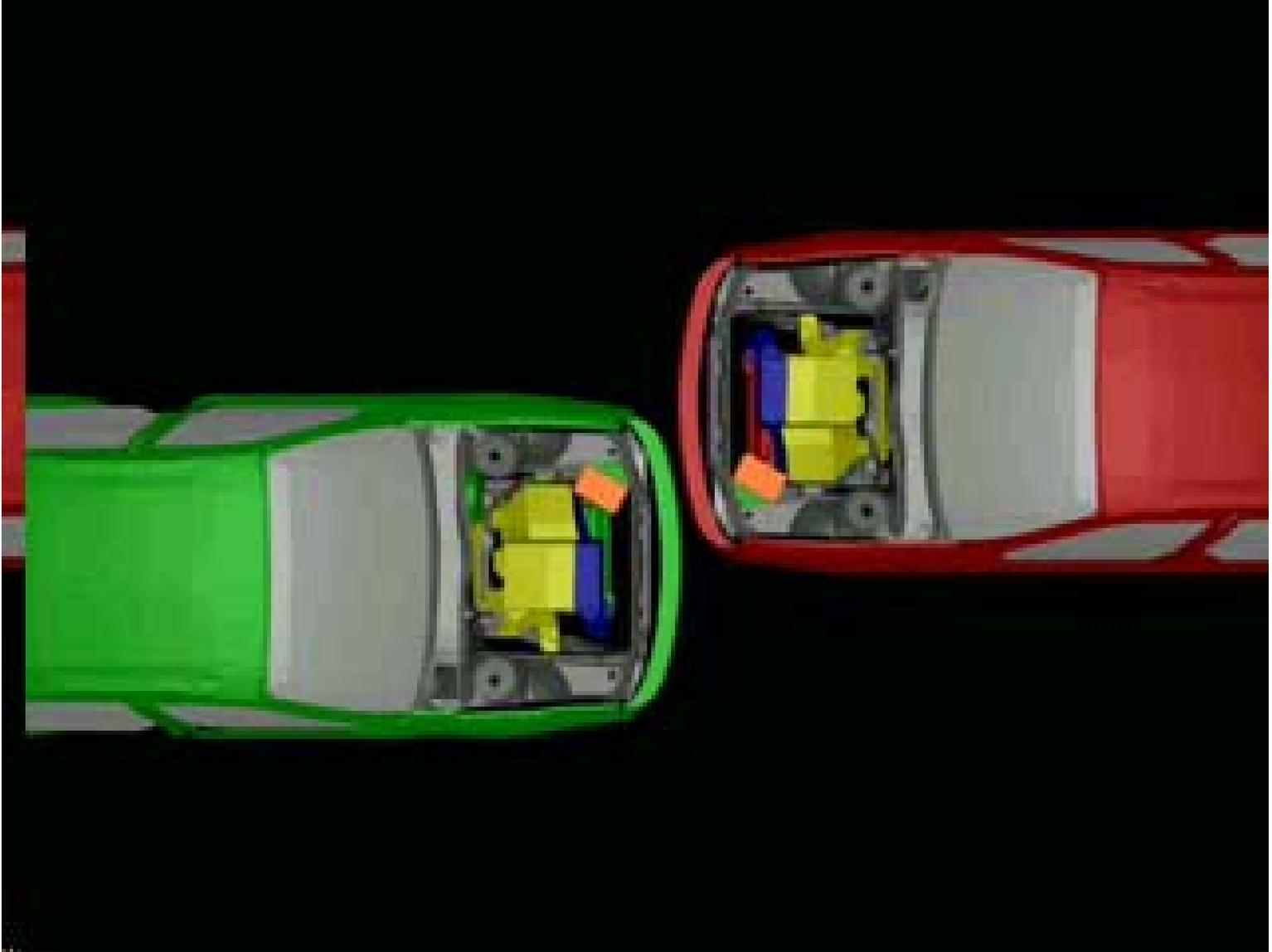
宇宙を探る



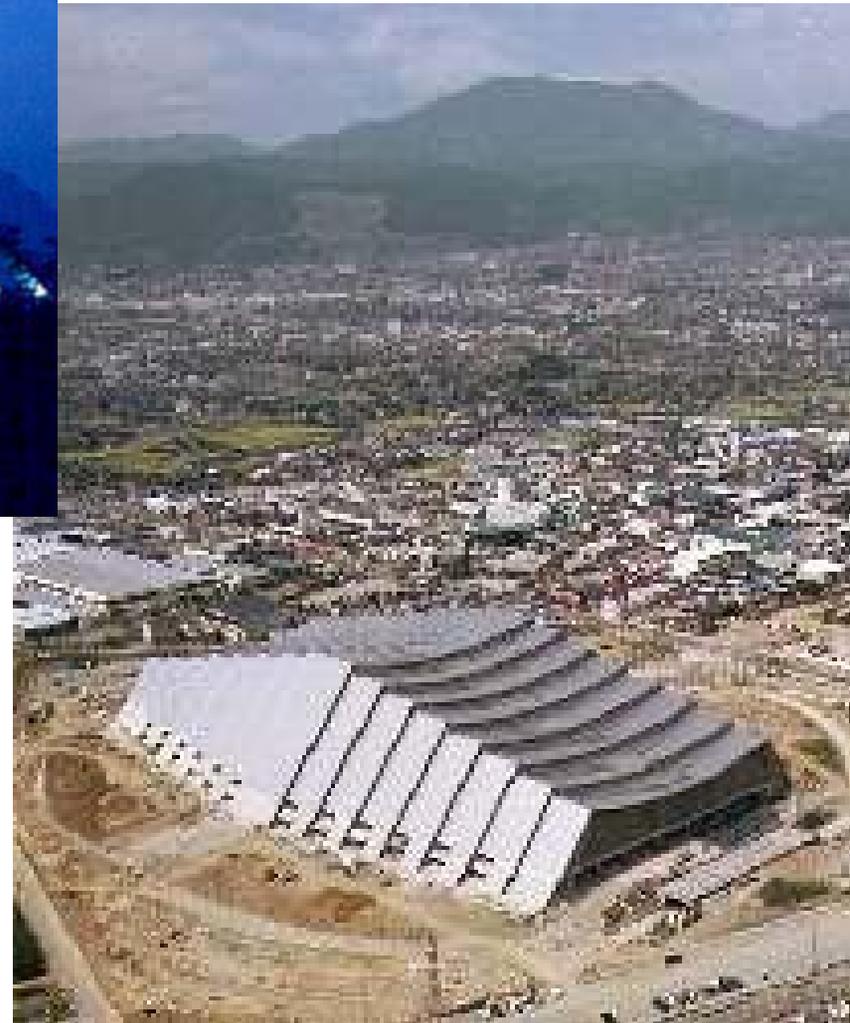
太陽黒点の爆発



工学への応用(製品開発)



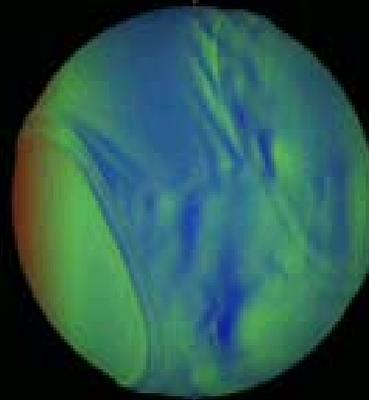
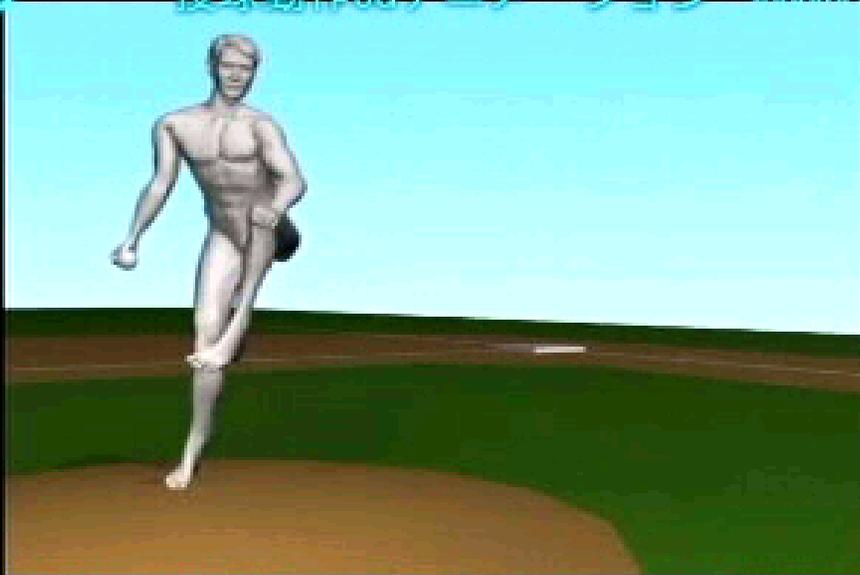
ドームの設計



スポーツ(野球)

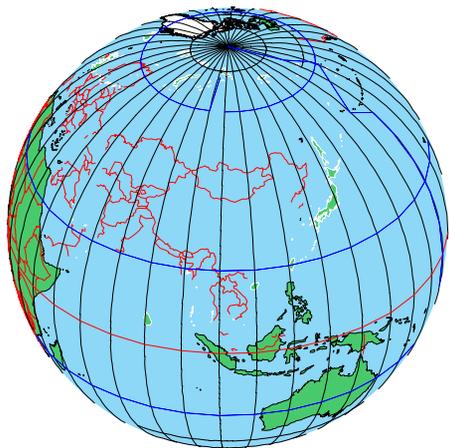


投球動作CGアニメーション 2000.01.19



Laboratories of Image Information Science and Technology

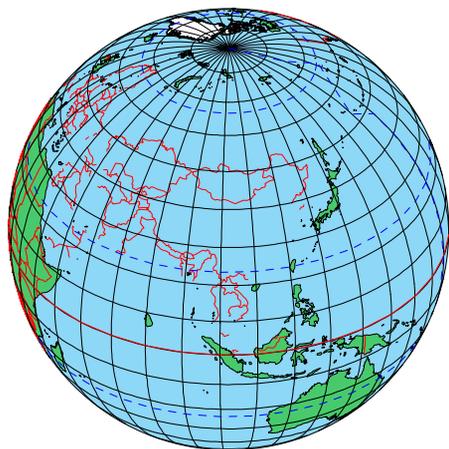
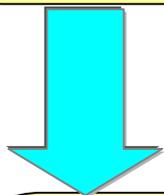
膨大な計算の例 - 地球温暖化シミュレーション(125年間) -



計算メッシュ: 300Km × 300Km × 18レベル(高さ)
1.5 × 10⁵ 格子点

演算数: 約1.5 × 10¹⁶

スーパーコンピュータ(100GFLOPS): 約5日
パソコン: 約4年!



計算メッシュ: 50Km × 50Km × 50レベル(高さ)
1.5 × 10⁷ 格子点

演算数: 約1.5 × 10¹⁸

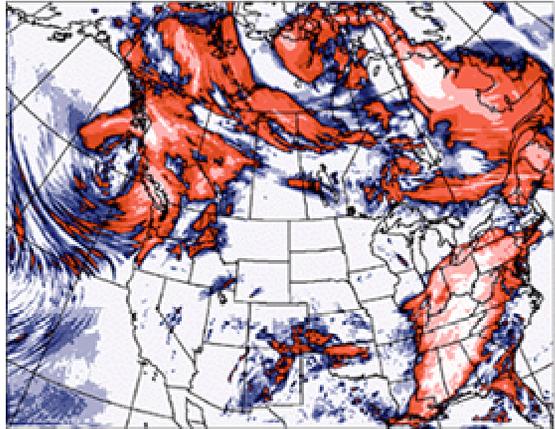
スーパーコンピュータ(100GFLOPS): 約1.4年
(1TFLOPS): 約50日
(10TFLOPS): 約5日

パソコン: 約400年!

地球シミュレータ計画(文部科学省)

スーパーコンピュータによるシミュレーションは:

- 見えないものを見る
- 出来ない実験を代替
- 未来を予測する

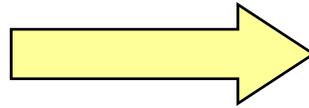


(North American 24hours Precipitation)

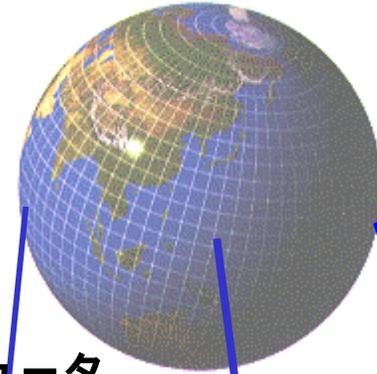


NEC SX-6/8A

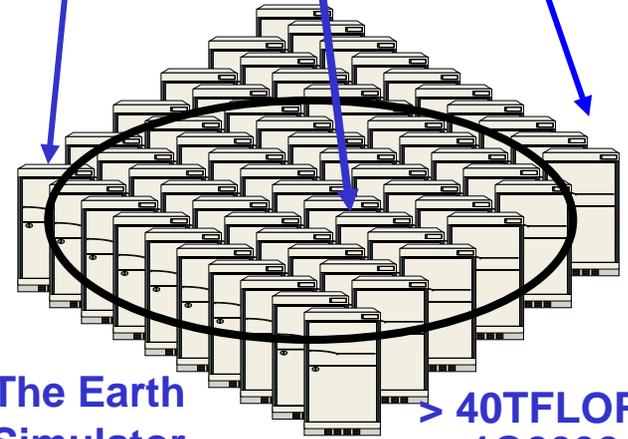
しかし、通常のスーパーコンピュータでは、地球環境問題の本格的解析には性能不足



Power
x 640



Each CPU executes their share of computation



The Earth Simulator

> 40TFLOPS
1Q2002

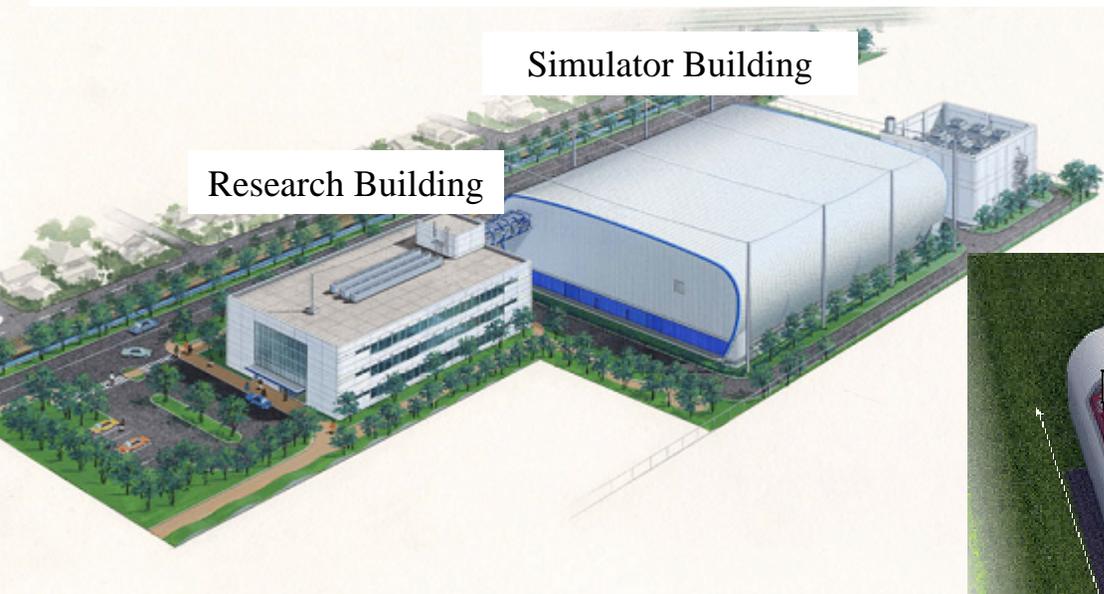
文部科学省、原研、宇宙開発事業団、海洋科学センターのプロジェクト

地球シミュレータ完成、稼働開始-2002年3月

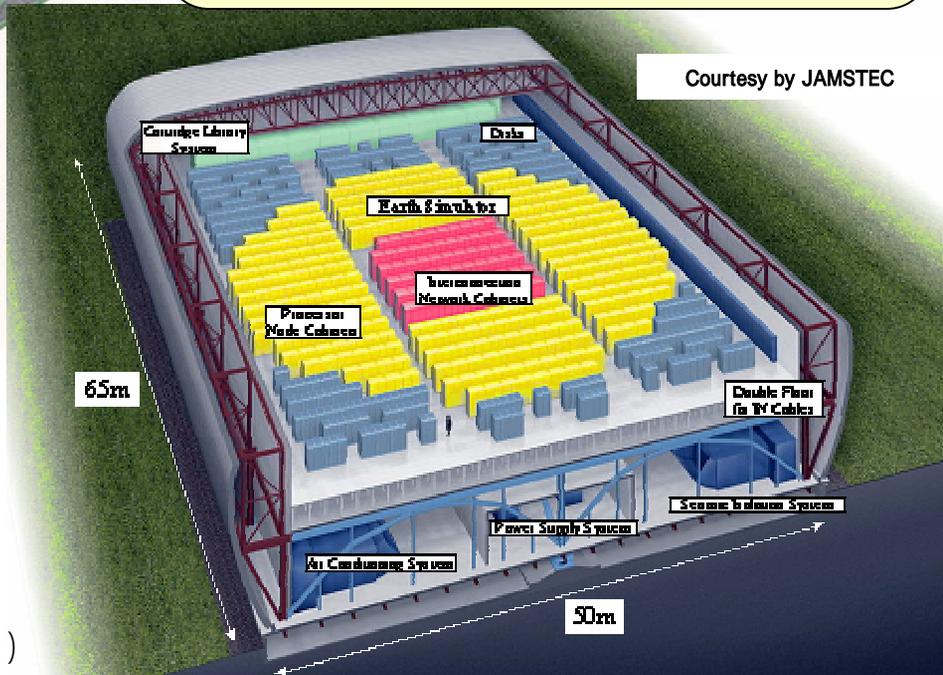
- 世界最速 (これまでの最速 = 米国パ'モア研究所のIBM機の × 5 倍の性能)
場所:横浜市

Cc

Earth Simulator Facilities



**New Linpack Record -
35.8TFLOPS (5 X previous
#1 ASCI White = 7.2TF)**



- 5120 プロセッサ
- 640 ノード
- 40 TFLOPS ピーク性能

Japanese Computer Is World's Fastest, as U.S. Falls Back

By JOHN MARKOFF



AN FRANCISCO, April 19 · A Japanese laboratory has built the world's fastest computer, a machine so powerful that it matches the raw processing power of the 20 fastest American computers combined and far

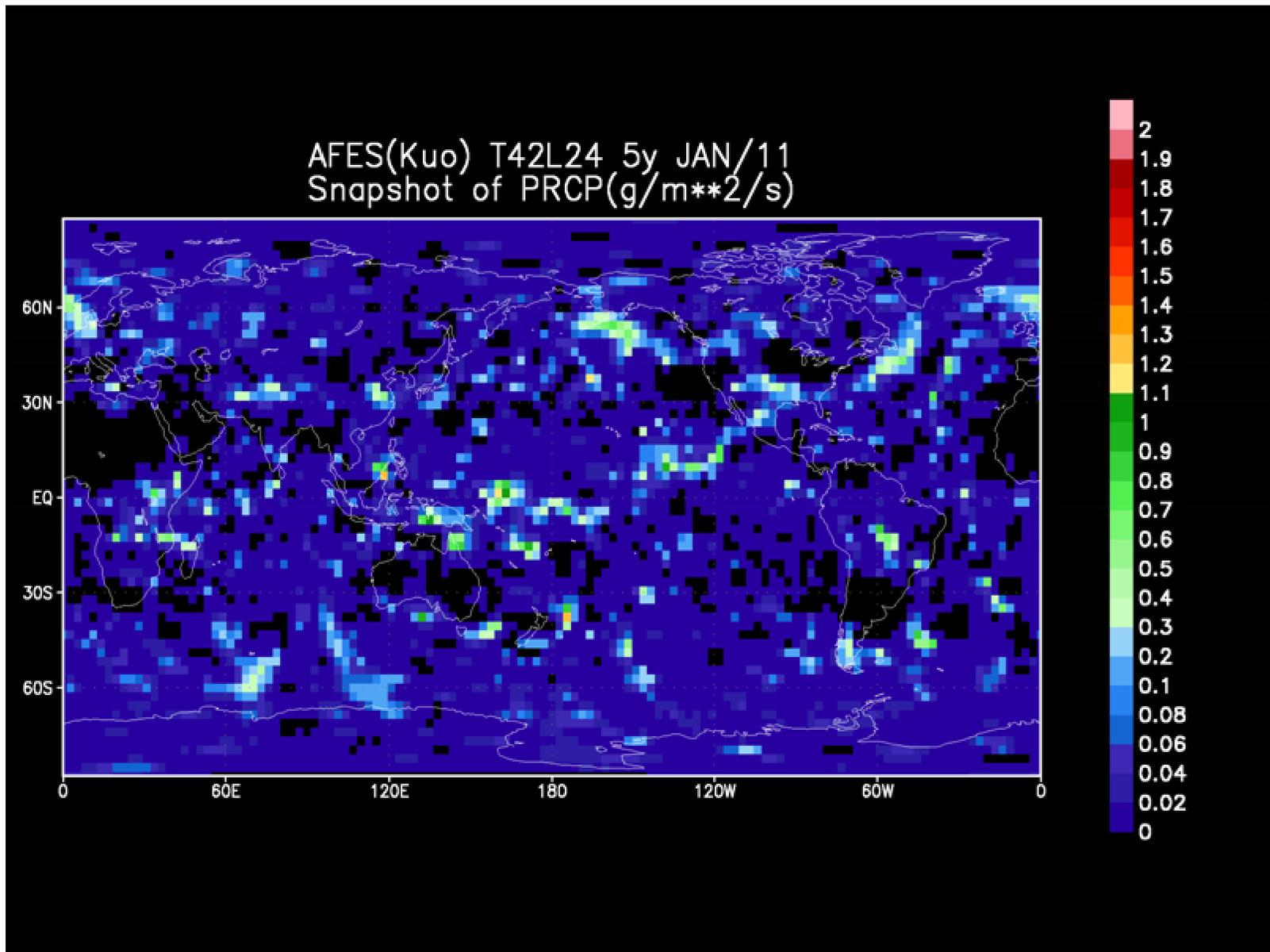
outstrips the previous leader, an IBM-built machine.

The achievement, which was reported today by an American scientist who tracks the performance of the world's most powerful computers, is evidence that a technology race that most American engineers thought they were winning handily is far from over. American companies have built the fastest computers for most of the last decade.

The accomplishment is also a vivid statement of contrasting scientific and technology priorities in the United States and Japan. The Japanese machine was built to analyze climate change, including global warming, as well as weather and earthquake patterns. By contrast, the United States has predominantly focused its efforts on building powerful computers for simulating weapons, while its efforts have lagged in scientific areas like climate modeling.

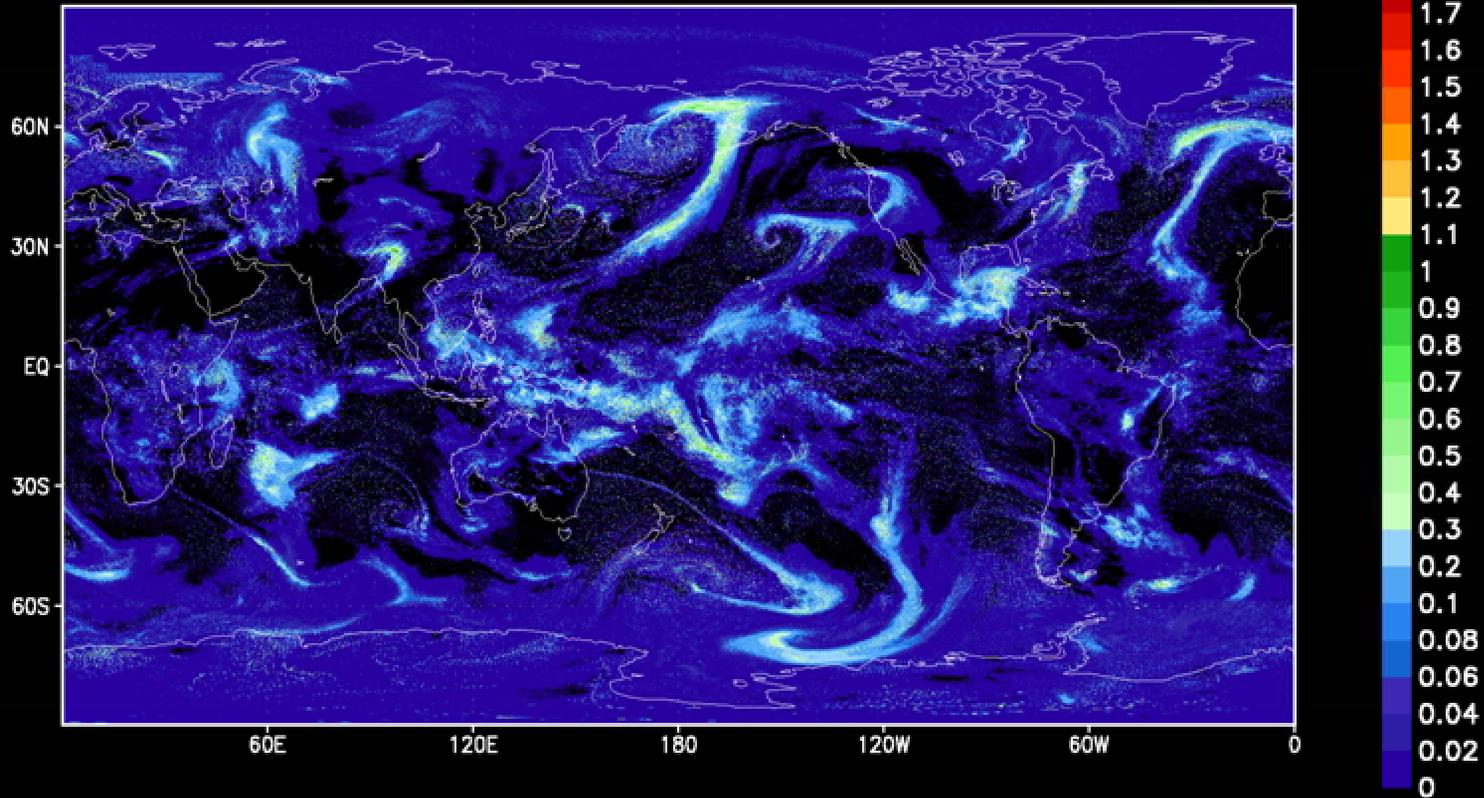
- ・ 日本のコンピュータが世界最高速、米国は後塵
- ・ 米国の技術が追い抜かれた証左
- ・ 旧ソ連のスプートニク以来の驚き
- ・ 世界の気象研究に非常に大きな貢献

降水量分布(312km,T42L24)



降水量分布 (10.4km,T1279L24)

AFES(Kuo) T1279L96 5y JAN/07 12Z hour
Snapshot of Precipitation PRCP(g/m**2/s)



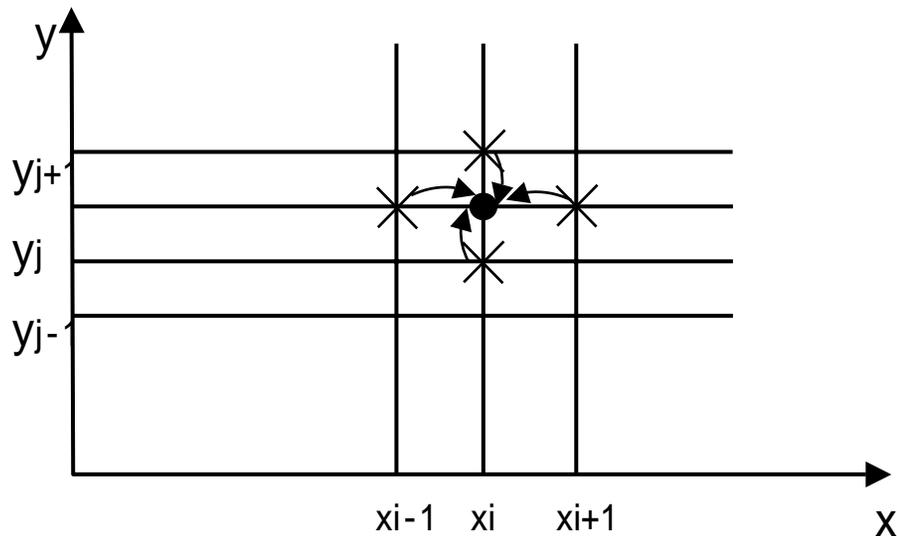
どんなプログラムを作るのか？
- スーパーコンピュータのプログラム -

科学技術計算の典型例 (偏微分方程式)

・ラプラス方程式: $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ 応用例: 流体解析

・有限差分法 (逐次近似法)

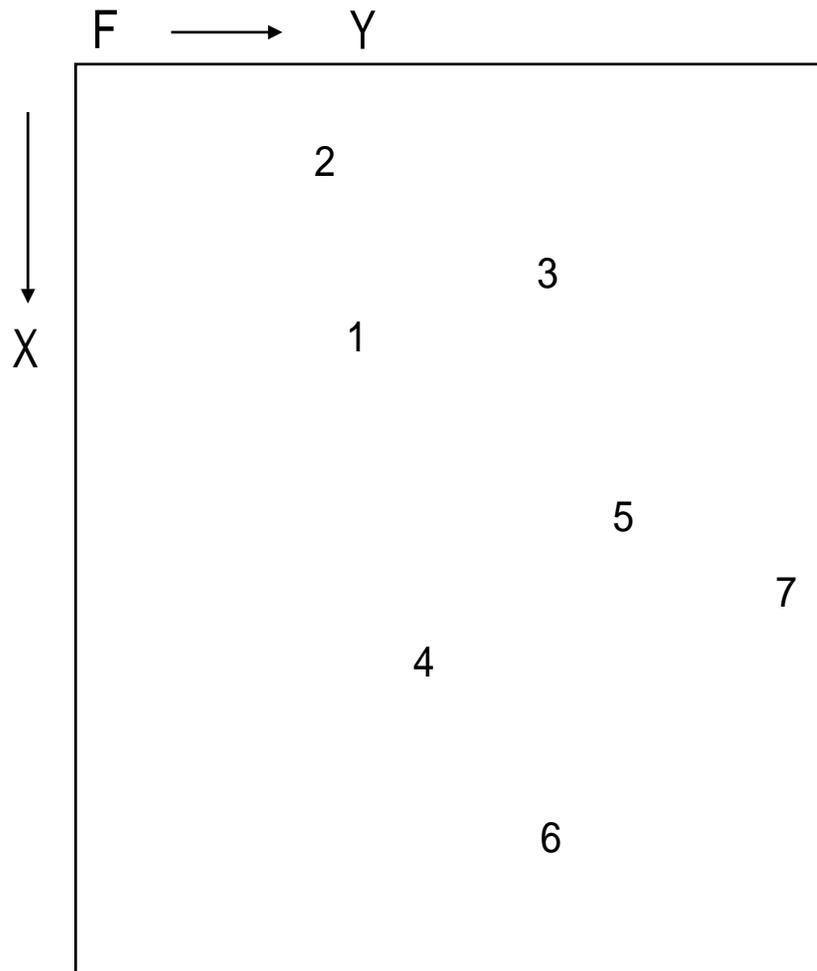
$$u_{i,j}^{(k+1)} = (u_{i+1,j}^{(k)} + u_{i-1,j}^{(k)} + u_{i,j+1}^{(k)} + u_{i,j-1}^{(k)}) / 4$$



・プログラム例

```
DO J = 2,N
  DO I = 2,N
    V(I,J) = 0.25 * (U(I+1,J)+U(I-1,J)+U(I,J+1)+U(I,J-1))
  ENDDO
ENDDO
```

粒子シミュレーション(モンテカルロシミュレーション)

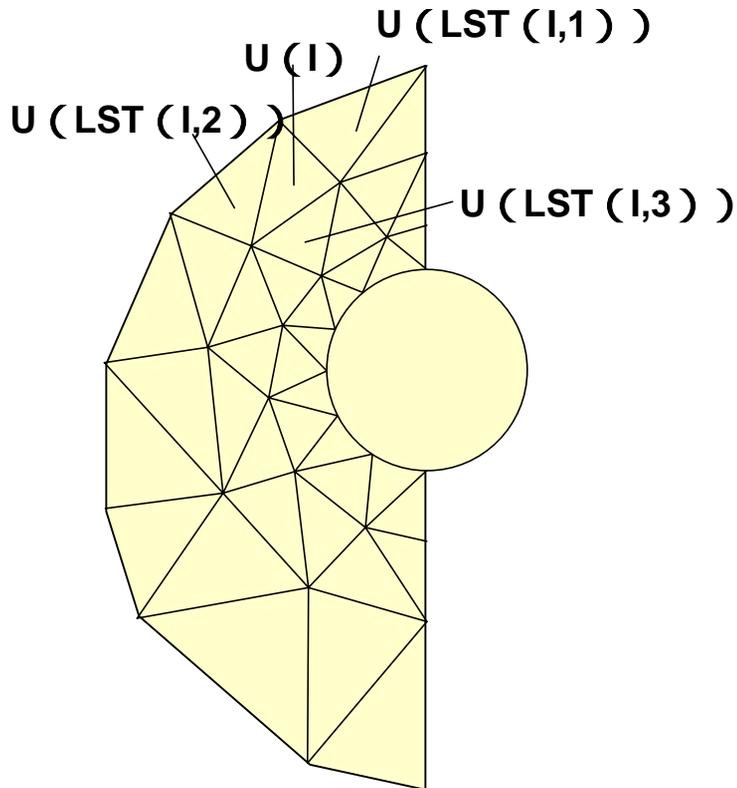


応用例: プラズマ物理・天体

$$P(I)=F(IX(I), IY(I))+B(I)$$

非構造格子

応用例：構造解析・衝突解析

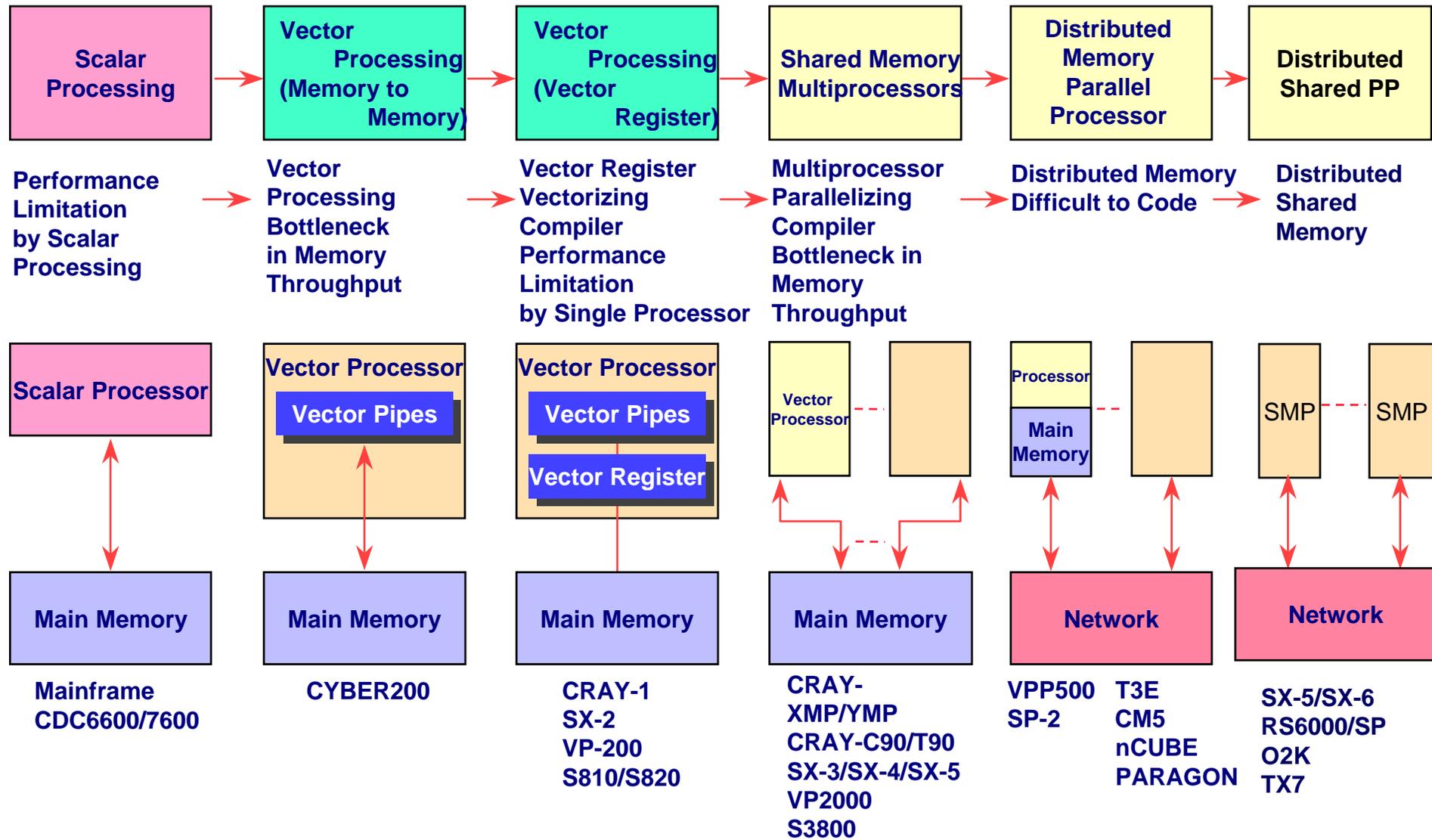


$$\begin{aligned} UU(I) &= COEF(I,1) * U(I) \\ &+ COEF(I,2) * U(LST(I,1)) \\ &+ COEF(I,3) * U(LST(I,2)) \\ &+ COEF(I,4) * U(LST(I,3)) \end{aligned}$$

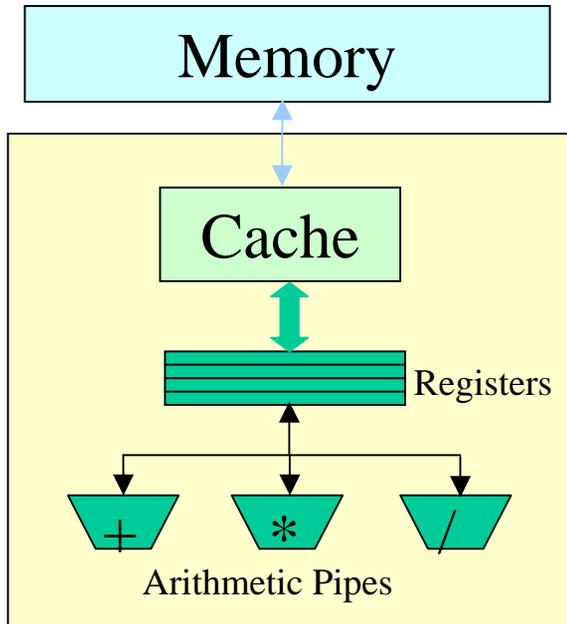
どうすれば速く計算できるか？

- スーパーコンピュータの仕組み -

スーパーコンピュータの構成

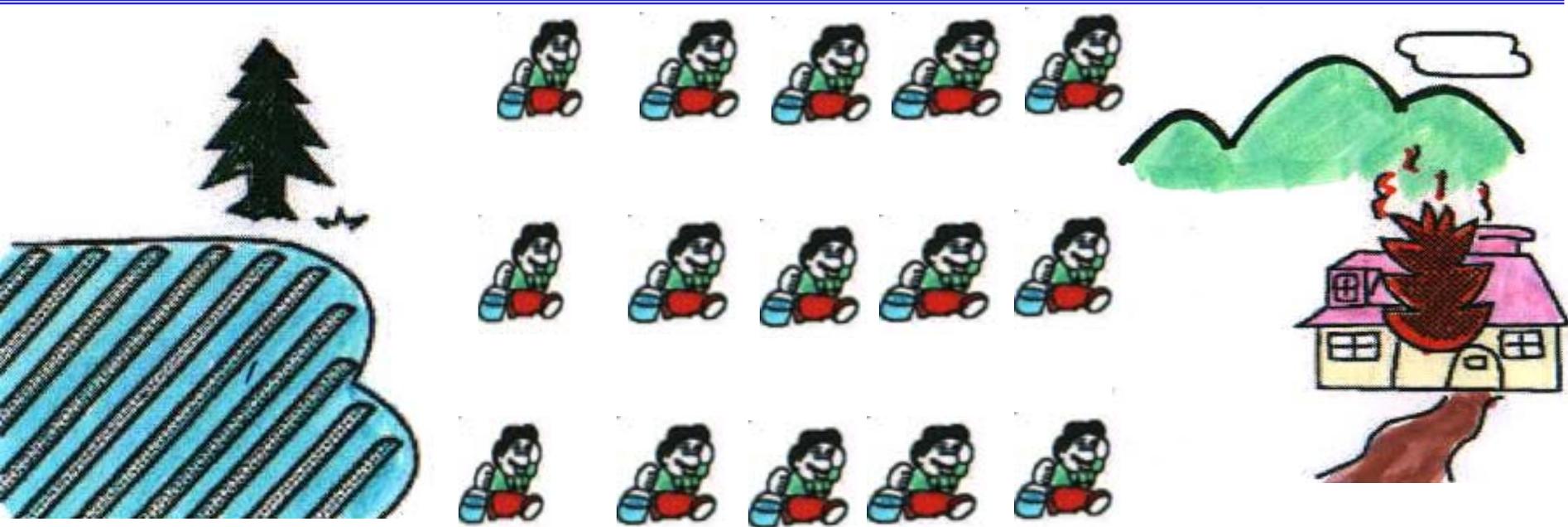


スカラ処理

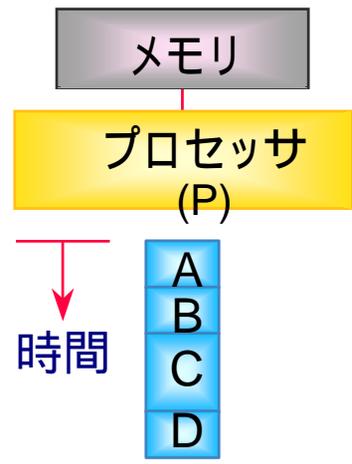


μ -Processor

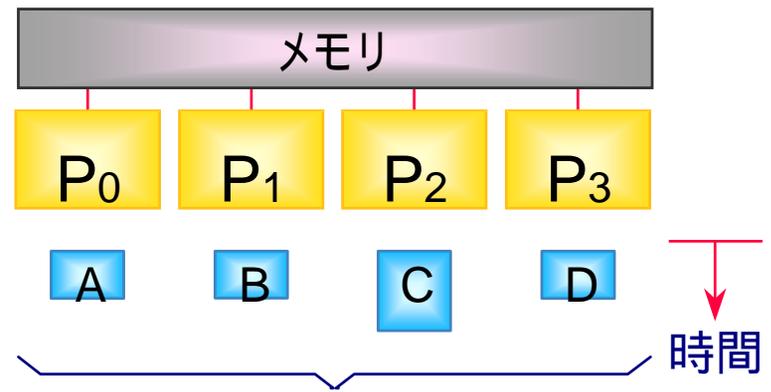
並列処理



一台のプロセッサ(CPU)



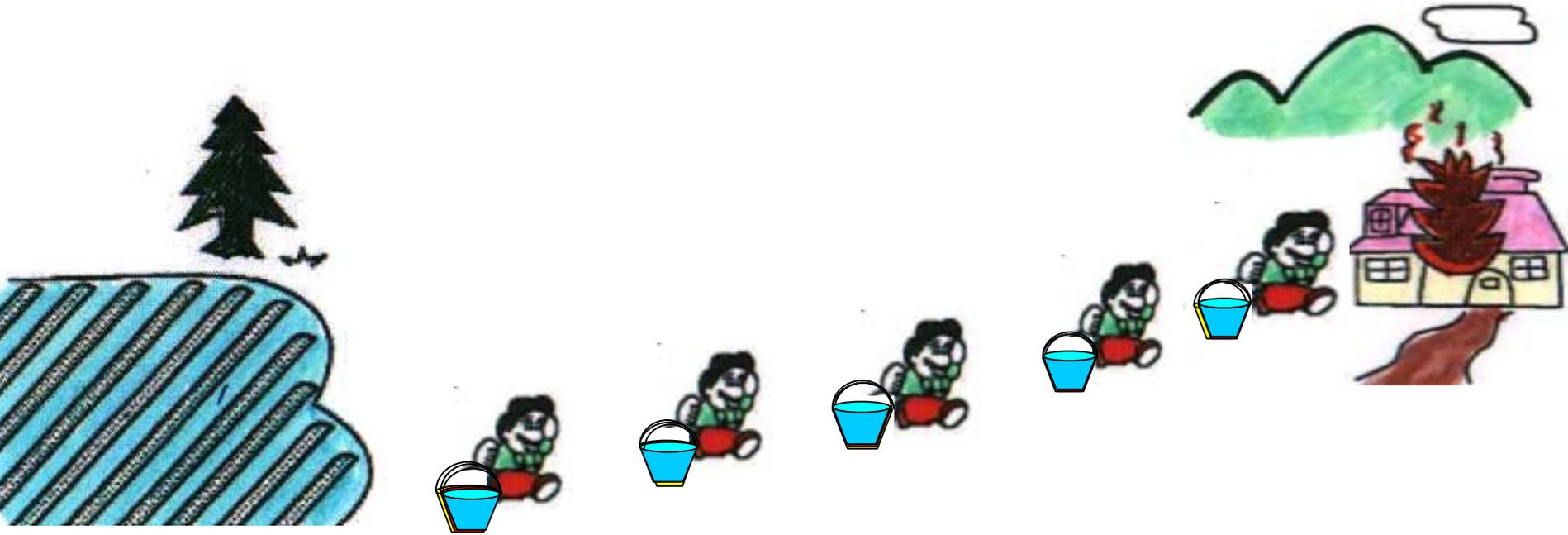
複数のプロセッサ



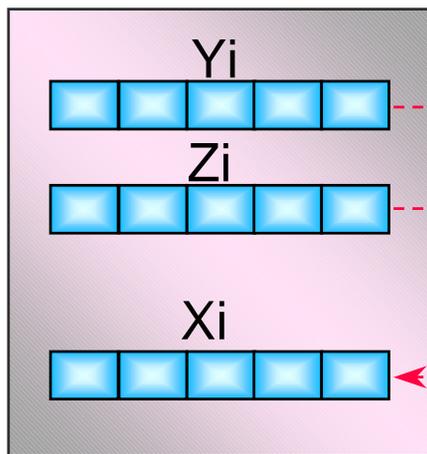
複数の処理部分に分割

プログラム

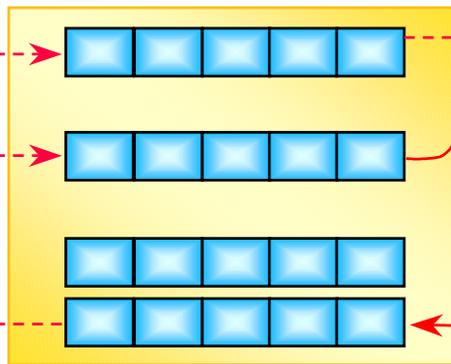
パイプライン処理



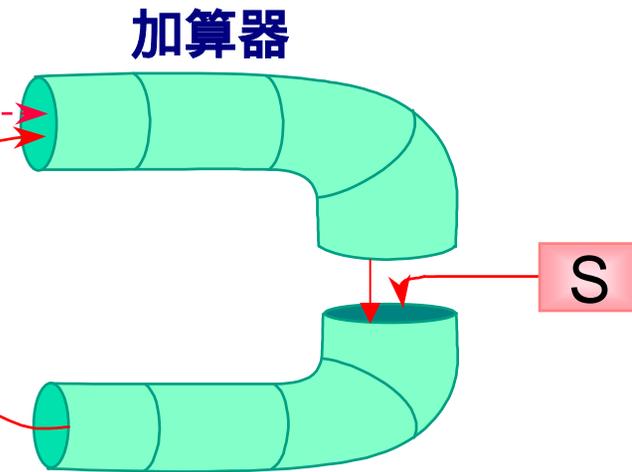
$$X_i = (Y_i + Z_i) * S$$



メモリ



レジスタ



加算器

乗算器

高速化には何が重要か？

・各CPUの計算速度

- 演算の高速化：クロック周波数の高速化
- 回路の高速化：高集積化・高密度実装

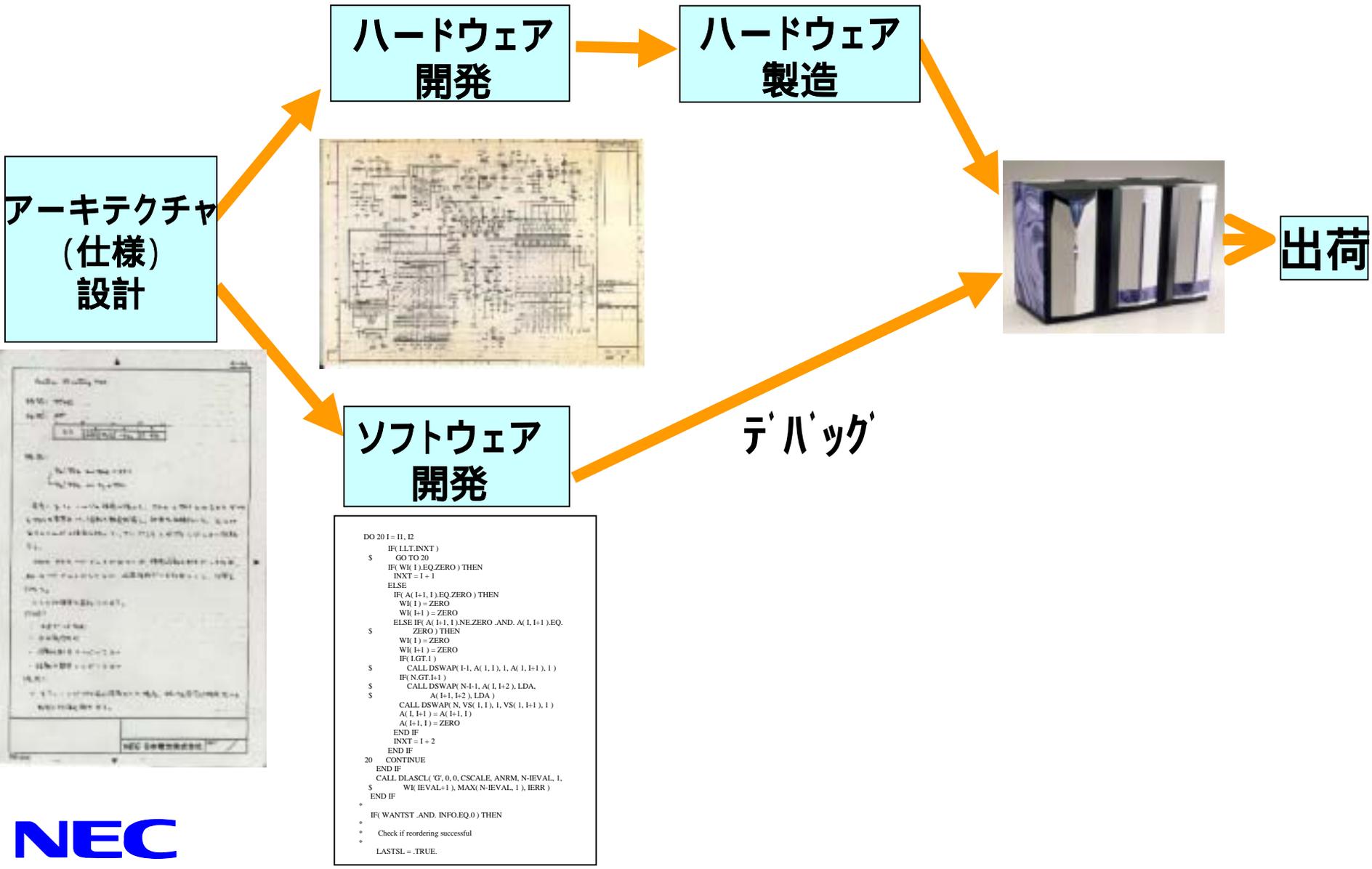
・各CPUがすべて同じ速度で計算（怠け者を作らない）

- 同期と負荷の均等化

・データの取り出し・格納

- メモリとCPUのデータ転送速度（バンド幅）
- データ幅の増大・アクセス時間の短縮

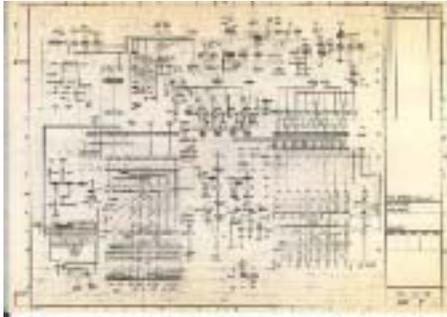
スーパーコンピュータの開発



アーキテクチャ
(仕様)
設計

ハードウェア
開発

ハードウェア
製造



出荷

ソフトウェア
開発

デバッグ



```
DO 20 I= 11, 12
  IF (L.T.INXT)
  GO TO 20
  IF (W(I).EQ.ZERO) THEN
    INXT = I + 1
  ELSE
    IF (A(I+1, I).EQ.ZERO) THEN
      W(I) = ZERO
      W(I+1) = ZERO
    ELSE IF (A(I+1, I).NE.ZERO .AND. A(I, I+1).EQ.
      ZERO) THEN
      W(I) = ZERO
      W(I+1) = ZERO
      IF (LGT.I)
      CALL DSWAP(I-1, A(I, I), I, A(I, I+1), I)
      IF (N.GT.I+1)
      CALL DSWAP(N+1, A(I, I+2), LDA,
      A(I+1, I+2), LDA)
      CALL DSWAP(N, VS(I, I), I, VS(I, I+1), I)
      A(I, I+1) = A(I+1, I)
      A(I+1, I) = ZERO
      INXT = I + 2
    END IF
  END IF
20 CONTINUE
END IF
CALL DLASCL( G, 0, 0, CSCALE, ANRM, N-IEVAL, I,
S W(IEVAL-1), MAX(N-IEVAL, I), IERR)
END IF
*
IF (WANTST .AND. INFO.EQ.0) THEN
*
* Check if reordering successful
*
LASTSL = .TRUE.
```

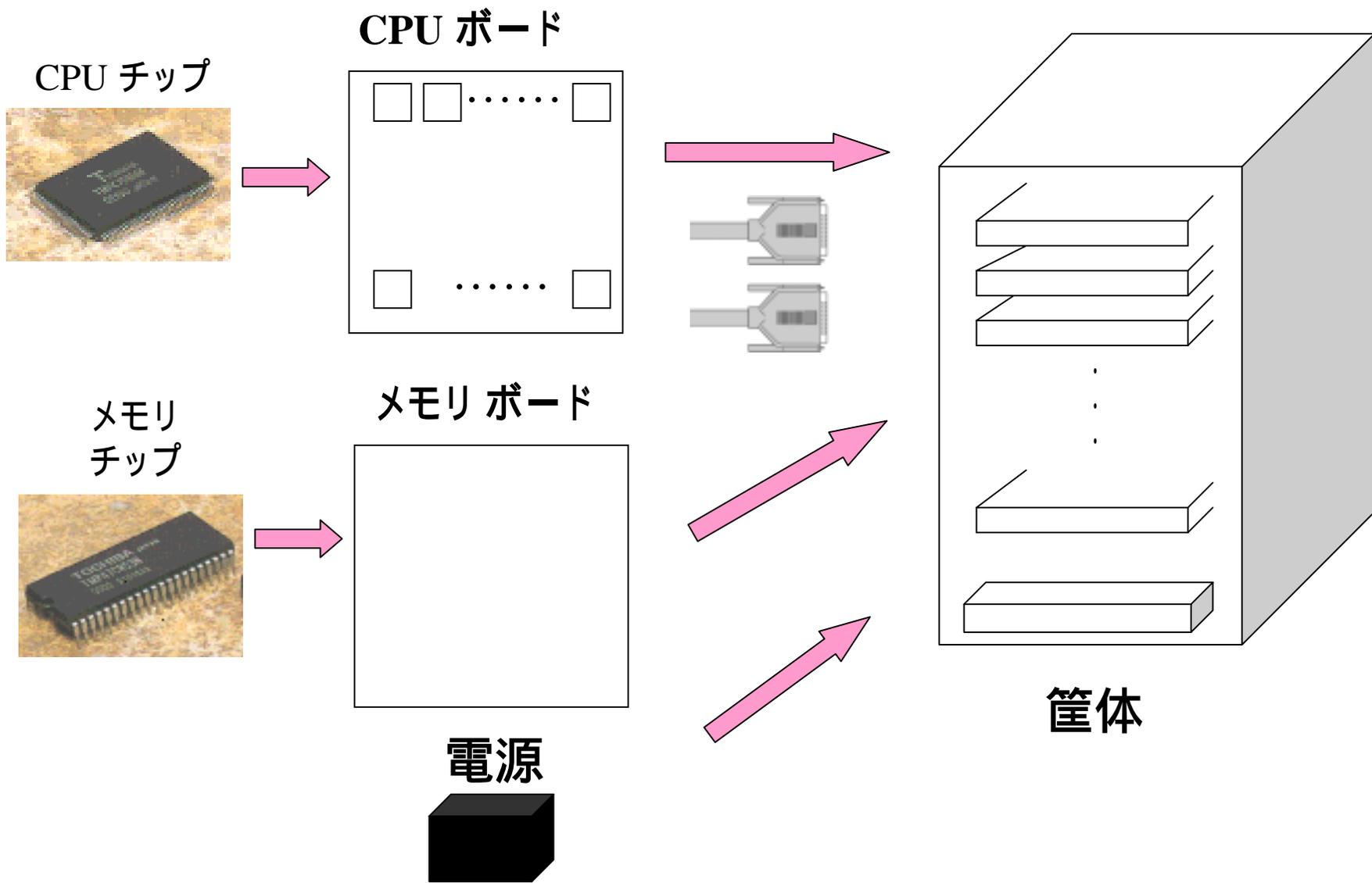
Nine Lessons Learned in the Design of CDC6600 (N.R.Lincoln)

**It's Really not as much Fun Building a Supercomputer as it is
Simply inventing one
(High Speed Computer and Algorithm Organization ,1977)**

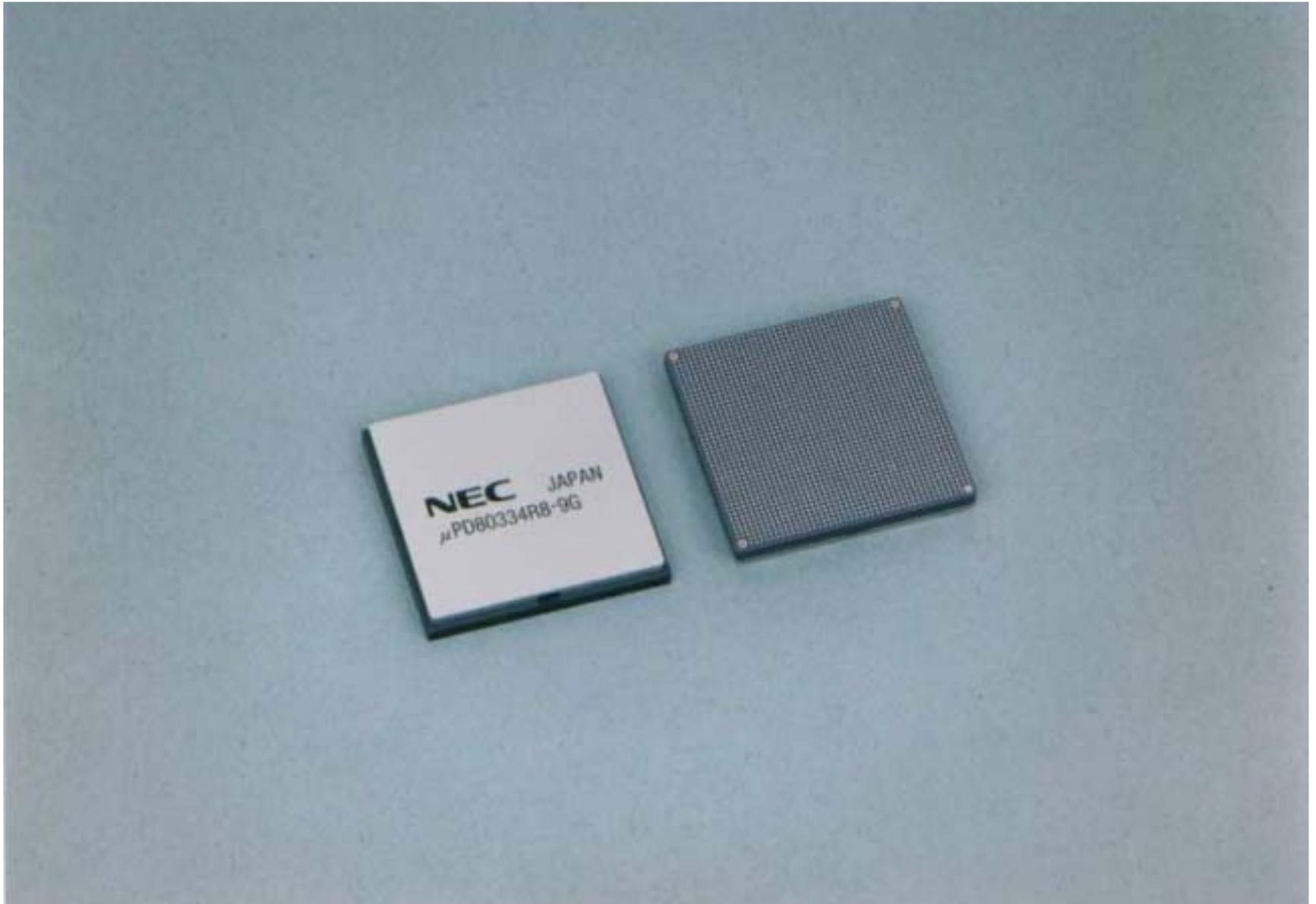
Lesson 2

- Circuit design and system architecture are only pieces in a large puzzle called “supercomputer CPU.” A major limitation on the feasibility of a given supercomputer project could well be the mechanical, power, packaging and cooling requirements of the overall electronic design.**

スーパーコンピュータを作る



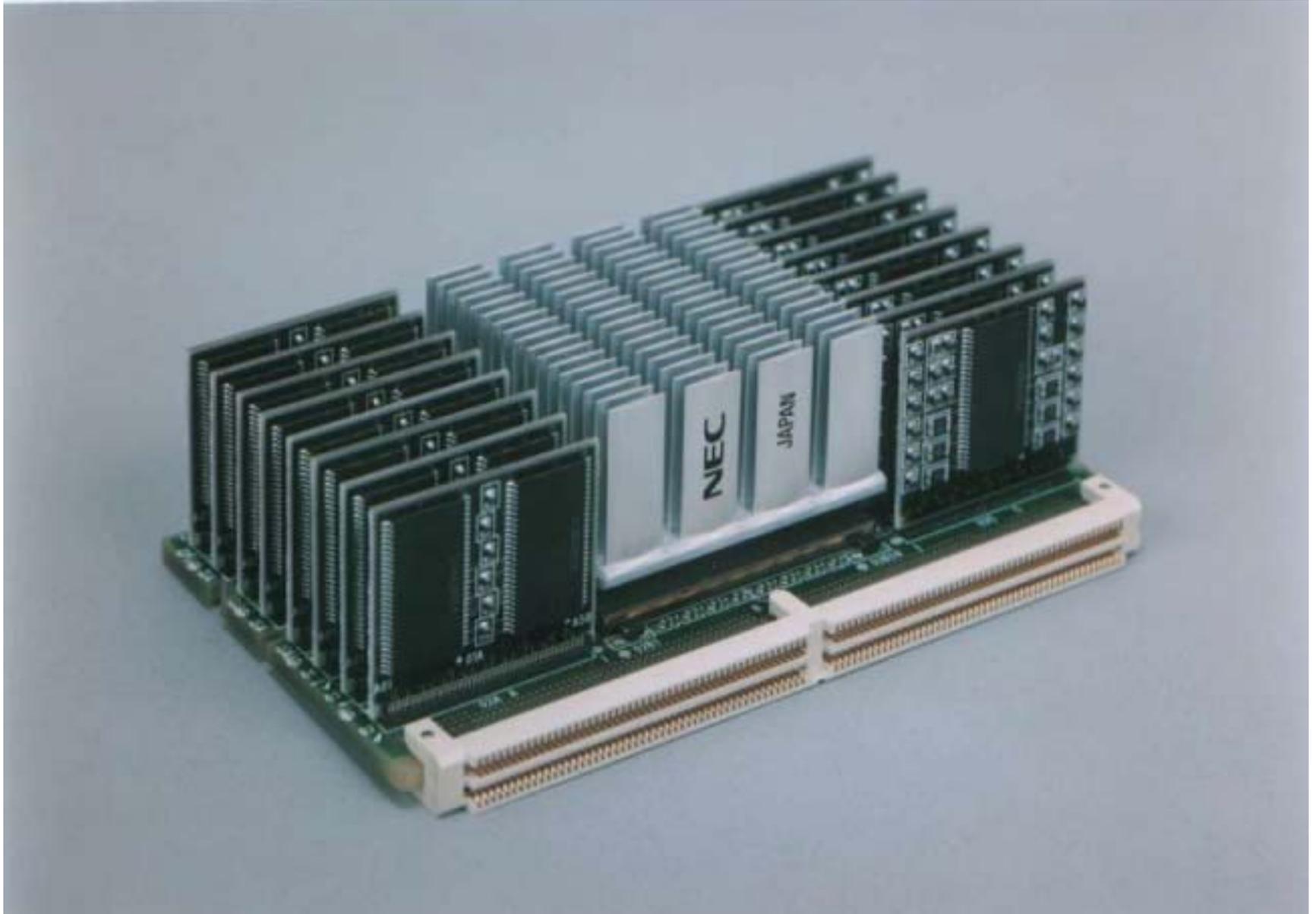
CPU チップ



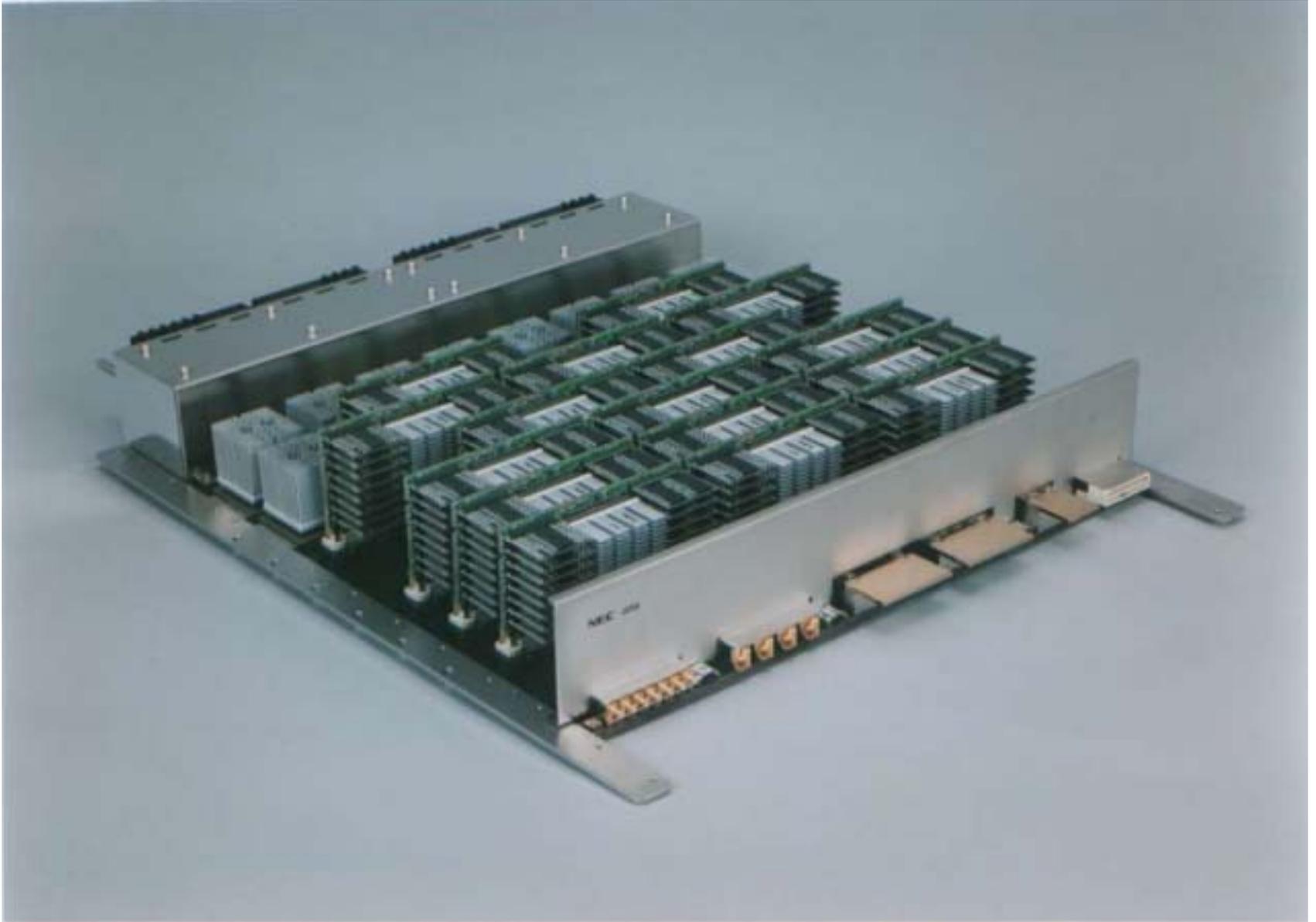
CPU ボード



メモリモジュール



メモリボード



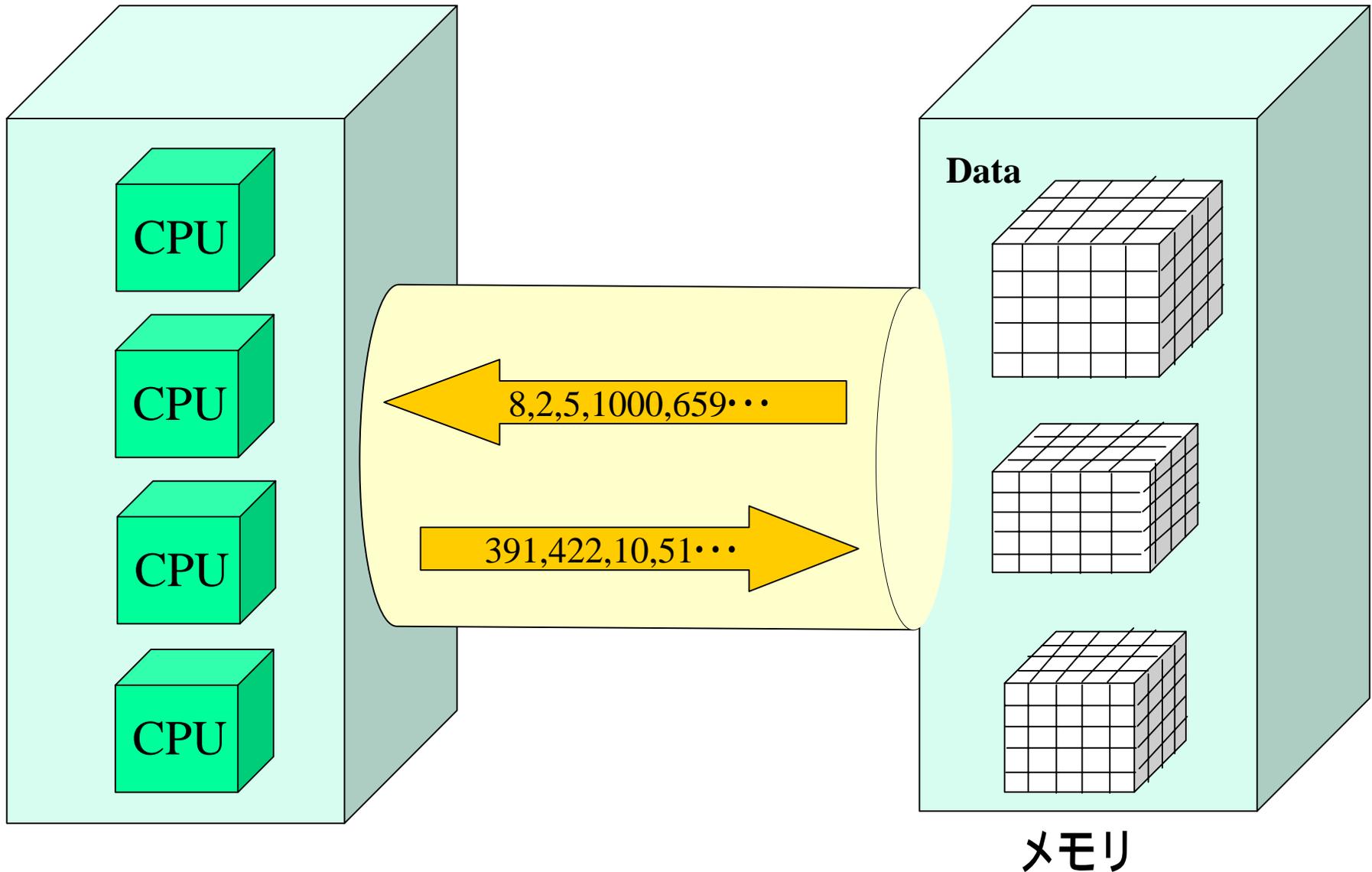
筐体



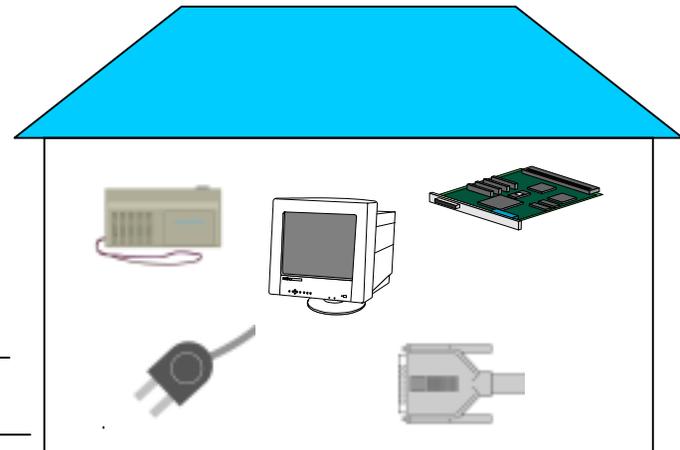
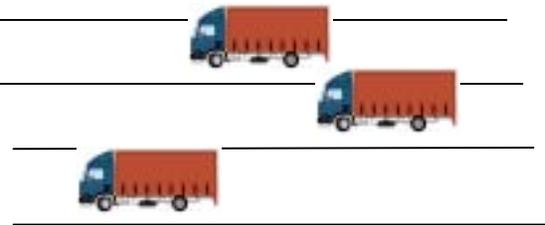
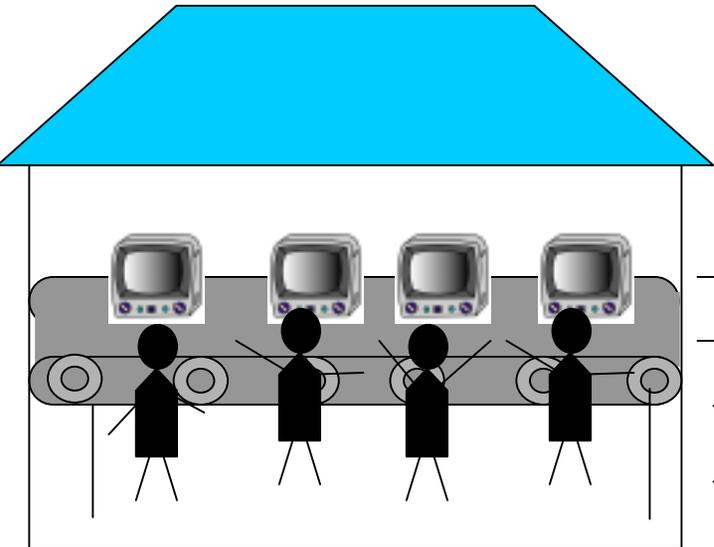
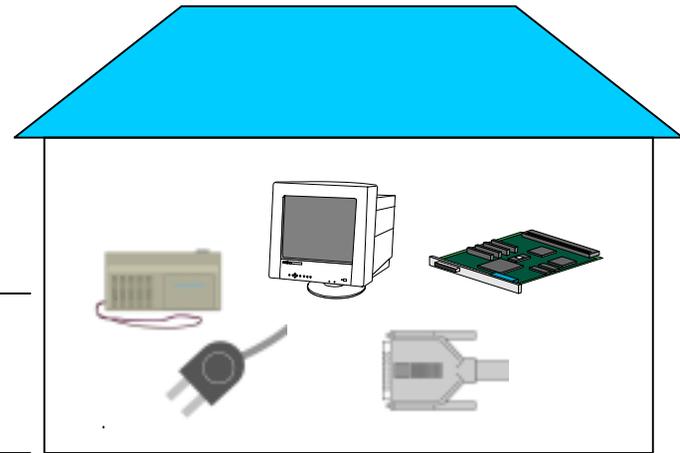
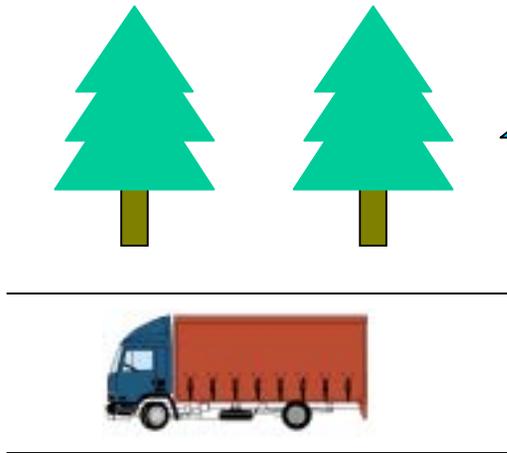
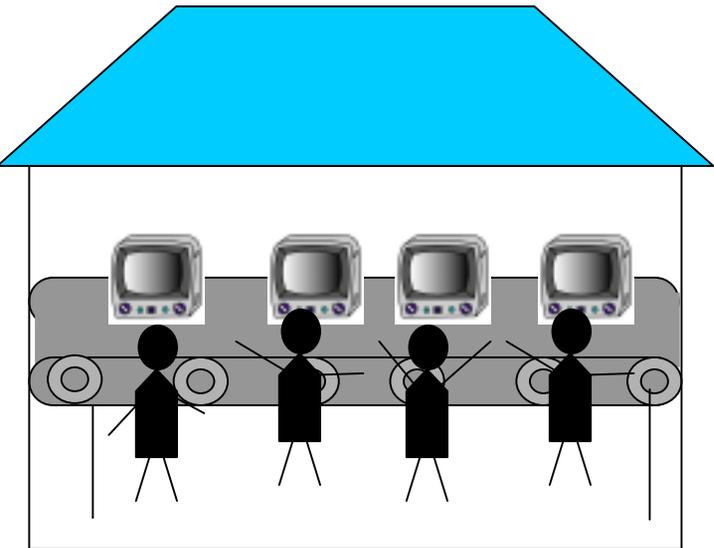
筐体内部



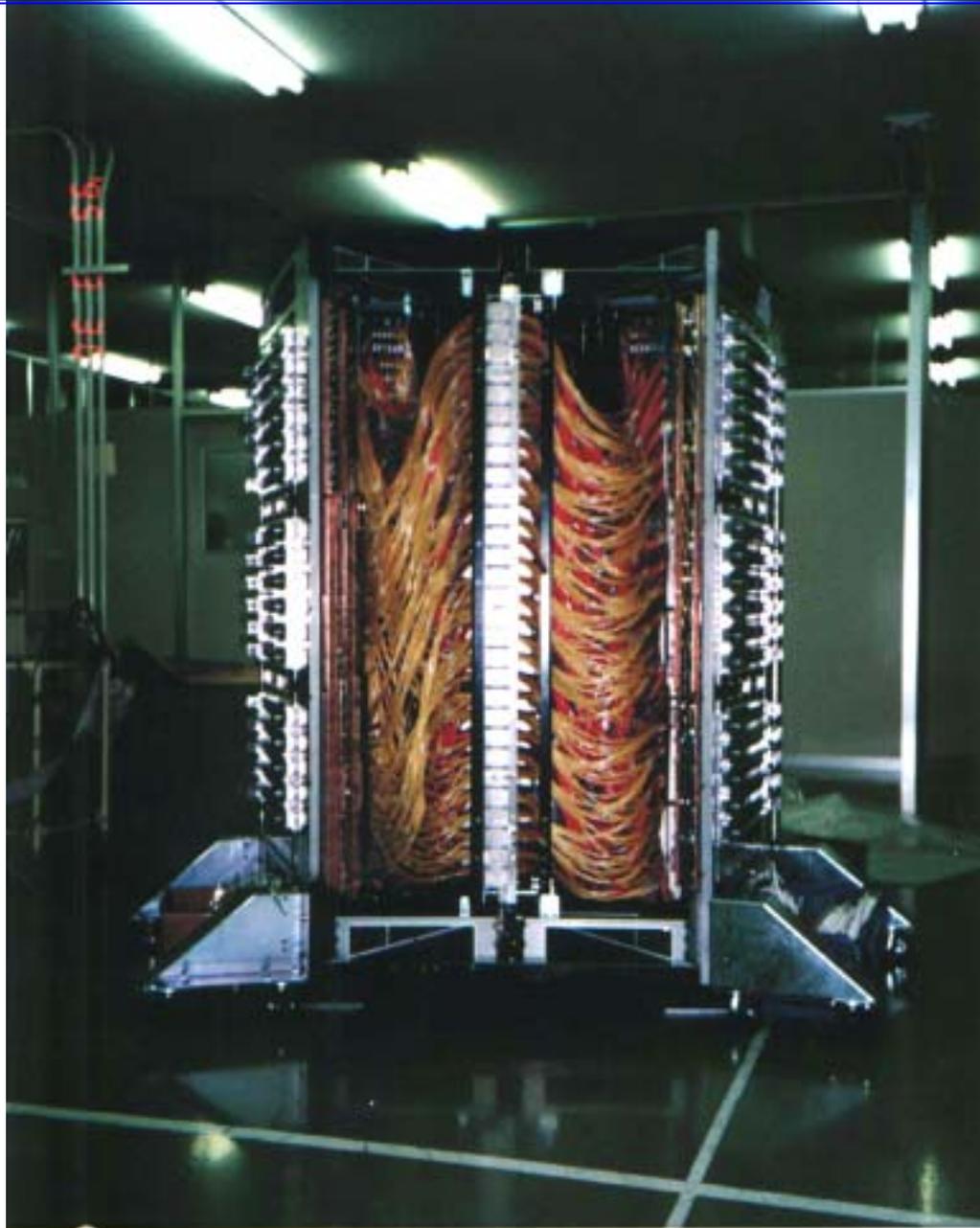
メモリのバンド幅



組立工程と部品供給

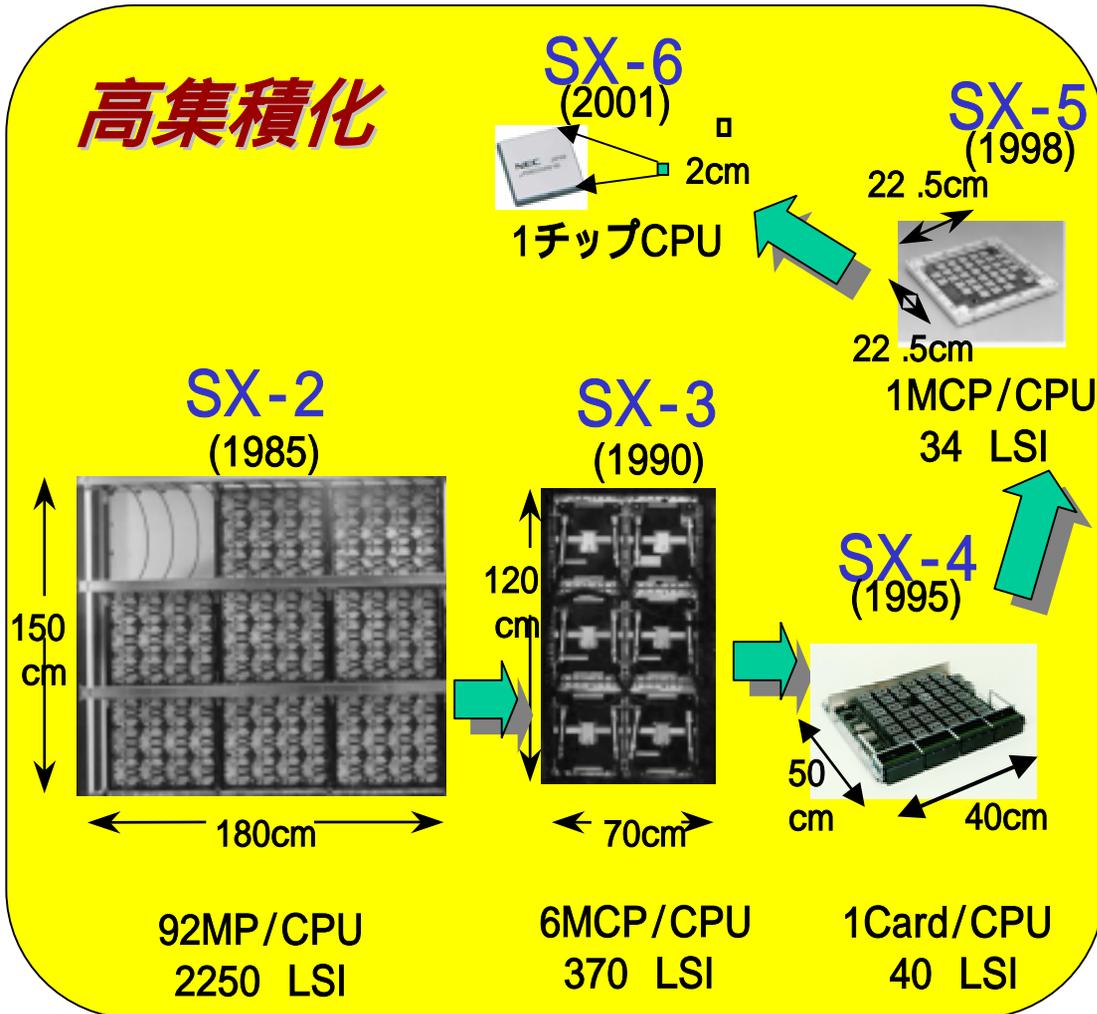


CPUとメモリの配線

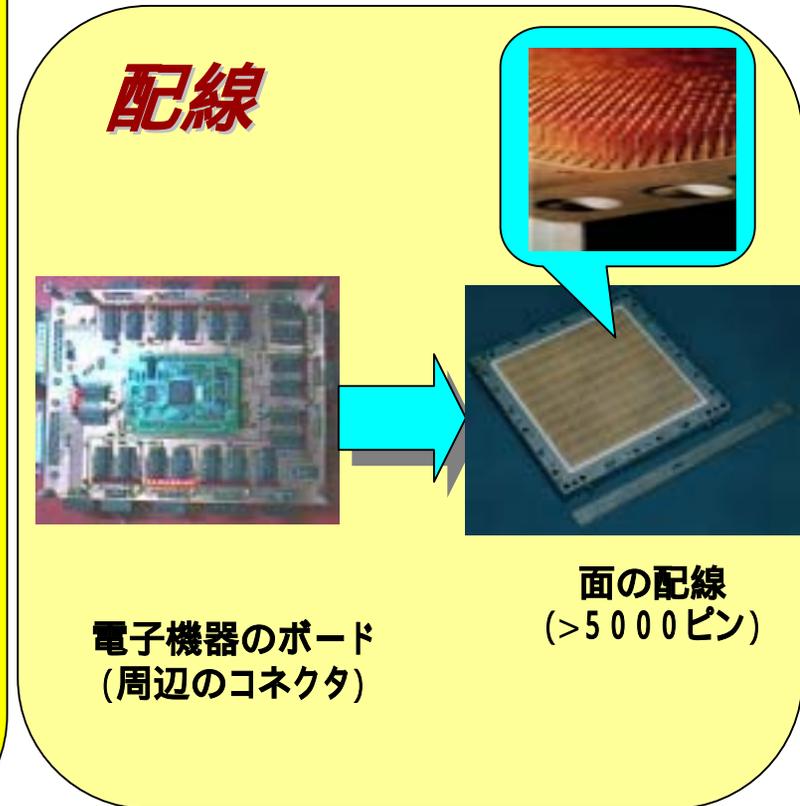


高集積化と配線 (線から面へ)

高集積化

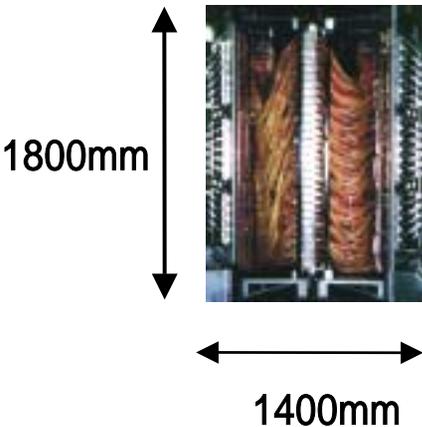


配線



SXにおける布線筐体の変遷

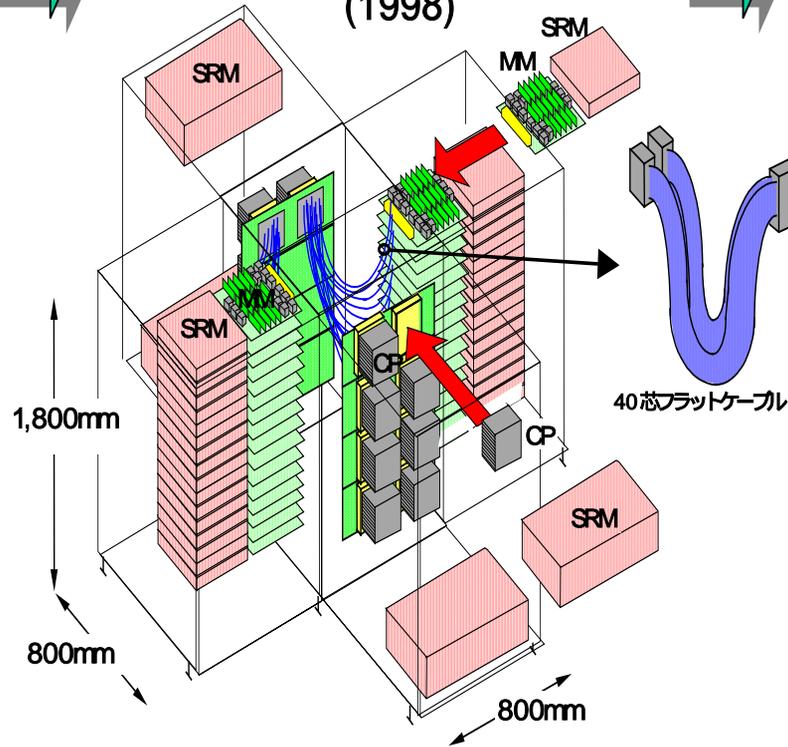
SX-4
(1995)



65



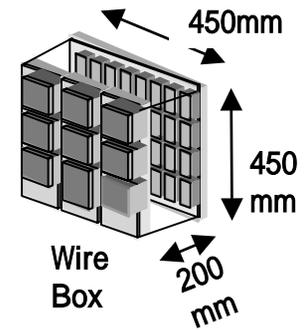
SX-5
(1998)



25

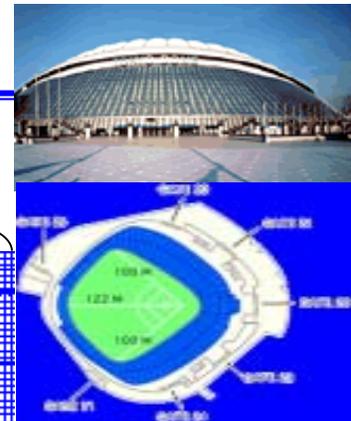


SX-6
(2001)

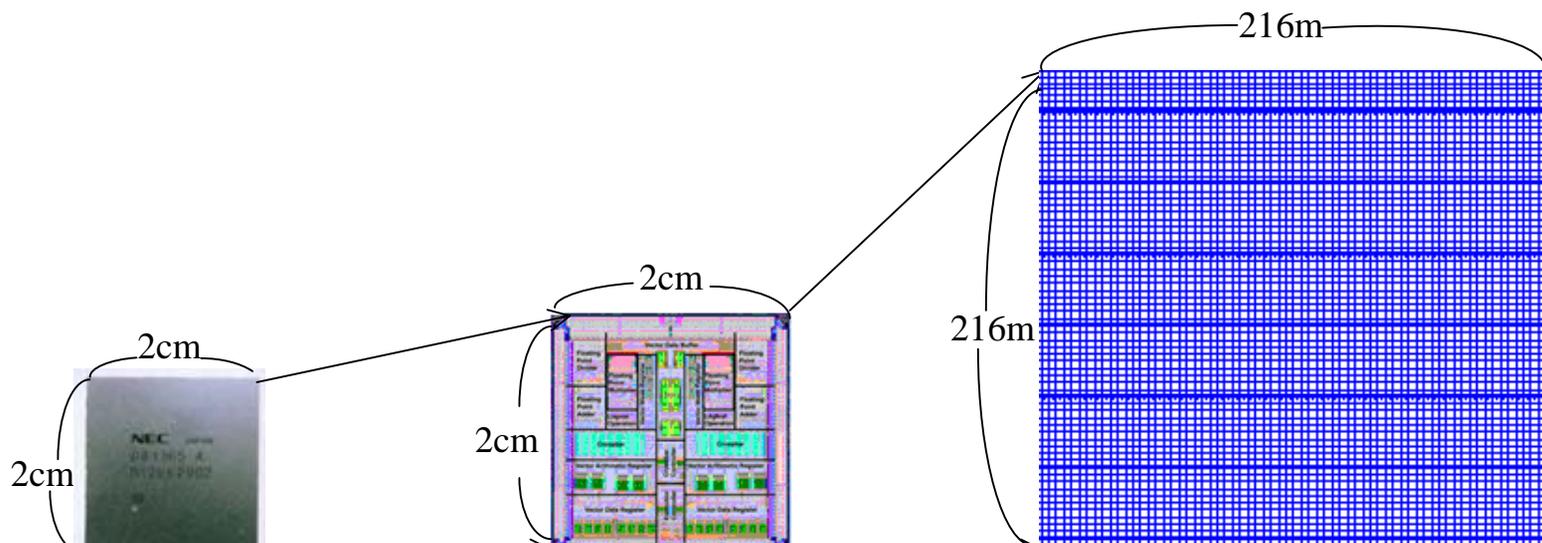


1

超LSIと実装の技術



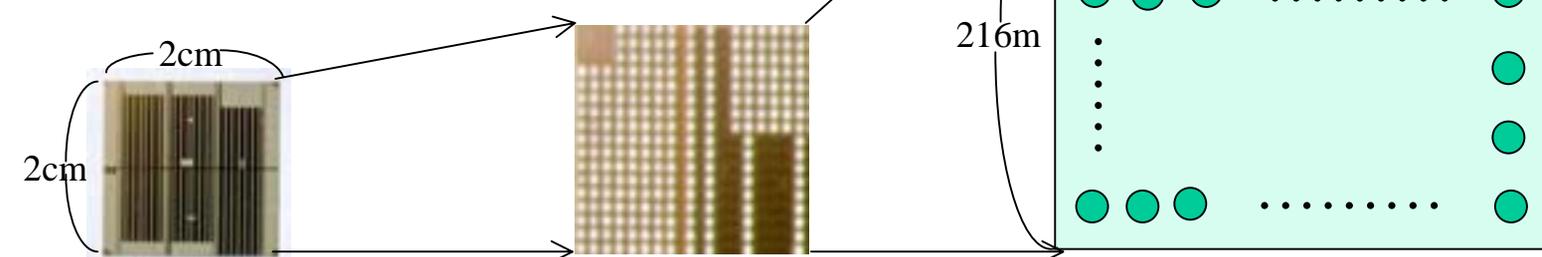
東京ドーム



超LSI

内部の回路と配線
(0.15 μ m線幅)

1.5mmの線を描くことに相当



超LSI

外部接続用のパッド
(直径0.1mm, 約5,200個)

2mおきに直径1mのベースをしきつめることに相当

NECのスーパーコンピュータ

- 歴史・激しい性能競争・主なユーザ・仲間 -

NEC/スーパーコンピュータの歴史

82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 00 01

プロジェクト開始

SX-4出荷

SX-5出荷

開発史

SX-1/2発表
(1GFLOPS超)
富士通・日立発表

SX-2出荷

SX-3発表
(UNIX)
SX-3出荷

SX-4発表
(CMOS)

SX-5発表

SX-6発表
(1Chip)

PC9801出荷

の計算
(キネブック)

日本IBMと提携

- ・大阪大/東北大
- ・海洋科学研究センター
- ・核融合科学研究所
- ・物質・材料研究機構
- ・トヨタ中央研究所
- ・日産自動車

国内市場と
市場開放

スーパー-301条問題
(不公正貿易)

S/C導入手続き
(MITI通達)

S/C導入手続き改定
(透明・無差別・公開)

核融合研調達

(クレイが苦情申立)

日米貿易摩擦(車,半導体)

ダンピング決定

グローバル
展開

HNSX(米国)
設立

ESS(欧州)設立

NCARダンピング問題

最高裁却下
Cray社と提携
(SXを北米で販売)

米政府圧力(MIT/LLNL/LANL/NASA/EPA等)

スイスCSCS

独DLR

Volvo

Daimler/
Chrysler

仏
IDRIS

伊
INGV

HARC
(米国初)

蘭NLR
(欧州初)

シンガポール
NSRC

独
ケルン大

カナダ
気象庁

ブラジル
気象庁

VW
GTRI

米
気象庁

デンマーク
豪・気象庁/
CSIRO

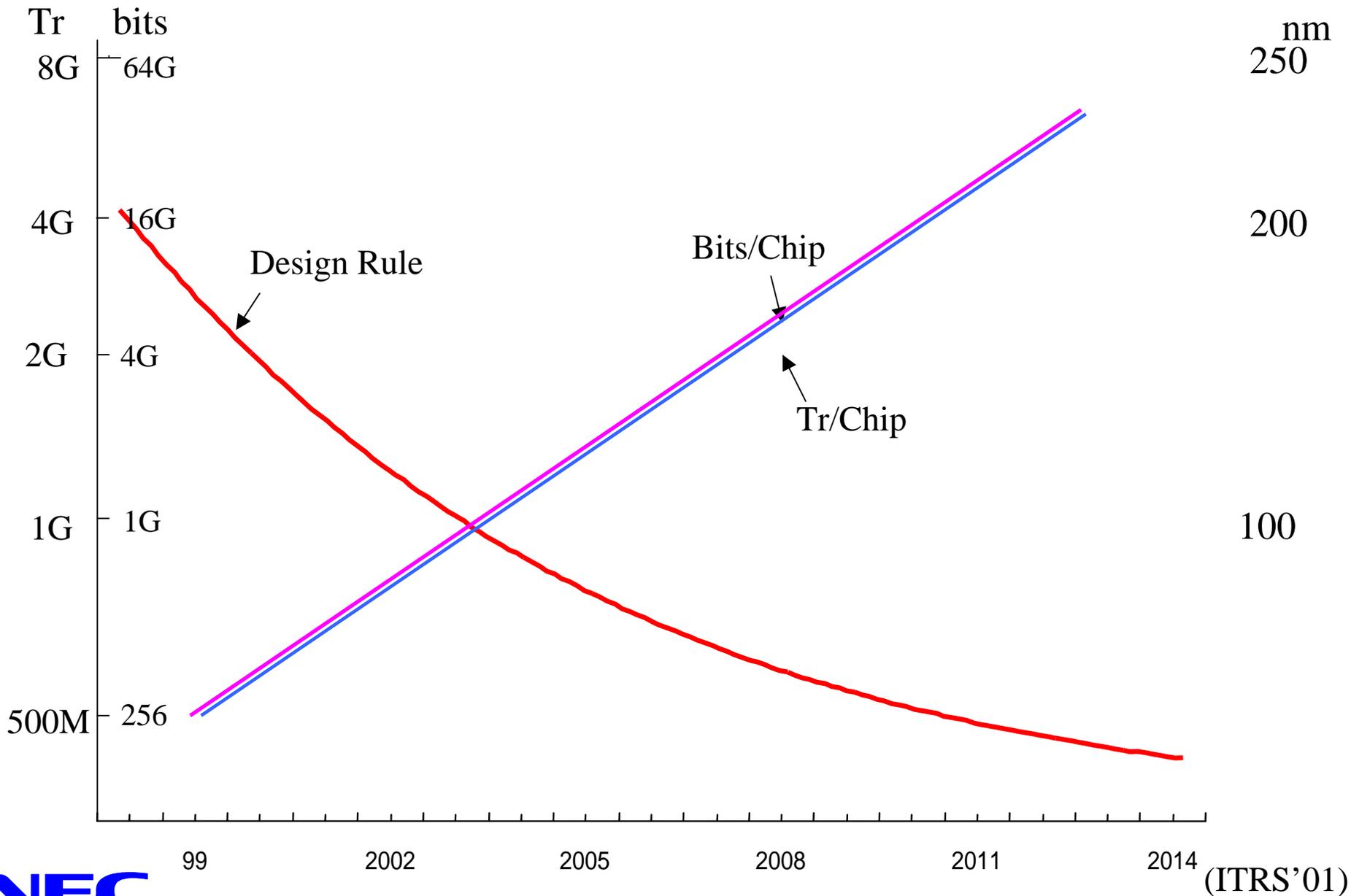
韓国気象庁
(KMA)

Renault

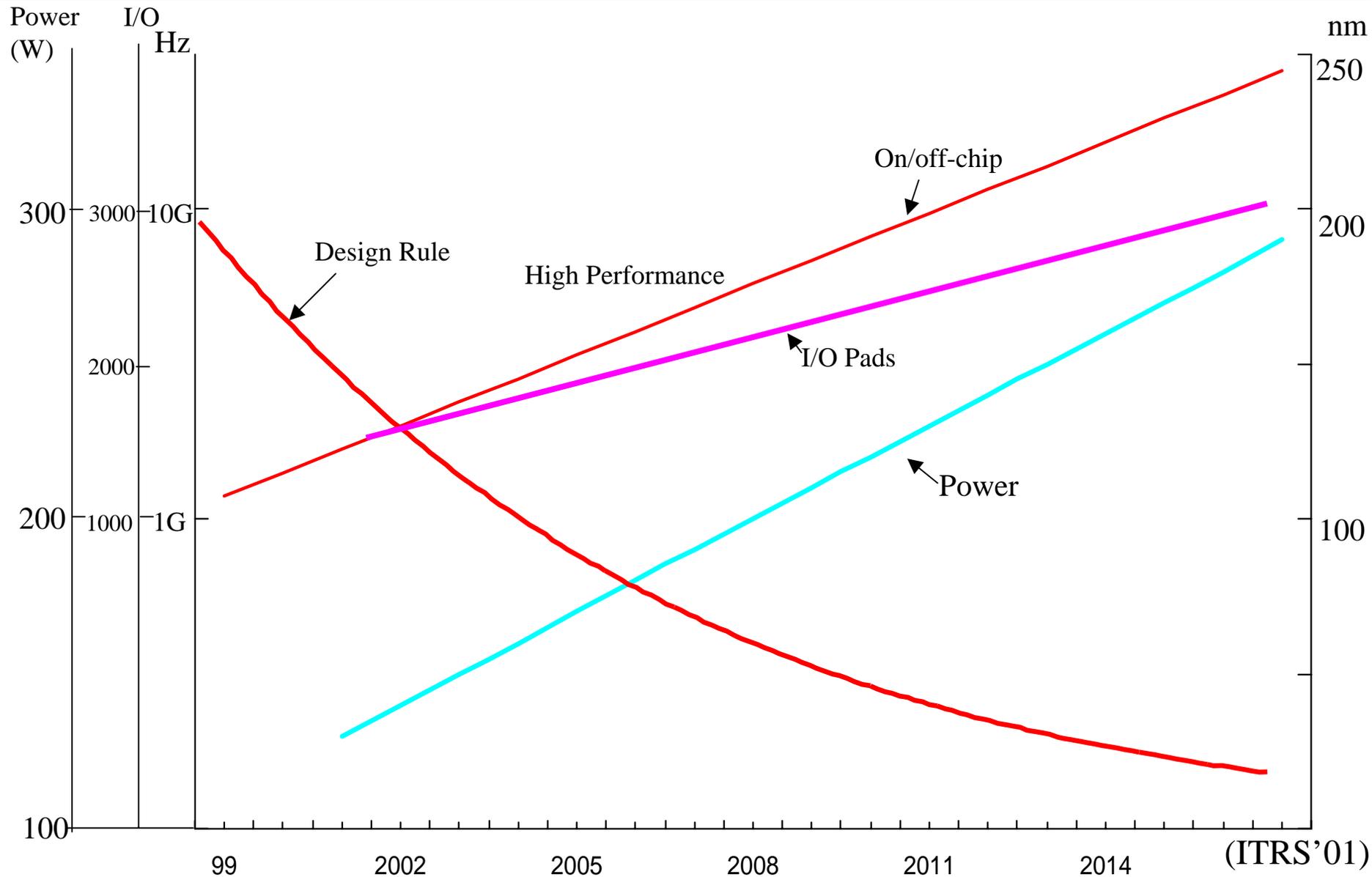
これからどうなるか？

- 将来のスーパーコンピュータ -

Memory Chip and Tr in μ -Processor



Clock Frequencies · I/O Pads · Power Dissipation



Logic Technology Roadmap

ITRS '99

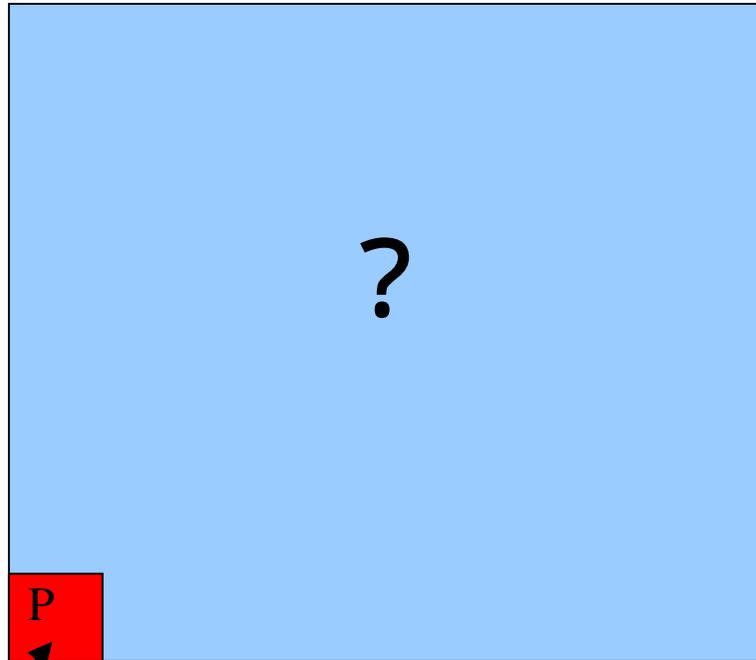
YEAR	1999	2000	2001	2002	2003	2004	2005	2008	2011	2014
MPU Gate Length (nm)	140	120	100	85	80	70	65	45	32	22
ASIC Gate Length (nm)	180	165	150	130	120	110	100	70	50	35
Nominal I_{on} at 25 °C ($\mu A/\mu m$)	750/350	750/350	750/350	750/350	750/350	750/350	750/350	750/350	750/350	750/350
[NMOS/PMOS] high-performance										
Maximum I_{off} at 25 °C (pA/ μm)	5	7	8	10	13	16	20	40	80	160
(For minimum L device) low power										
Equivalent physical oxide thickness	1.9-2.5	1.9-2.5	1.5-1.9	1.5-1.9	1.5-1.9	1.2-1.5	1.0-1.5	0.8-1.2	0.6-0.8	0.5-0.6
Tox (nm)										
L_{gate} 3 σ variation (nm)	14	12	10	8.5	8	7	6.5	5	3.2	2.2
(dense and isolated lines)										
Gate electrode sheet R_s (Ω/\square)	4-6	4-6	4-6	4-6	4-6	4-6	4-6	4-6	4-6	4-6
Silicide thickness (nm)	55	45	40	34	32	28	25	20	15	12
Contact silicide sheet R_s (Ω/\square)	2.7	3.3	3.8	4.4	4.7	5.4	6.0	7.5	10.0	12.5
Drain extension X_j (nm)	42-70	36-60	30-50	25-43	24-40	20-35	20-33	16-26	11-19	8-13
Number of metal levels	6-7	6-7	7	7-8	8	8	8-9	9	9-10	10
Local wiring pitch (nm)	500	450	405	365	330	295	265	185	130	95
Intermediate wiring pitch (nm)	640	575	520	465	420	375	340	240	165	115
Minimum global wiring pitch (nm)	1050	945	850	765	690	620	560	390	275	190
Conductor effective resistivity	2.2	2.2	2.2	2.2	2.2	2.2	2.2	1.8	<1.8	<1.8
Cu wiring ($\mu\Omega\text{-cm}$)										
Barrier/cladding thickness (for Cu wiring) (nm)	17	16	14	13	12	11	10	0	0	0
Interlevel metal insulator	3.5-4.0	3.5-4.0	2.7-3.5	2.7-3.5	2.2-2.7	2.2-2.7	1.6-2.2	1.5	<1.5	<1.5
effective dielectric constant (k)										

ITRS '01

YEAR OF PRODUCTION	2001	2002	2003	2004	2005	2006	2007	2010	2013	2016
DRAM 1/2 PITCH(nm)	130	115	100	90	80	70	65	45	32	22
MPU/ASIC1/2PITCH(nm)	150	130	107	90	80	70	65	50	35	25
MPU PRINTED GATE LENGTH(nm)	90	75	65	53	45	40	35	25	18	13
MPU PHYSICAL GATE LENGTH(nm)	65	53	45	37	32	28	25	18	13	9
Physical gate length high-performance(HP)(nm)[1]	65	53	45	37	32	28	25	18	13	9
Equivalent physical oxide thickness for high-performance T_{ox} (EOT)(nm)[2]	1.3-1.6	1.2-1.5	1.1-1.6	0.9-1.4	0.8-1.3	0.7-1.2	0.6-1.1	0.5-0.8	0.4-0.6	0.4-0.5
Gate depletion and quantum effects electrical thickness adjustment factor(nm)[3]	0.8	0.8	0.8	0.8	0.8	0.8	0.5	0.5	0.5	0.5
T_{ox} electrical equivalent(nm)[4]	2.3	2.1	2	2	1.9	1.9	1.4	1.2	1	0.9
Nominal power supply voltage(V_{dd})(V)[5]	1.2	1.1	1	1	0.9	0.9	0.7	0.6	0.5	0.4
Nominal high-performance NMOS sub-threshold leakage current, $I_{sd,leak}$ (at 25 °C) ($\mu A - \mu m$)[6]	0.01	0.03	0.07	0.1	0.3	0.7	1	3	7	10
Nominal high-performance NMOS saturation drive current, I_{dd} (at V_{dd} , at 25 °C) ($\mu A - \mu m$)[7]	900	900	900	900	900	900	900	1200	1500	1500
Required percent current-drive "mobility/transconductance improvement"[8]	0%	0%	0%	0%	0%	0%	0%	30%	70%	100%
Parasitic source/drain resistance(R_{sd})(ohm- μm)[9]	190	180	180	180	180	170	140	110	90	80
Parasitic source/drain resistance(R_{sd})percent of ideal channel resistance(V_{dd}/I_{dd})[10]	16%	16%	17%	18%	19%	19%	20%	25%	30%	35%
Parasitic capacitance percent of ideal gate capacitance[11]	19%	22%	24%	27%	29%	32%	27%	31%	36%	42%
High-performance NMOS device $f(C_{gate} * V_{dd} / I_{dd} - NMOS)$ (ps)[12]	1.6	1.3	1.1	0.99	0.83	0.76	0.68	0.39	0.22	0.15
Relative device performance[13]	1	1.2	1.5	1.6	2	2.1	2.5	4.3	7.2	10.7
Energy per($W/L_{gate}=3$)device switching transition $(C_{gate} * (3 * L_{gate}) * V^2)$ (fJ/Device)[14]	0.347	0.212	0.137	0.099	0.065	0.052	0.032	0.015	0.007	0.002
Static power dissipation per($W/L_{gate}=3$)device (Watts/Device)[15]	0.5E-09	6.7E-09	1.0E-08	1.1E-08	2.6E-08	5.3E-08	5.3E-08	9.7E-08	1.4E-07	1.1E-07

How to Utilize Chip Area? (~ 2010)

Chip Size: 6.2cm^2
($0.07\ \mu\text{m}$ Rule)



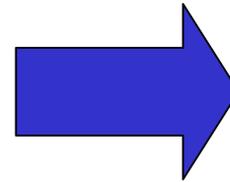
μ -P Core: 0.1cm^2
(5MTr)

10年後のパソコン

現在のスーパーコンピュータ



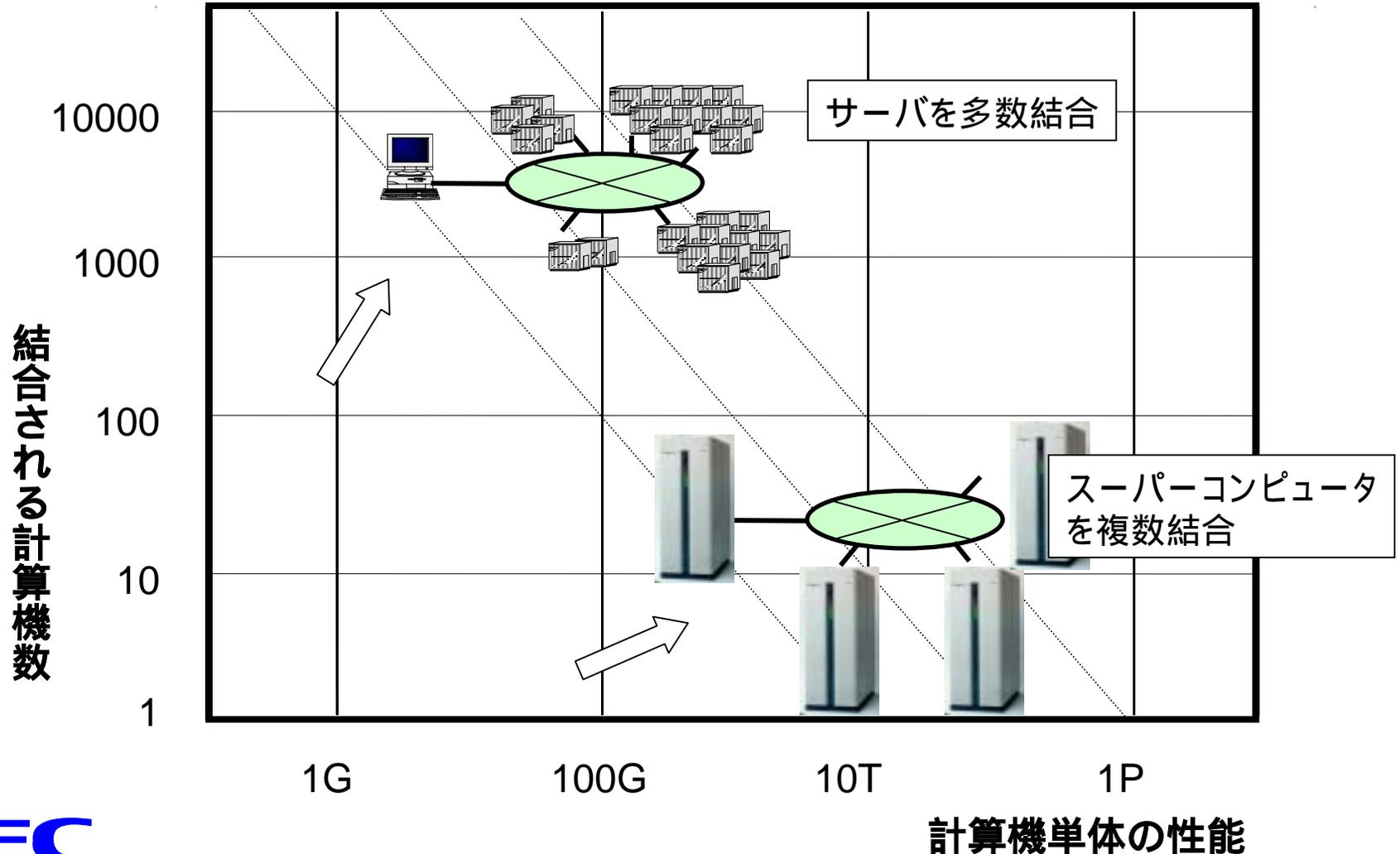
10年後のパソコン



今後のスーパーコンピュータ

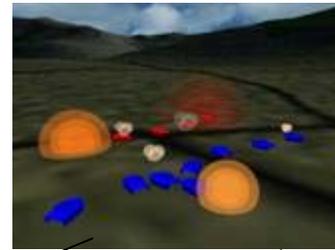
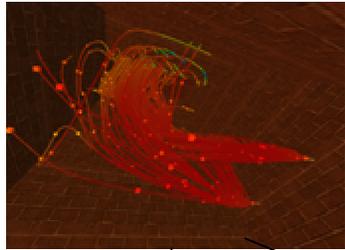
さらなる高速化をめざすために

- ・ 計算機単体の性能向上
- ・ 結合計算機数の拡張



グリッド・コンピューティング

- ・ 高速ネットワークで計算機資源を接続
 - 計算機資源の場所を意識せず
 - 必要なときに必要なサービスを得られる

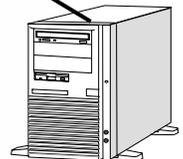
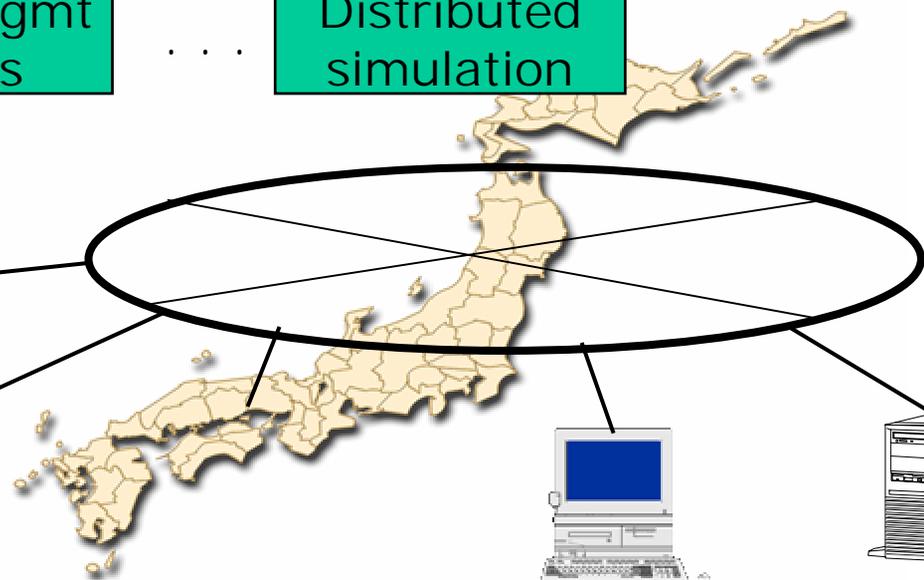


Collaboration
Tools

Data Mgmt
Tools

...

Distributed
simulation



END