HPC Workshop

"Hardware and software for large-scale biological computing in the next decade"

The Japanese Next Generation Supercomputer Project

Tadashi Watanabe
Project Leader
Next-Generation Supercomputer R&D Center, RIKEN
December 14, 2007

Contents

- The Next Generation Supercomputer project
- Technical Challenges for Peta-scale System

The Next Generation Supercomputer project

Six Goals of Japan's "3rd Science and Technology Basic Plan" and Next-Generation Supercomputer Project

<Goal 1>
Discovery & Creation of Knowledge toward the future

Milky Way formation process

by RIKEN

Planet formation process

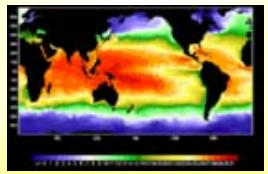
by RIKEN

Aurora outbreak process

by JAMSTEC

< Goal 3 >
Sustainable Development - Consistent with Economy and Environment -

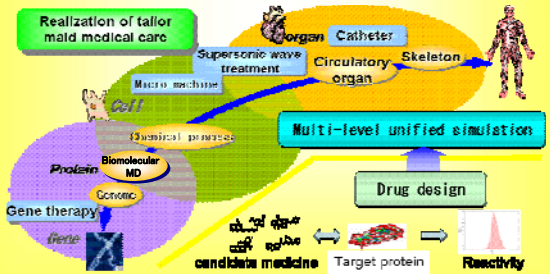
An influence prediction of El Nino phenomenon



by JAMSTEC

< Goal 5 >
Good Health over Lifetime

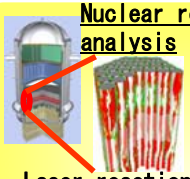
Multi-level unified simulation



by Univ. of Tokyo and RIKEN

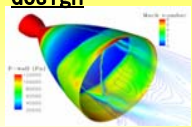
Development and Application of Next-Generation Supercomputer

Nuclear reactor analysis



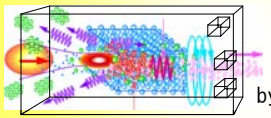
by JAEA

Rocket engine design



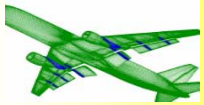
by JAXA

Laser reaction analysis



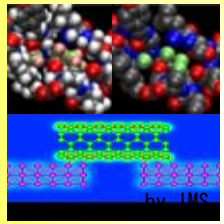
by JAEA

Plane development



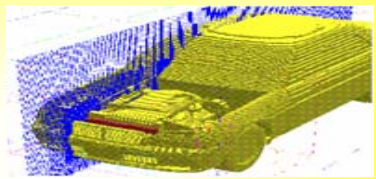
by JAXA

Nano technology



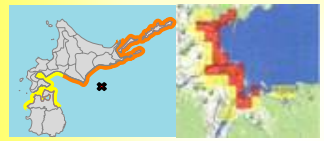
by IMS

Car development



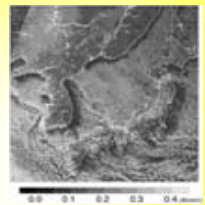
by NISSAN

Tsunami damage prediction



by Tohoku Univ.

Clouds analysis



by MRI

< Goal 2 >
Breakthroughs in Advanced Science and Technology

< Goal 4 >
Innovator Japan - Strength in Economy & Industry -

< Goal 6 >
Safe and secure Nation

Development & Application of Next-Generation Supercomputer Project by MEXT

FY2006: 3,547Million yen / FY2007: *7,736*Million yen
FY2006～FY2012 (total budget expected） about 110billion yen

1. Purpose of policy

Development and implementation of the world's most advanced and high-performance Next-Generation Supercomputer, and to develop and disseminate its usage technologies, as one of Japan's "Key Technologies of National Importance" (National Infrastructure).

2. Expected effects

As an important tool for simulation, supercomputing needs to be developed further. This project aims to bring the Next-Generation Supercomputer to completion in 2012.
In order to maintain world-leading position in variety of areas, the following academic-industrial collaboration activities will be conducted under the initiative of MEXT.

(1) Development and implementation of the world's most advanced high-performance Next-Generation supercomputer
(2) Development and dissemination of software that makes optimum use of the supercomputer
(3) Establishment of the world's most advanced and highest standard supercomputing Center of Excellence, which includes the Next-Generation Supercomputer

3. Project Framework

- Integrated development of computer and software
- Establishment of nationwide academic-industrial collaborative structure, with RIKEN as the project headquarters
- A new law has been introduced for the framework of usage and administration

Policy and Outline of A Next Generation Supercomputer Project

Purpose of policy:

development, installation and application of an advanced high performance supercomputer system, as one of Japan's "Key Technologies of National Importance"

Total Budget:

about 115 billion Yen (~ $ 1 billion)

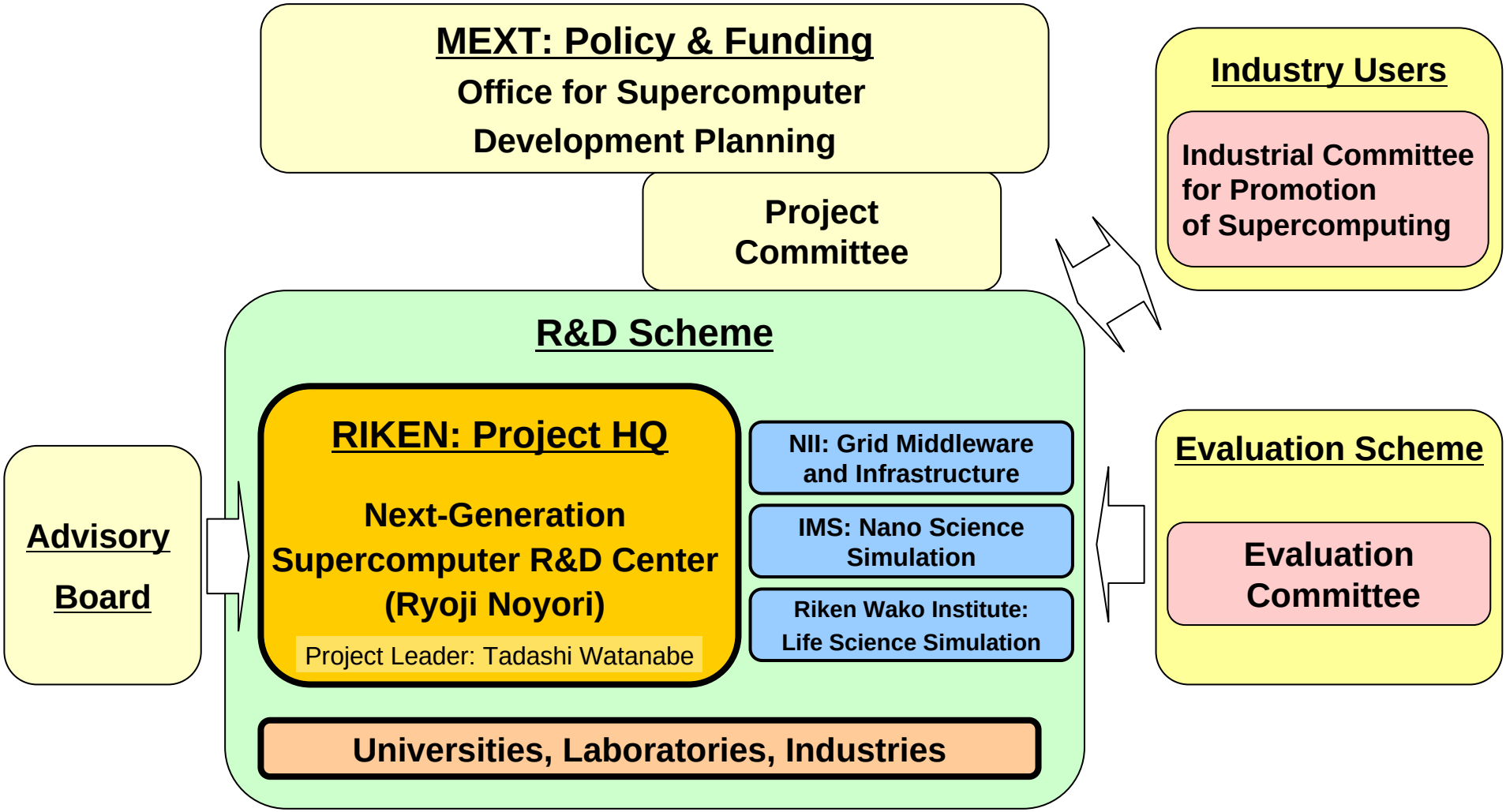
Period of Project:

FY2006 – FY2012

Goals of the Next Generation Supercomputer Project

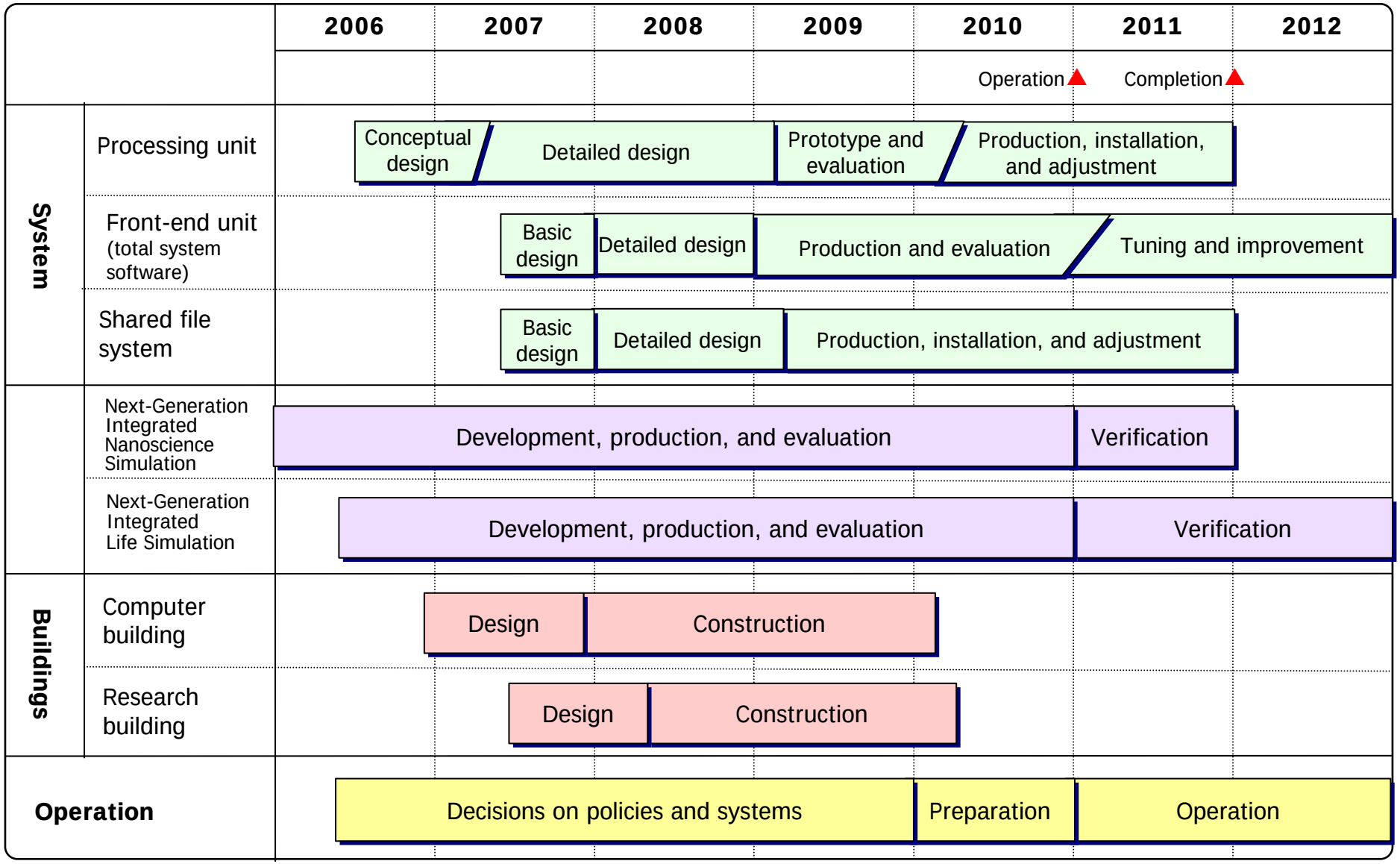
1. Development and installation of the most advanced high performance supercomputer system
2. Development and wide use of application software to utilize the supercomputer to the maximum extent
3. Provision of flexible computing environment by sharing the next generation supercomputer through connection with other supercomputers located at universities and research institutes
4. Establishment of “Advanced Computational Science and Technology Center (tentative name)”

Project Organization



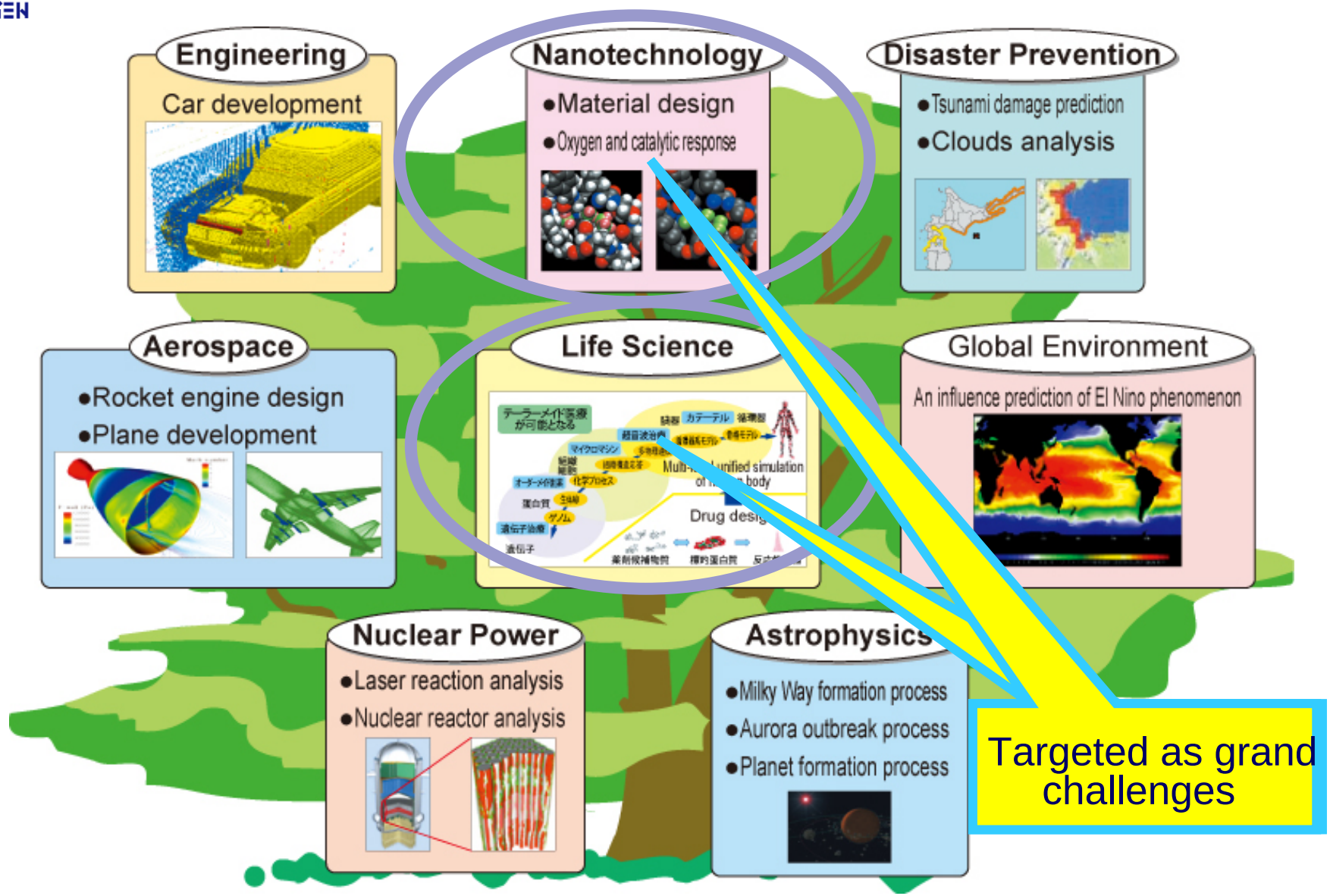
(Note) NII: National Institute of Informatics, IMS: Institute for Molecular Science

SCHEDULE


2006
2007
2008
2009
2010
2011
2012
Operation
Completion
System
Processing unit
Conceptual design
Detailed design
Prototype and evaluation
Production, installation, and adjustment
Front-end unit (total system software)
Basic design
Detailed design
Production and evaluation
Tuning and improvement
Shared file system
Basic design
Detailed design
Production, installation, and adjustment
Next-Generation Integrated Nanoscience Simulation
Development, production, and evaluation
Verification
Next-Generation Integrated Life Simulation
Development, production, and evaluation
Verification
Buildings
Computer building
Design
Construction
Research building
Design
Construction
Operation
Decisions on policies and systems
Preparation
Operation

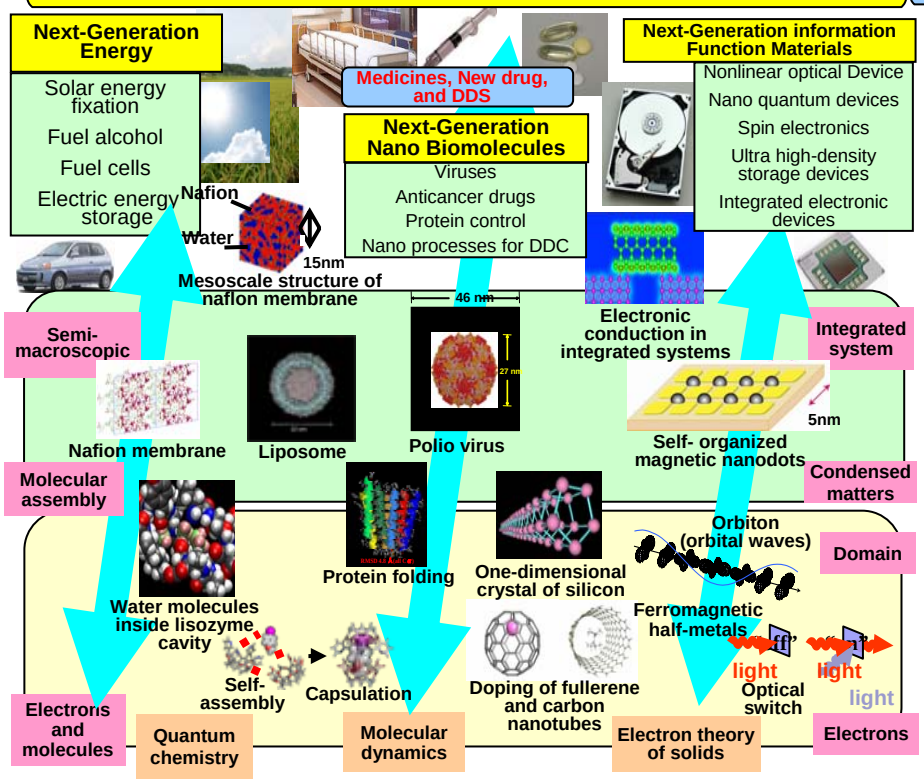

Shinkansen-Line
Shin-Kobe Station
Mt. Rokko
Ashiya
Sannomiya
Kobe Medical Industry
Development Project
Core Facilities
Port Island
About 5km from Sannomiya
12 min. by Portliner
Next-Generation
Supercomputer
Portliner
Kobe Sky Bridge
To Akashi / Awaji-Island
To Osaka
Kobe Airport
Photo: June, 2006

Major Applications of Next Generation Supercomputer



Grand Challenge Applications

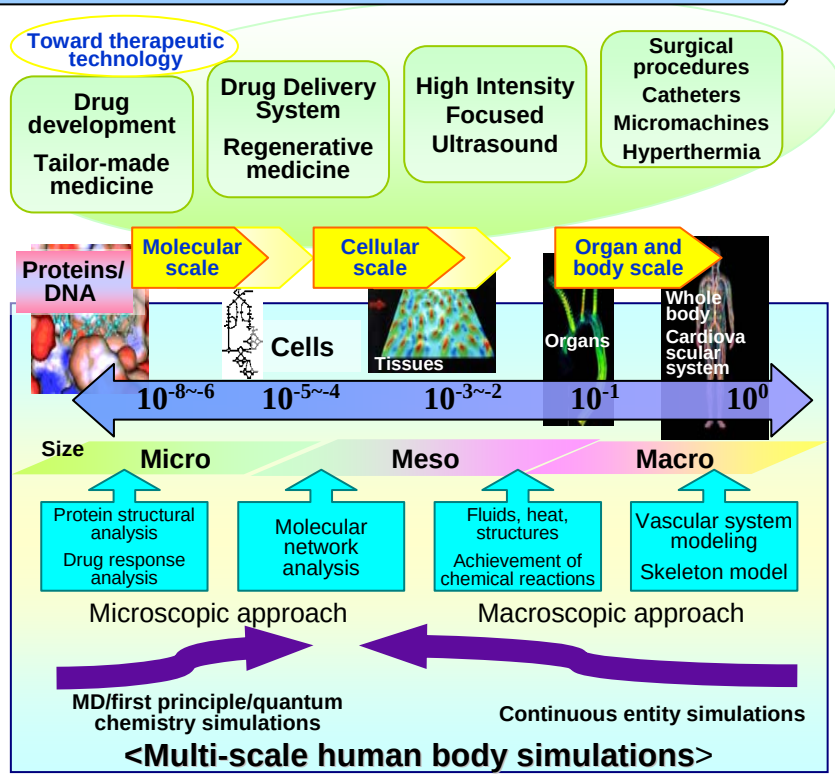
Next-Generation Integrated Nano-Science Simulation Software (2006–2010)



Base site: Institute for Molecular Science

To create next-generation nano-materials (new semiconductor materials, etc.) by integrating theories (such as quantum chemistry, statistical dynamics and solid electron theory) and simulation techniques in the fields of new-generation information functions/materials, nano-biomaterials, and energy

Next-Generation Integrated Life-Science Simulation Software (2006–2012)



Base site: RIKEN Wako Institute

To provide new tools for breakthroughs against various problems in life science by means of petaflops-class simulation technology, leading to comprehensive understanding of biological phenomena and the development of new drugs/medical devices and diagnostic/therapeutic methods

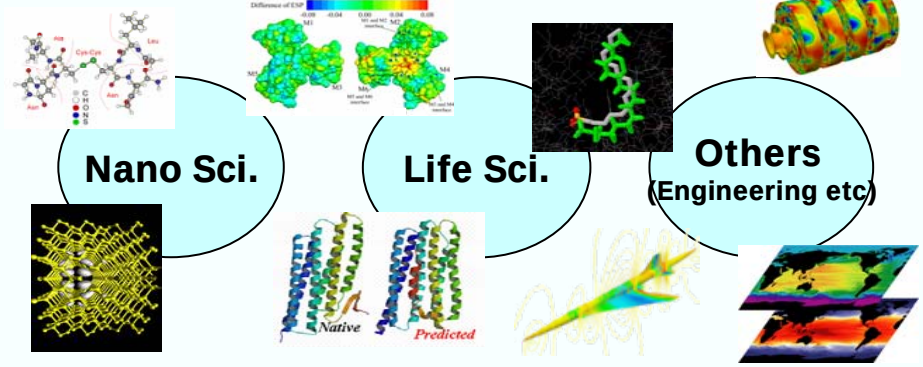
Promotion Program of Supercomputers for Industries

Industrial Committee for Super Computing Promotion

- established in 2006
- participated by more than 160 companies from various industries
- promote high performance computing technologies in industry, particularly for the use of supercomputers
- give proposals, opinions and advices to the project

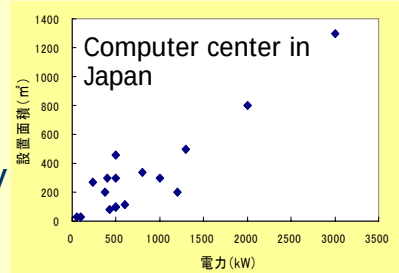
System Design

Requirement from Grand Challenges



Requirements from Computer Centers

- Power, Space
- Reliability, operability
- Cost (development, manufacturing, maintenance)



Popular numerical scheme

FMO, MO, RISM, Real space density functional method, Monte Carlo, FFT, Searching algorism, FEM, FVM, Linear Solver

Optimal system

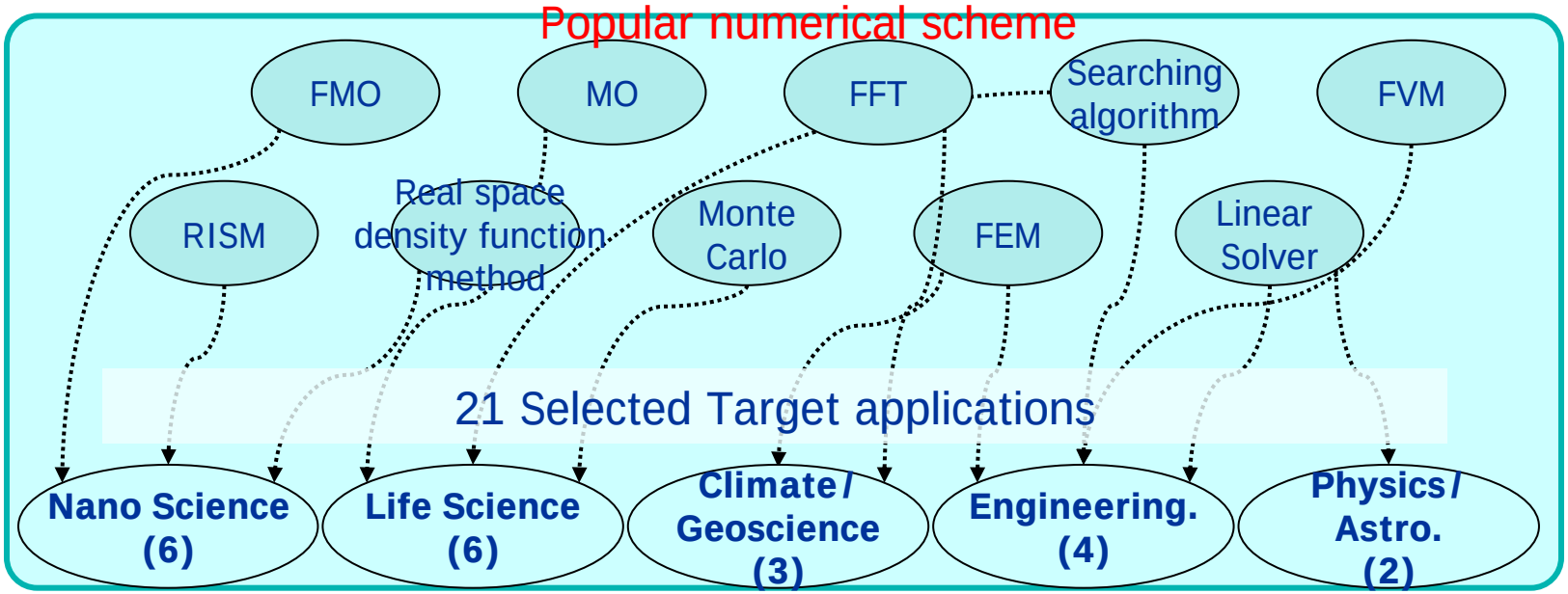
21 Selected Target applications

- Nano Science (6)
- Life Science (6)
- Climate/ Geoscience (3)
- Engineering. (4)
- Physics/ Astro. (2)

Technology Limit

- Other Project Watch
- Technology Survey
- Operation & Utilization
- Essential Element technologies
  - LSI process
  - Power saving SOI Low-k High-k
  - Light transm tech.
  - Software OS, compiler
- Spin off to the consumer electronics

Target applications for system evaluation



- 21 application programs have been identified by the application committee as candidates to be used for system evaluation.

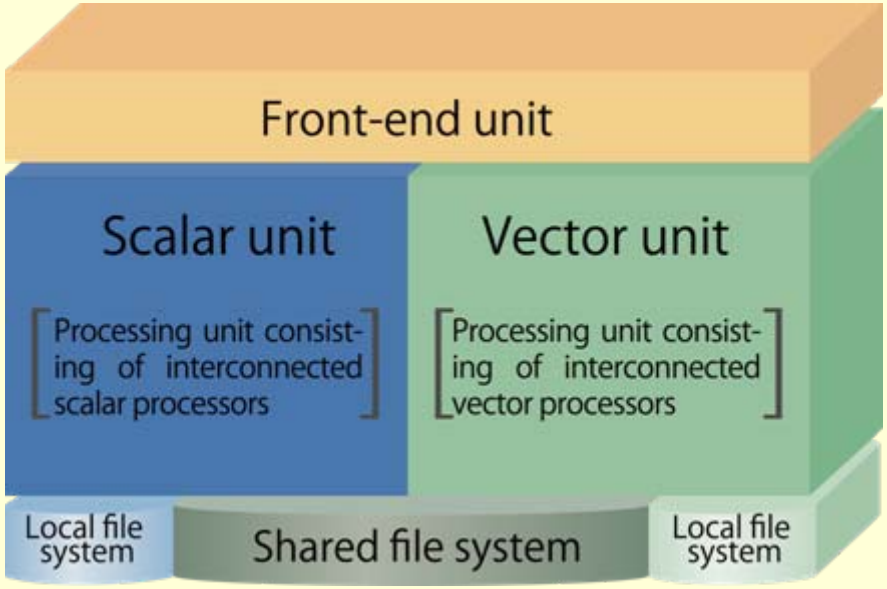
- Those applications include :
  protein folding, molecular orbital calculation, molecular dynamics, real space DFT, atmospheric circulation model with cloud resolution, lattice QCD, and compressed fluid dynamics.

The Next-Generation Supercomputer project

The Next-Generation Supercomputer project started in 2006 which is being carried out by RIKEN, with partners in industry, universities, and the government, under an initiative by MEXT (the Ministry of Education, Culture, Sports, Science and Technology).

Due to be ready in 2012, the peta-scale computing by the new supercomputer will ensure that Japan continues to lead the world in science and technology, academic research, industry, and medicine.

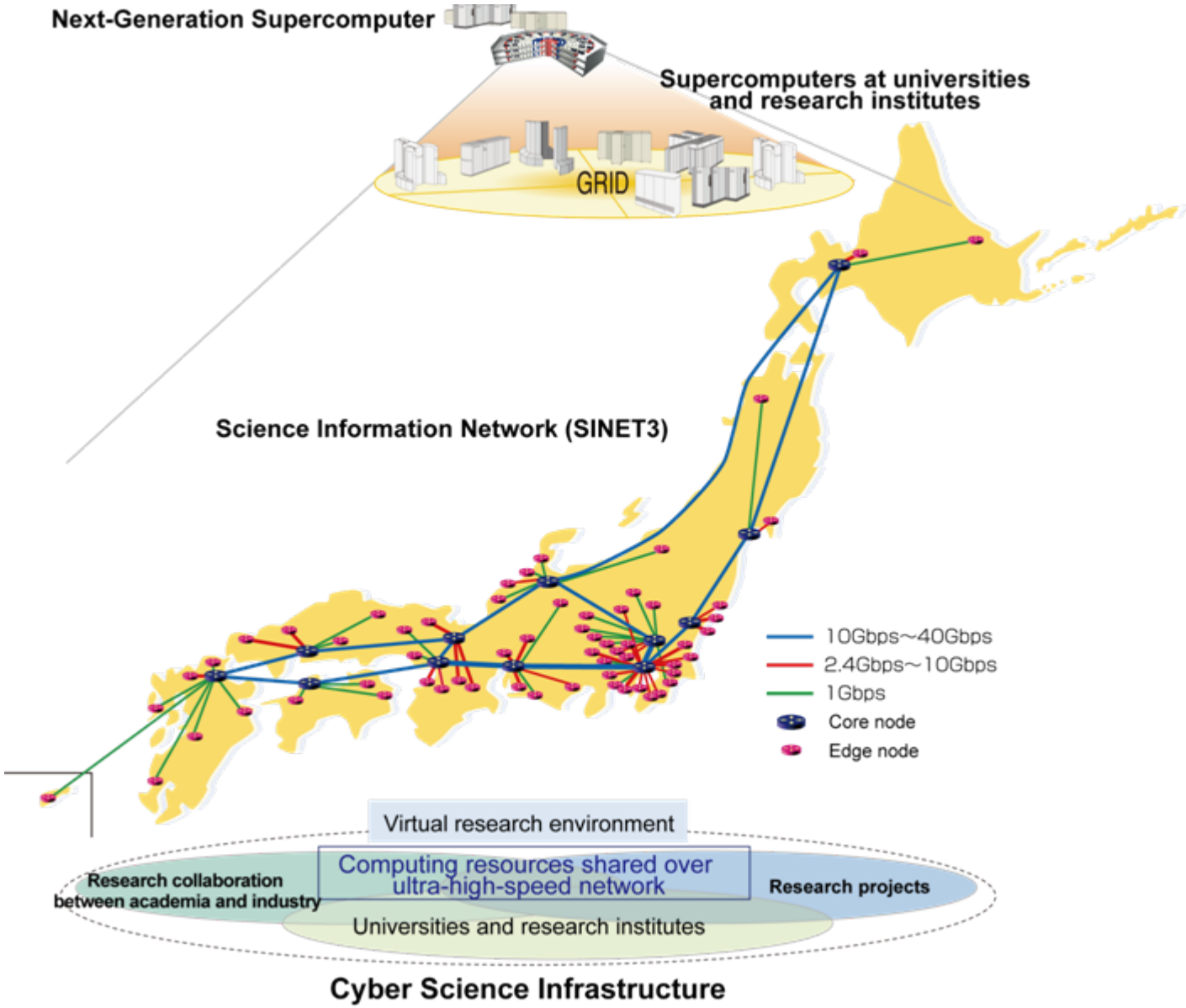
[System configuration]



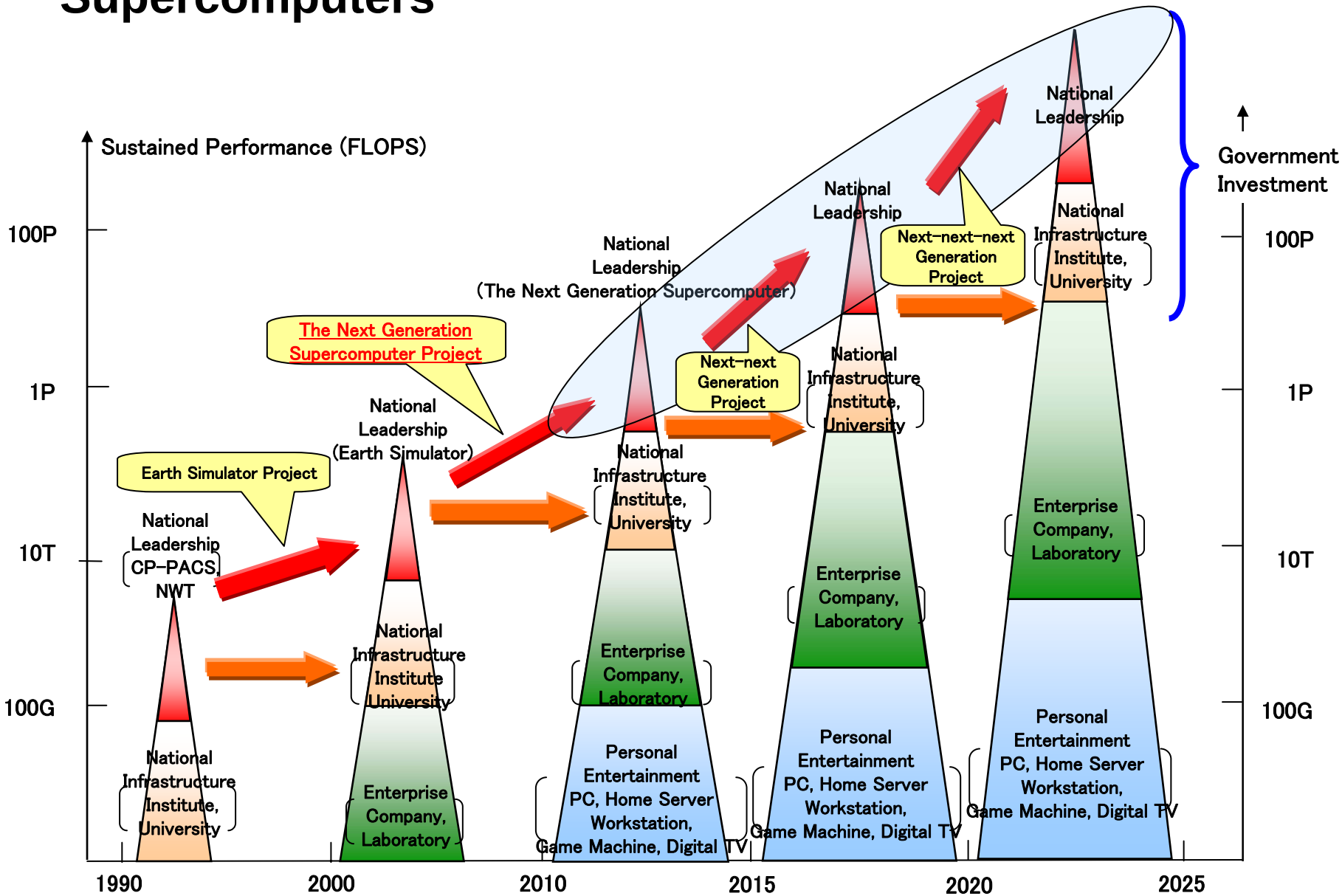
The Next-Generation Supercomputer will be hybrid general-purpose supercomputer that provides the optimum computing environment for a wide range of simulations.

- Calculations will be performed in processing units that are suitable for the particular simulation.
- Parallel processing in a hybrid configuration of scalar and vector units will make larger and more complex simulations possible.

Cyber Science Infrastructure plan

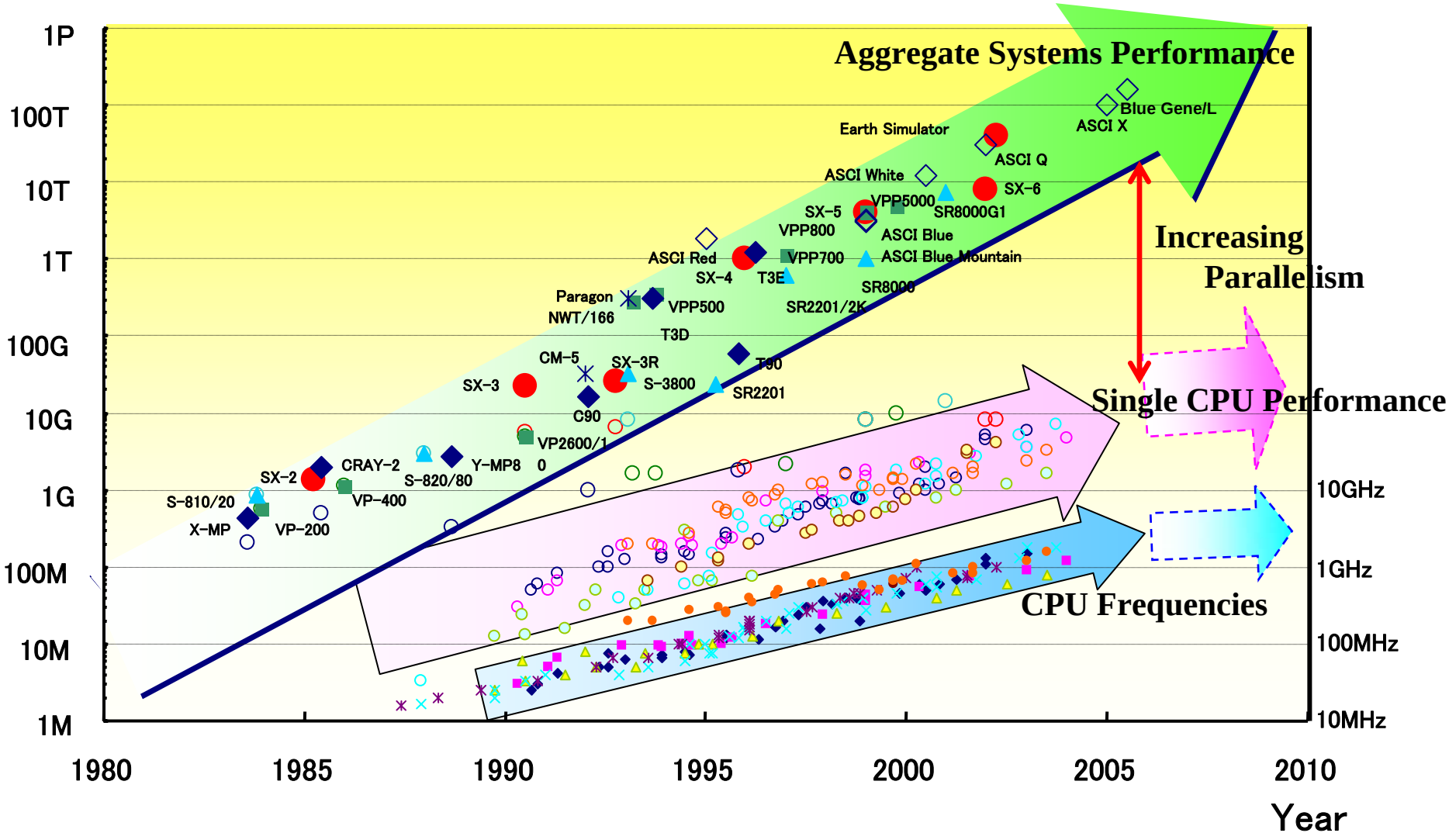


MEXT's Vision for Continuous Development of Supercomputers

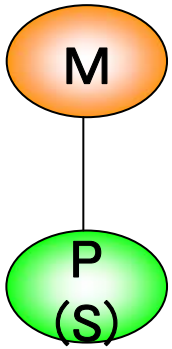
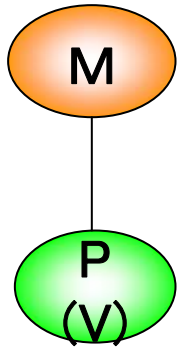
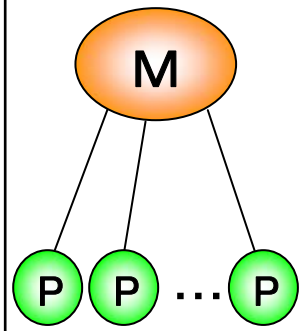
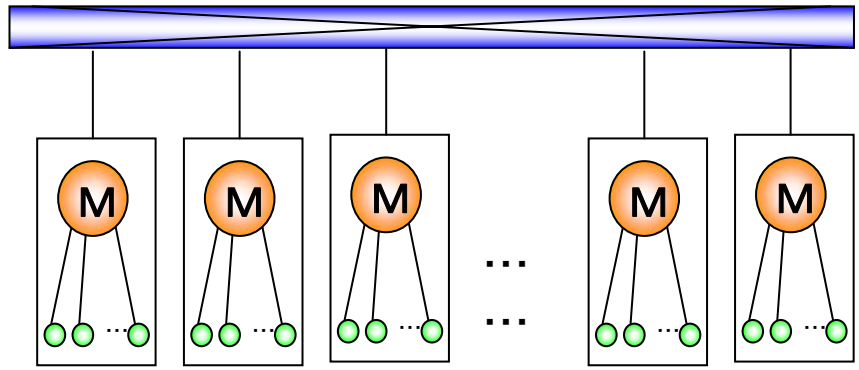


Technical Challenges for Peta-scale System

History of High Performance Computers



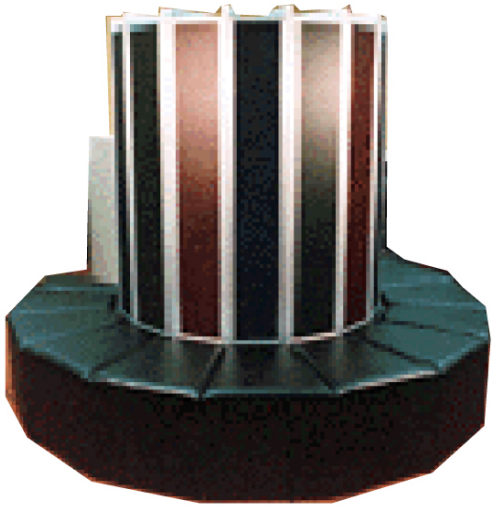
HPC Architecture

	Single CPU		Shared Memory	(Shared-)Distributed Memory		
System Configuration						
# of CPU	1CPU	1CPU	>10 CPUs	>100CPUs	>1000CPUs	＞10,000CPUs
CPU Architecture	Scalar	Vector	Scalar/ Vector	Scalar/ Vector	←	←
Language	Fortran	Vector Fortran	Parallel/Vector Fortran	Parallel/Vector Fortran/C, MPI	←	←
Tuning	Compiler	Vector Tuning	Parallel/Vector Tuning	Parallel Tuning	←	←

P :Processor, P (S) :Scalar, P (V) : Vector

M :Memory

Increasing Power and Foot-Print



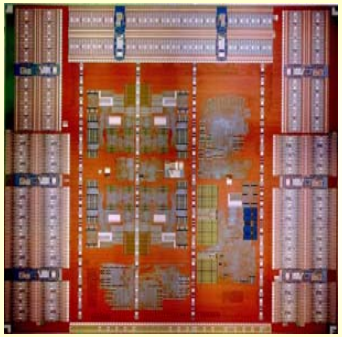
Cray-1

Earth Simulator

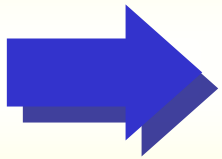
Sourced from http://www.thocp.net/hardware/cray_1.htm

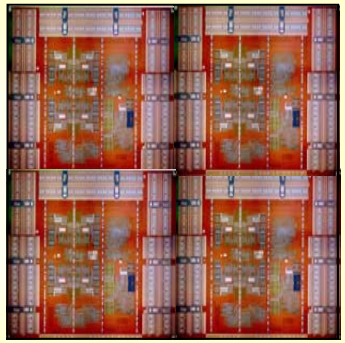
Increasing Parallelism

Device Technology



Single Core

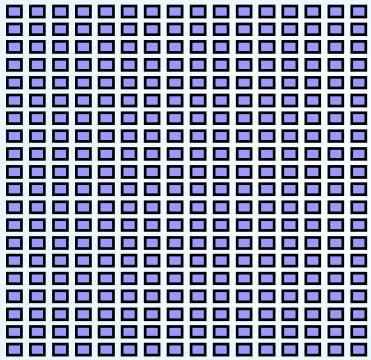


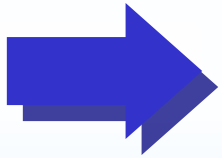


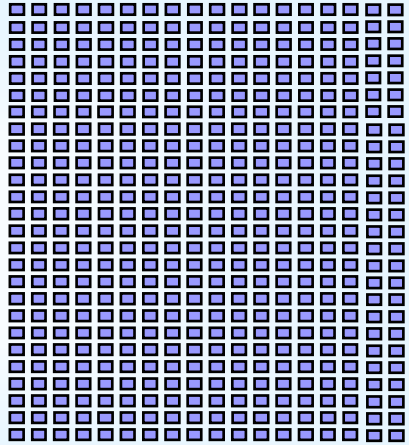
Multi Core

System Technology

(Parallelism)







More Parallel

Increasing Complexity of Programming

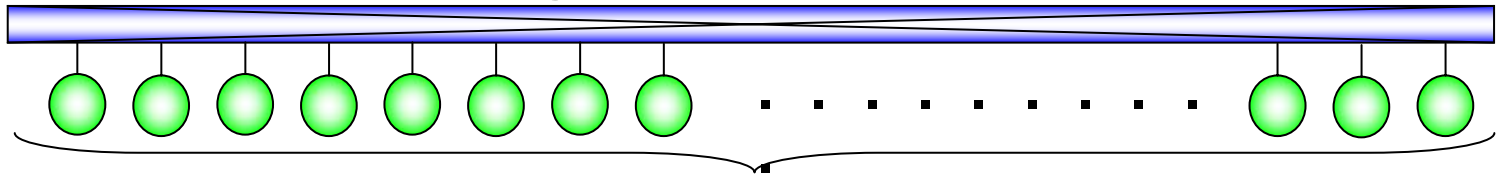
Serial Code

```
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
```

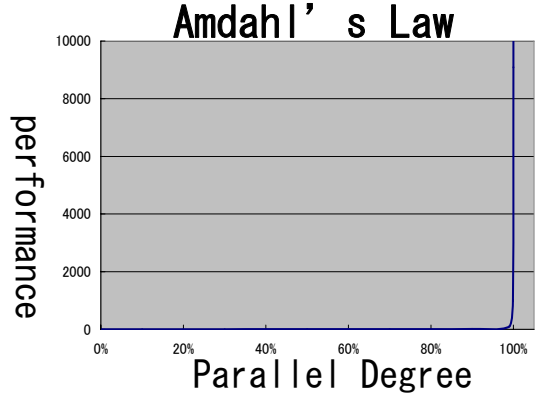
MPI

```
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
```

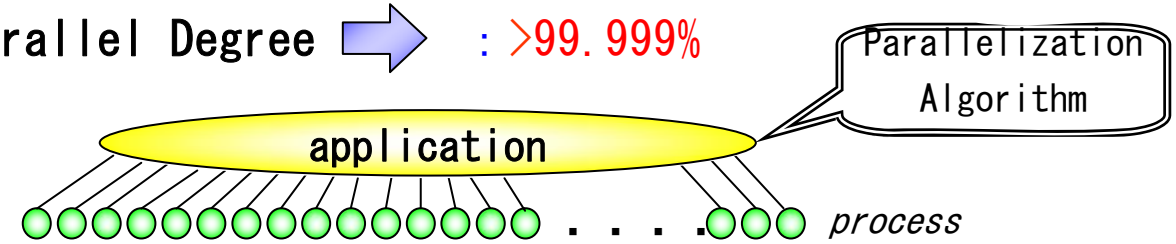
Tuning of Parallelization



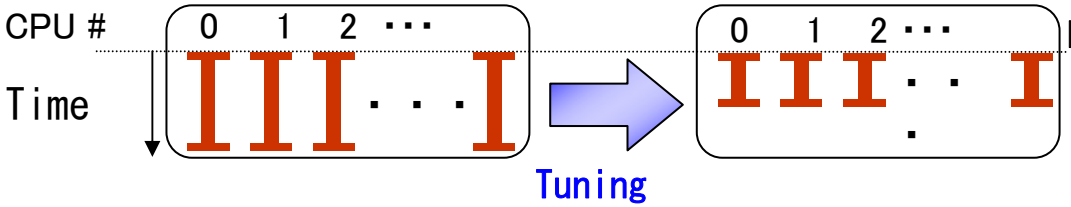
> 100,000 processes



①Parallel Degree ⇨ : >99.999%



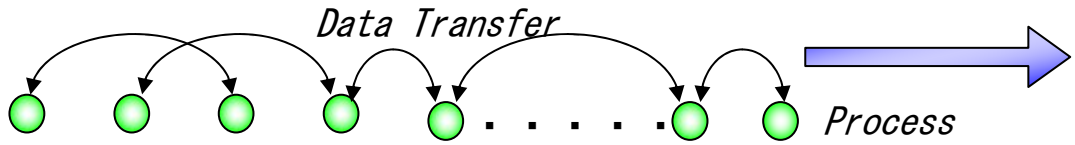
② CPU (Process) Tuning



■Tuning Method

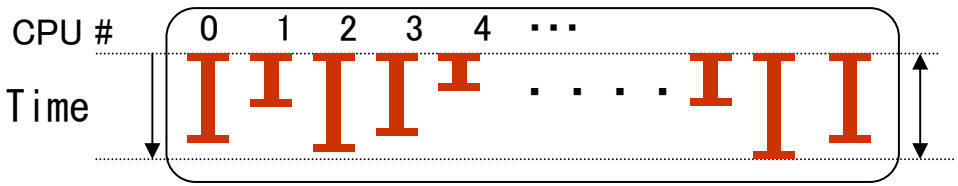
-Cache/Software Pipelining
-Vector Tuning
-・・・

③ Communications



⇨ Decrease Data Transfer

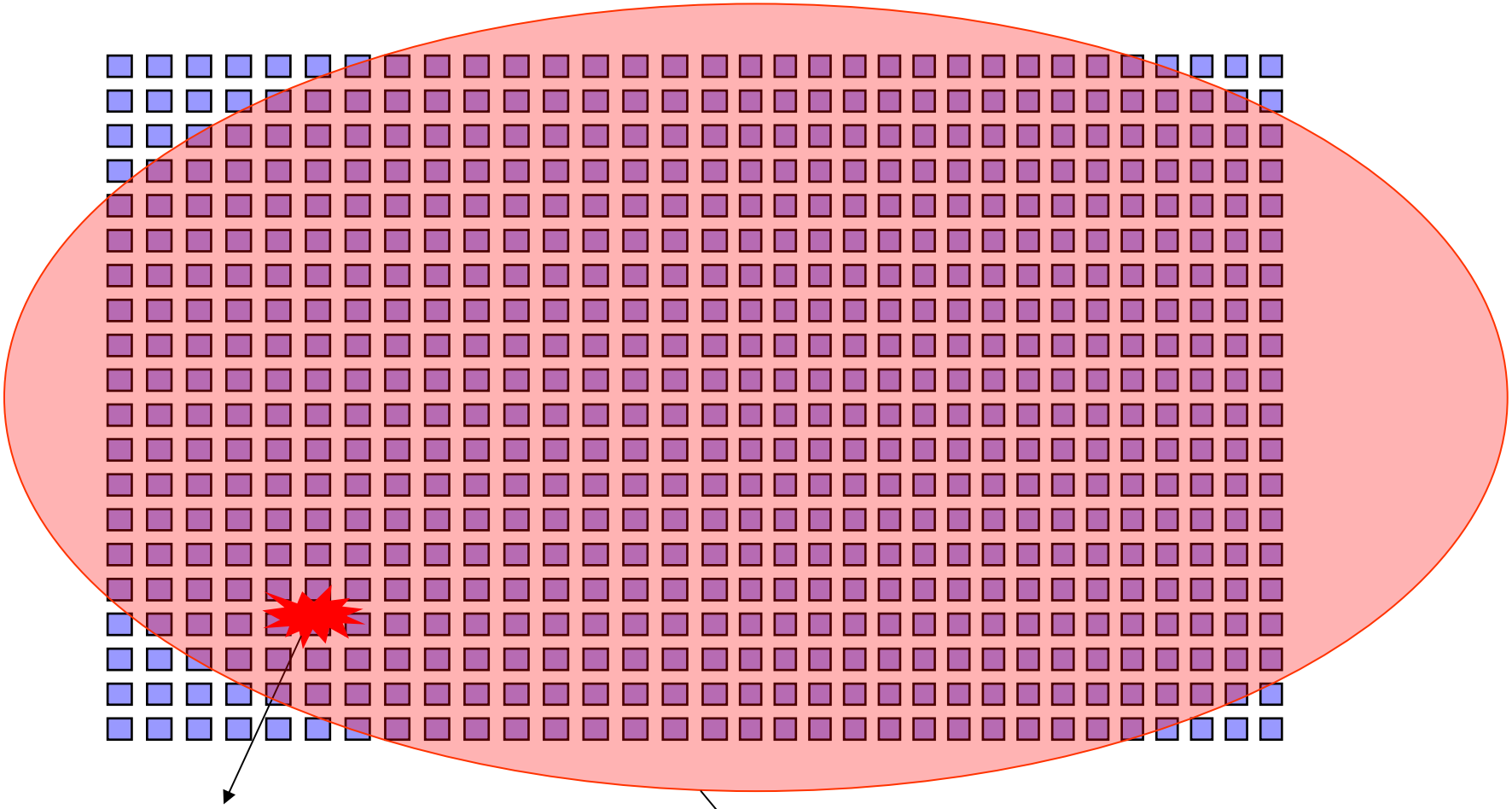
④ Load Balance



Bottle Neck by Most Time Consuming Process

System Reliability

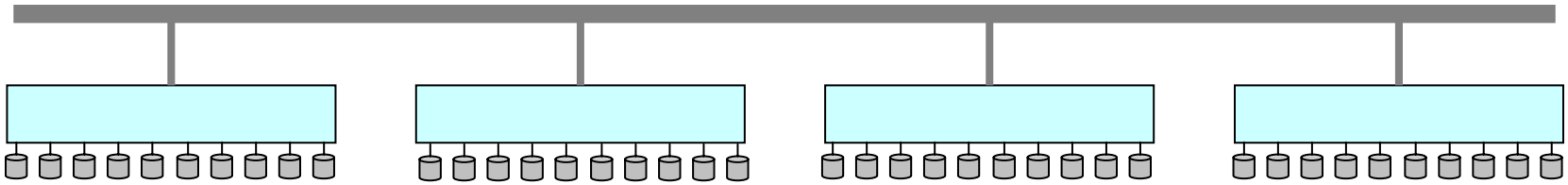
- Minimize Total System Down
- Maintain Data Integrity (Error Checking)

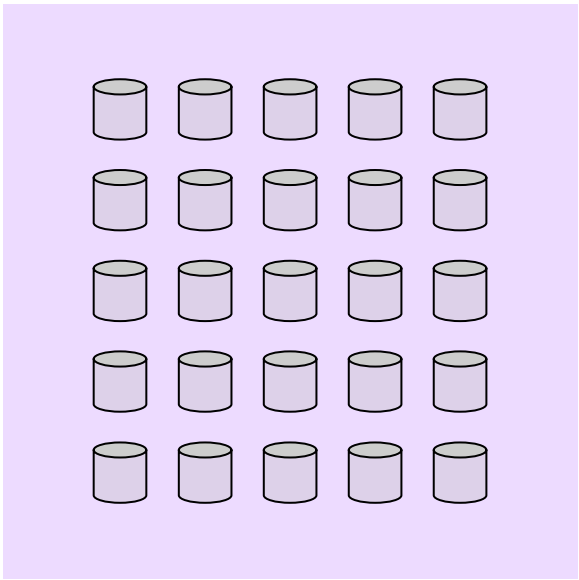


A Single Point of Failure Must not Cause Total System Crash.

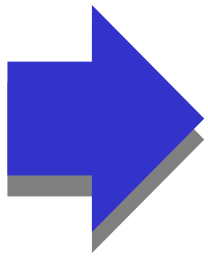
Solid and Intermittent Errors of Data Paths, Control, Arithmetic Units and Memory Must be Strictly Checked

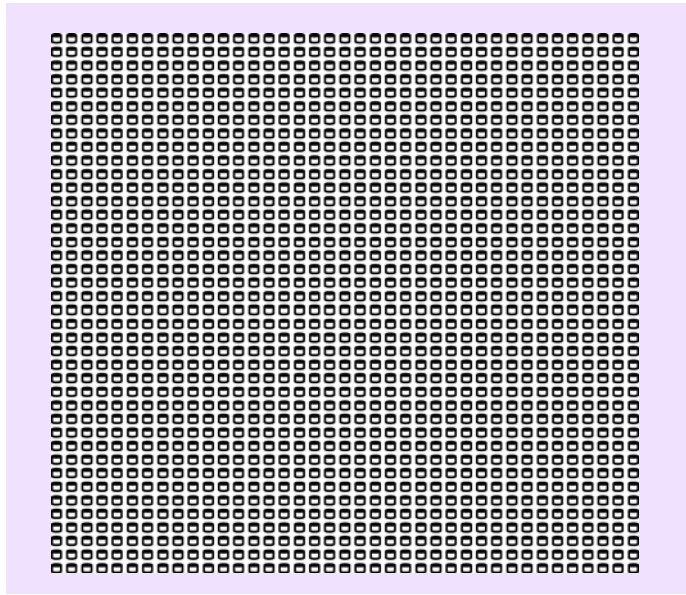
Increasing Storage (Data Explosion)





Peta Bytes





Exa Bytes !

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