Serendipity in Human Biology

~ High performance computing and advanced computational science ~

(Parallel Processing Technology meets Computational Science)

Taisuke Boku
taisuke@cs.tsukuba.ac.jp
Department of Computer Science
Grad. School of Systems and Information Engineering
University of Tsukuba

My Career Path

- In high school days, my dream was “someday, I want to make my own computer system”.
- Got Bachelor, Master and PhD degrees from Faculty of Science and Technology, Graduate School of Keio University
- When I was still a PhD candidate (D3), I became an assistant professor of the same university
  - As assistant professor: Department of Physics
  - As PhD student: Department of Electrical Engineering
- Then moved to Institute of Electronics and Information Engineering (Department of Computer Science today) as lecturer about 20 years ago
- At the same time, Center for Computational Physics (former Center for Computational Sciences) was established and I’ve been working there also for 20+ years
- Currently, Professor at Department of Computer Science and Deputy Director of Center for Computational Sciences

What was a goodness in my career ?

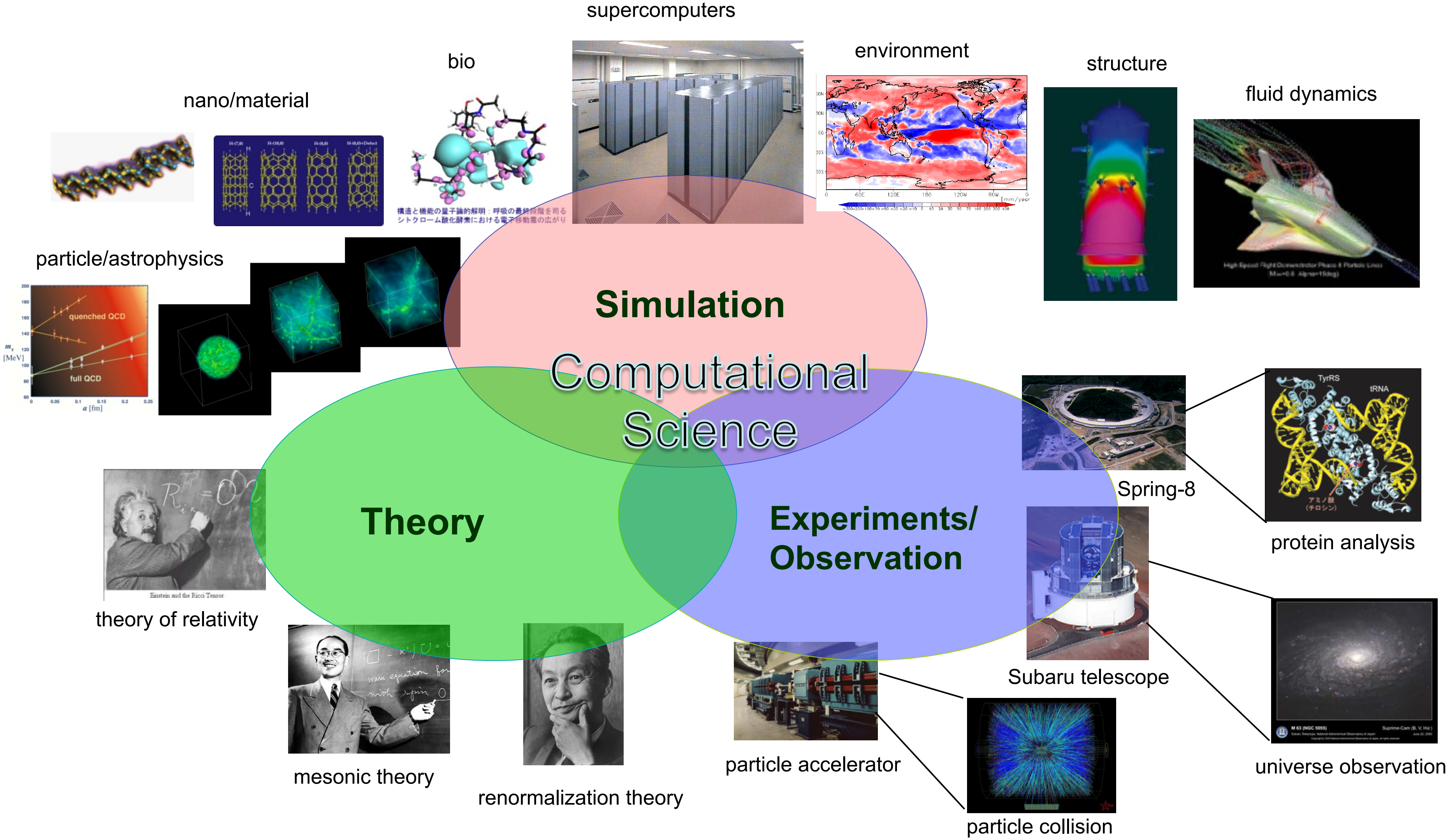
- Working with professional physicists while I was a PhD student was very good opportunity for me to consider what is the really effective and required as supercomputer features
- After getting PhD degree on computer science, I joined to Center for Computational Physics, newly established center in U. Tsukuba
  ⇒ my career as parallel processing researcher who knows physics greatly helped me to get the position
- For 20+ years, I have been working with domain scientists such as particle physics, astrophysics, material science, geo science, life science, ...
  ⇒ good target for real working supercomputers
- For CP-PACS project (1992-1997), I was responsible for interconnection network design and evaluation of the machine which became the world #1 supercomputer (Nov. 1996)
- In 2007, a follow-up project launched as PACS-CS supercomputer, where I was fully responsible for design and implementation
  ⇒ My Dream Comes True!

What is HPC (High Performance Computing) ?

- Today's science (domain science) is driven by three elements
  - Experiment
  - Theory
  - Computation (Simulation)
- In many of these problems, computation performance and capacity are required to be larger and larger
  - Floating point operation speed
  - Memory capacity (amount)
  - Memory bandwidth (memory speed)
  - Network bandwidth (network speed)
  - Disk (2nd storage) capacity
- "High Performance" does not mean only the speed but also capacity and bandwidth

Computational Science

- In all fields of science, “simulation” covers “experiments” and “theory” ⇒ “computational science”



HPC: High Performance Computing

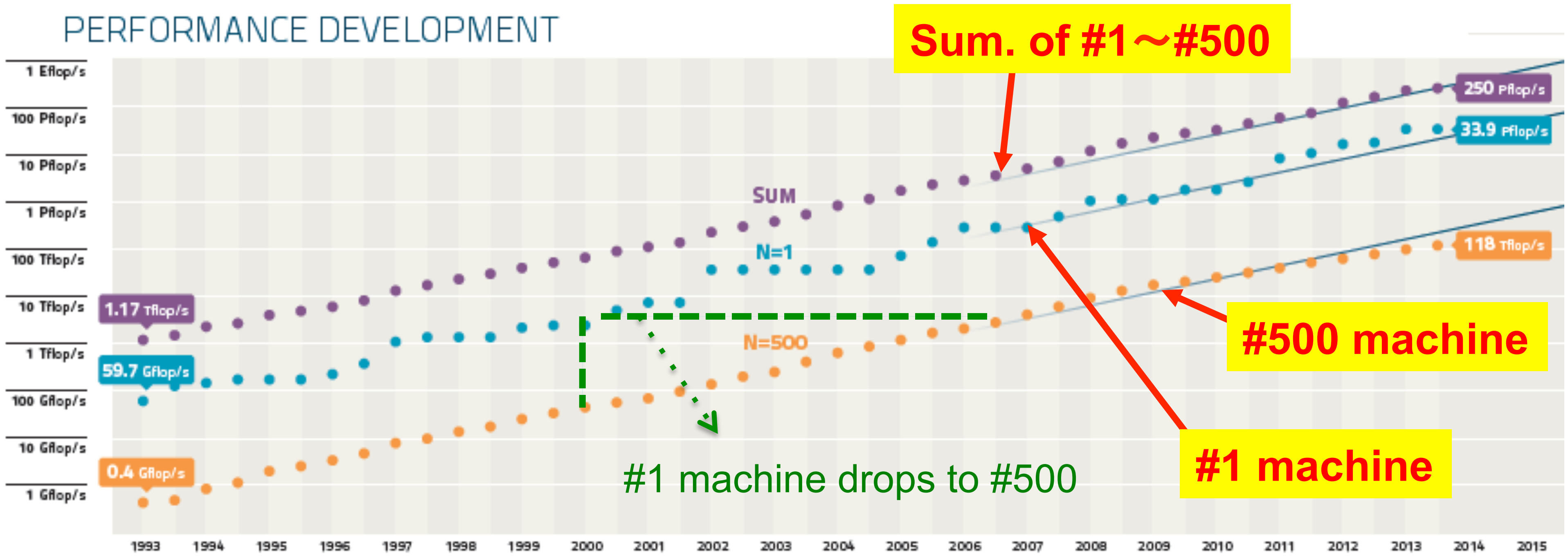
- Solving scientific/engineering problems with very high performance computing resources such as supercomputers
  - Large scale computer simulation for computational sciences (particle physics, astrophysics, material science, drag design)
  - Large scale engineering simulation (car/aircraft design, plant design, disaster simulation)
  - Handling large amount of data (I/O)
- Large amount of scientific calculation
  - FLOPS: Floating Point Operations Per Second
  - Today's peak performance = 55 PFLOPS (5.5×10^{16} FLOPS)
- HPC and Computational Science
  - HPC technology must be always with Real World Applications
  - Application -> HPC: providing real example as driving force of technology and hints for way to go
  - HPC -> Application: providing computation power/capacity to drive advanced scientific simulation
- Both fields of scientists have to collaborate and work together

Performance metric on computation and communication

- computation performance (mainly floating point)
  - FLOP: (number of) Floating point Operations
    number of floating point operations in the processing
    ex) for(i=0; i<100; i++) a[i] = b[i] * c + b[i];
    ⇒ 200FLOP
  - FLOPS: Floating point Operations Per Second
    floating point operations per second -> Performance
    ex) computing the above calculation in 2 micro-sec.-> 100 MFLOPS
    K: 10^3 M: 10^6 G: 10^9 T: 10^{12} P: 10^{15} E: 10^{18}
- communication performance
  - B/s (Byte/sec):
    data transfer amount per second
    ex) theoretical peak performance of Infiniband 4xQDR = 4 GB/s
    sometimes, with bps (bit per second)
    Caution: not always 1Byte=8bit !!

TOP500 List (HPL performance: ~ peak performance)

www.top500.org



	NAME	SPECS	SITE	COUNTRY	CORES	R_{MAX} PFLOP/S	POWER MW
1	**Tianhe-2 (Milkyway-2)**	NUDT, Intel Ivy Bridge (12C, 2.2 GHz) & Xeon Phi (57C, 1.1 GHz), Custom interconnect	NSCC Guangzhou	China	3,120,000	**33.9**	17.8
2	**Titan**	Cray XK7, Operon 6274 (16C 2.2 GHz) + Nvidia Kepler GPU, Custom interconnect	DOE/SC/ORNL	USA	560,640	**17.6**	8.2
3	**Sequoia**	IBM BlueGene/Q, Power BQC (16C 1.60 GHz), Custom interconnect	DOE/NNSA/LLNL	USA	1,572,864	**17.2**	7.9
4	**K computer**	Fujitsu SPARC64 VIIIfx (8C, 2.0GHz), Custom interconnect	RIKEN AICS	Japan	705,024	**10.5**	12.7
5	**Mira**	IBM BlueGene/Q, Power BQC (16C, 1.60 GHz), Custom interconnect	DOE/SC/ANL	USA	786,432	**8.59**	3.95

What is parallel processing ?

- "Decomposing single problem with in a number of processes and solving it to enhance the performance and/or increase problem size"
  - "Solving single problem" ⇒ differs from distributed processing
  - "Problem decomposing (parallelizing)" ⇒ careful for efficiency
  - "Improved" issues ⇒ not just speed, but also problem size, computing accuracy, etc. (various metrics)

- resources to contribute for parallel processing
  - CPU, memory, disk, network, etc. ⇒ all the computation resources may contribute for improvement
  - hereafter, we call these processes to be mapped to multiple CPUs as "parallel processes"

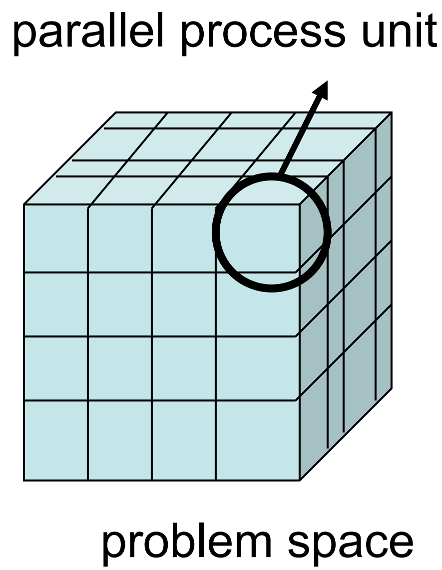
High performance computing (HPC) and parallel processing

- Requirement for numerical computing performance in scientific computation and large amount of data processing
- Computing order to increase the problem size N is not O(N) (linear)
  - HF matrix calculation on molecular orbital computation: $O(N^4)$
  - 3-dimensional fluid dynamic (climate simulation etc.)
    when spatial resolution on 1-dimension is N, computing operation's order is $O(N^3)$
  - matrix calculation (linear equation)
    for direct method (Gaussian elimination etc.) for N variables of equation, computing operation's order is $O(N^3)$, but reduced up to $O(\alpha N^2)$
  - n-body problem (gravity calculation in astrophysics)
    force computation for N particles requires computing operations with $O(N^2)$
  - FFT (Fast Fourier Transform): $O(N \log N)$
- There is no "enough performance and amount" for the requirement
  ⇒ large scale scientific computation does no more exist without Parallel Processing

Example of data parallel computing

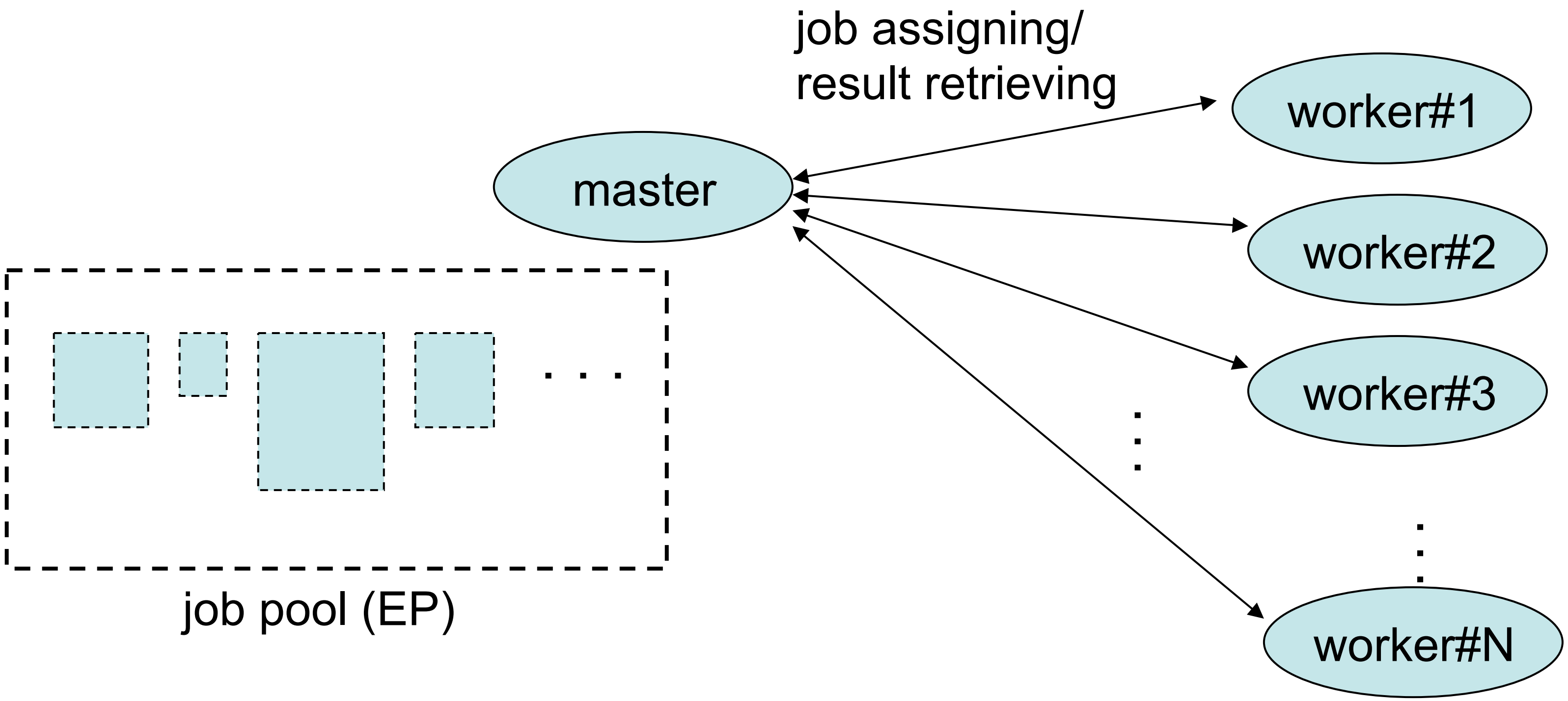
- domain decomposition
  - Calculating points are uniformly distributed in some dimensions of space, and partitioning them into orthogonal blocks to be parallelized
  - There are some communication required to exchange data
    ex) for PDE with explicit method, surface points data are exchanged with neighbors
  - ex (1-dimension))

```
for(t=0; t < T; t++){
   for(i=0; i < N; i++)
      a[i] = b[i-1] + 2*b[i] + b[i+1];
   for(i=0; i < N; i++)
      b[i] = a[i];
}
```



Master/Worker Parallelization

- Especially effective when the loads of processes are not balanced and it is difficult to keep load balance
- Each process should be EP



CCS at University of Tsukuba

- Center for Computational Sciences
- Established in 1992
  - 12 years as Center for Computational Physics
  - Reorganized as Center for Computational Sciences in 2004
- Daily collaborative researches with two kinds of researchers (about 30 in total)
  - Computational Scientists
    who have NEEDS (applications)
  - Computer Scientists
    who have SEEDS (system & solution)

CCS（cont’d）

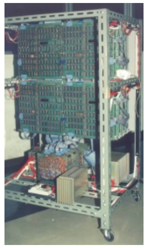
- Application field divisions
  - Particle Physics
  - Astrophysics and Nuclear Physics
  - Quantum Condensed Matter Physics
  - Life Science
  - Biological Science
  - Global Environment Science
- Computer system field divisions
  - High Performance Computing System
  - Computational Informatics
- Not a general “Computer Service Center”
  - CCS is a research oriented center
  - Developing our original supercomputers (not just procuring vendor’s machines)
  - Daily collaboration among application and system fields

History of parallel computer PAX(PACS) in U-Tsukuba

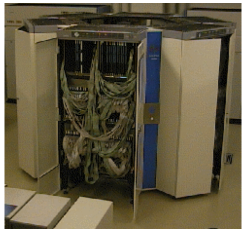
1978
1: PACS-9

Started by Prof. Hoshino and Kawai

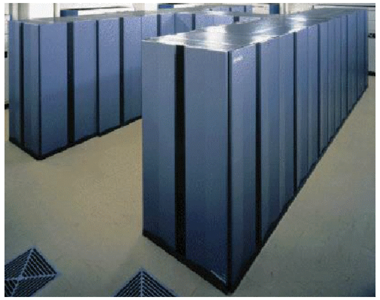
1980
2: PAXS-32



1989
5: QCDPAX



1996
6: CP-PACS



Top1 in Top500 List

2006
7: PACS-CS

Service out in last Sep.

Year	Name	Performance
1978	PACS-9	7KFLOPS
1980	PAXS-32	500KFLOPS
1983	PAX-128	4MFLOPS
1984	PAX-32J	3MFLOPS
1989	QCDPAX	14GFLOPS
1996	CP-PACS	614GFLOPS
2006	PACS-CS	14.3TFLOPS
2012	HA-PACS	800TFLOPS

- Cooperation with Computational Scientists and Computer Engineers
- Target performance driven by application
- Continuous development with experience accumulation

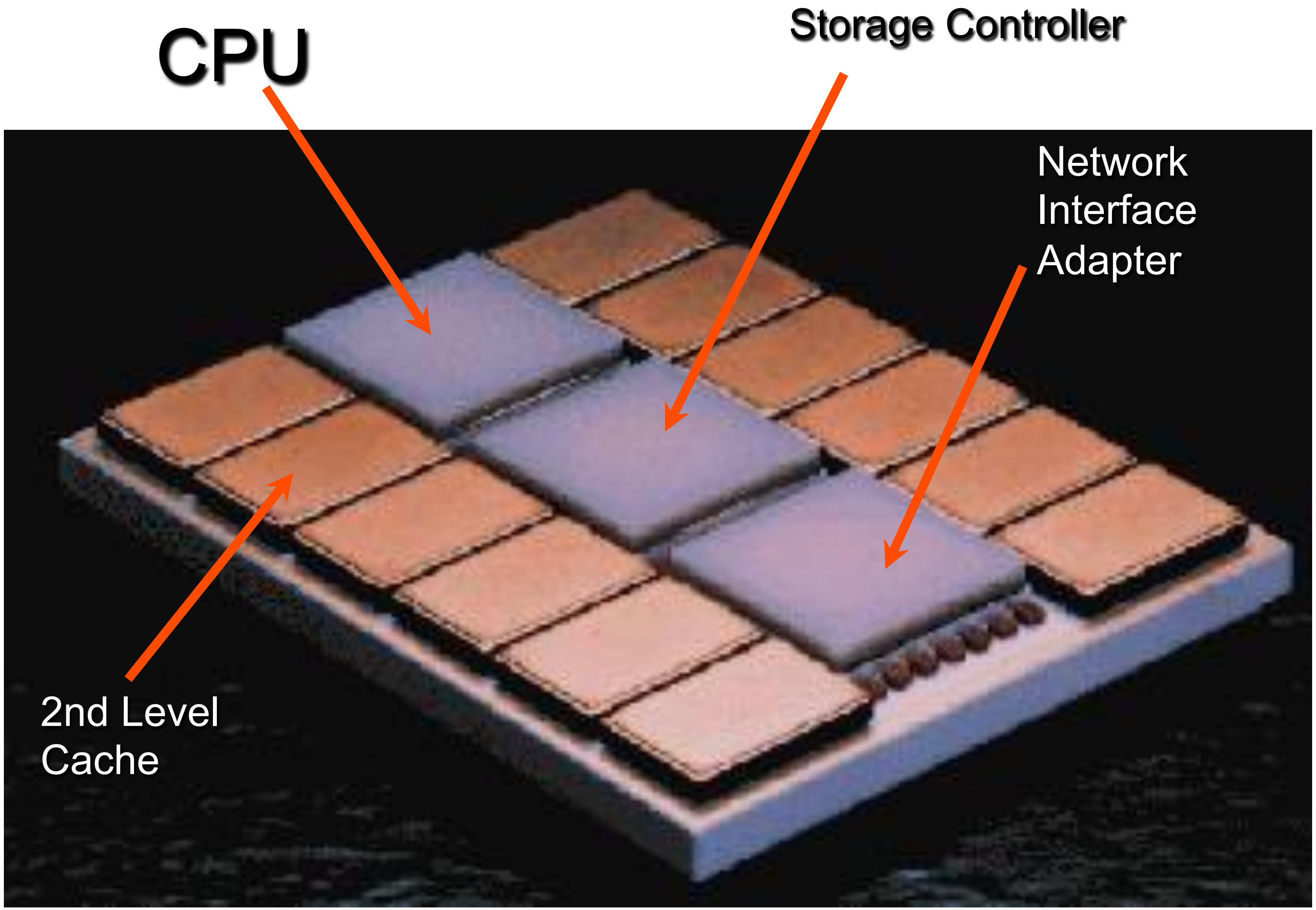
PAX (PACS) Series

- MPP system R&D continued at U. Tsukuba for more than 30 years
- Coupling of need from applications and seeds from the latest HPC technology, the machines have been developed and operated with the effort by application users on programming
  → a sort of application oriented machine
  (not for a single application)
- HA-PACS is the first system in the series to introduce accelerating devices (GPUs)
- CCS has been focusing on the accelerating devices for ultra high performance to provide to “high-end” users who require extreme computing facilities

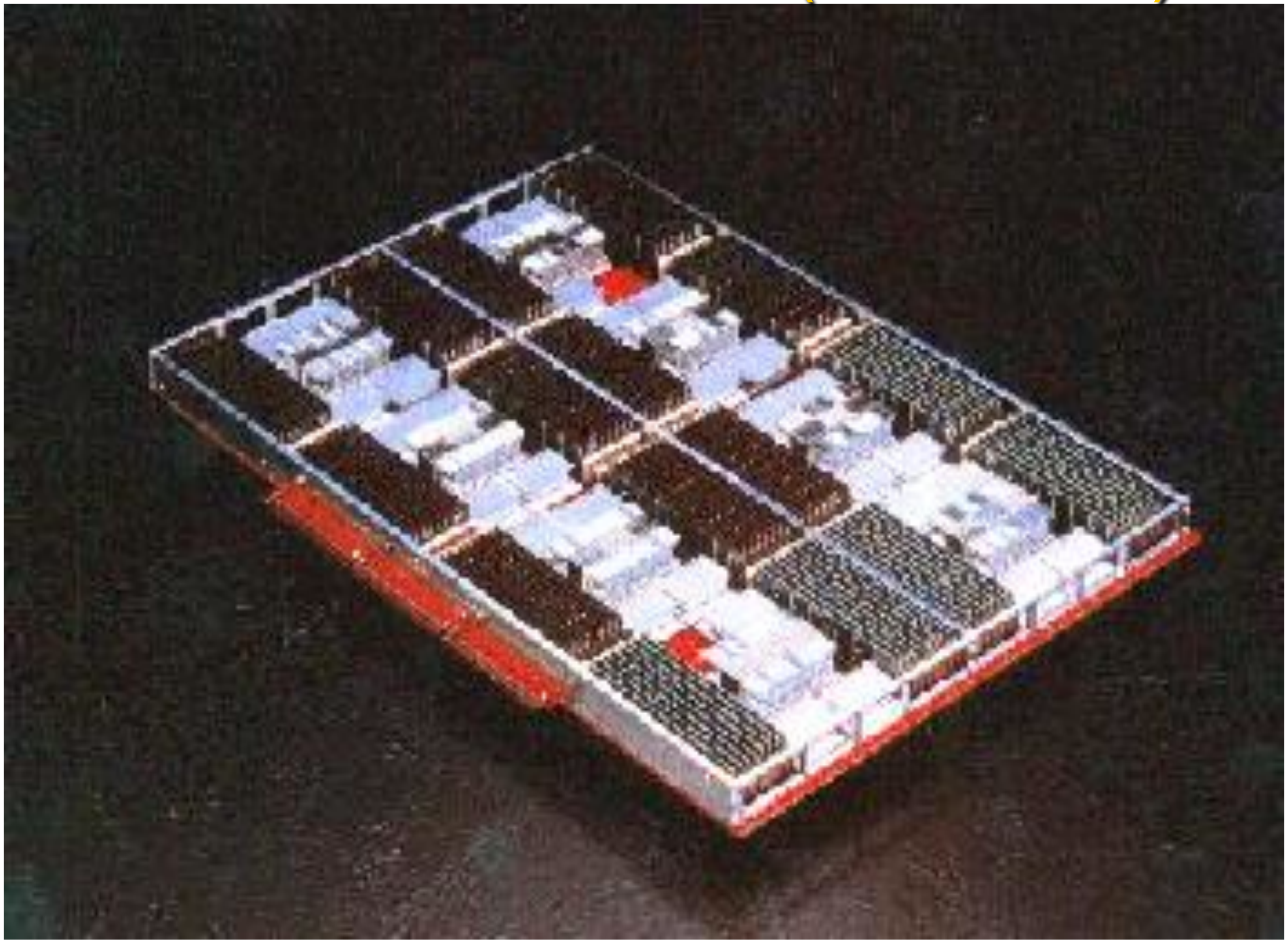
CP-PACS (1996 Univ-Tsukuba)

- First large scale massively parallel supercomputer developed in Japan
  - ✓ Scalar processor with pseudo vector
  - ✓ Flexible and high performance network
- Collaboration with physics and computer science.
- Collaboration with university and vender (Hitachi), Hitachi developed SR-2201 based on CP-PACS
- Scientific breakthrough in particle physics and astrophysics
  - ✓ First principle calculation for QCD
  - ✓ General simulation model for field (fluid, electromagnetic field, wave function, etc)

Multi-Chip Module on Computation Node of CP-PACS



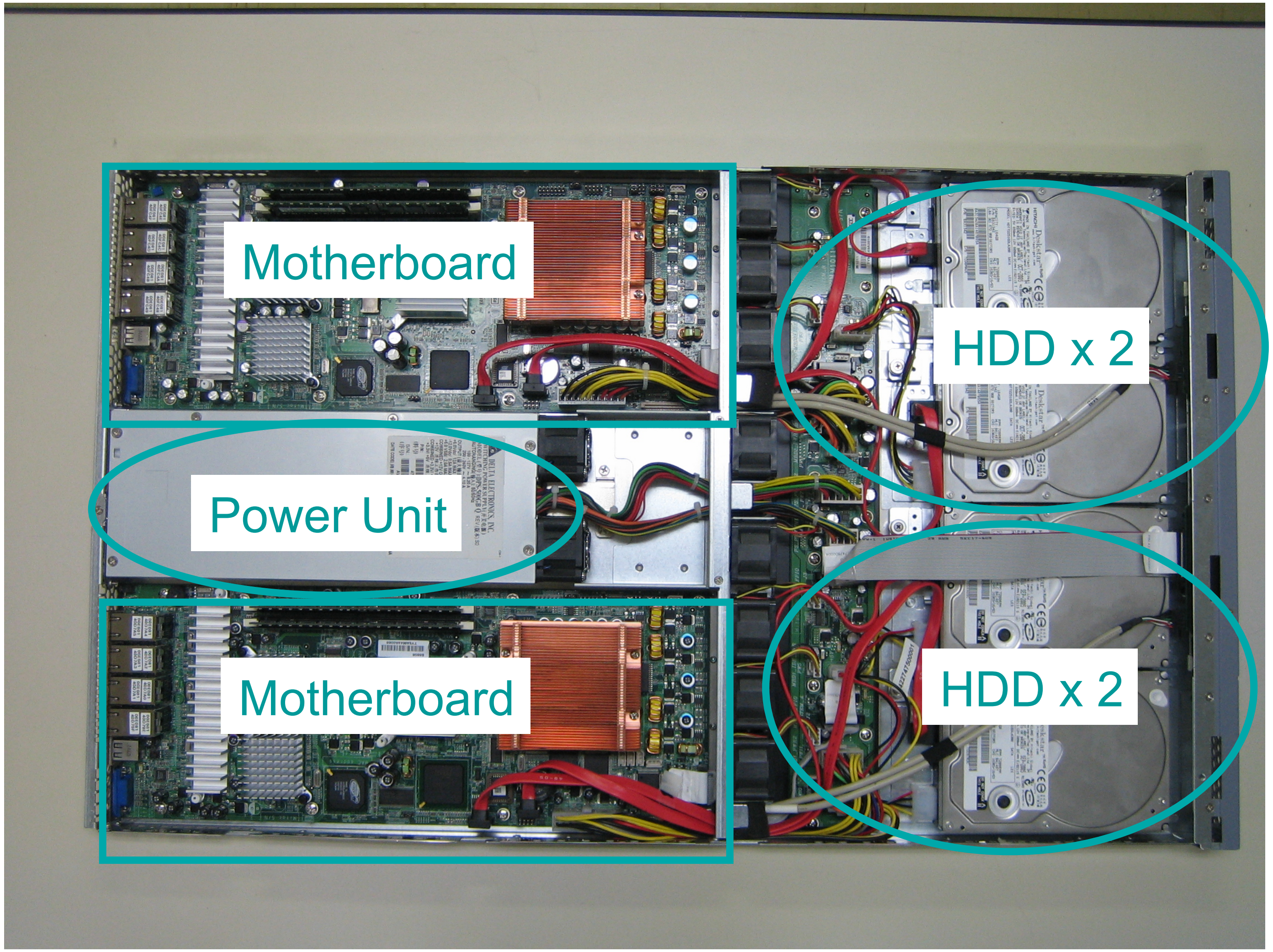
CP-PACS PU board (with 8 PU)



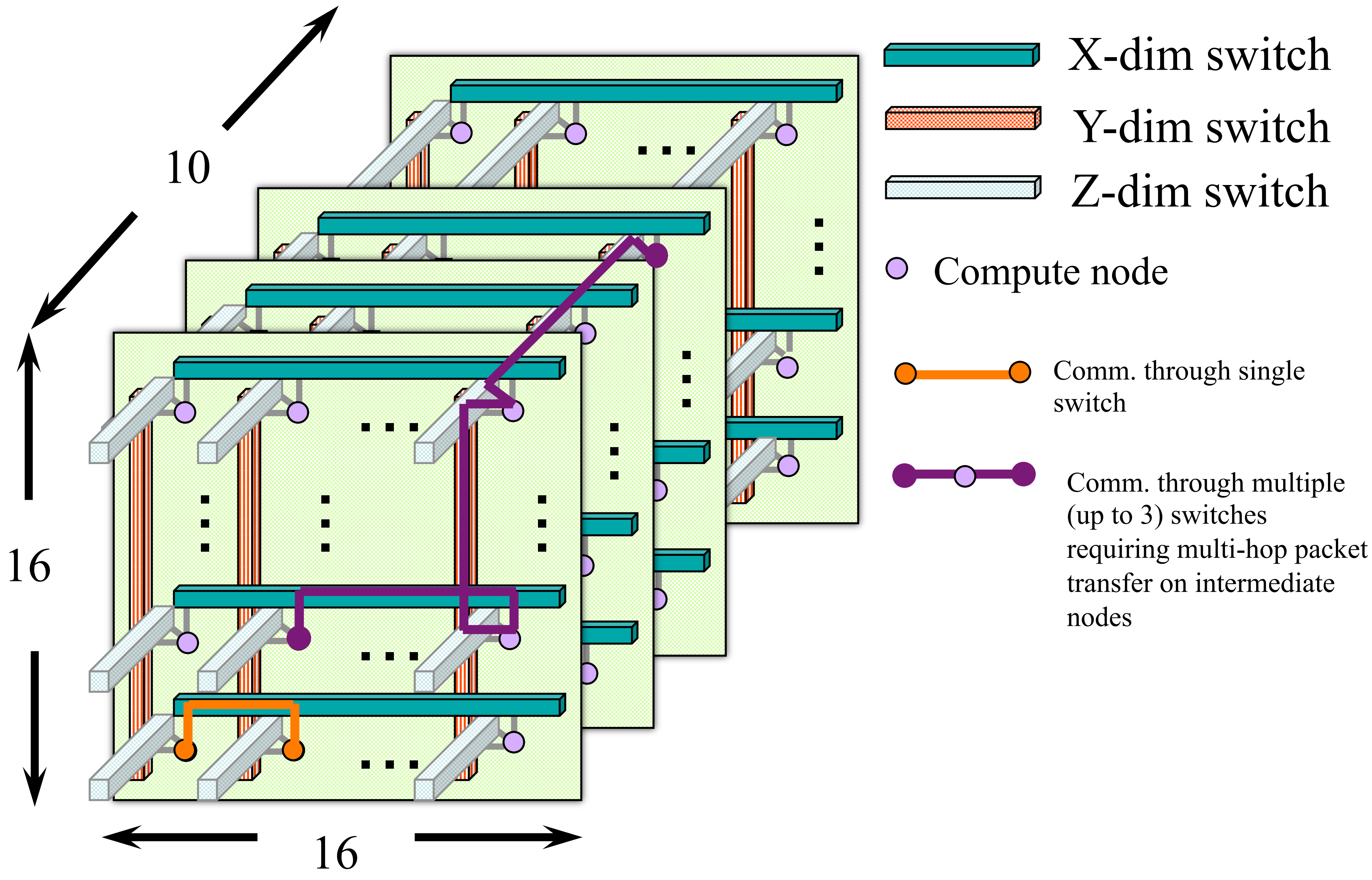
Full system of PACS-CS (2560 nodes)

#34 at TOP500 on June 2006 (Linpack: 10.5 TFLOPS)

Unit chassis (19inch x 1U)



3D-HXB (16x16x10=2560 node)



HA-PACS Base Cluster + TCA
(TCA part starts operation on Nov. 1st 2013)

- HA-PACS Base Cluster = 2.99 TFlops x 268 node = 802 TFlops
- HA-PACS/TCA = 5.69 TFlops x 64 node = 364 TFlops
- TOTAL: 1.166 PFlops

HA-PACS/TCA computation node inside

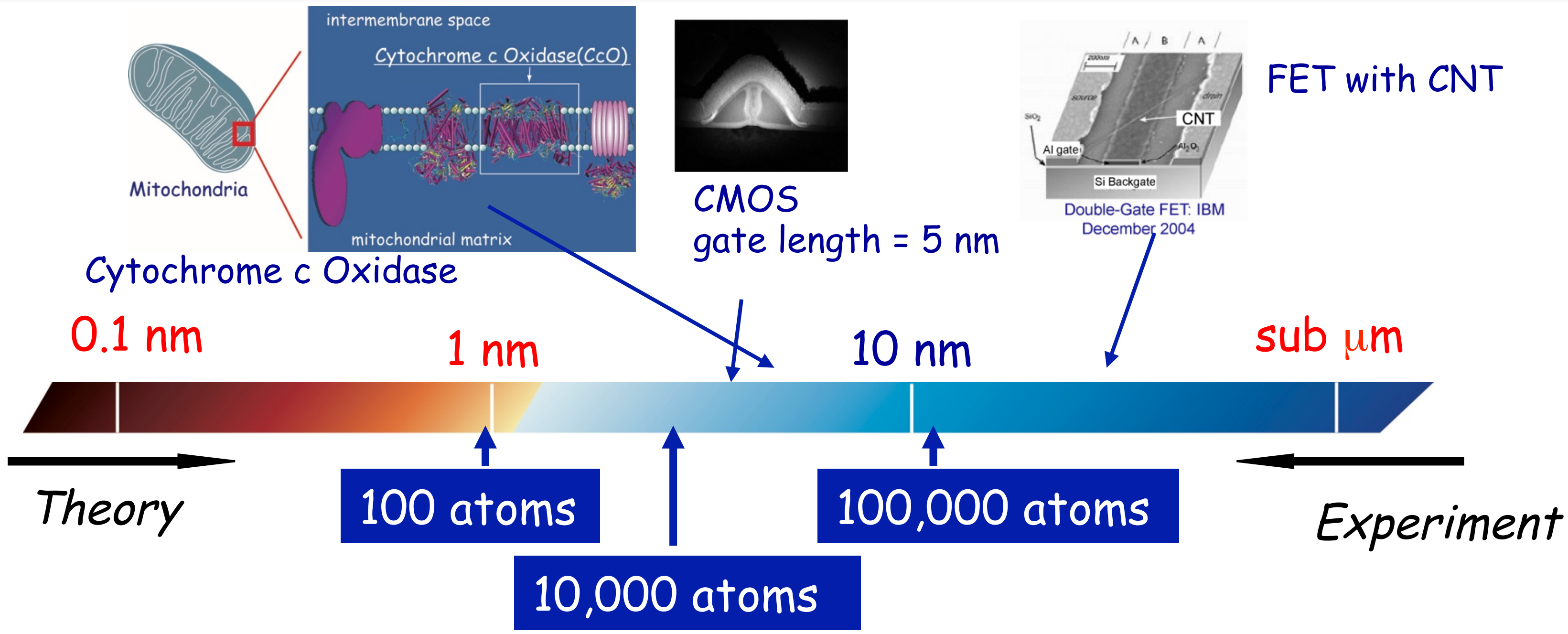
HPC and Computational Science in U. Tsukuba

- Center for Computational Sciences strongly proceeds the collaborative research between HPC researchers (system) and domain scientists (application)
- CCS has been developing large scale parallel processing systems under collaboration of these researchers for 20 years (origin is the start of PACS/PAX series 30+ years ago)
- CCS is recognized as a pioneer and leading institute for such a style of research in Japan
- We will support any of computational science research in our university
- All professors in HPC Division of CCS belong to "High Performance Computing System Lab." at Department of Computer Science

Why top-level HPC today is so difficult

- High end HPC system (supercomputers) are too complicated, too large and too hard to utilize (program) with high efficiency
- Advanced simulation is also too complicated to be understood by computer scientists and hard to tune without detailed knowledge on code behavior
- Finding out the problem on performance is essential for big success, and it requires a tight collaboration between system side and application side researchers

Large-scale first-principles calculations in nano science



Large-scale DFT calculations and experiments meet together in Nano World !

- Materials = Nuclei + Electrons (Quantum Objects)
- Density Functional Theory (DFT) = Efficient framework based on first-principles of quantum theory, but limited to N =100～1000-atom systems before

Challenge: 10,000 ～ 100,000-atom calculations overcoming N^3 scaling to reveal nano-scale world!

2011 ACM Gordon Bell Prize for
Peak Performance Award

"First-principle calculation of electronic states of a silicon nanowire with 100,000 atoms on the K computer"

What was good in this collaboration ?

- Collaboration with computer scientists (HPC) and computational scientists (CM = condensed matter science)
- High cache-hit ratio techniques are provided by HPC people and CM people rewrote their codes
  => “introducing common sense for HPC but not for MC”
- “Real Space” method is provided by MC people when HPC people just watched the original code only, resulting large reduction of memory capacity
  => “back to principle of computation”
- “Orthogonalization” is straight forward numerical solution, but our HPC team’s solution was not
  => “think differently”
- Really effective and practical high performance computation must be performed with both “system” and “application” researchers
  => “co-design”

Message

- Keep your dream and always prepare to realize it.
- Think what is needed for your work.
- Collaboration across fields helps the work although sometime it is tough.
- Straight forwarding is the basic research style, but think or watch things from “different angle”.
- “ 1 + 1 != 2 (possibly)”