Report from Japan

Hirofumi Tomita, Mitsuhisa Sato, Yutaka Ishikawa

Riken AICS

U. of Tsukuba
Riken AICS

U. of Tokyo
Riken AICS

February 27, 2014

Feasibility Studies on Future HPC R&D in Japan FY2012 – FY2013

1 application study team

Hirofumi Tomita

Satoshi Matsuoka

RIKEN AICS and TITECH
Collaboration with application fields

- Identification of scientific and social issues to be solve in the future
- Drawing science road map until 2020
- Selection of the applications that plays key roles in the roadmap
- Review of the architectures using those applications

3 system study teams

Hiroaki Kobayashi

Mitsuhisa Sato

Yutaka Ishikawa

Tohoku Univ. and NEC

U. of Tsukuba, Titech, and Hitachi

U. of Tokyo, Kyushu U., Fujitsu, Hitachi, and NEC

- Design of computer systems solving scientific and social issues
- Identification of R&D issues to realize the systems
- Review of the system using the application codes
- Estimation of the system's cost

Feasibility Studies on Future HPC R&D in Japan FY2012 – FY2013

1 application study team

Hirofmi Tomita

Satoshi Matsuoka

RIKEN AICS and TITECH
Collaboration with application fields

- Identification o
  to be solve in t
- Drawing scienc
- Selection of the
  roles in the roa
- Review of the a
  applications

3 system study teams

Hiroaki Kobayashi

Mitsuhisa Sato

Yutaka Ishikawa

Tohoku Univ. and NEC

U. of Tsukuba, Titech, and Hitachi

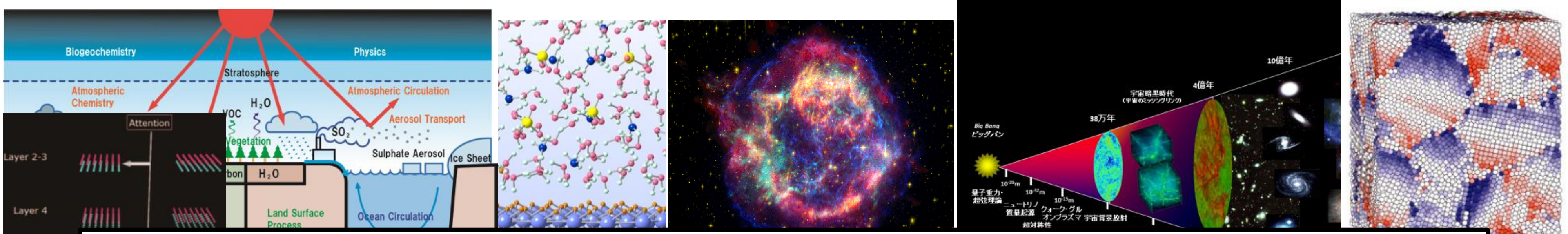
U. of Tokyo, Kyushu U., Fujitsu, Hitachi, and NEC

r systems solving
l issues
D issues to realize the
em using the
ystem's cost

Based on these studies, RIKEN proposes the successor of K computer, postK

Three Talks

- Applications
  - Hirofumi Tomita, RIKEN
- Accelerator
  - Mitsuhisa Sato, University of Tsukuba
- General Purpose
  - Yutaka Ishikawa, University of Tokyo

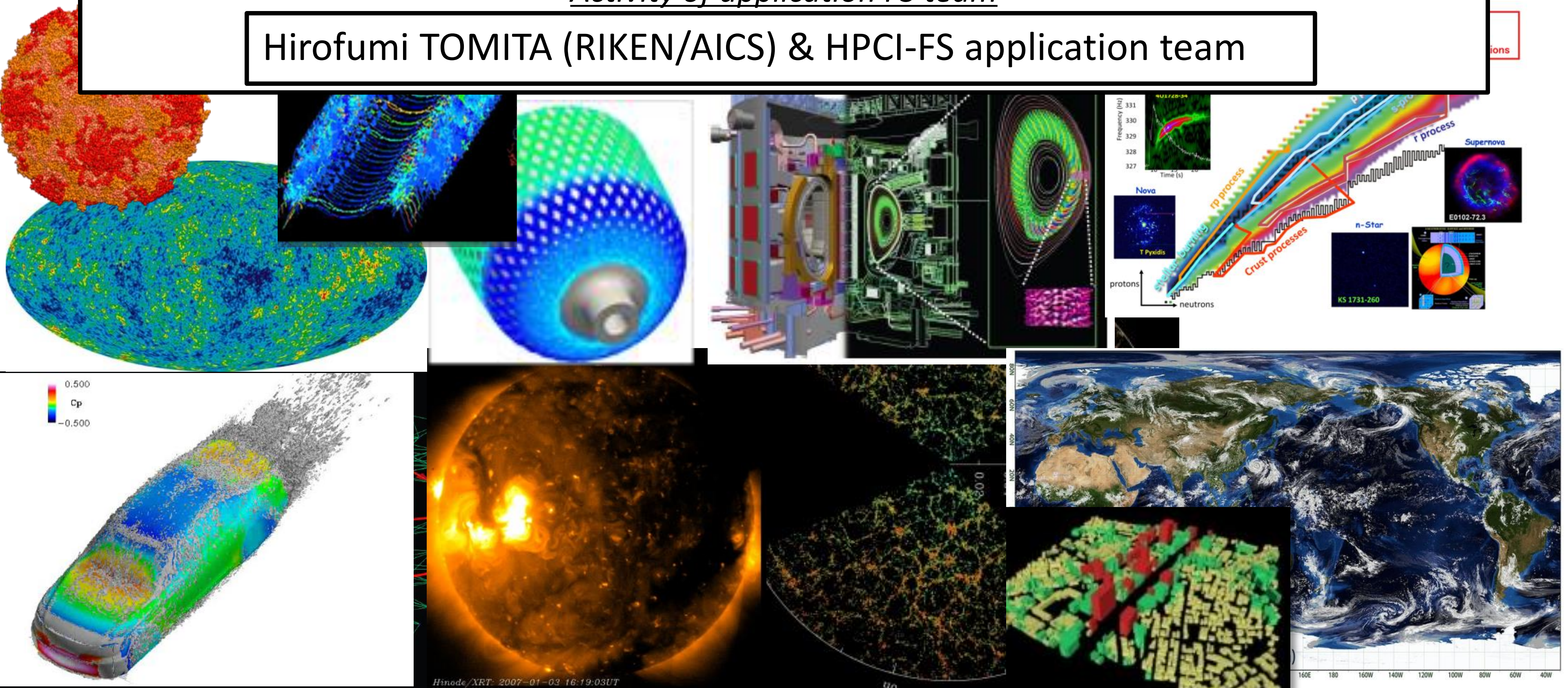


Future social and scientific subjects resolved by science-driven HPC

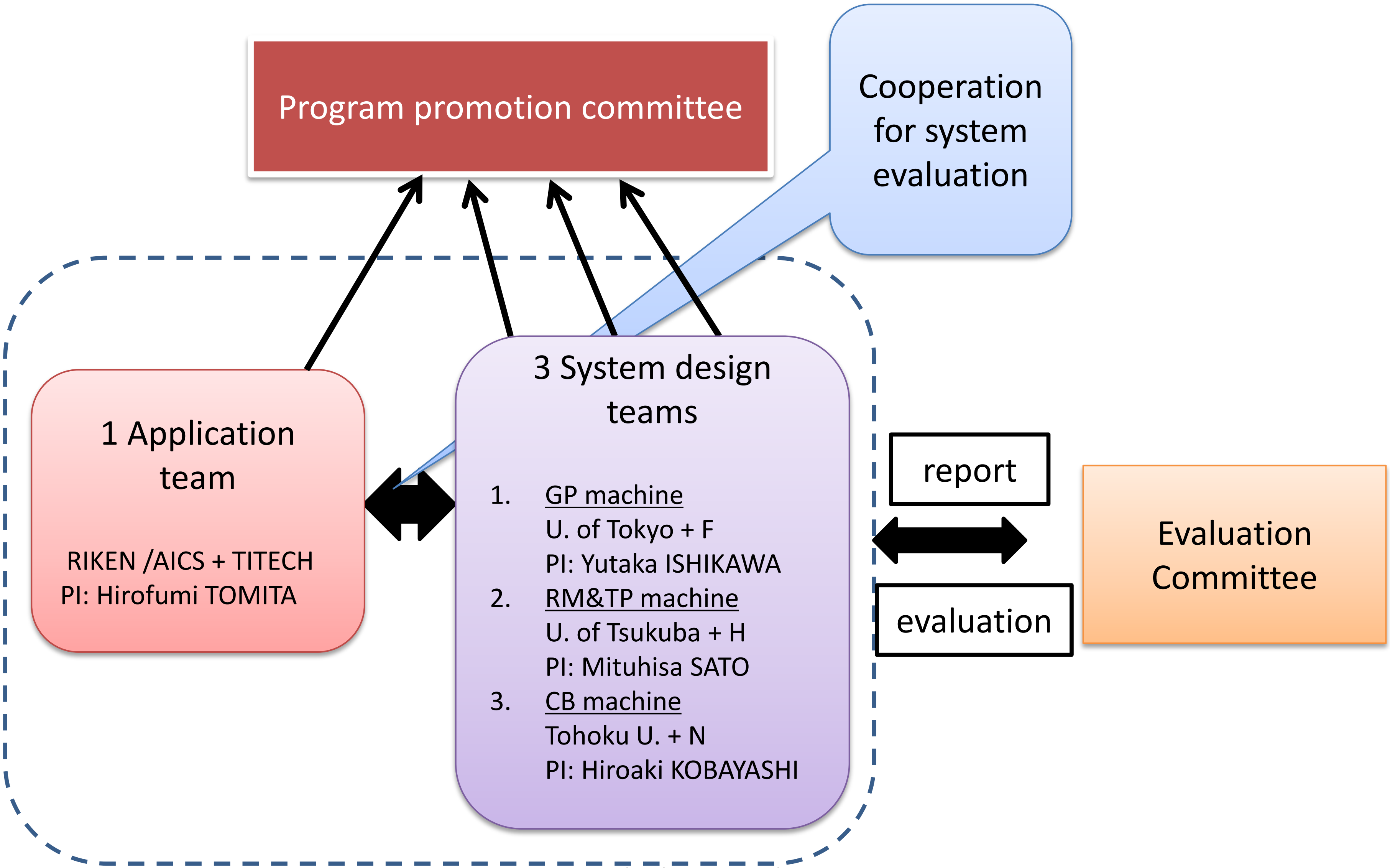
Report on Feasibility Study on Future HPC Infrastructure

Activity of application FS team

Hirofumi TOMITA (RIKEN/AICS) & HPCI-FS application team



Organized structure of HPCI-FS project



Position & Role of Application team

<u>Main mission of application FS team</u>

- Extraction of scientific and social subjects and discuss about contribution from HPC
- Proposed-system evaluation in cooperation with 3 system-FS teams.

From Application side,
we can image the exaflops machine, but,
only consideration of an exaflops machine is not always our purpose.
Out porpose : extraction of social/scientifc subjects and consider the efficient way for resolving these subjects with system teams

HPCI Feasibility Study from the application community

- Start from July 2012
  - Actually, discussion started in FY2011(FY2011 application working group)
- Discuss the contribution from HPC *to social and science field* with many disciplines people
  - Not only existing major field(life science, material science, earth science, manufacturing, Fundamental physics)
  - But also **a new discipline** (may use the HPC intensively near the future)
    - e.g. social sciences
- Many institutes, universities, and enterprises are attending
  - **More than 100 research scientists in 36 organization** related to HPC application

Application FS Activity

- General meetings are held once by 3 months, including 3 system FS teams.
  - What is discussed about?
    1. What is the current social subjects and scientific problem?
       - Extraction of the social and scientific subjects
         - » e.g. disaster prevention, drug design, and so on
    2. How and what can the computational science contribute directly to such subject?
    3. Subsequently, how does the science change essentially or qualitatively?
    4. How can we evaluate the proposed-systems from the application side?
- Open to other communities (inside & outside)
  - Not close within the HPC community!
    - Discussion with people of the experimental, observational, and theoretical people
    - Discussion with the communities of each discipline
      - E.g. special session are organized

General consensus through the discussion

- To resolve the social subjects,
  - it is needed that more tightly collaboration over the existing disciplines (& new discipline)
    - Shuffling and reconstruction of research scheme according to the social subjects
- On the other hand, conventional scheme, which has been sophisticated so far in each of science areas, also should be enhanced.
  - Get comprehensive knowledge by understanding of nature.

Important point : the two ways (basic science research and social outreach) should be balancing
=direct contribution and indirect contribution

Science roadmap using the future HPC system from the viewpoint of computational science

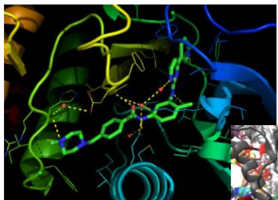
- Full version (tentative):
  http://hpci-aplfs.aics.riken.jp/document/roadmap/roadmap_1309.pdf
  - Over 300 pages:
    - Including future scope of each discipline
- Overview version (tentative):
  http://hpci-aplfs.aics.riken.jp/document/roadmap/roadmap_abst_1309.pdf
  - Just 30 pages:
    - Focusing on the subjects that should be resolved as soon as possible
    - Including summary of each disicpline's scope
- Within this FY, the final version is fixed!

The English version is available soon!

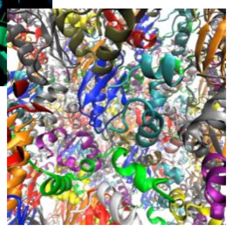
Four social core-subjects

2. Social and Scientific Problems in Computational Sciences

Drug Discovery・Health Care

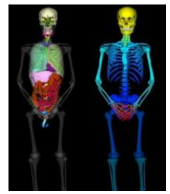


Interaction between anti-cancer drug Glivec and protein



Drug design in cell environment

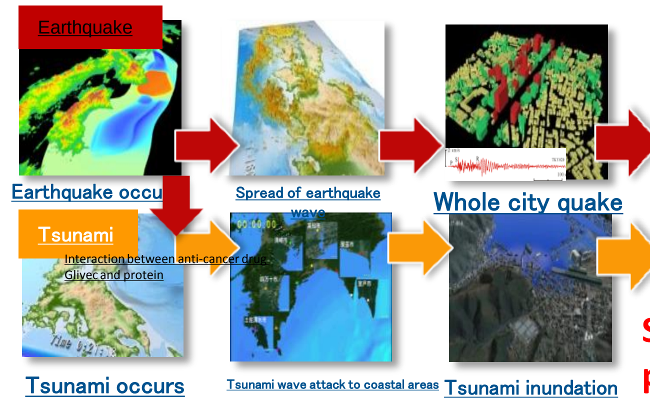
Detailed simulation of organs （京）



Whole body scale simulation

Innovation of drug design and medical technology

Disaster Mitigation and Reduction

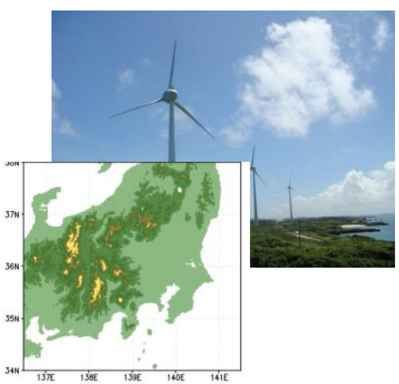


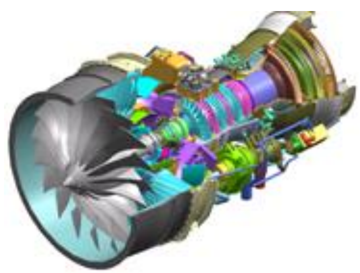
Estimation on indirect damage

Collaboration with Social Science

Systematic disaster prediction based on scientific studies

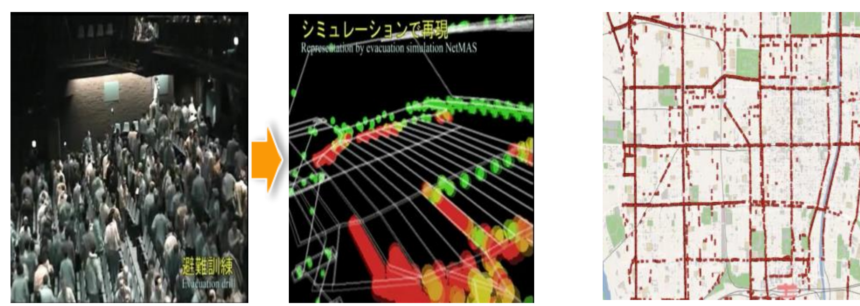
Energy & Environmental problem





Promoting advanced energy technologies in harmony with environment

Social Science



Social forecasting system (integration of each compnent, trafic, social econimics, searching the evacuation route at disaster)　: seeds research?

Enhancement of new science
by collaboration beyond each of discipline

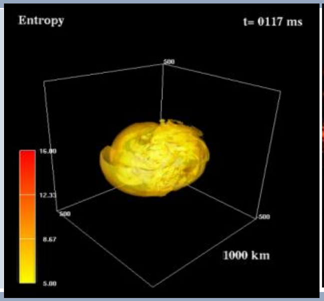
3. Creation of new sciences from multi-disciplinary collaboration

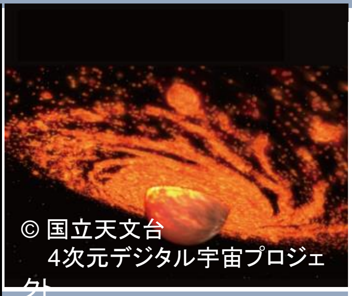
3.1 Cooperation with basic sciences and Unified understanding

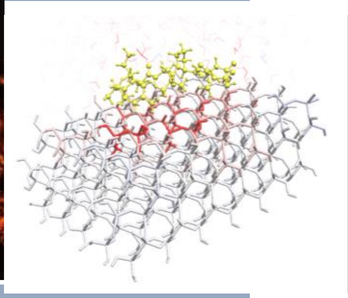
(1) New theoretical paradigm (elementary particle physics, nuclear physics and astrophysics)

(2) Planetary science (astrophysics and earth science)

(3) Analysis based on biomolecules, complex structure (life science & material science)





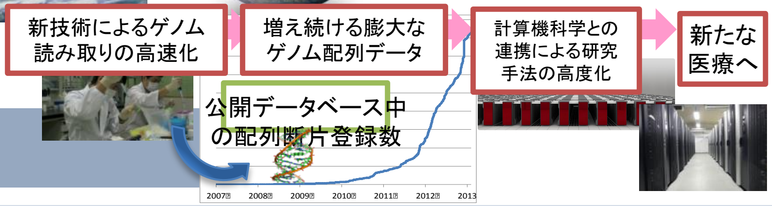


3.2 Efficient use of big data

(1) Efficient use of satellite and observation data

(meteorology & computer science)

(2) Genomic analysis (lifescience & compuer science support)



3.3 Cooperation with experimental facilities

Cooperation with life Science and large scale experimental facilities such as SACLA*

* SACLA(SPring-8 Angstrom Compact free electron LAser

e.g. Imaging of 4D structure

- Structural analysis non-crystalline particle and fine crystal
- Analysis hierarchic dynamics of bioparticle

4. Deep understating for nature in each of sciences

- Basic study from each of disciplines. typically,
  - life science,
  - material science,
  - earth science,
  - manufacturing, engineering,
  - fundamental physics,
  - social science
- See the full version of roadmap
  - Full version (tentative):
    http://hpci-aplfs.aics.riken.jp/document/roadmap/roadmap_1309.pdf
- English version is coming soon!

Mini-App Benchmarks activity

- To evaluate the system in a general sense, Mini applications may be useful. (philosophically, independent of specific architecture)
- Now, we prepare
  - About 30 full applications from many computational sciences with agreement on copyright and licenses
  - Detailed studies of several full applications
    - Molecular dynamics（MARBLE、MODYLAS）
    - Gene network estimation（SiGN-L1)

Plan of preparation by the end of this FY

MD
    Two variants: FFT-based and FMM-based ones (“Marble” and “Modylas”, respectively)
QCD
    “CCS-QCD”
CFD
    “FFVC”
Genome sequence matching
    “NGS Analyzer”
First principles density functional theory
    “CONQUEST” NTChem
Quantum Monte Carlo
    “ALPS/looper”
Climate simulations
    Spectral and icosahedral models (“NICAM” and “DCPAM”)
Brain
    NEURON k+

Summary

- Computational roadmap
  - Until the end of March, we will finish.
  - But, continuing that we will discuss it after FS terms.
- Mini-application
  - For machine evaluation, we will enhanced much more, extending other types of application.
- Efficient use of BIG DATA is one of keys for future HPC.
  - e.g.

    severe rainfall prediction with data assimilation using the phased array (quick scan facility)

    - Possibility of real-time prediction of each of cumuli(10km scale)

Feasibility Study
on Future HPC Infrastructure

-- Study on exascale heterogeneous systems with accelerators --

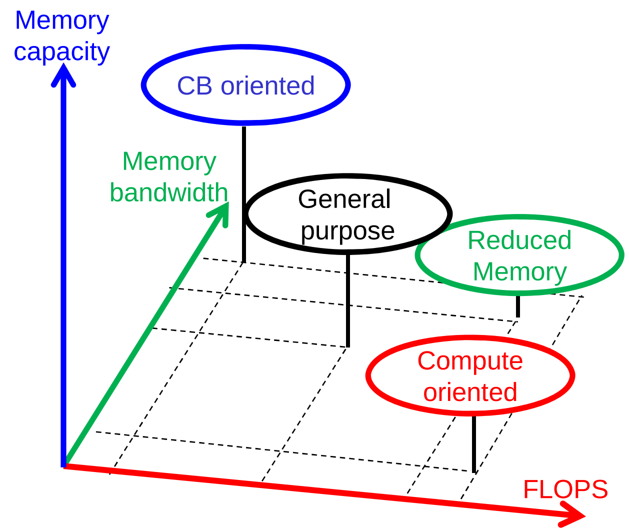
Mitsuhisa Sato

Professor, Center for Computational Sciences, University of Tsukuba /
Team Leader of Programming Environment Research Team, AICS, Riken

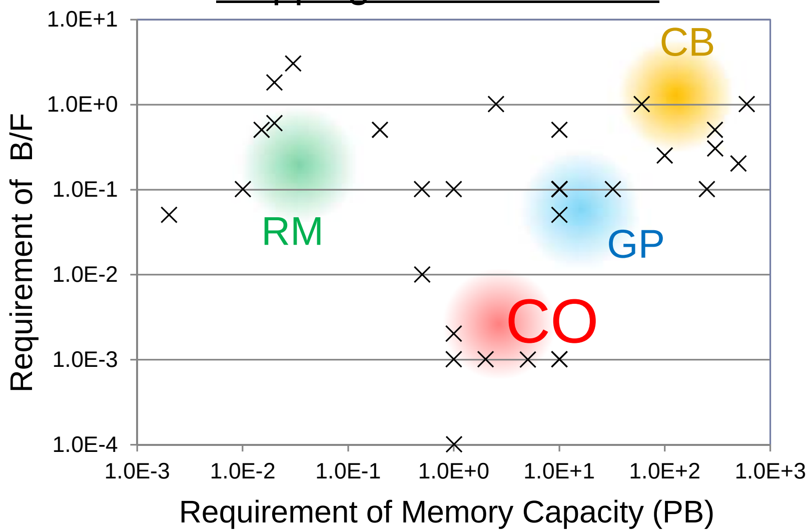
(From SDHPC white paper)

Alternatives of Exascale Architecture

- Four types of architectures are identified for exascale:
  - General Purpose (GP)
    - Ordinary CPU-based MPPs
    - e.g.) K-Computer, GPU, Blue Gene, x86-based PC-clusters
  - Capacity-Bandwidth oriented (CB)
    - With expensive memory-I/F rather than computing capability
    - e.g.) Vector machines
  - Reduced Memory (RM)
    - With embedded (main) memory
    - e.g.) SoC, MD-GRAPE4, Anton
  - Compute Oriented (CO)
    - Many processing units
    - e.g.) ClearSpeed, GRAPE-DR, GPU?

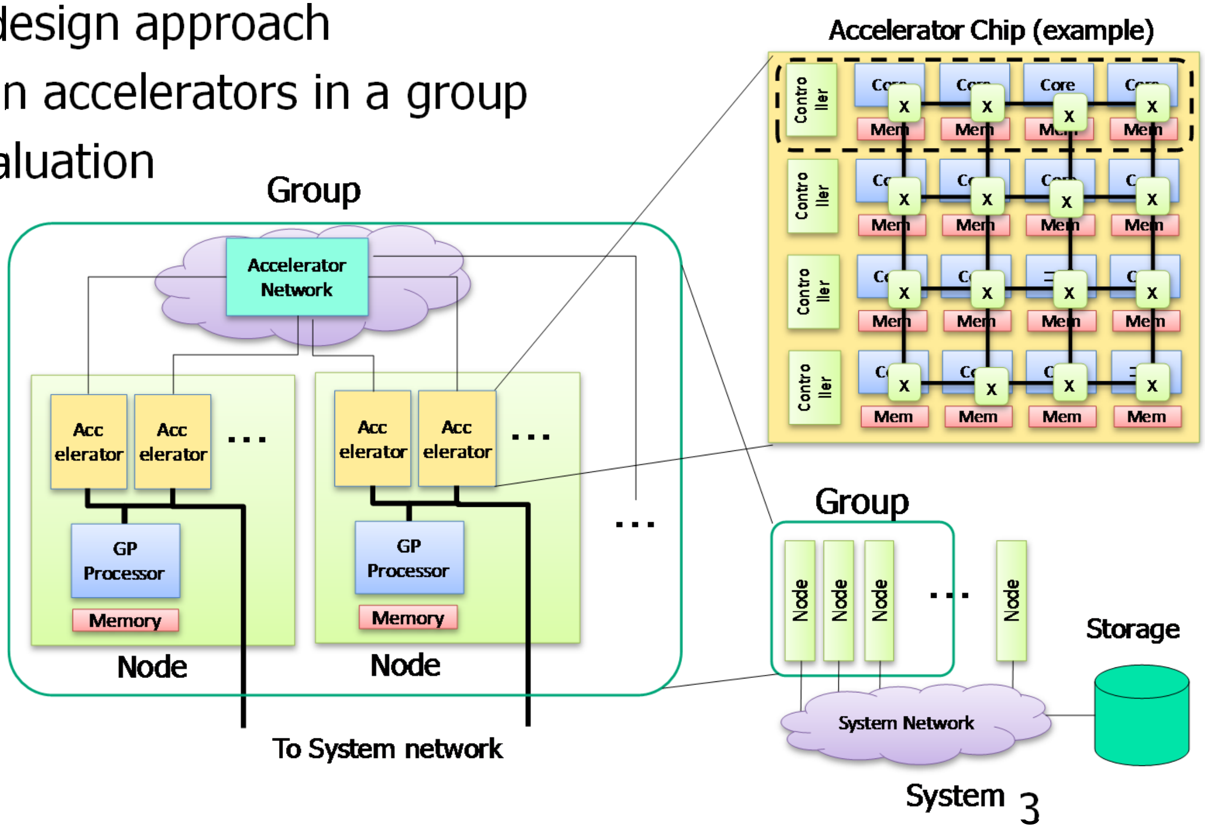


Mapping of Architectures



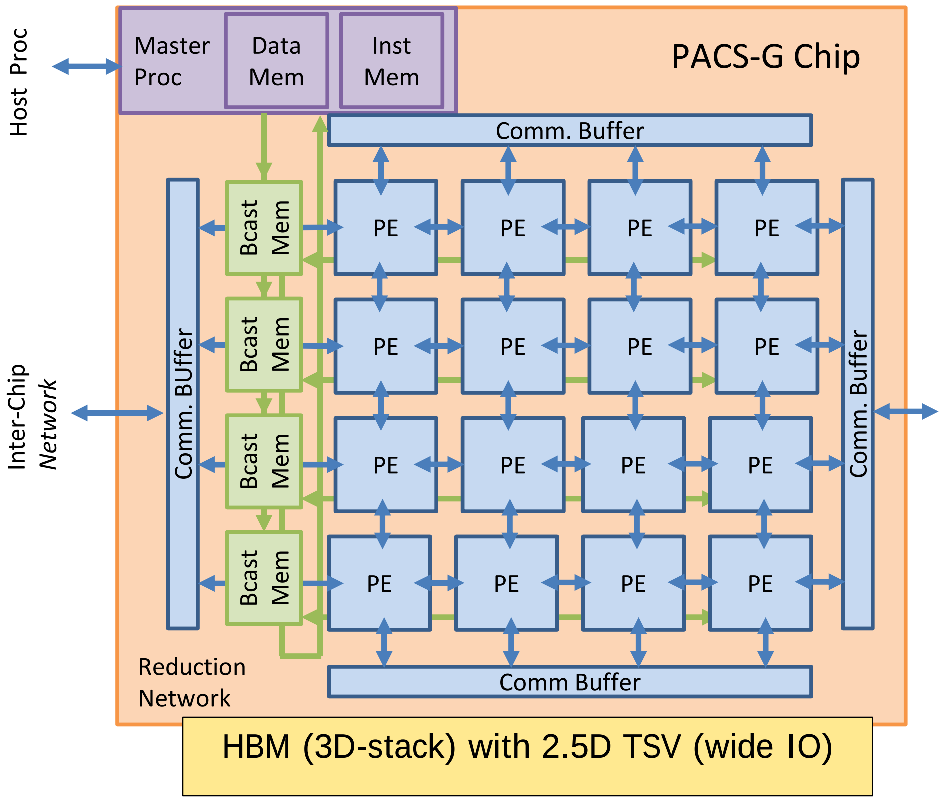
Study on exascale heterogeneous systems with accelerators (U Tsukuba proposal)

- Two keys for exascale computing
  - Power and strong-scaling
- We study "exascale" heterogeneous systems with accelerators of many-cores. We are interested in:
  - Architecture of accelerators, core and memory architecture
  - Application-oriented, co-design approach
  - Direct connection between accelerators in a group
  - Power estimation and evaluation
  - Programming model and computational science applications
  - Requirement for general-purpose system
  - etc …



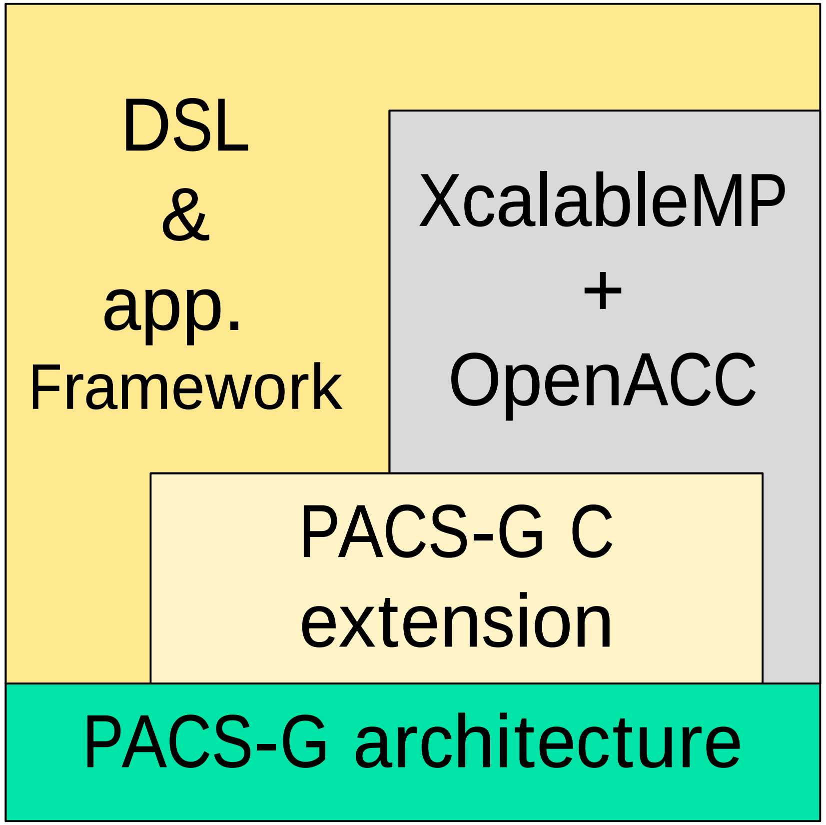
PACS-G: a straw man architecture

- "extreme" SIMD architecture, for compute oriented apps (N-body, MD), and stencil apps.
  - ALL PEs are controlled in SIMD manner by Master processor
- On-chip network and On-chip memory in each PEs (Software-controllable)
- We expect 10nm technology available in the range of year 2018-2020, Chip-dai size: 20mm x 20mm
- 2048～4096 cores (PE)
- Mainly working on on-chip memory and,
- with module (global) memory by HBM (3D-stack/wide IO DRAM memory via 2.5D TSV), bandwidth 1TB/s, size 16-32GB/chip (block access only, no random access)
- No external memory (DIM/DDR)
- (direct) Network between processors
- > 50GF/W expected (at system level)



Programming models for PACS-G

- PACS-G C extension for low-level programing
- XcalableMP (subset/extension) + OpenACC for directive based programming for stencil apps.
  - to make it easy to port existing codes
- Domain-Specific Language (DSL) and application framework
  - e.g. DSL for particle-based apps.
- (OpenCL?)



Current status

- Performance estimation by co-design process
  - 2012 (done): QCD, N-body, MD, HMD
  - 2013: earth quake sim, NICAM (climate), FMO (chemistry) ⇒RSDFT, CONQUEST
- Development of simulators (clock-level/instruction level) for more precious and quantitative performance evaluation

- Programming models and compiler development
  - PACS-G C extension
  - XMP for PACS-G (and OpenACC)
- (Re-)Design and investigation of network topology
  - For both on-chip network and inter-chip network
  - 2D mesh is sufficient? or, other alternative?
- Precise and more detail estimation of power consumptions
- Application development for PACS-G for quantitative performance evaluation, using our programming models. …
- …. Further study on the architecture for beyond stencil and particle apps.

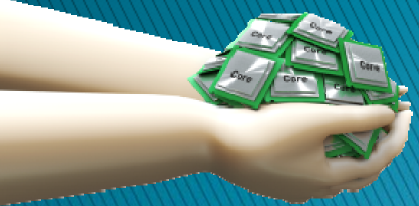
Summary and Concluding Remarks

- Issues for exascale computing
  - Power and Strong-scaling
  - Solution: Accelerated Computing

- PACS-G: the "extreme SIMD" architecture
  - Co-design for compute oriented apps (N-body, MD), and stencil apps.
  - Aiming to high performance(> 10TF/chip) , but also to good performance/power.
  - On-chip memory for each PE using 10nm silicon technology at 2018-2020.
  - dedicated inter-chip network and HBM with 2.5 TSV
  - Programming models and applications are under development for performance evaluation.

Report on Feasibility Study on Future HPC Infrastructure -- Approach from General Purpose CPU Architecture --

Yutaka Ishikawa
University of Tokyo



Feasibility Study on
Advanced and Efficient Latency Core-based
Architecture for Future HPCI R&D

Target System and Approach

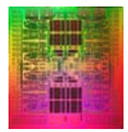
Feature of Target System:

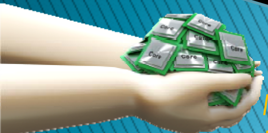
- ✓Deployment around 2018/2020
- ✓Power consumption up to 30MW
- ✓2000 m^2 space

Approach:

- ✓Material and climate sciences have been studied as the first target applications in FY2012, and more applications, e.g., life science, cosmology, and manufacturing technologies, are being considered in FY2013.
- ✓Approach from evolution of the K architecture
- ✓System Software Stack is designed for both the proposed and commodity-based machines

Post-petascale Machine





Organization

Approach:

- ✓Material and climate sciences have been studied as the first target applications in FY2012, and more applications, life science, cosmology, and manufacturing technologies, are being considered in FY2013.
- ✓Approach from evolution of the K architecture
- ✓System Software Stack is designed for both the proposed and commodity-based machines

PI: Yutaka Ishikawa, U. of Tokyo
- ➢ Organization
- ➢ System Software Stack
- ➢ Performance Prediction and Tuning

Applications

System Software Stack
(MPI, parallel file I/O, PGAS, Batch Job Scheduler, Debugging and Tuning Tools)

Commodity-based Supercomputer

Next-Gen General Purpose Supercomputer

Co-PI: Kei Hiraki, U. of Tokyo
- ➢ Architecture Evaluation, Compiler, and Low power technologies

Co-PI: Mutsumi Aoyagi, Kyushu U.
- ➢ Network Evaluation Environment

Co-PI: Naoki Shinjo, Fujitsu
- ➢ Processor, Node, Interconnect Architecture and System Software Stack

Co-PI: Yuichi Nakamura, NEC
- ➢ System Software Stack

Co-PI: Tsuneo Iida, Hitachi
- ➢ Storage Architecture and System Software Stack

What we have done

- Tightly coupled design of architecture by architects, software developers, and application developers.

Type-A (v0-v2)	Low power
Type-B (v0-v6)	Based on K (Reference)
Type-C (v0-v4)	High Clock
Type-D (v4-v10)	High IPC
Type-E (v6-v8)	Matrix multiply
Type-F (v8)	Heterogeneous
Type-G (v10)	Pipeline Improvement
Type-H (v10)	Large Cache & Power Knob

1 Cycle / 2 months

Application Code with performance counter instrumentation

Fujitsu FX10

Performance Parameters

Prediction Tools

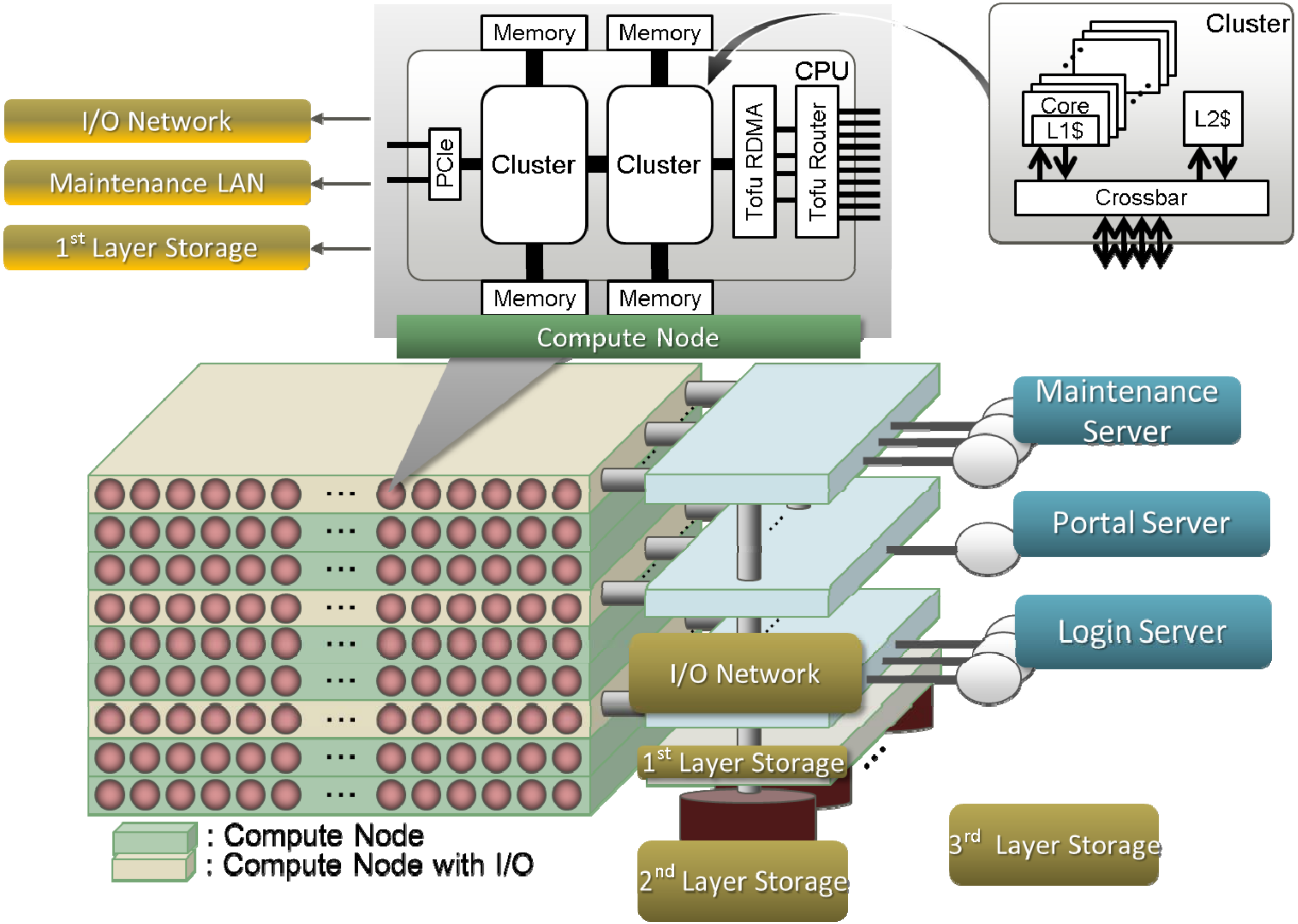
Performance Prediction

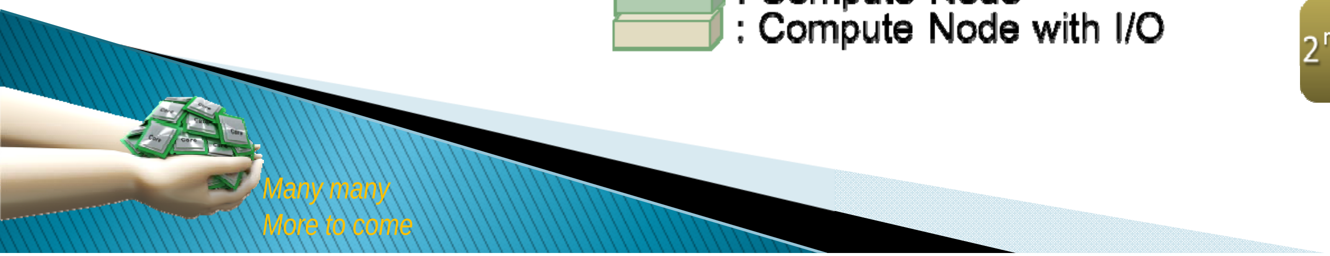
Evaluation and Tuning of Apps.

Architecture Design:
Processor, Node, Network, and Storage

Many many
More to come

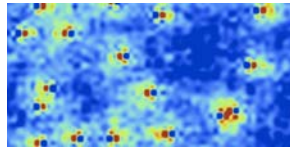
An Overview of Architecture

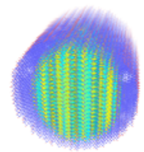


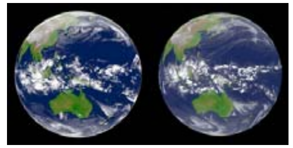


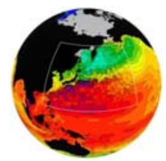
Part of Target Applications

	Scientific Challenges with Exaflops Machine	Exascale Target Problem Size	Required Time	Predicted Time
ALPS	Clarify exotic quantum phases and novel functions of strongly correlated and nano magnetics materials (10mn) at very low temperatures.	Nano-scale magnetics (10nm) simulation at very low temperatures. （N~300000, T~0.001）	1 day	Within 2.5 hours
RSDFT	Prediction of channel characteristics in transistors by Si nano-wire with 10,000 - 100,000 atoms.	Simulation with order 300,000 atoms.	1 Month	Within 38 days
NICAM	High resolution simulation for long-time weather forecasting, typhoon and heavy rain forecasting in global warming.	Problem size: glevel=14, vlayer=100 (vertical 400m, horizontal 250m) 1 month weather prediction	10 simultaneous jobs within 10 days	2 jobs within 17 days
COCO	Real-time simulation for reaction between small-flow in coast and ocean current in open sea.	Problem size: 38,400 x 38,400 x 100 for 10 years	1 Month	Within 21 days









Other benchmarks, Mini-apps, requested by application FS are being evaluated

System Software Stack: An Overview

Yellow: Mainly Concern

Real-Time/Big-data Visualization

Math. Libraries

Tuning and Debugging Tools

Batch Job System

Hierarchical File System

Parallel File System

Programming Languages/Models

Hierarchical Memory Management

Process/Thread

Parallel File I/O

Communication MPI, ….

Parallel Process Spawn

Low Level Communication

Power Management

• • •

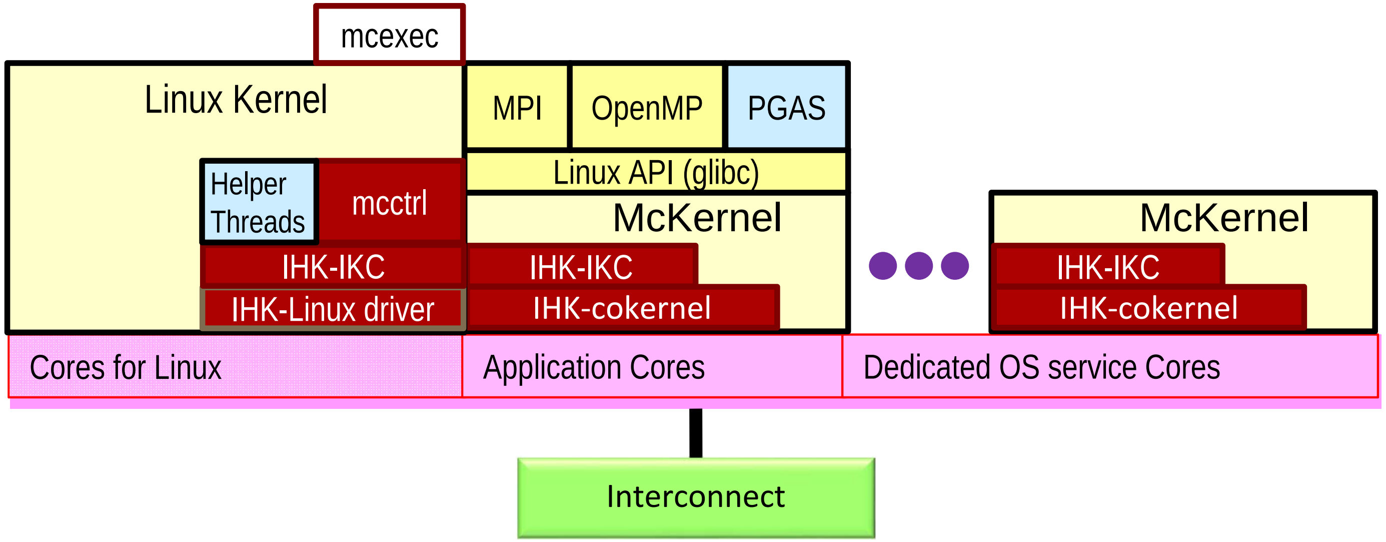
Hetero Operating System: Linux and light-weight micro kernel (McKernel)

Commodity Machine, Proposed Machine, …

Infiniband	RoCE	BGQ	Tofu	Fabric With manycore

Energy Consumption Model
Energy Control Model

Design and Implementation: IHK and McKernel



- IHK (Interface for Heterogeneous Kernel)
    - IHK-Linux driver: provides the functions of co-processors, such as booting, memory copy, and interrupt
    - IHK-cokernel: Abstracts the hardware functions of the manycore devices
    - IHK-IKC: providing communication between the Linux and co-kernel
- McKernel
    - lightweight micro kernel running on co-processors

Features implemented so far in KNC
- glibc and pthread
    - Thread and memory management
    - File I/O, delegated to Linux in host
    - Memory map and dynamic link library
- Process launcher in host
- Direct Communication with InfiniBand
- MPI library (not fully) running on Xeon Phi
- OpenMP environment with Intel compiler

Concluding Remarks

- Based on this study, RIKEN has designed a strawman of postK machine
- The system software stack proposed is not only for the postK machine, but will be also used for commodity supercomputers
  - Xeon Phi implementation is heading
- International collaboration for system software development has been considered