RENKEI:A Light-weight Grid Middleware for E-Science Community

June 27, 2011

Kenichi Miura, Ph.D.

三浦　謙一

Center for Grid Research and Development
National Institute of Informatics
Tokyo, Japan

Outline

- Overview of RENKEI Project
- the Next Generation Supercomputer Project
  (K Computer)

Cyber-Science Infrastructure for R & D

RENKEI

Cyber-Science Infrastructure（CSI)

NAREGI Outputs

NII-REO (Repository of Electronic Journals and Online Publications

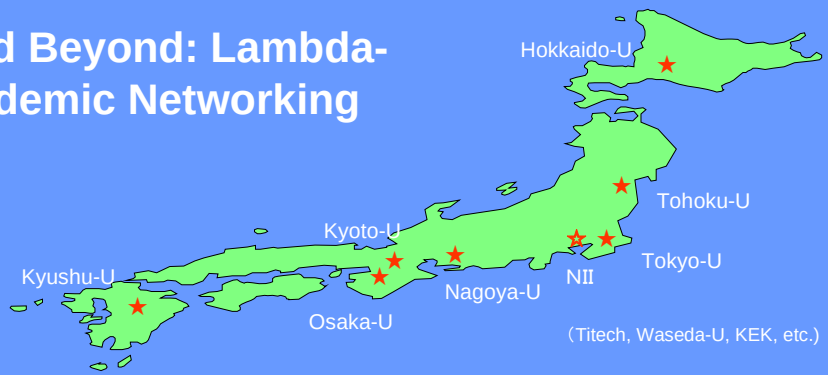
GeNii (Global Environment for Networked Intellectual Information)

Virtual Labs
Live Collaborations

Deployment of NAREGI Middleware

UPKI: National Research PKI Infrastructure

SINET4a nd Beyond: Lambda-based Academic Networking Backbone



Restructuring Univ. IT Research Resources
Extensive On-Line Publications of Results

Industry/Societal Feedback

International Infrastructural Collaboration

Science Information Network (SINET4)

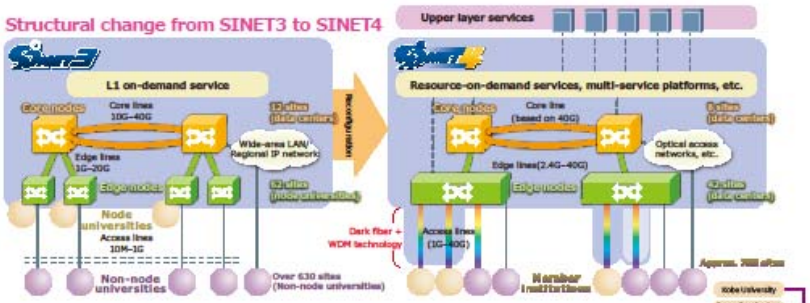
SINET4: Japanese Academic Optical Backbone Network

to research networks such as Internet2 in the U.S. and GÉANT2 in Europe to facilitate dissemination of research collaborations over networks.
SINET4 began operations in April 2011, and it replaces the previous SINET3. SINET4 plays an important role as nent of the Cyber Science Infrastructure(CSI).

Information and

the core compo-

Structural change from SINET3 to SINET4



Characteristics of SINET4

①Higher network speed The effective network bandwidth has been increased and the rerouting function has been improved by reconfiguring the network and adopting solutions including dark fiber and WDM technology. This has made the network even more cost effective.

②Provision of diverse services SINET4 inherits all of SINET3's services, with services such as resource-on-demand strengthened and expanded.

③Higher edge node stability SINET4 positions both edge nodes and core nodes in data centers, improving the reliability of the network including its availability, maintainability, and security.

④Establishment of an environment for high-speed access lines By undertaking shared procurement of access lines, a faster access system has been created for member institutions other than those on site (SINET3 node universities). In addition, installment of nodes is scheduled to be completed within FY2011 in all prefectures in Japan.

⑤Upper layer deployment Installation of interfaces and service-providing platforms to support the upper layer is being considered.

Participating SINET Institutions
(as of April 1, 2011)

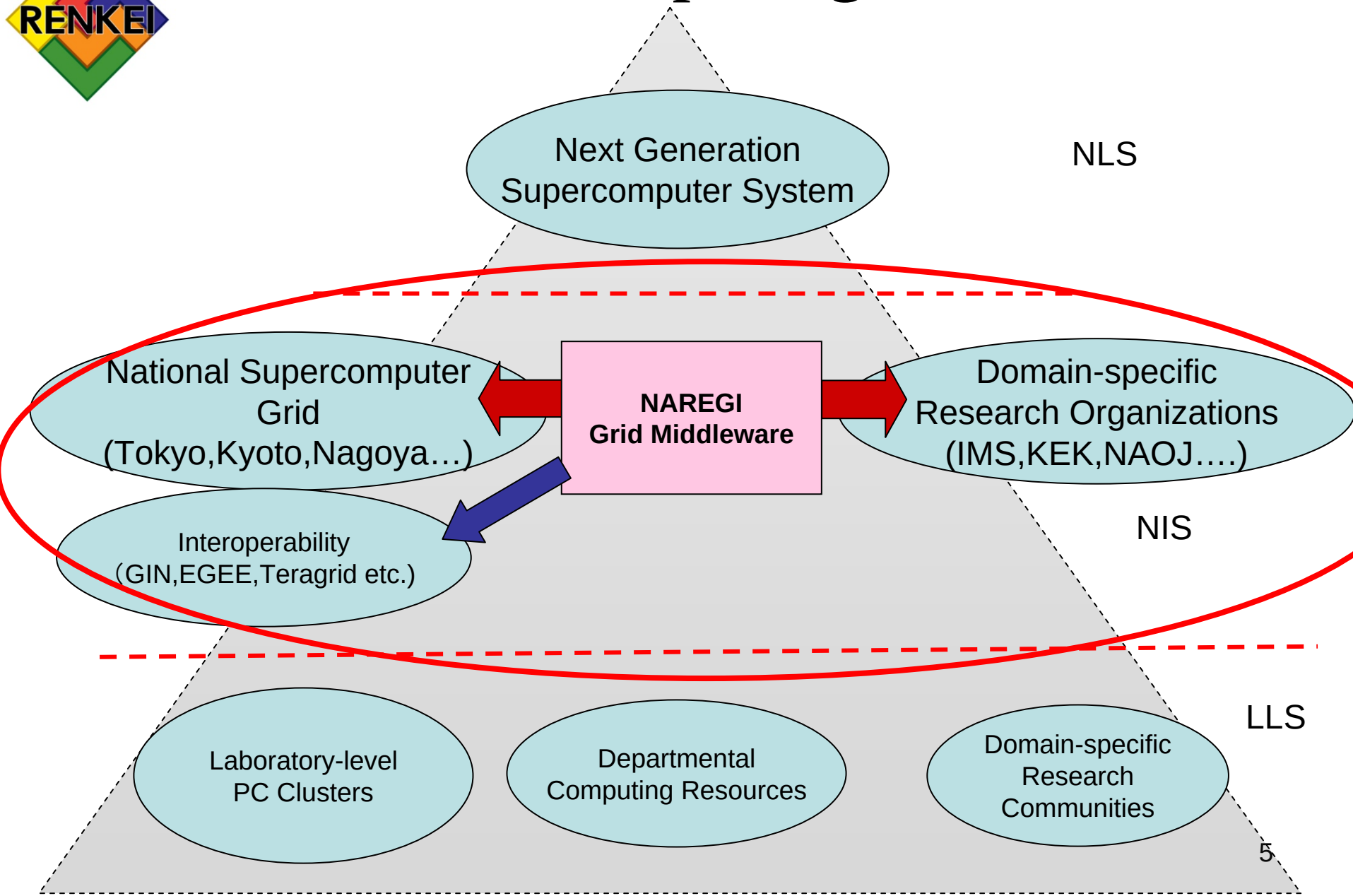
National universities	86
Municipal universities	59
Private universities	297
Junior colleges	61
Technical colleges	51
Inter-University Institutes	16
Others	170
Total	740

SINET4 Services (Classification bynetwork layer and QoS)

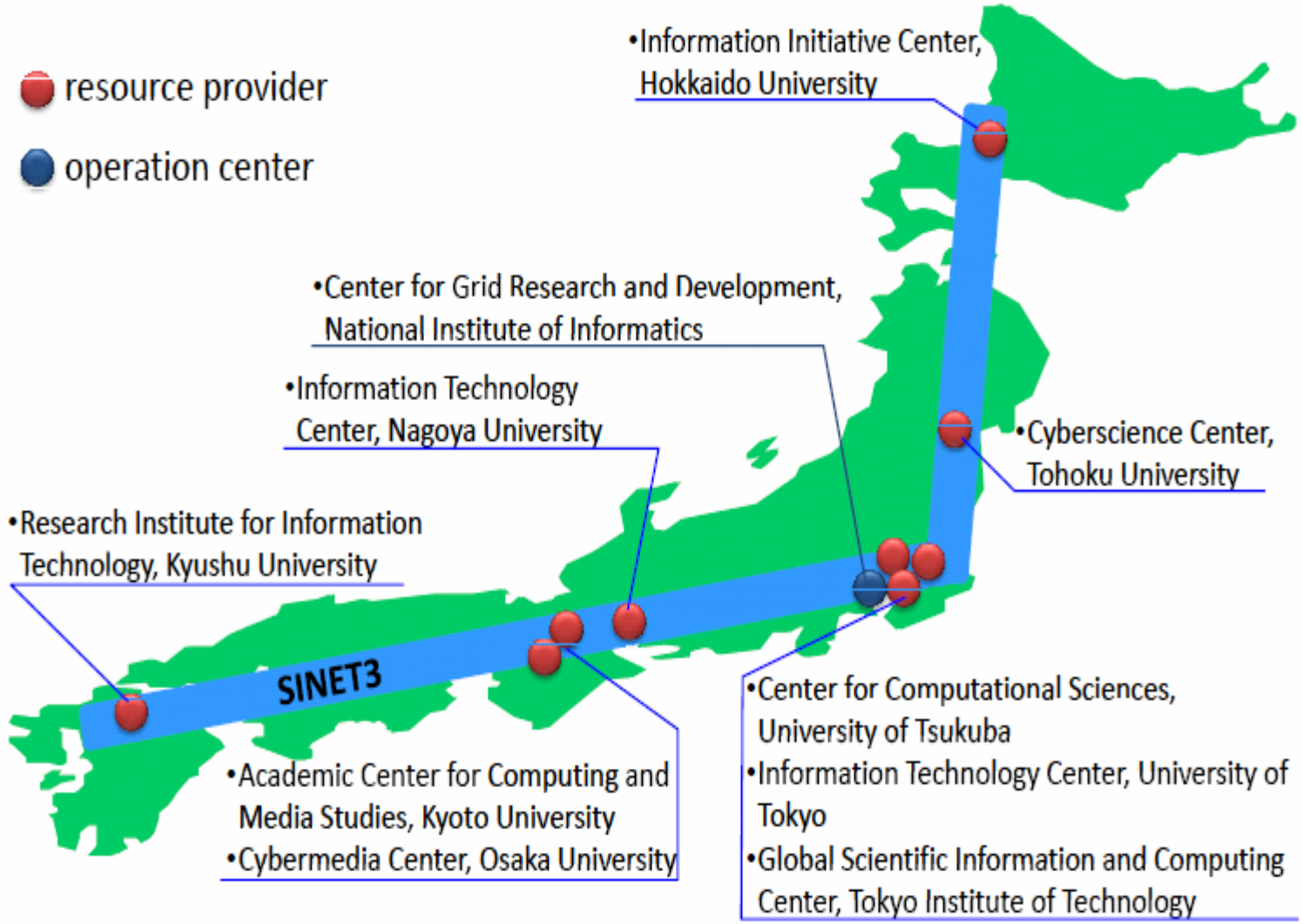
QoS	IP (L3)	Ethernet (L2)	Lambda/ Dedicated (L1)
QoS-guaranteed			On-demand BW-specified L1VPN Lambda L1VPN
High Priority	L3VPN(QoS) Multicast (QoS) Application-based QoS	On-demand VPLS(QoS) L2VPN(QoS)	
Best Effort	L3VPN Multicast Multi-homing IPv4 IPv6	On-demand VPLS L2VPN	

International network collaboration

Hierarchical Computing Environment

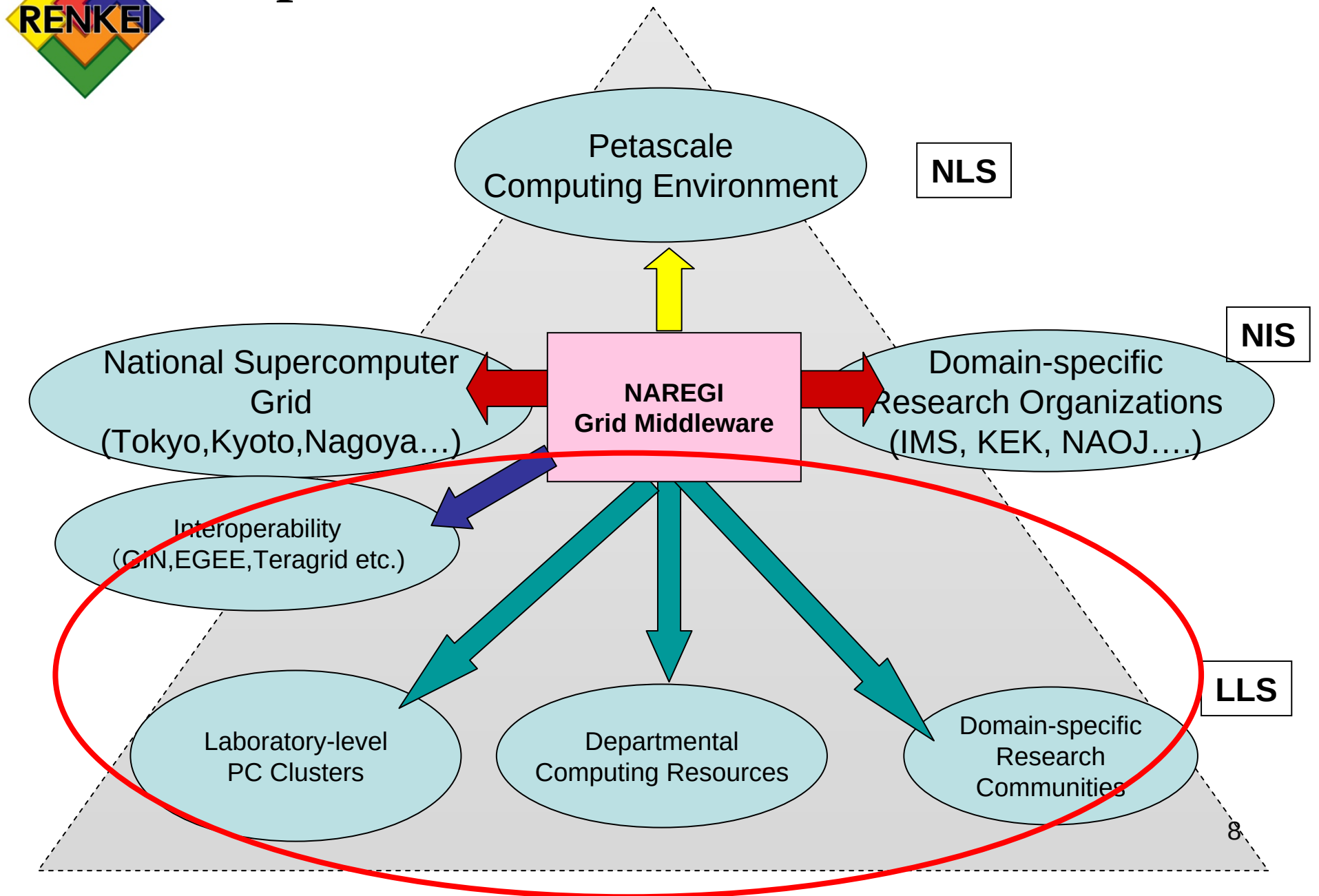


Deployment of NAREGI Grid Middleware to 9 Supercomputer Centers (in progress)



RENKEI Project: Resource Collaboration Technologies for e-Science Communities (FY2008-2011)

Expansion Plan of NAREGI Grid



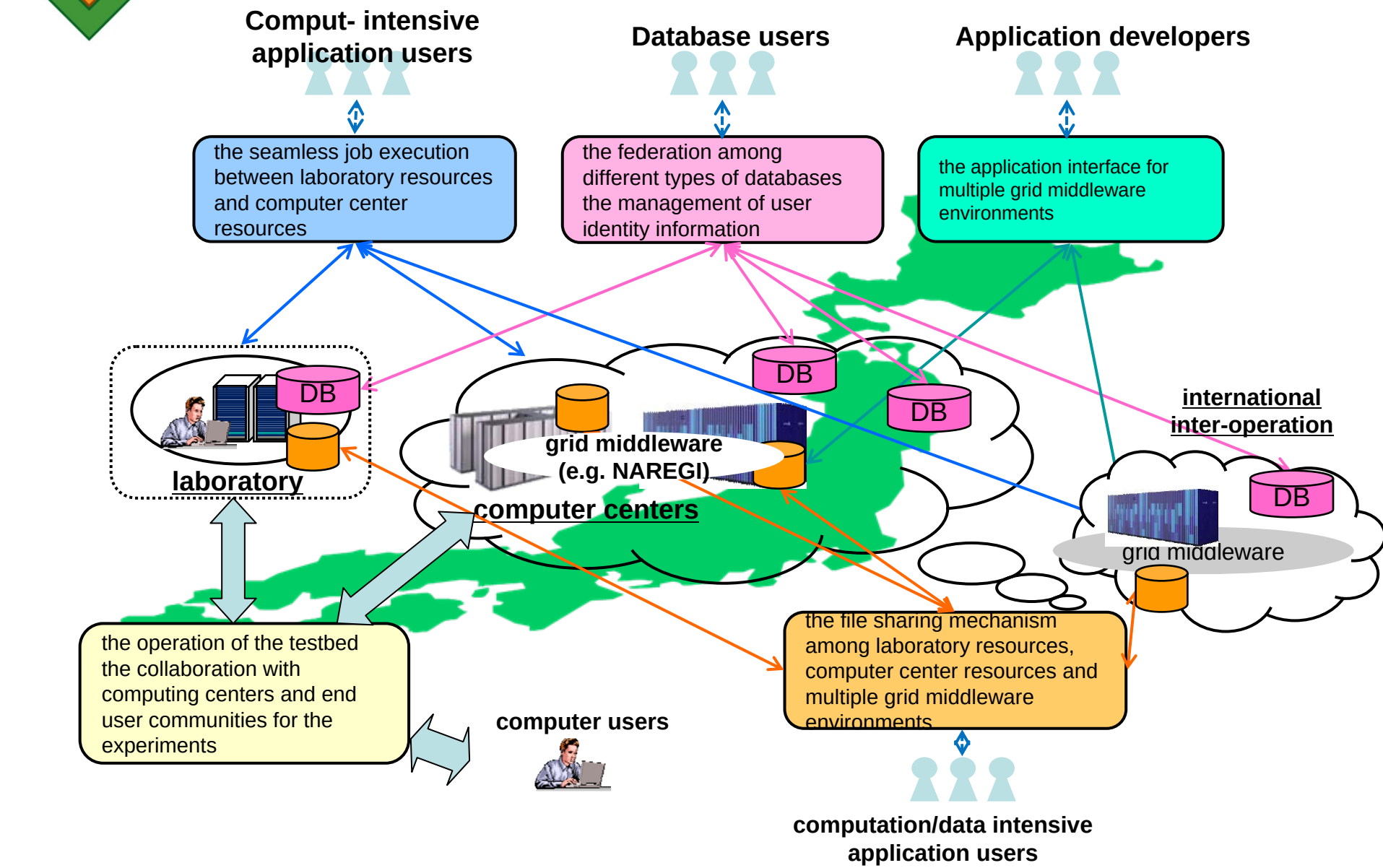
Description of RENKEI Project

The “RENKEI Project” is a new R&D project, which started in September 2008 under the auspices of MEXT*. In this project, a new **light-weight grid middleware and software tools** will be developed in order to provide the connection between the NAREGI Grid environment and wider research communities.

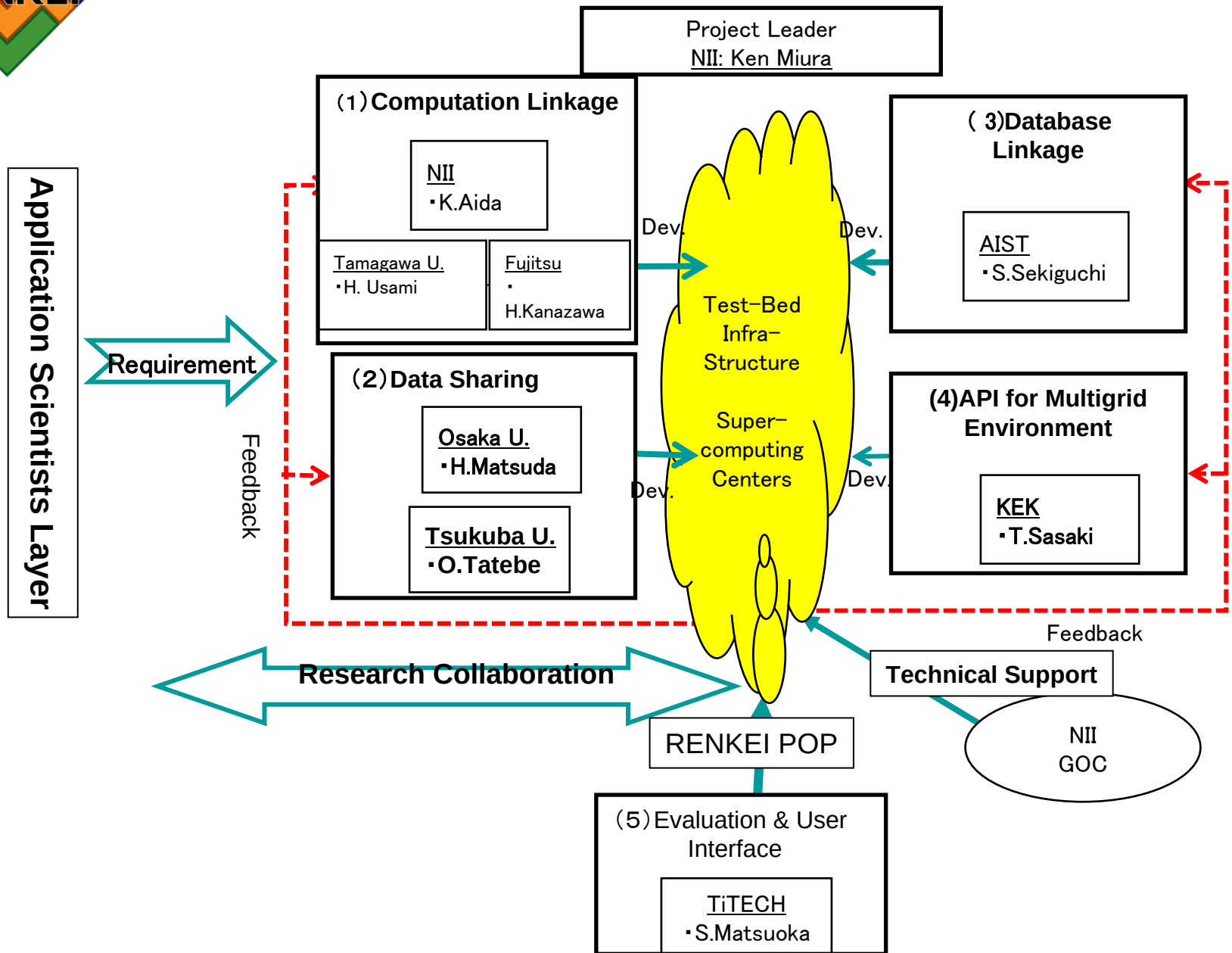
In particular, technology for **the flexible and seamless access** between the national computing center level and the departmental/laboratory level resources, such as computers, storage and databases is highly emphasized. Also, this newly developed grid environment will be made interoperable with the major international grids.

***MEXT: the Ministry of Education, Culture, Sports, Science and Technology**

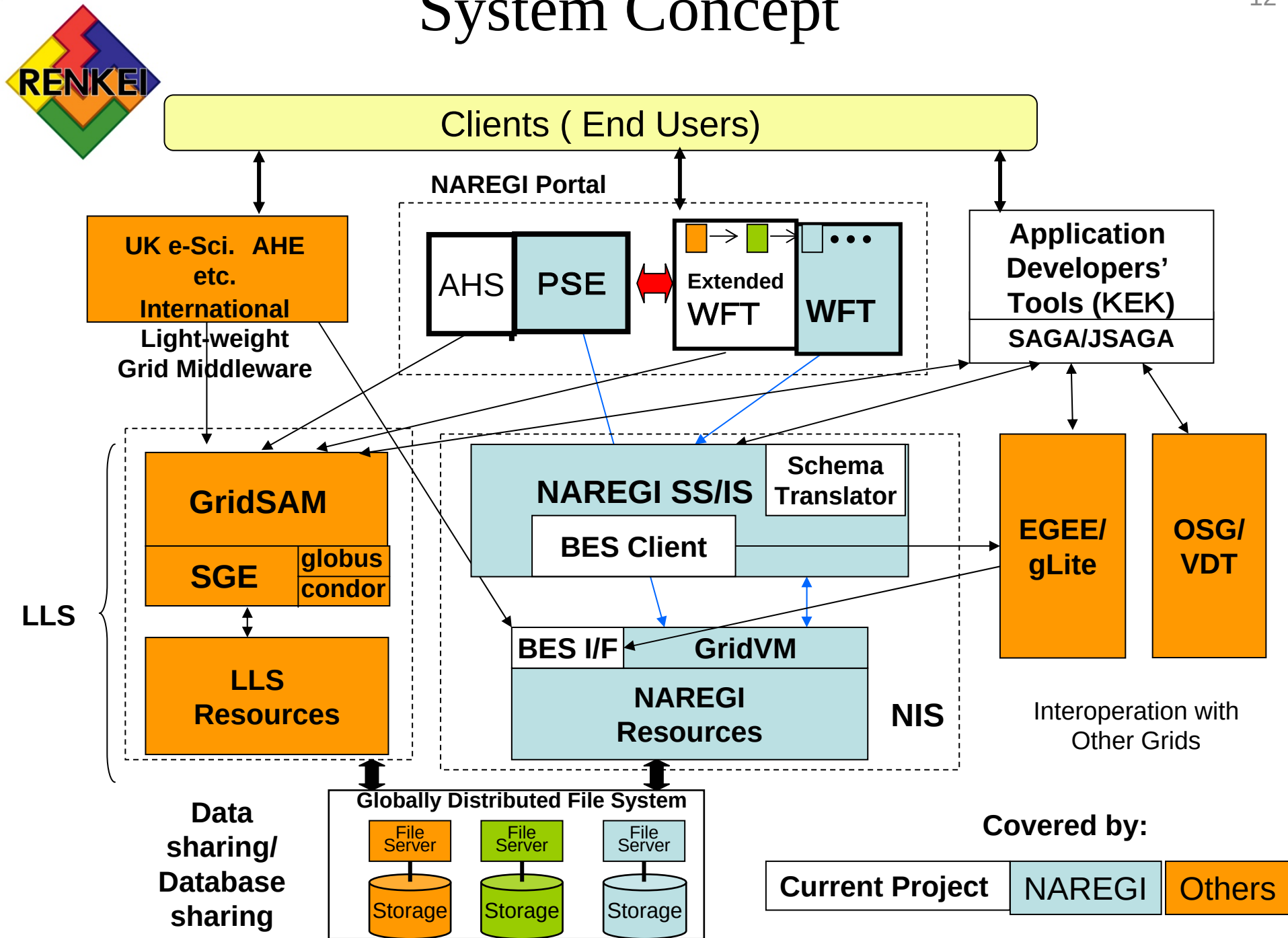
Overview of RENKEI Project



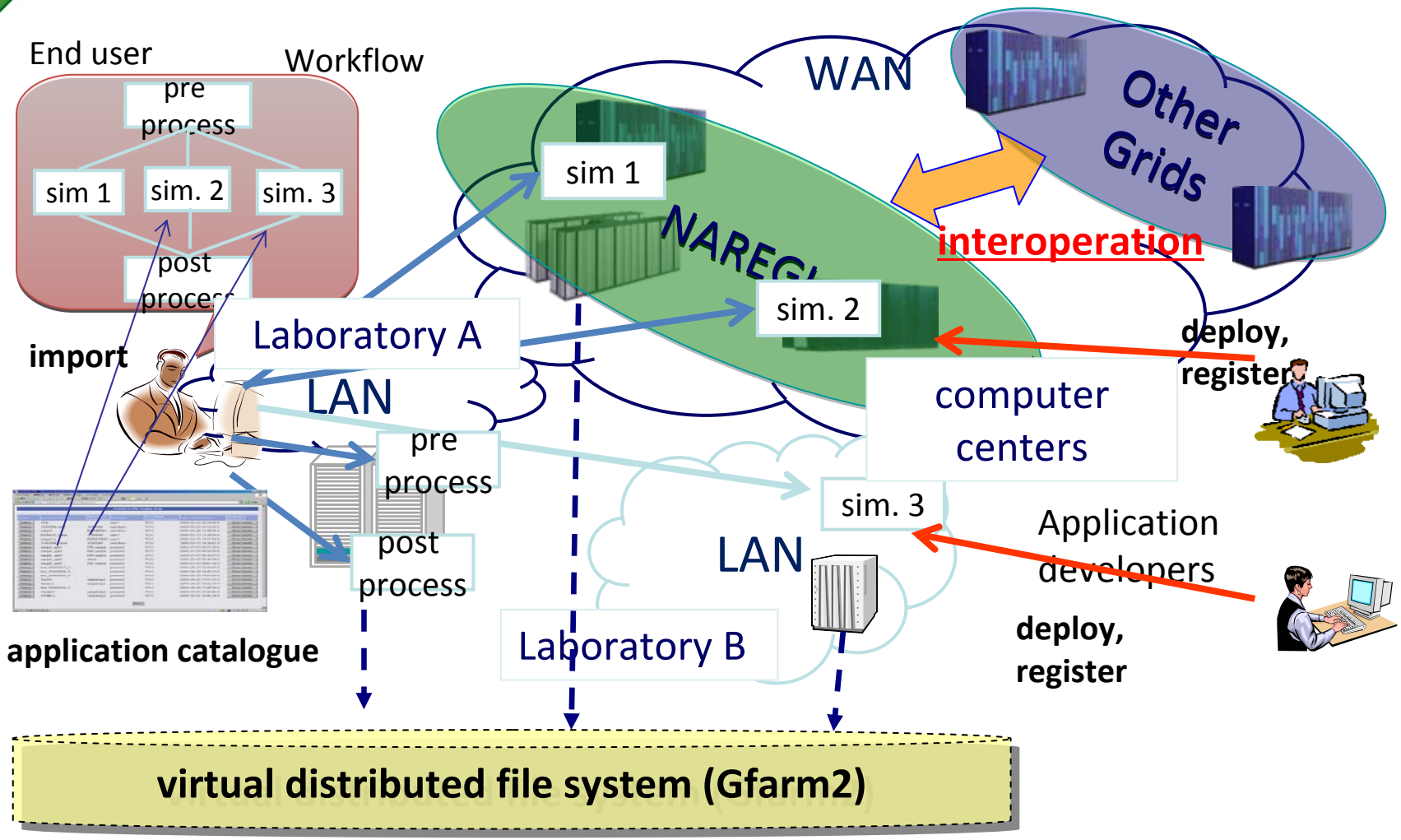
Organization



System Concept

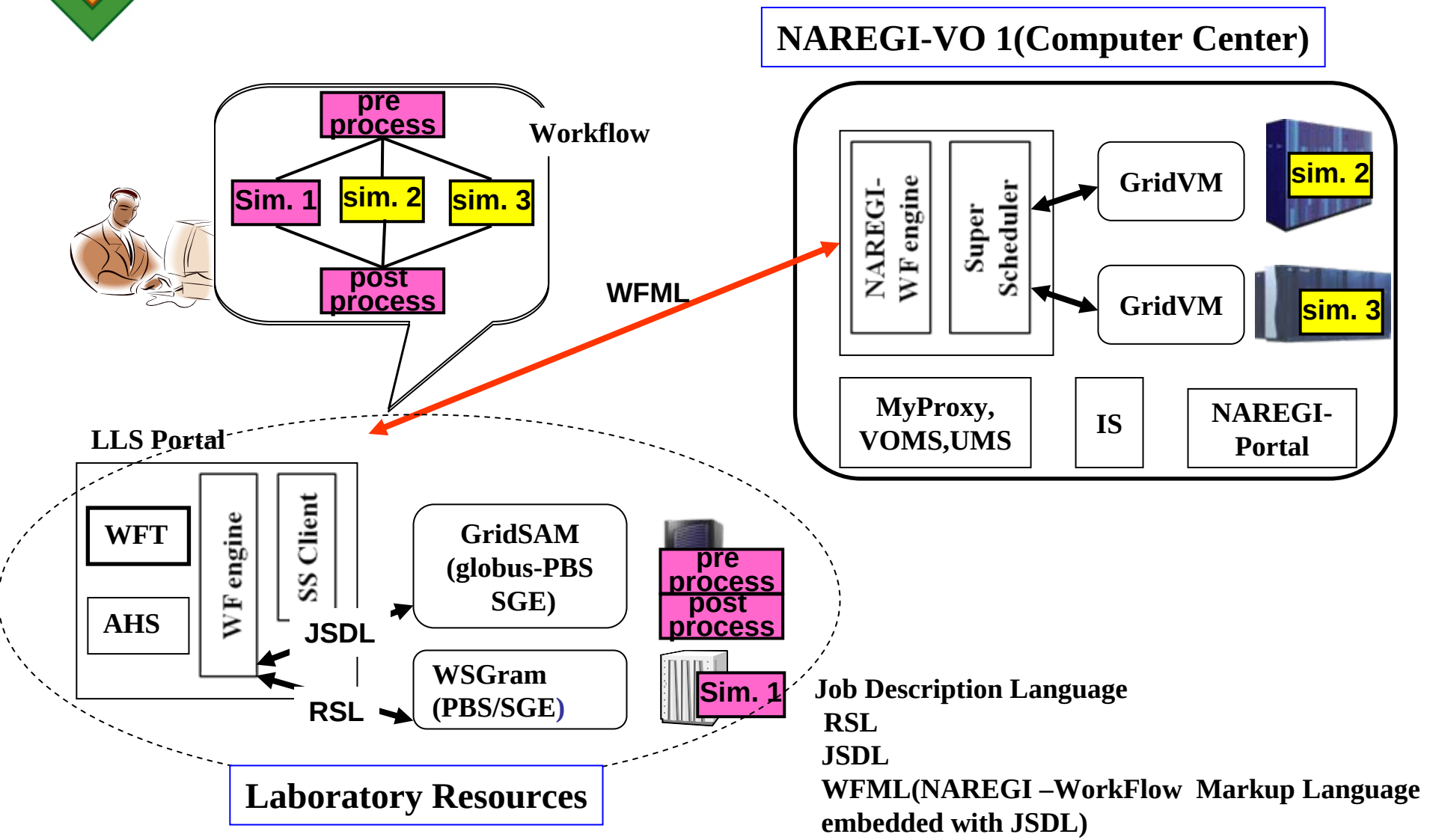


Seamless Connection between NIS and LLS



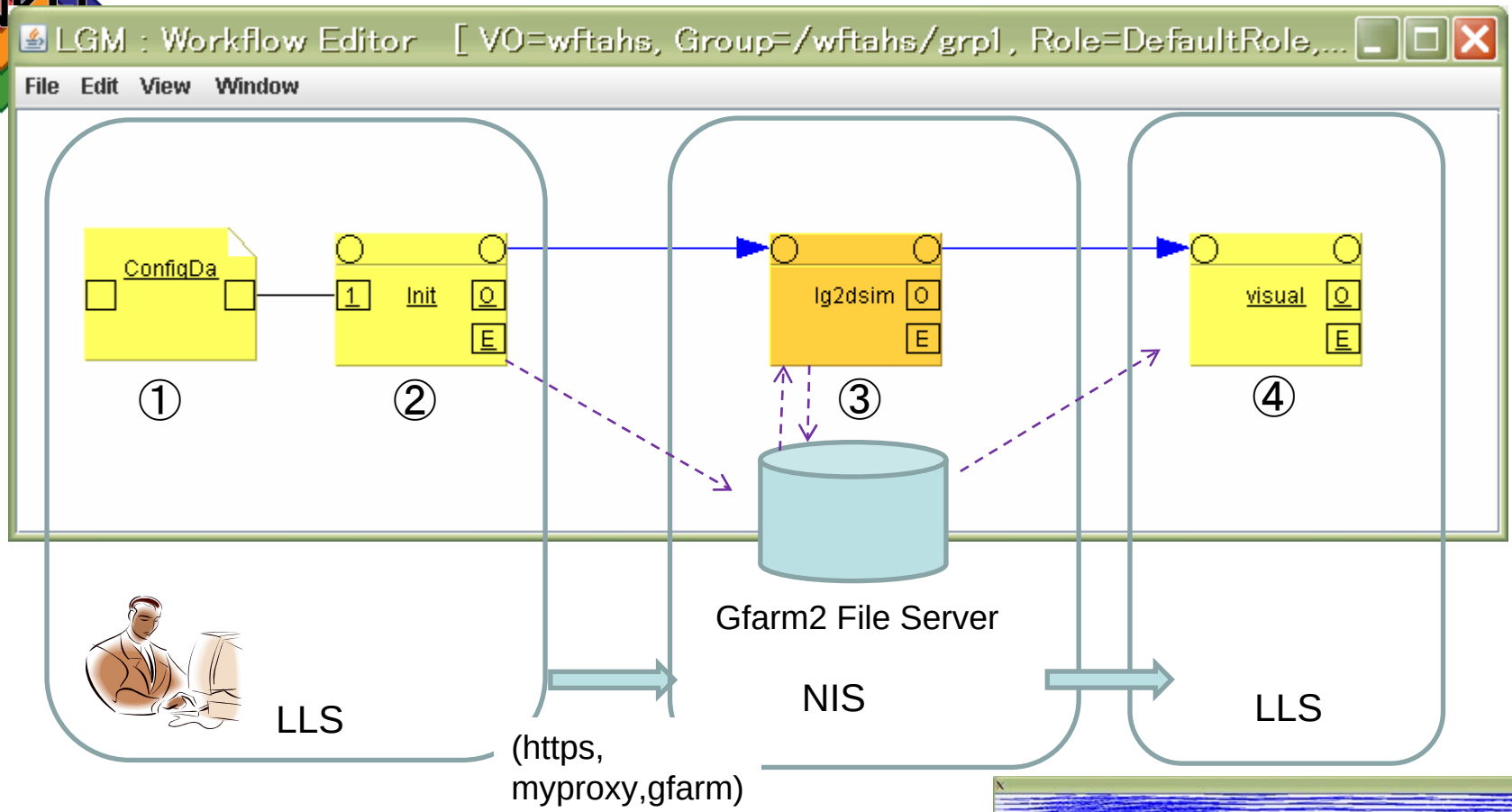
NIS: National Infrastructure System, LLS: Laboratory Level System

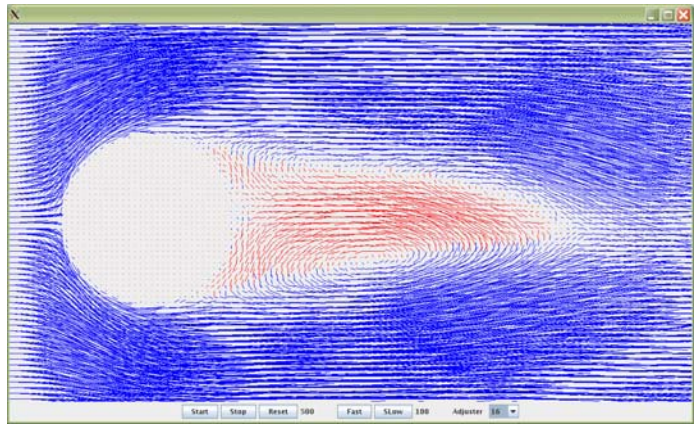
Extended Workflow System



Job Description Language
RSL
JSDL
WFML(NAREGI –WorkFlow Markup Language embedded with JSDL)

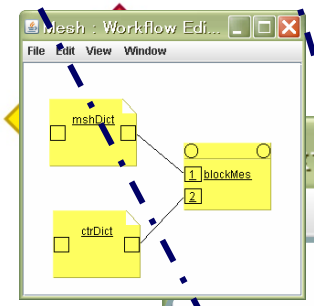
Case 1：Lattice Gas Method





① Initial Configuration for Data Input
② Generation of Initial Data
③Time-step Simulation of LGM
④Visualization in LLS

Case 2：OpenFOAM



flow : Workflow Editor [VO=wftahs, Group=/wftahs/grp1, Ro...

View Window

Mesh

SetField

Mesh GeneatioWF

Initialization WF

decomp

paraview

Domain Decomp.WF

Vis. prg
paraview

Two-phase flowprg(interFoam)

interFoa

Gfarm2 File Server

(https,
myproxy,gfarm)

reconstru

Domain Decomp
Unification WF

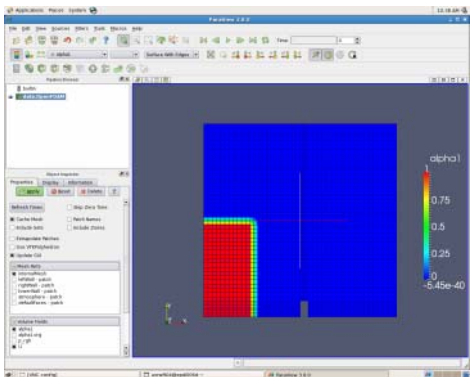
paraview

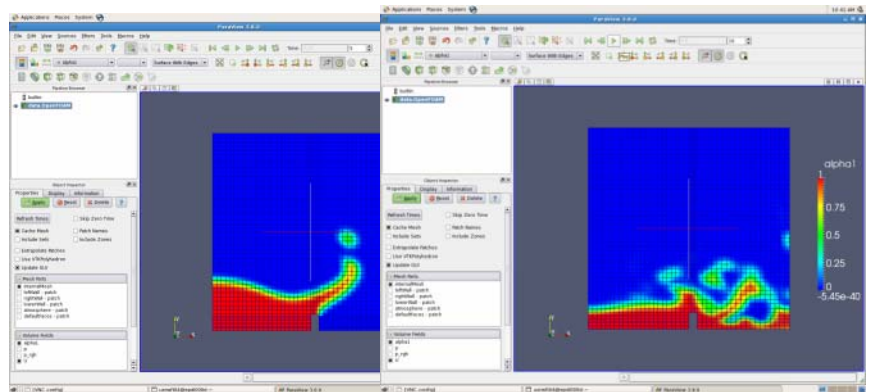
Vis.prg
paraview

LLS

NIS(Tsubame2)

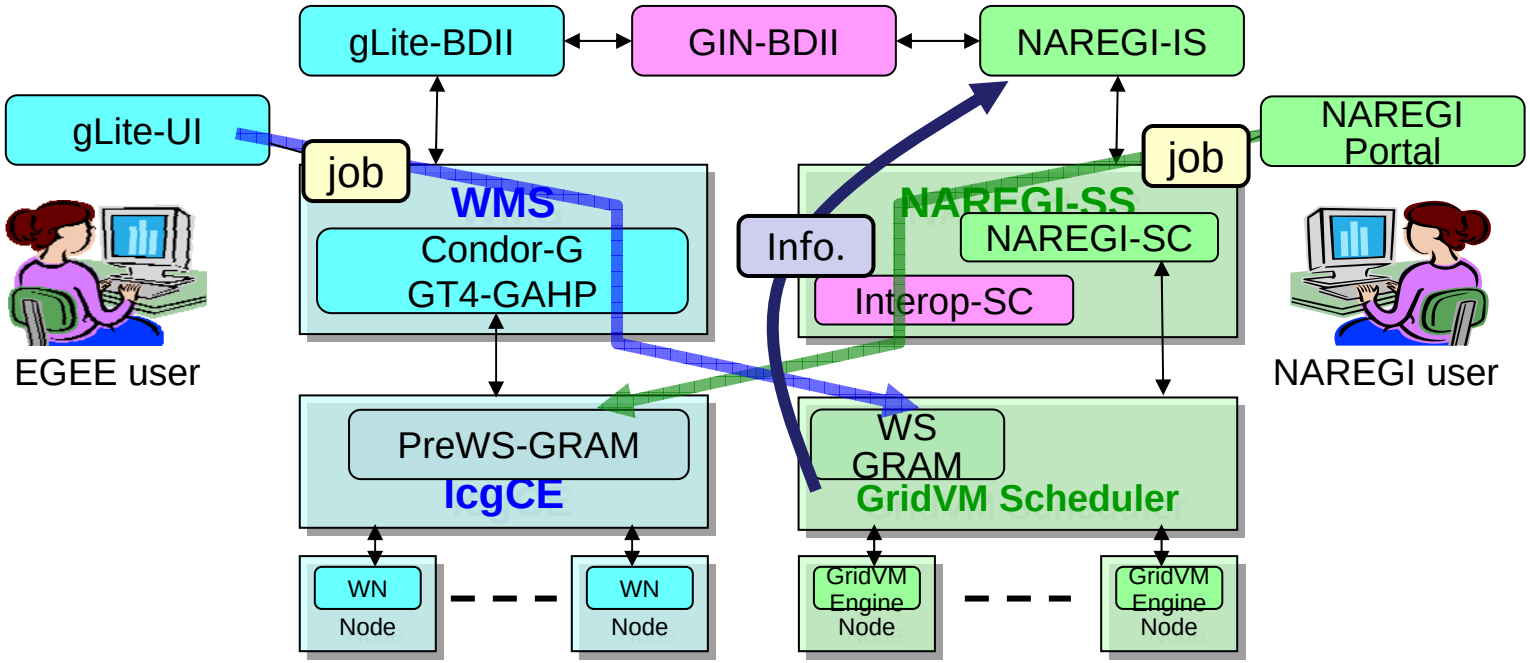
LLS





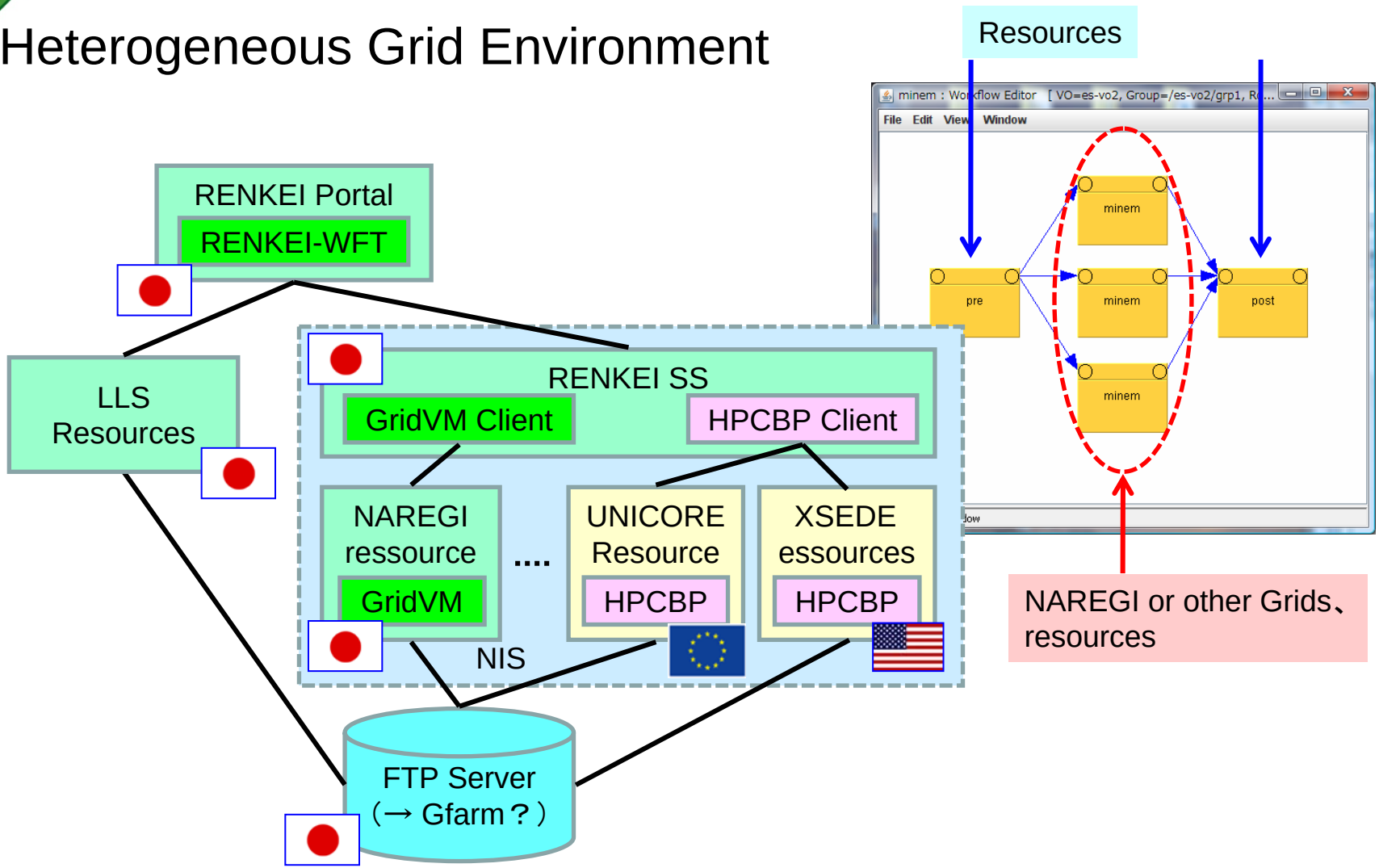
Interoperation between Two Different Grid Middleware

- Objective
  - Mutual job execution between NAREGI and gLite
- FY2008
  - Scheduler
    - NAREGI (SS) → gLite (lcgCE) , gLite (WMS) → WS-GRAM
  - Information Service
    - Collection of WS-GRAM Resources／usage Info,and Storing to IS

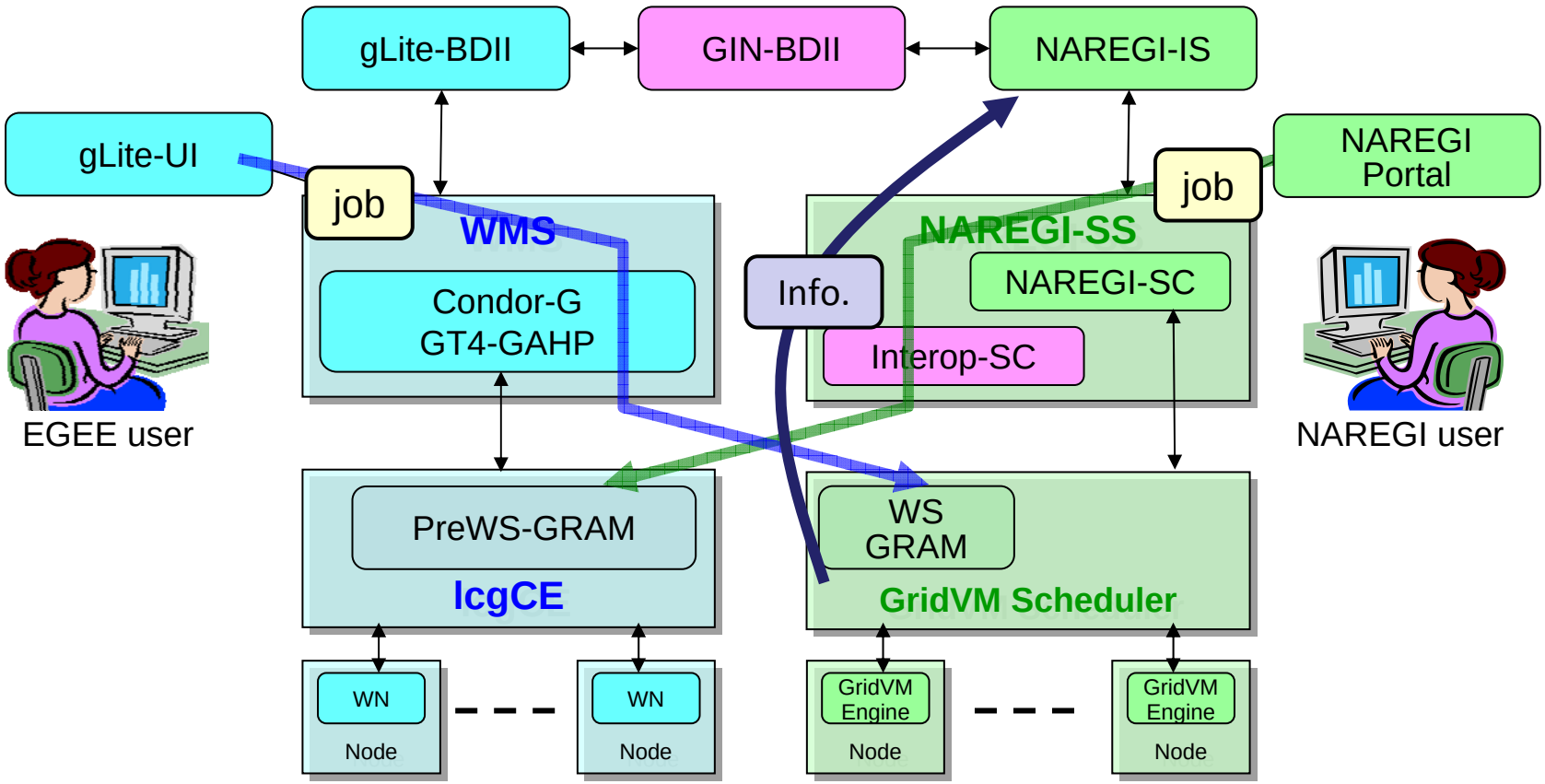


Interoperation Experiment

- Heterogeneous Grid Environment

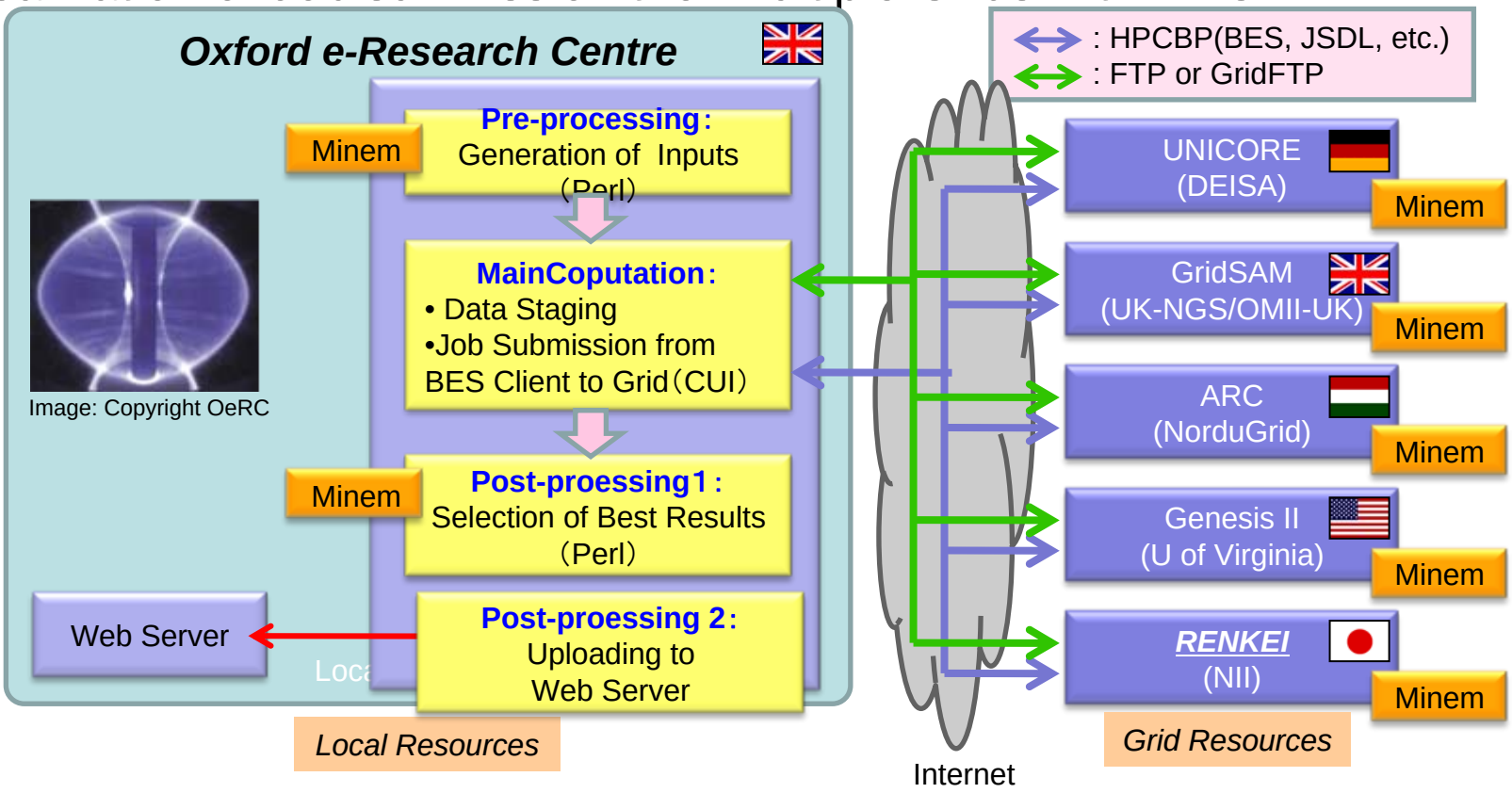


Interoperation between two different Grid Middleware

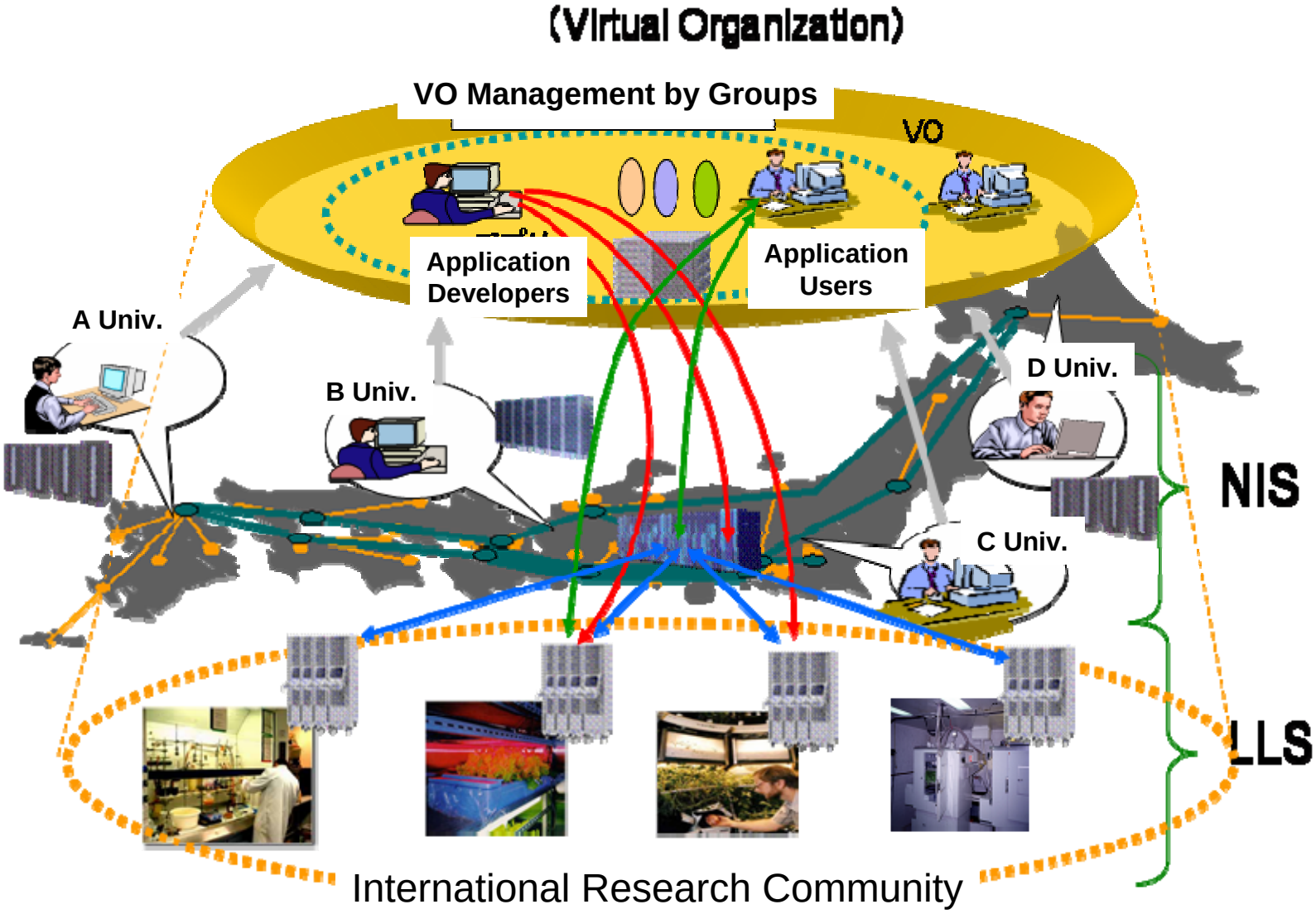


Interop Demo @ 5th IEEE eScience(2009)

- Application：Minem（Plasma Charge Minimization）
  - Minimization of Energy on the Surface of Sphere
  - Pre- Post – Processing on the Local Resources /Main computation over the Grid
- Realization of Job sunmission the multiple Grids with HPCBP

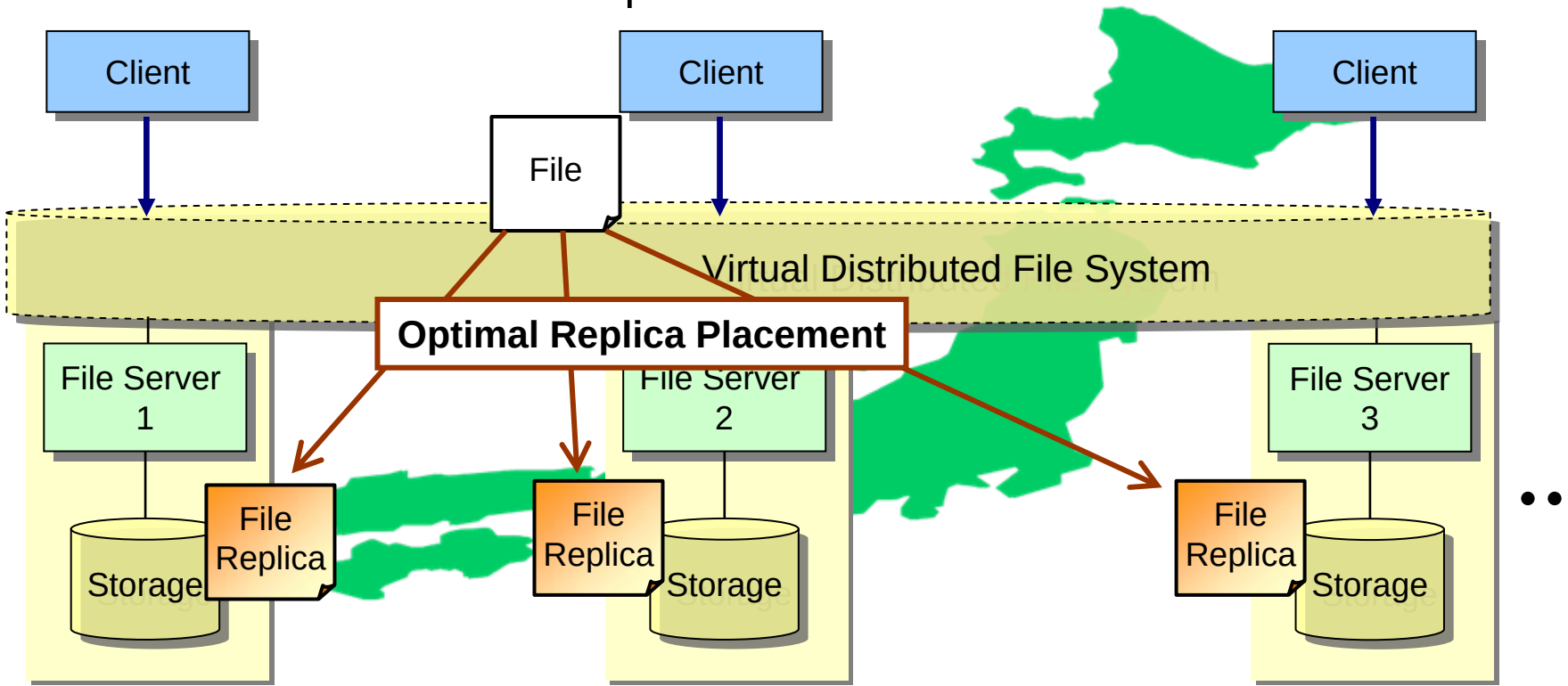


Applications and Knowledge Sharing with Applications Hosting Service (AHS)

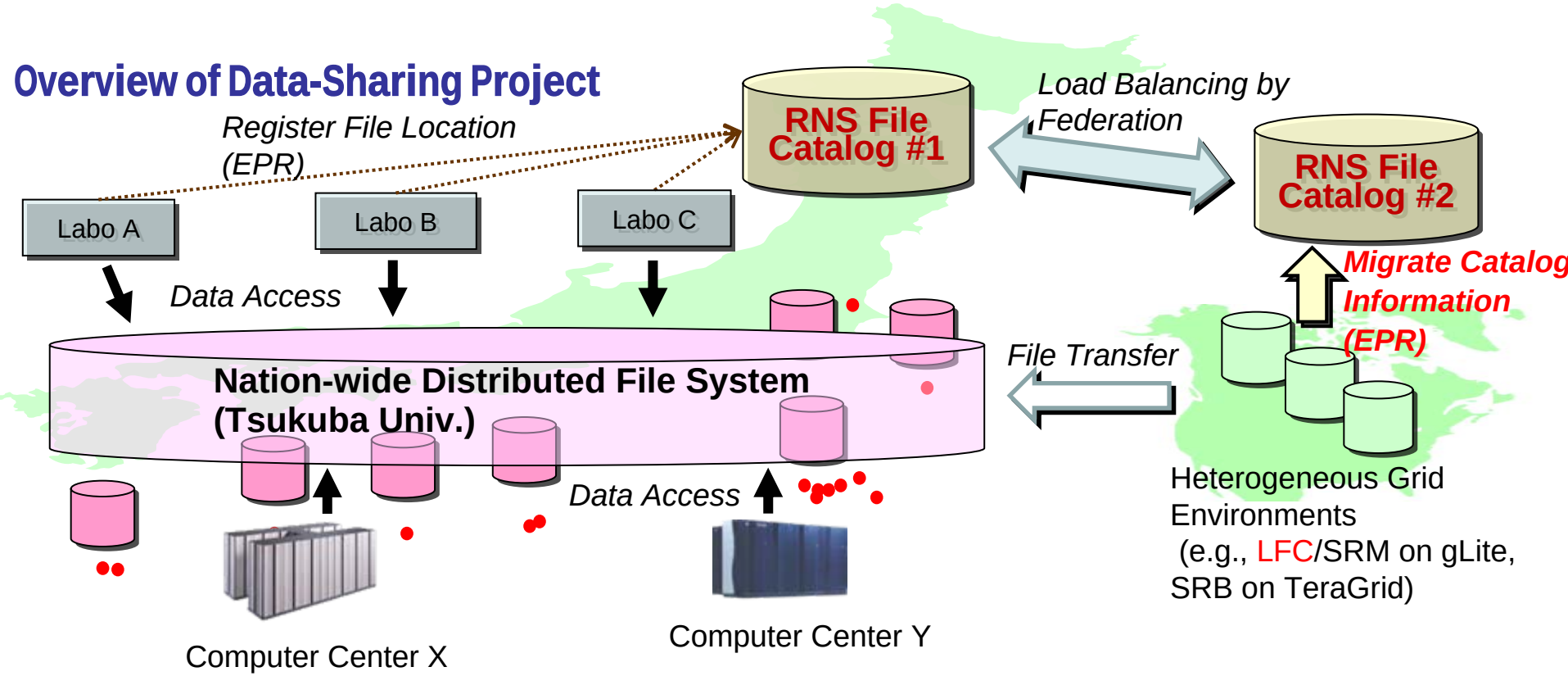


Nation-wide Distributed File System

- Goal: Development of distributed file system technology spread over nation-wide with comparative performance of local fileserver
- Research Topics:
  - Optimal automatic placement of file replicas based on **Gfarm 2.0**.
  - Fault tolerance with file replicas



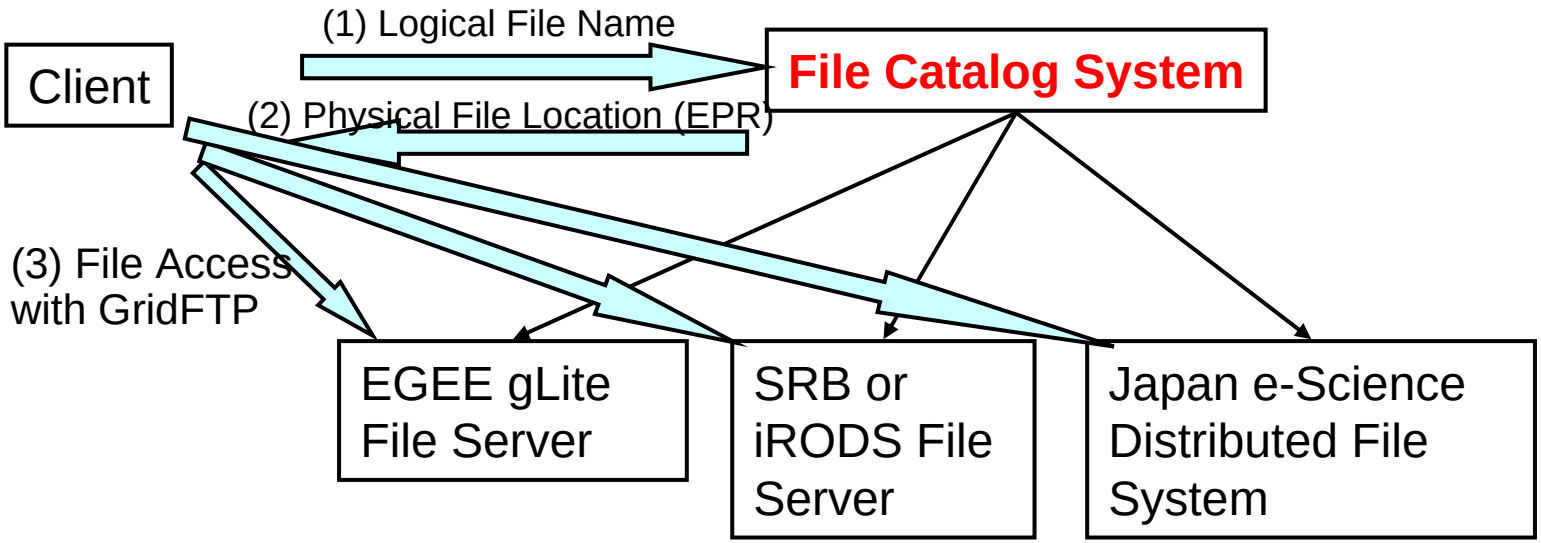
File Catalog Federation



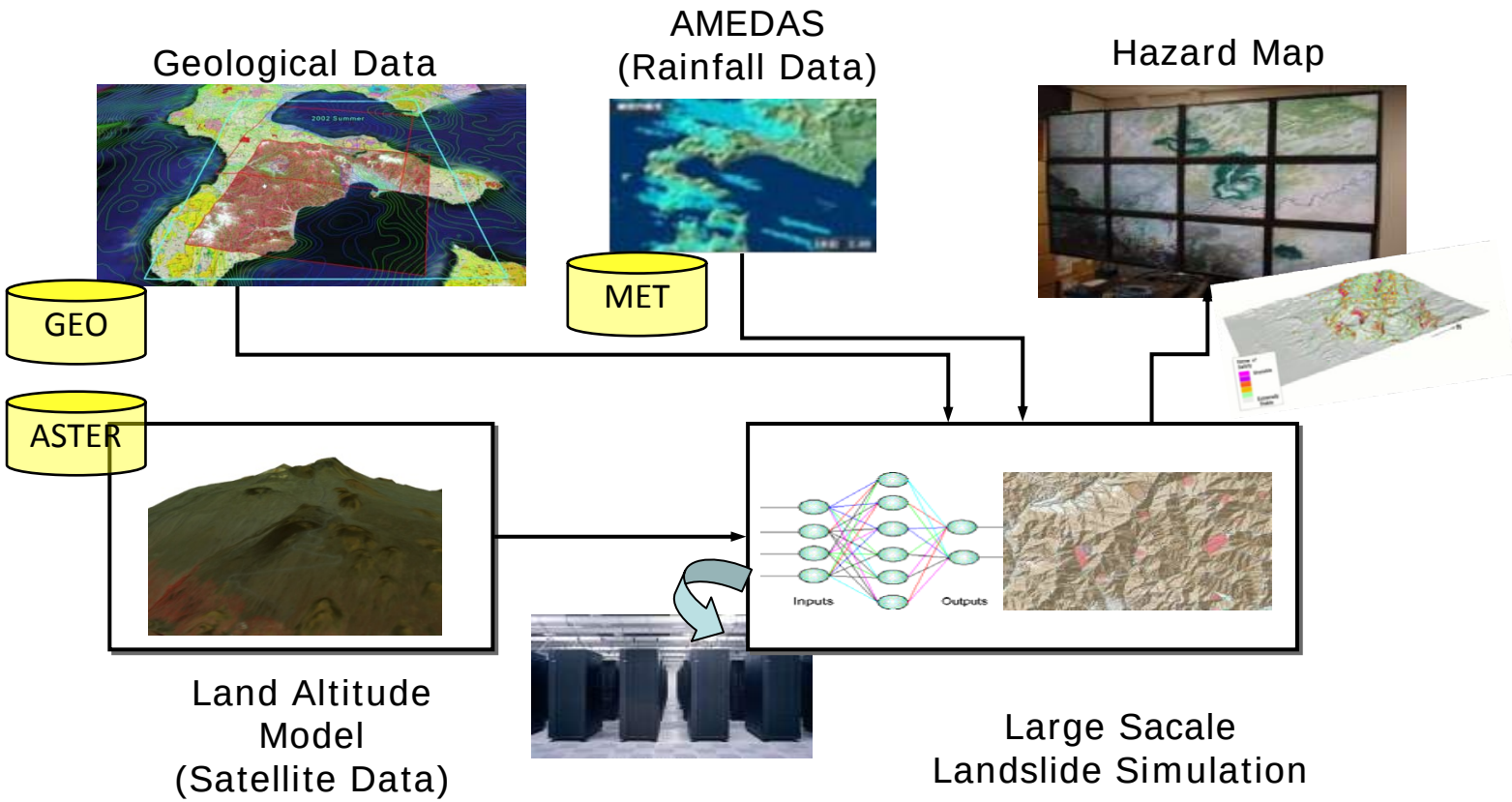
File Catalog Service

Goal: Development of interoperable file catalog service between heterogeneous Grid environments.

- Current file catalog systems (LFC (EGEE gLite), MCAT (SRB), etc.) do not have interoperability among each other.
- Development of standardized file catalog based on **RNS (Resource Namespace Service)** specifications (OGF).



Database Federation



KEK and Collaborating Organizations

gLite

- CERN
- Academia Sinica (Taiwan)
- Tohoku Univ.
- Tsukuba Univ.
- Nagoya Univ.
- Kobe Univ.
- Hiroshima Inst. Tec
- Etc.

NAREGI/RENKEI

- NII
- CC-IN2P3 (Lyon, France)
- NAOJ

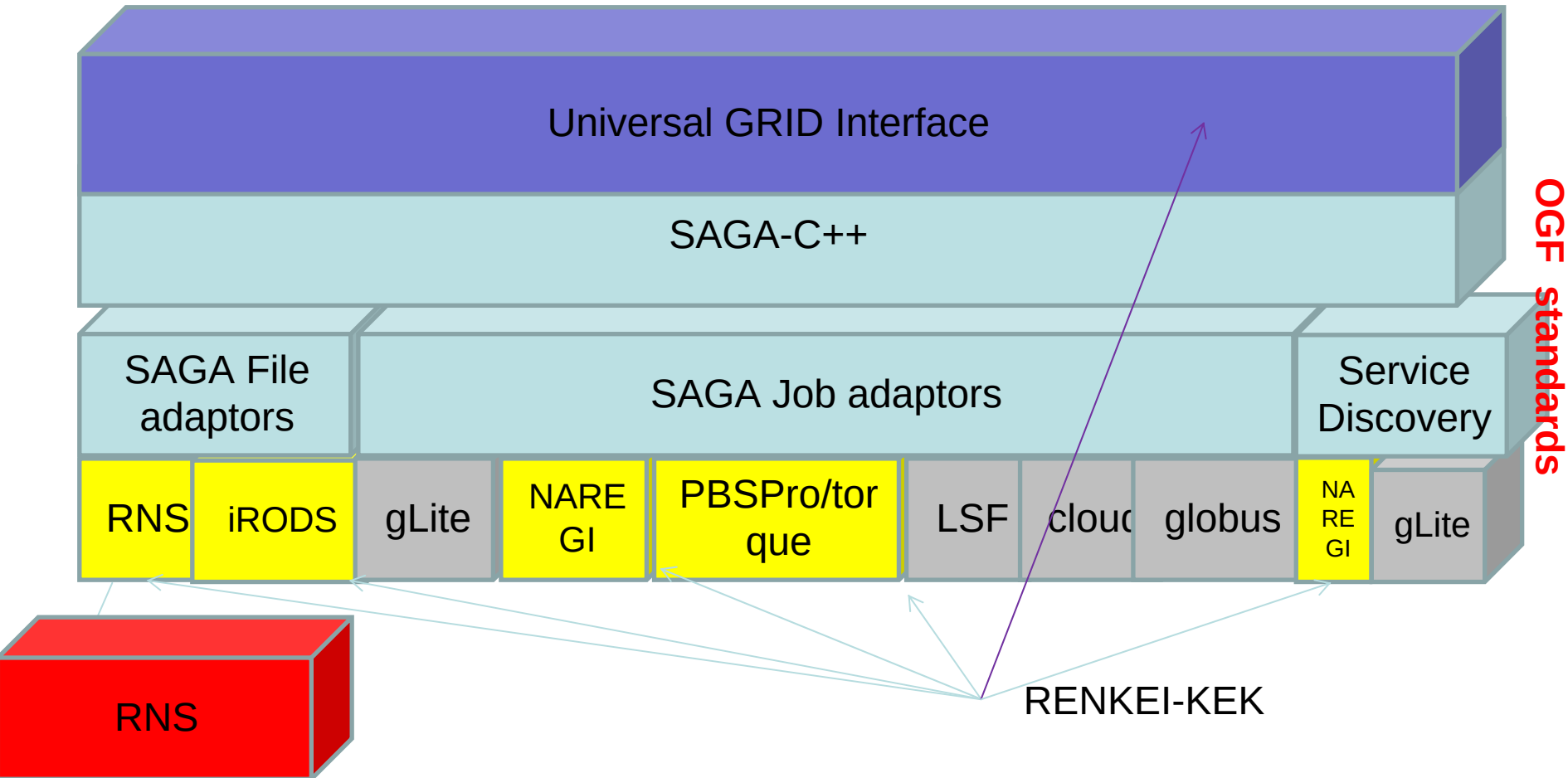
Interoperability

- **GIN**
- **JSAGA**

Universal GRID Interface(UGI)

Goal:

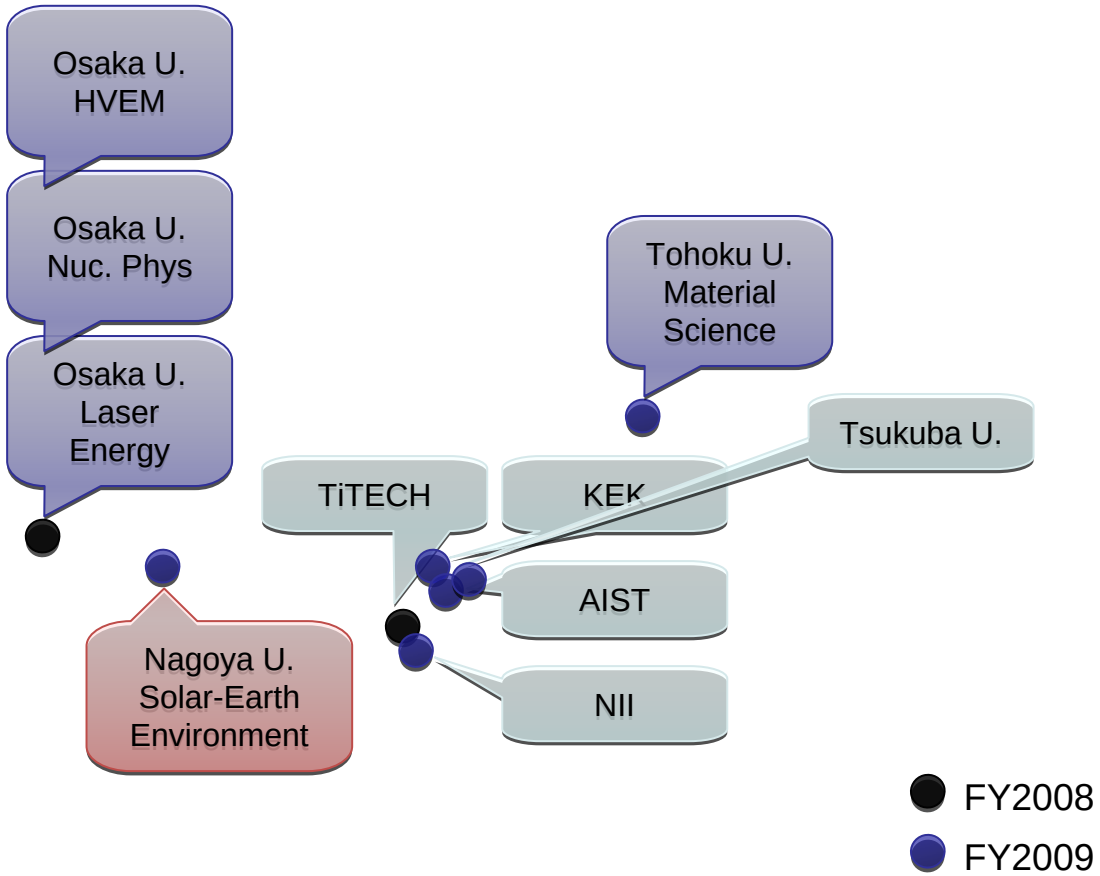
- Hide the differences of underlying middleware from users
- Single commands set will work for everything



RENKEI PoP (Testbed)

Planned Sites (10Gbps connections)

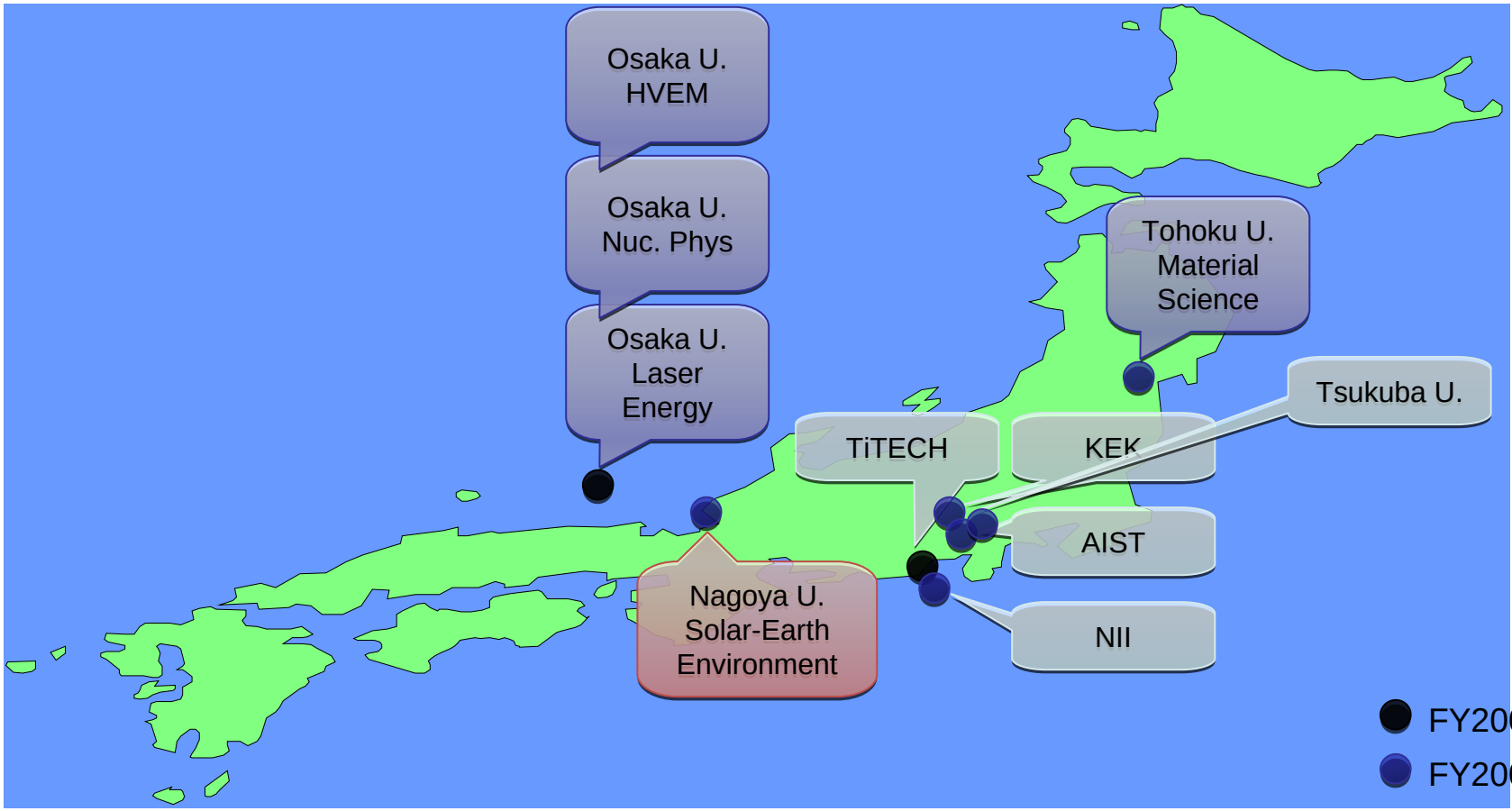
- 8sites,200TB(raw)・100TB(stable,with Apps.) – (Storage not included)



RENKEI PoP (Testbed)

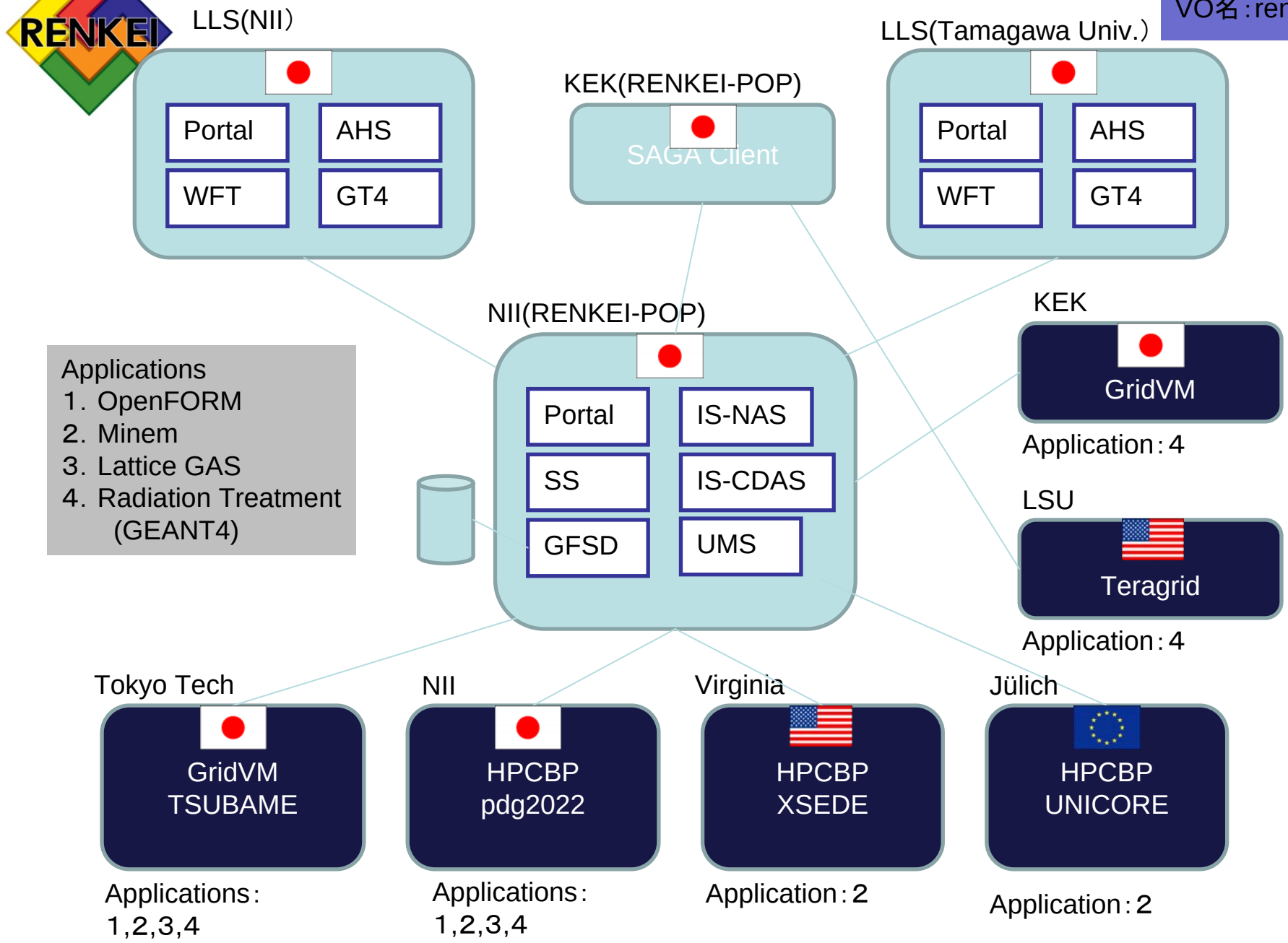
Planned Sites (10Gbps connections)

- 8sites,200TB(raw)・100TB(stable,with Apps.) – (Storage not included)



FY2010-2011 Test Bed Environment

VO名：renkei



The Next Generation Supercomputer Project

Total Budget:

about 115 billion Yen (1.15 billion US dollars)

Period of Project:

FY2006 – FY2012

Courtesy of RIKEN

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 (HPCI)

4. Establishment of “Institute for Advanced Computational Science”

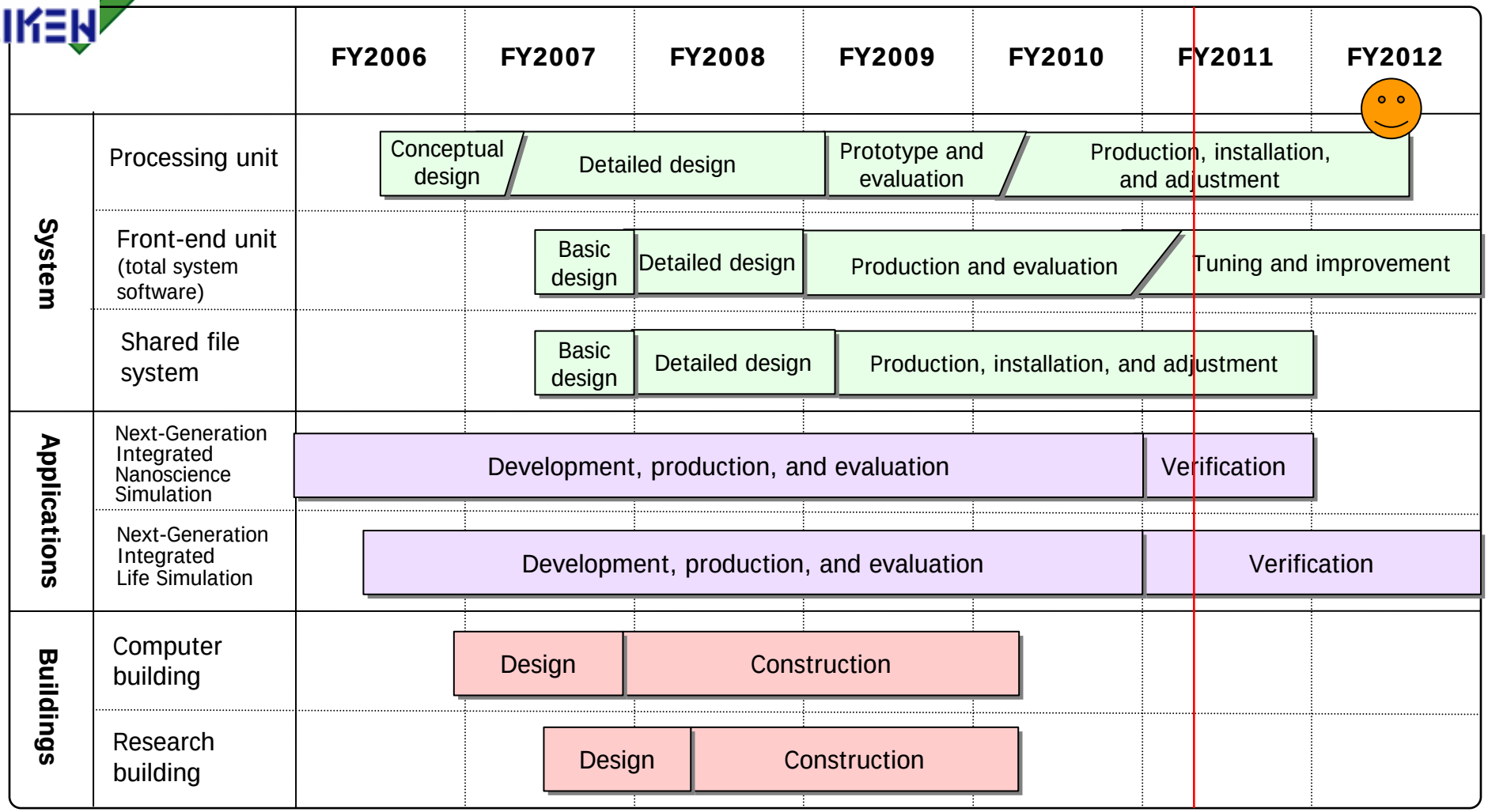
Courtesy of RIKEN

京

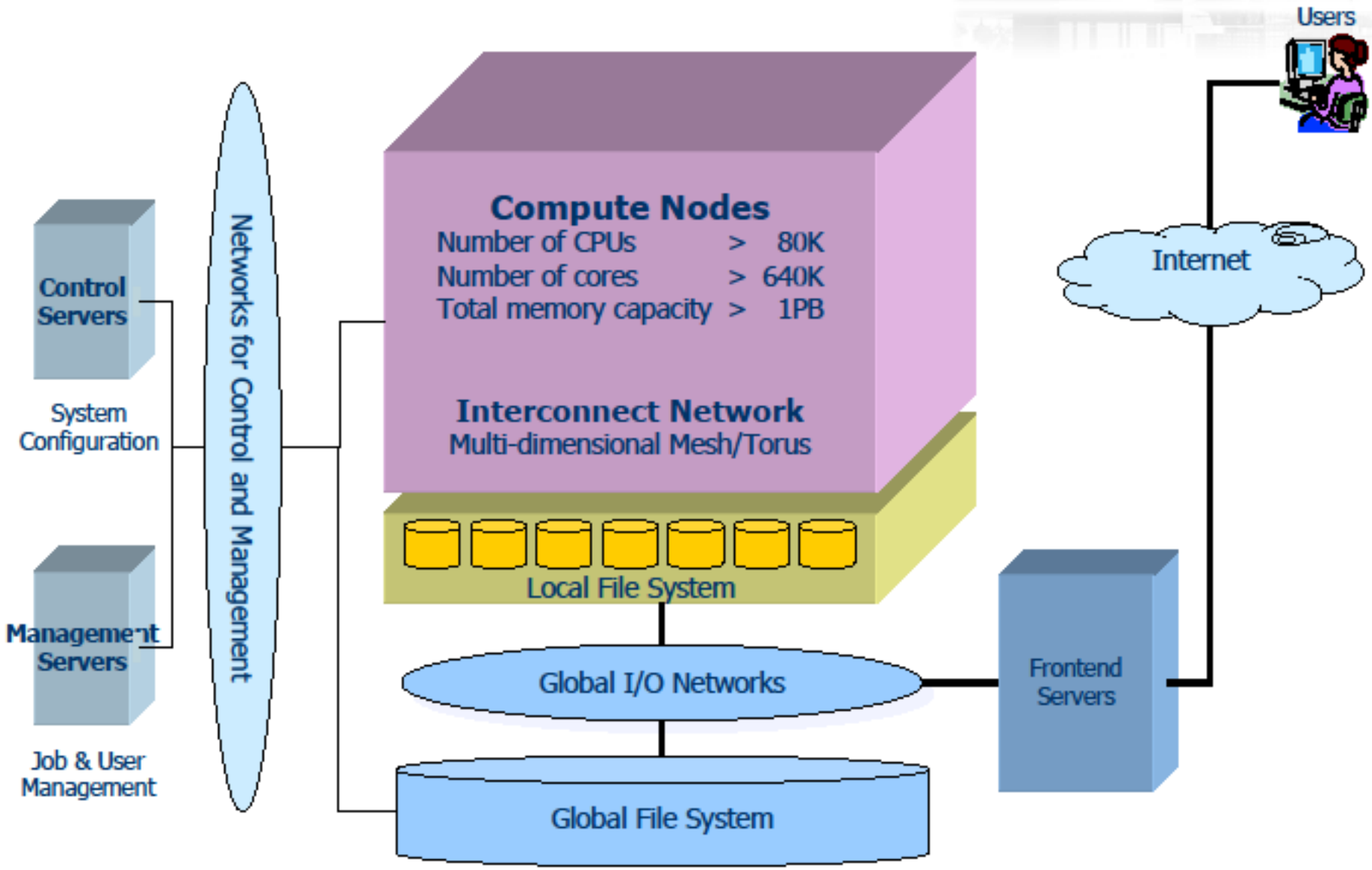
K computer

- "京 (Kei)"
  - 10^{16}, or 10 peta (flops system)
  - Arch (to a new era of computational science)

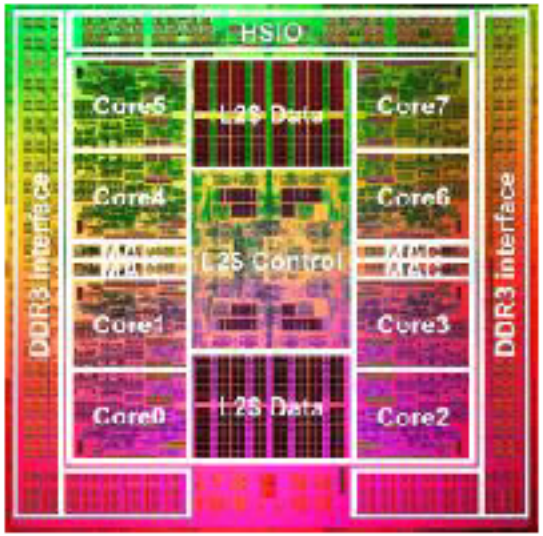
Schedule of Project



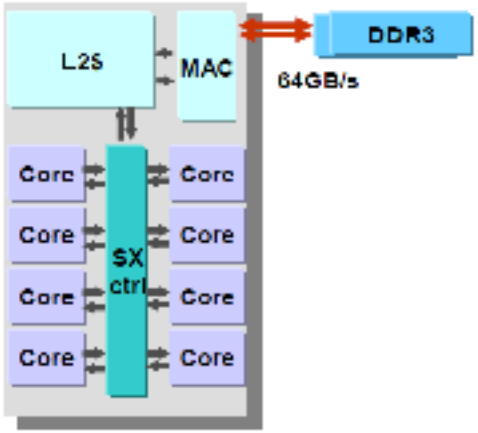
System Configuration



SPARC64™VIIIfx Chip Overview



Courtesy of FUJITSU Ltd.

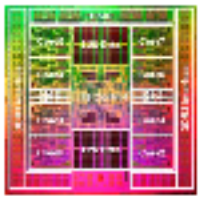


- Architecture Features
  - 8 cores
  - Shared 6MB L2$
  - Embedded memory controller
  - 2 GHz clock
- Fujitsu 45nm CMOS
  - 22.7mm x 22.6mm
  - 760 M transistors
  - 1,271 signal pins
- Performance (peak)
  - 128 GFlops
  - 64 GB/s memory bandwidth
- Power
  - 58W (typ.,@30℃)
  - Water cooling
    - Low leakage current and high reliability

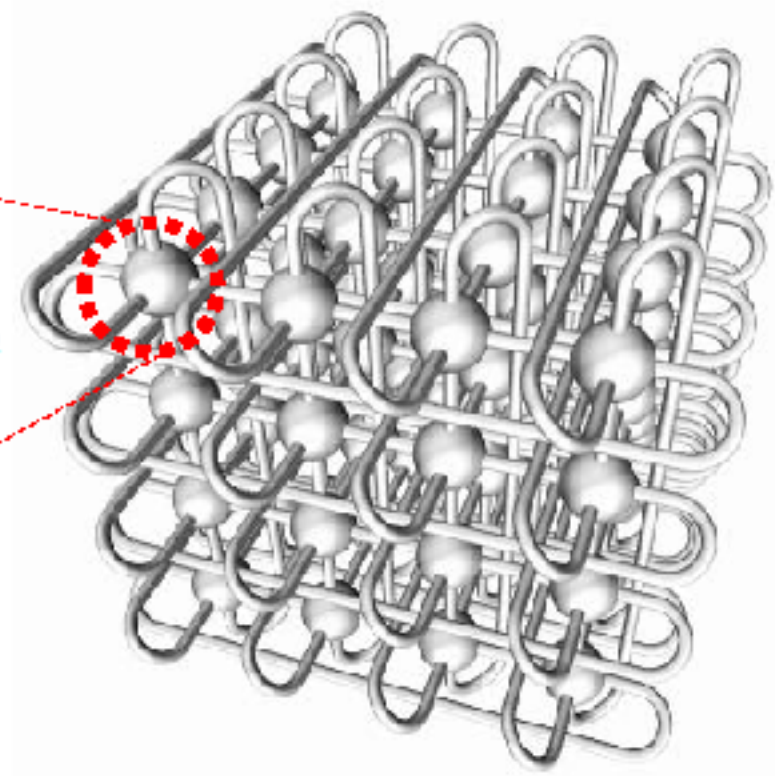
Compute nodes and network

- Compute nodes (CPUs): > 80,000
  - Number of cores: > 640,000
- Peak performance: > 10PFLOPS
- Memory: > 1PB (16GB/node)

- 6-dimensional mesh/torus network: Tofu
  - 10 connections to each adjacent node
- Peak bandwidth: 5GB/s x 2 for each connection
- Logically 3-dimensional torus network



SPARC64™ VIIIfx



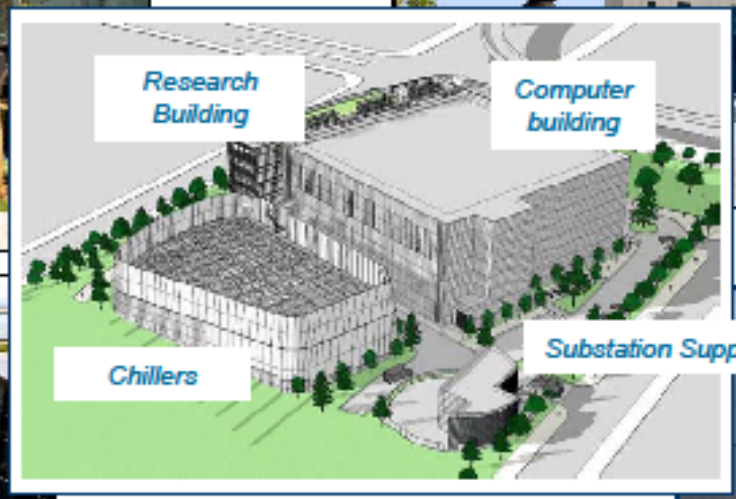
Courtesy of FUJITSU Ltd.

Hardware Basic Performance

- CPU
  - DGEMM (Matrix Multiply)
    - 123.6 GFlops (efficiency :96.6%) in 21,504x21,504x448
  - STREAM Triad (Memory Bandwidth)
    - 46.6 GB/s (peak : 64 GB/s)
  - Barrier sync. for inter-cores
    - 49 ns
- Network (ICC)
  - Throughput between 2 nodes
    - 4.75 GB/s (peak : 5 GB/s)
  - Latency
    - Max. 112 ns/hop
- LINPACK
  - RMax : 8.162PFlops (Efficiency : 93%)
  - RPeak: 8.774PFlops
  - Cores : 548,352
  - Nmax : 10,725,120
  - Time : about 28 Hours
  - Speed/W : 824.56MFlops/W

Major Applications of Next Generation Supercomputer

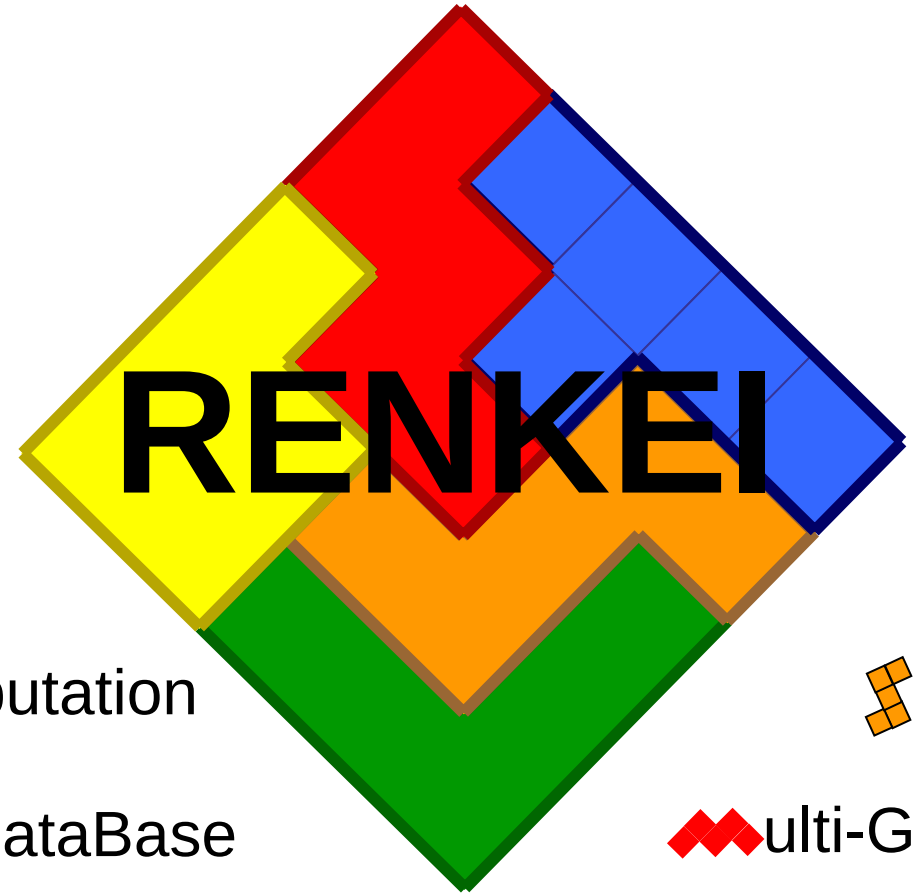
Layout of the buildings



Summary

- NAREGI Grid MiddlewareV.1.1.5 is being deployed to the national supercomputer centers as the important component of the Japanese Cyber Science Infrastructure Framework.
- A new project (RENKEI) started in FY 2008 to provide seamless access between NAREGI and the 3rd Tier resources.
- NI I is planned to provide the Networking access and security infrastructure for the Next Generation Supercomputer System

Resources Linkage for e-Science



omputation

hared Files

ataBase

ulti-Grid API

estBed

National Institute of Informatics (NII)
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan
URL: http://www.e-sciren.org/index-e.html