



Laboratoire d'Informatique Fondamentale de Lille



---

# Les grilles et l'écosystème nippon de calcul haute performance

---

Serge G. Petiton  
Serge.petiton@lifl.fr



# Plan

1. Introduction
2. Des projets nippons :
  - Naregi, *slides de Kenichi Miura (NII)*,
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. Le projet CNRS-JST NEGST
4. Conclusion



# Plan

1. **Introduction**
2. Des projets nippons :
  - Naregi, *slides de Kenichi Miura (NII)*,
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. Le projet CNRS-JST NEGST
4. Conclusion

# Introduction

- Plusieurs projets de Grilles au Japon depuis quelques années,
- Un projet majeur, « Naregi », mené par NII,
- Regroupement de chercheurs de divers domaines, dont les STIC, et des industriels,
- Des Grilles opérationnelles sur les campus, accessibles à l'ensemble des étudiants,
- Plusieurs collaborations importantes Franco-Japonaises autour des Grilles,
- Les Grilles sont au cœur de l'écosystème nippon de calcul haute performance, en particulier du projet Next-Gen,
- Grilles et Calculs Hautes Performances sont désormais intégrés sans aucune autre distinction.



# Plan

1. Introduction
2. **Des projets nippons :**
  - **Naregi, slides de Kenichi Miura (NII),**
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. Le projet CNRS-JST NEGST
4. Conclusion

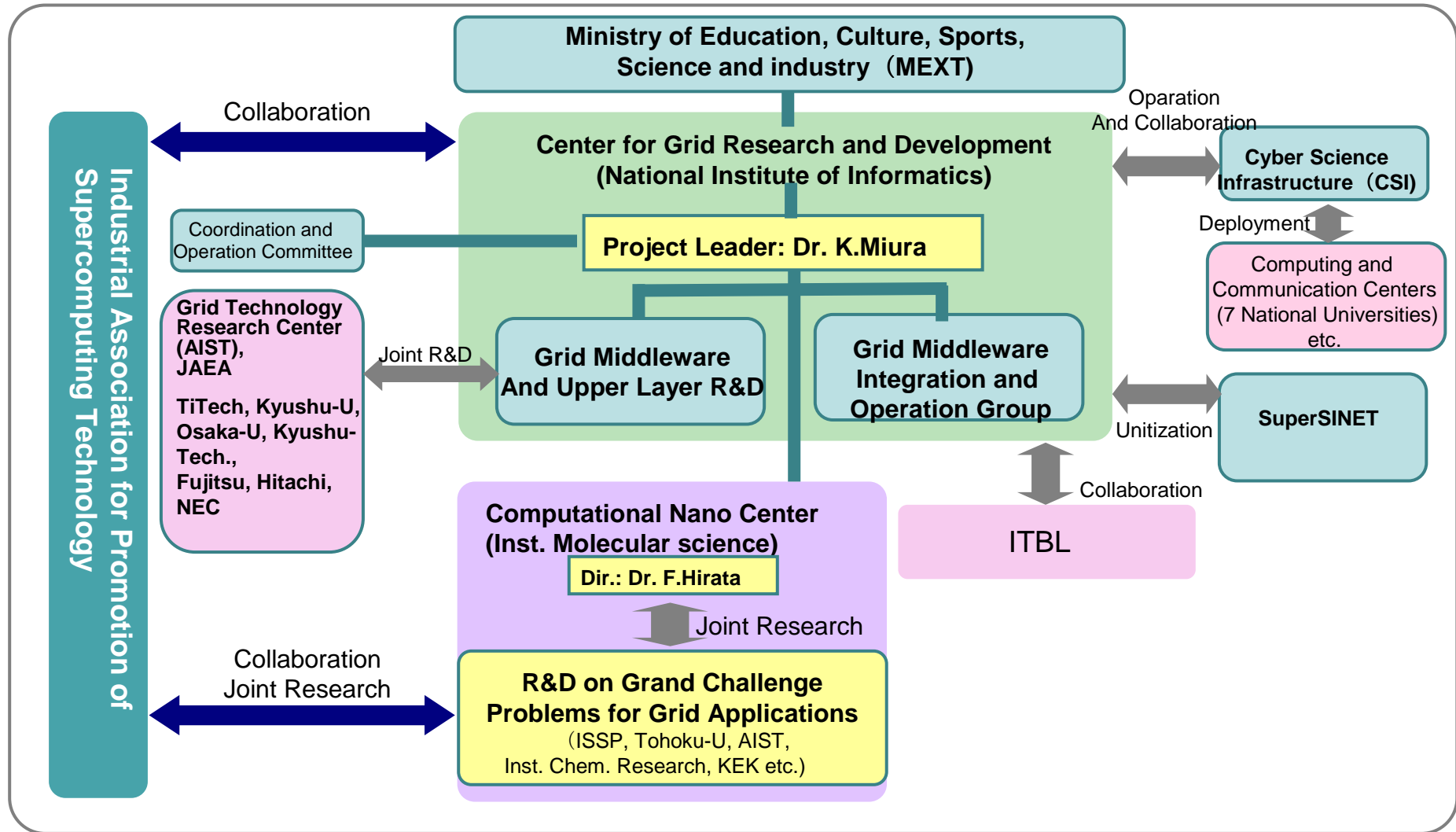
# National Research Grid Initiative (NAREGI) Project:Overview

- Started as an R&D project funded by MEXT (FY2003-FY2007)
- One of Japanese Government's Grid Computing Projects  
ITBL, Visualization Grid, GTRC, BioGrid etc.
- Collaboration of National Labs. Universities and Industry  
in the R&D activities (IT and Nano-science Apps.)
- NAREGI Testbed Computer Resources

# National Research Grid Initiative (NAREGI) Project:Goals

- (1) To develop a Grid Software System (R&D in Grid Middleware and Upper Layer) as the prototype of future Grid Infrastructure in scientific research in Japan
- (2) To provide a Testbed to prove that the High-end Grid Computing Environment can be practically utilized in the Nano-science Applications over the Super SINET.
- (3) To Participate in International Collaboration (U.S., Europe, Asian Pacific)
- (4) To Contribute to Standardization Activities, e.g., GGF

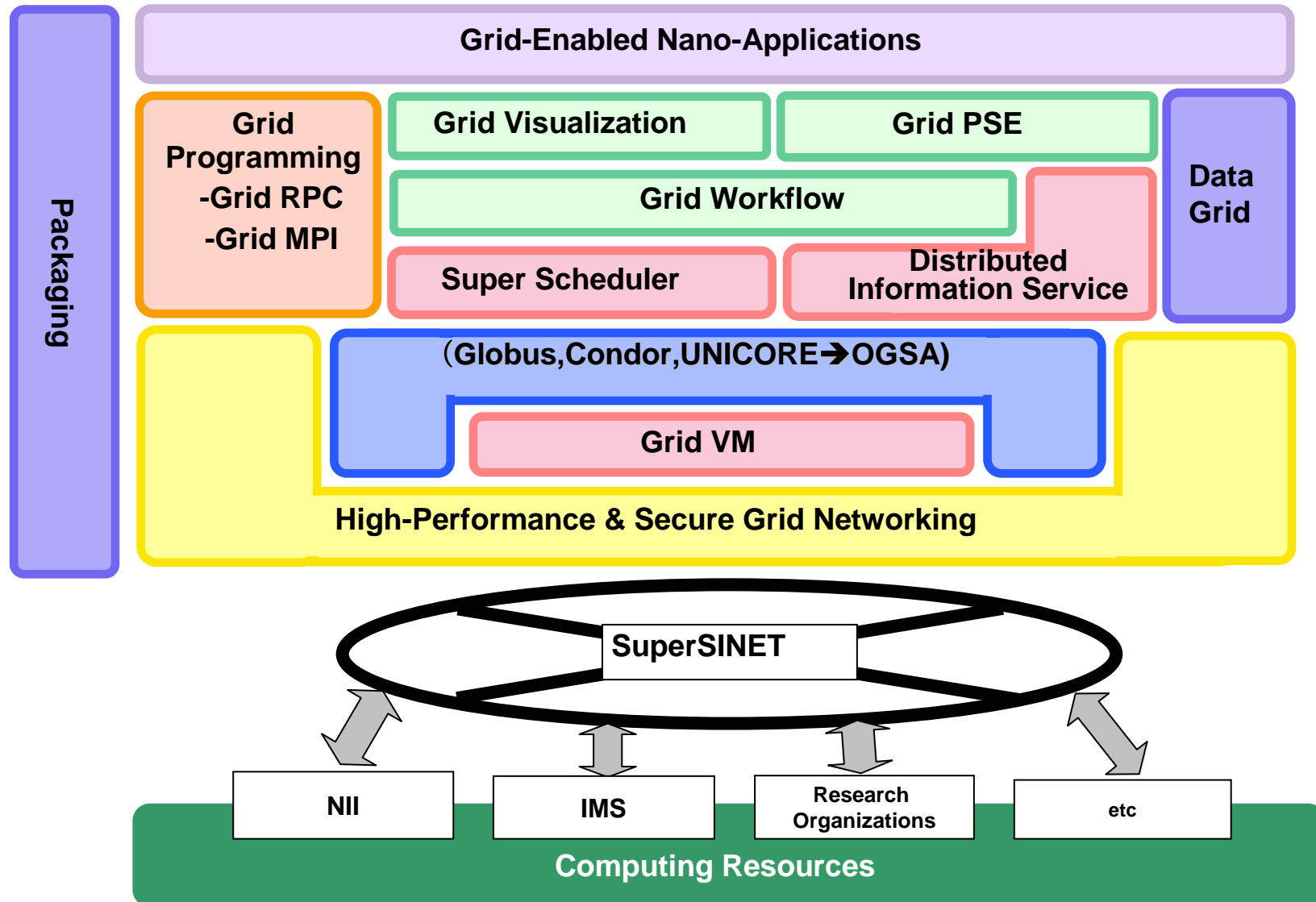
# Organization of NAREGI







# NAREGI Software Stack



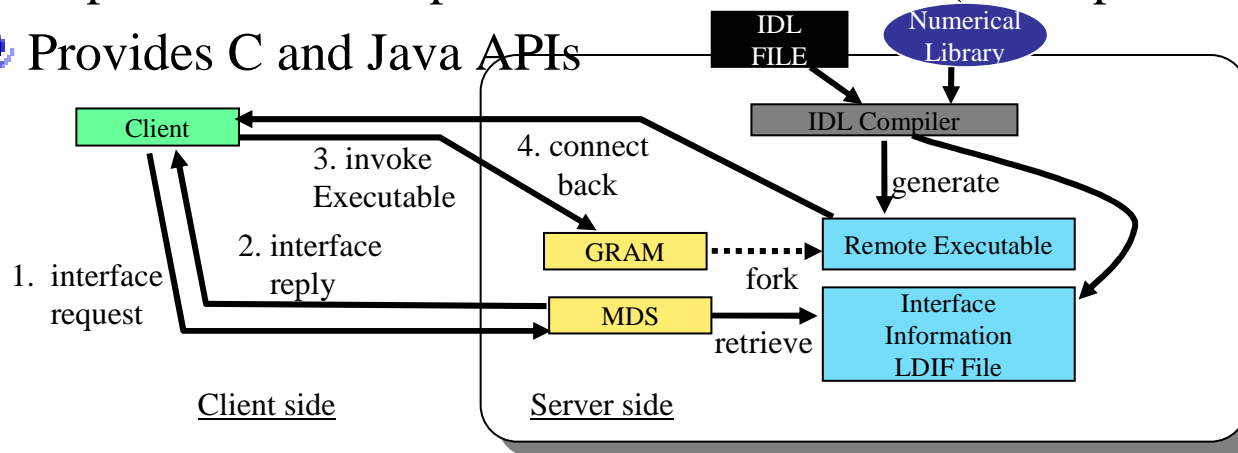
# R&D in Grid Software and Networking Area (Work Packages)

- WP-1: Lower and Middle-Tier Middleware for Resource Management:  
**Matsuoka** (Titech), Kohno(ECU), Aida (Titech)
- WP-2: Grid Programming Middleware:  
**Sekiguchi**(AIST), Ishikawa(AIST)
- WP-3: User-Level Grid Tools & PSE:  
**Usami** (NII), Kawata(Utsunomiya-u)
- WP-4: Data Grid Environment  
**Matsuda** (Osaka-u)
- WP-5: Networking, Security & User Management  
**Shimojo** (Osaka-u), Oie ( Kyushu Tech.),  
Imase (Osaka U.)
- WP-6: Grid-enabling tools for Nanoscience Applications :  
**Aoyagi** (Kyushu-u)

# WP-2:Grid Programming – GridRPC/Ninf-G2 (AIST/GTRC)

## GridRPC

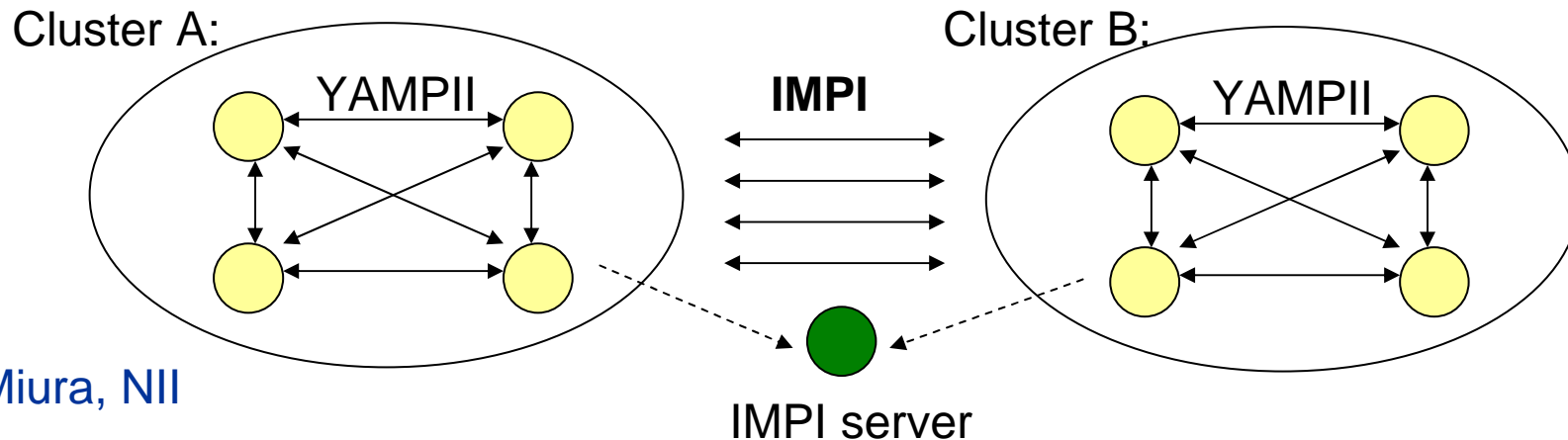
- ◆ Programming Model using RPC on the Grid
- ◆ High-level, tailored for Scientific Computing (c.f. SOAP-RPC)
- ◆ GridRPC API standardization by GGF GridRPC WG
- ◆ Ninf-G Version 2
  - ⊙ A reference implementation of GridRPC API
  - ⊙ Implemented on top of Globus Toolkit 2.0 (3.0 experimental)
  - ⊙ Provides C and Java APIs



# WP-2:Grid Programming

## -GridMPI (AIST and U-Tokyo)

- GridMPI is a library which enables MPI communication between parallel systems in the grid environment. This realizes;
    - Huge data size jobs which cannot be executed in a single cluster system
    - Multi-Physics jobs in the heterogeneous CPU architecture environment
- ① Interoperability:
    - IMPI (Interoperable MPI) compliance communication protocol
    - Strict adherence to MPI standard in implementation
  - ② High performance:
    - Simple implementation
    - Built-in wrapper to vendor-provided MPI library



# WP-3: User-Level Grid Tools & PSE

## ■ Grid PSE

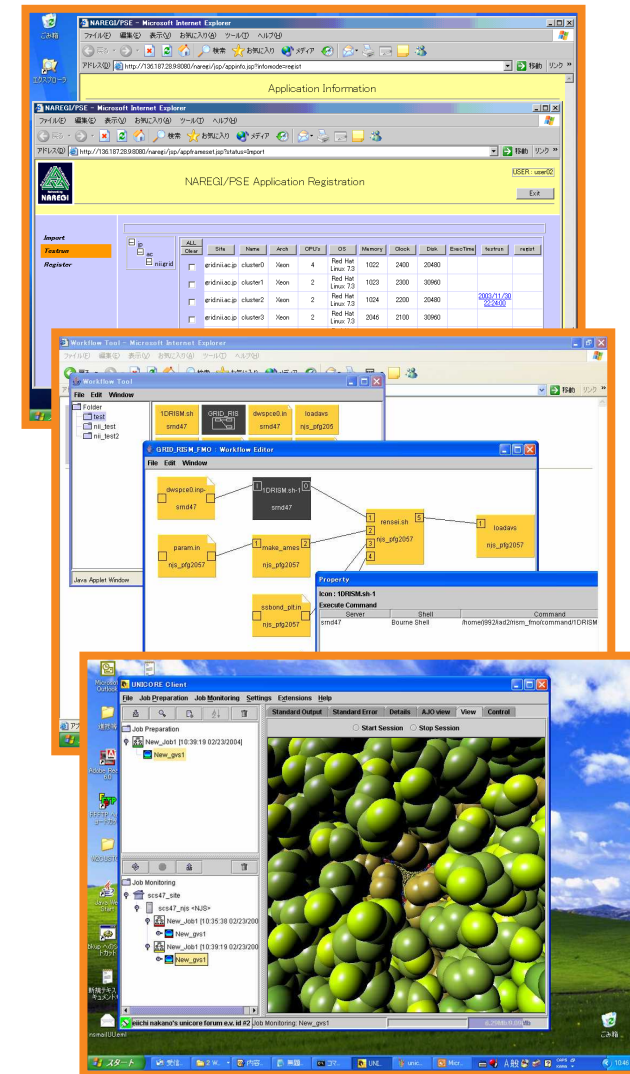
- Deployment of applications on the Grid
- Support for execution of deployed applications

## ■ Grid Workflow

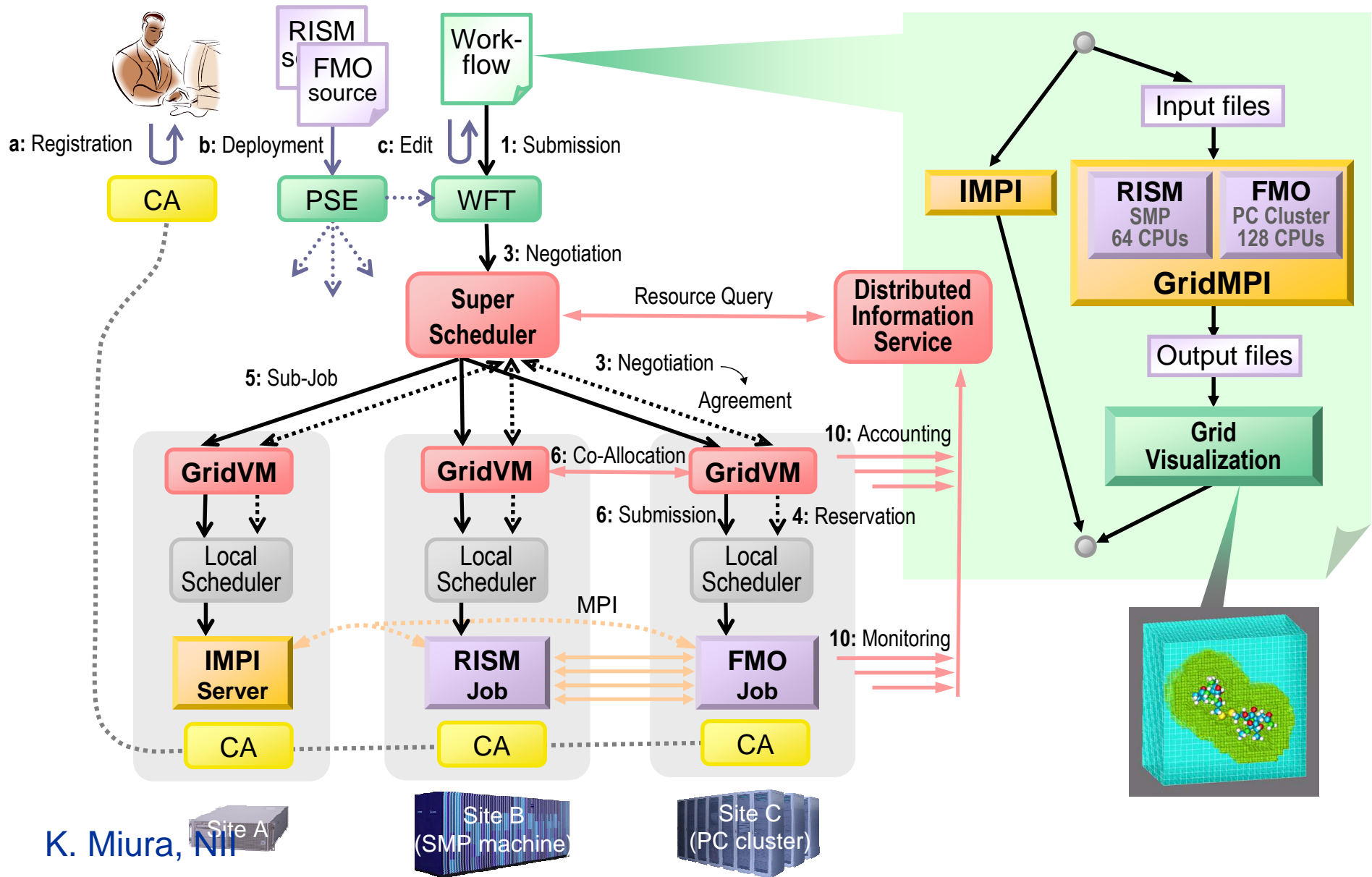
- Workflow language independent of specific Grid middleware
- GUI in task-flow representation

## ■ Grid Visualization

- Remote visualization of massive data distributed over the Grid
- General Grid services for visualization



# Scenario for Multi-sites MPI Job Execution



# Collaboration in Data Grid Area

- High Energy Physics
  - KEK
  - EGEE
- Astronomy
  - National Astronomical Observatory  
(Virtual Observatory)
- Bio-informatics
  - BioGrid Project

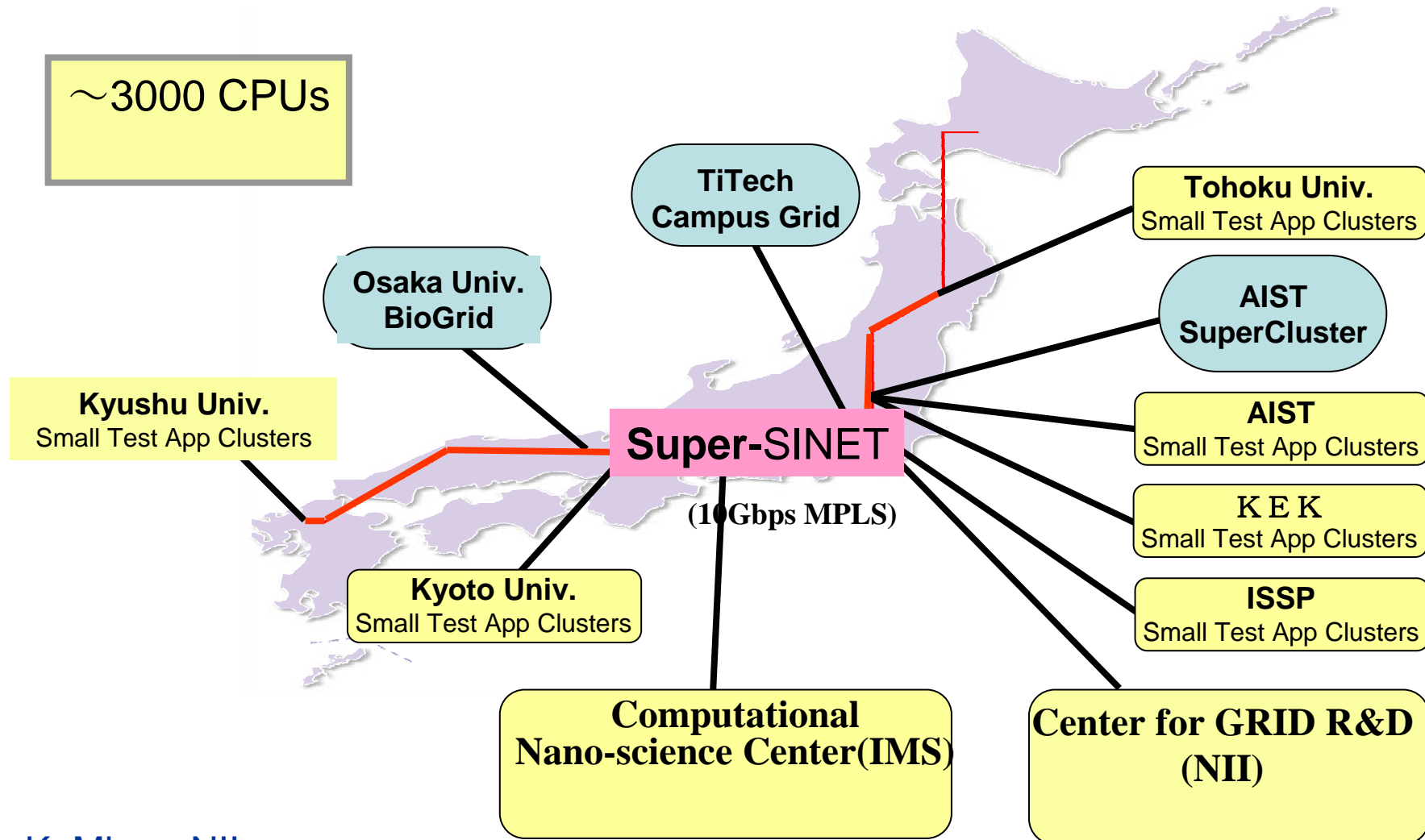


# NAREGI Version 1

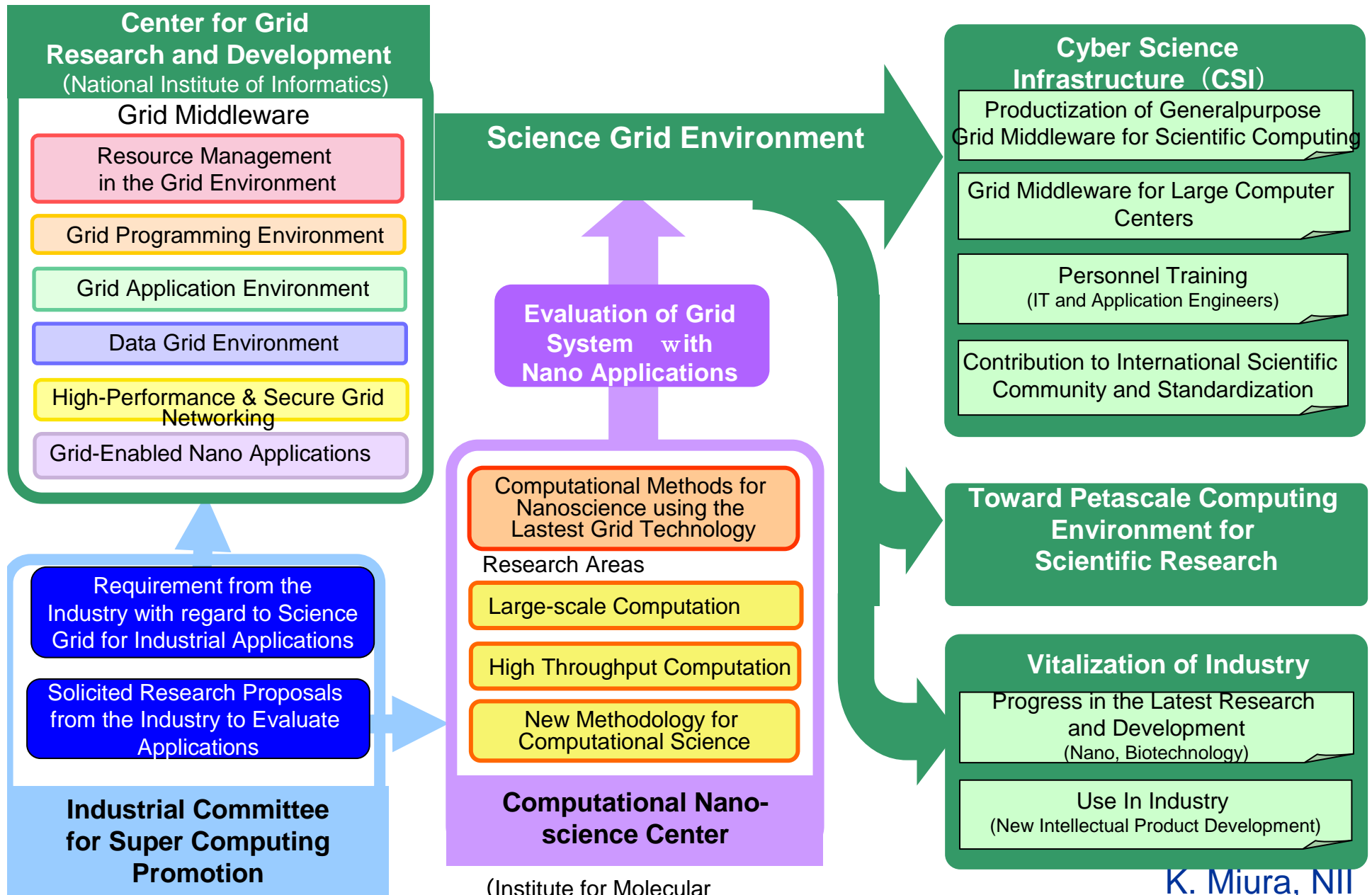
Operability, Robustness, Maintainability

- To be developed in FY2007
- More flexible scheduling methods
  - Reservation-based scheduling
  - Coexistence with locally scheduled jobs
  - Support of Non-reservation-based scheduling
  - Support of “Bulk submission” for parameter search type jobs
- Improvement in maintainability
  - More systematic logging using Information Service (IS)

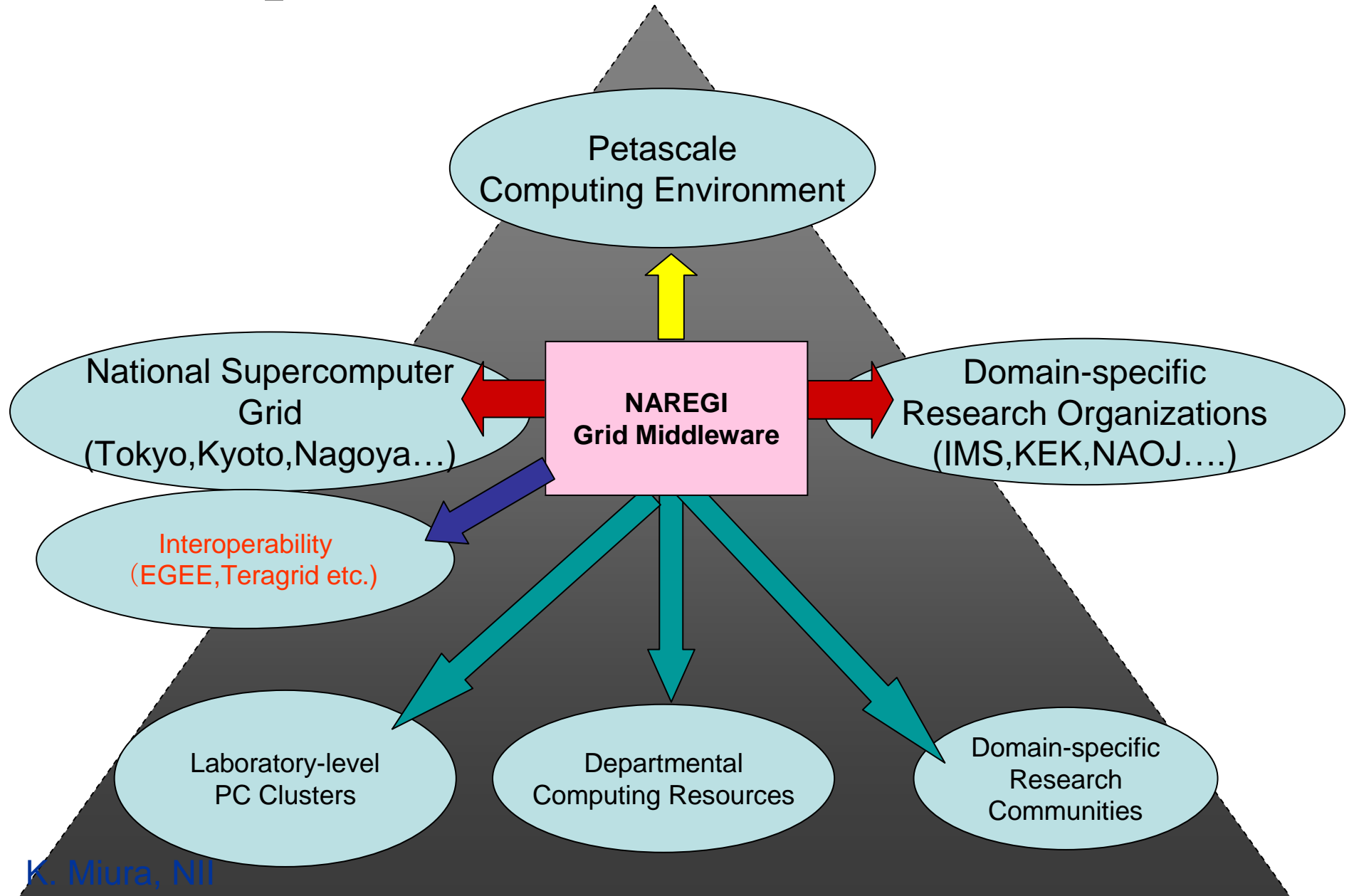
# NAREGI Phase 1 Testbed



# Future Direction of NAREGI Grid Middleware



# Expansion Plan of NAREGI Grid



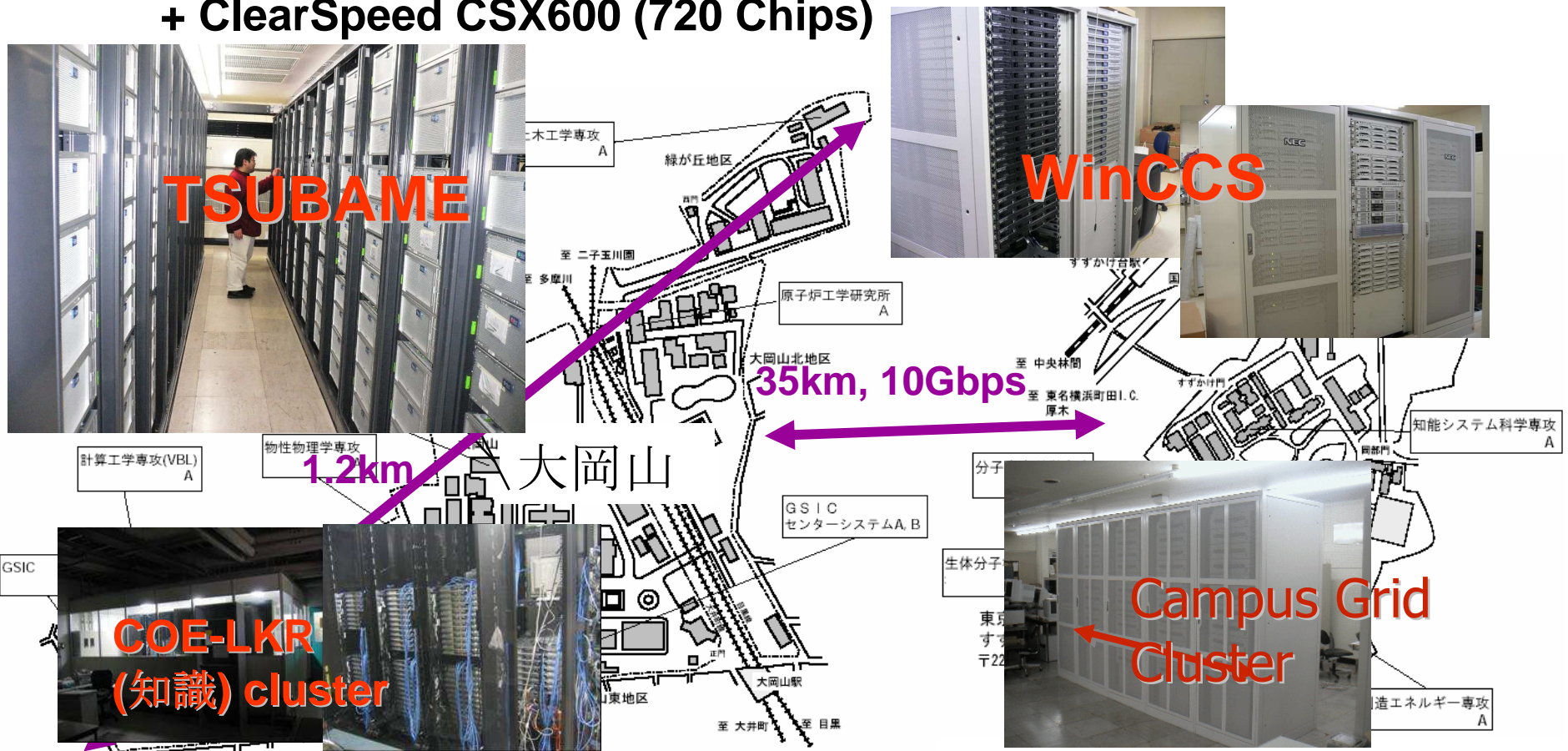


# Plan

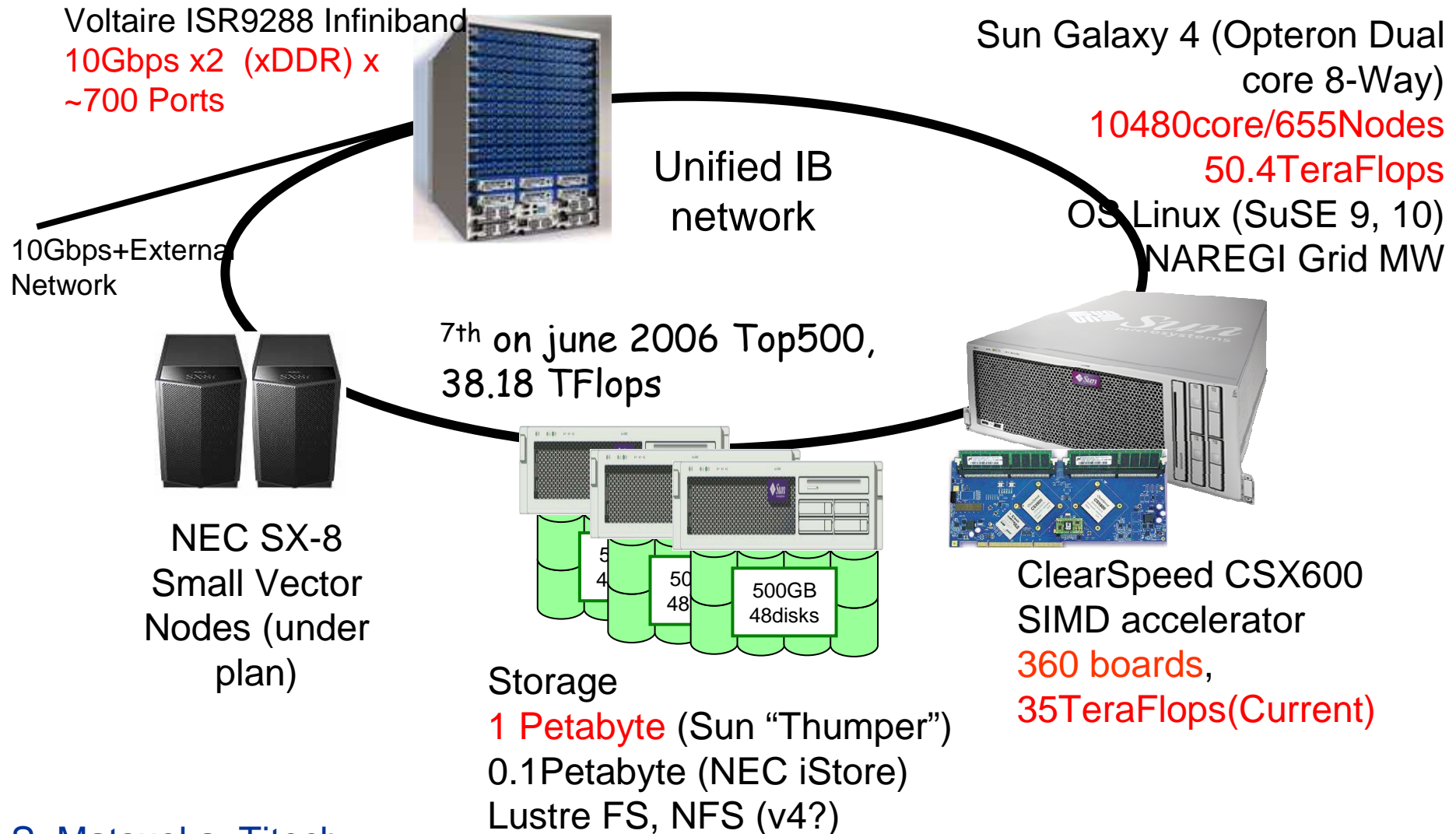
1. Introduction
2. **Des projets nippons :**
  - Naregi, *slides de Kenichi Miura (NII)*,
  - **Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)***
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. Le projet CNRS-JST NEGST
4. Conclusion

# Titech Supercomputing Grid

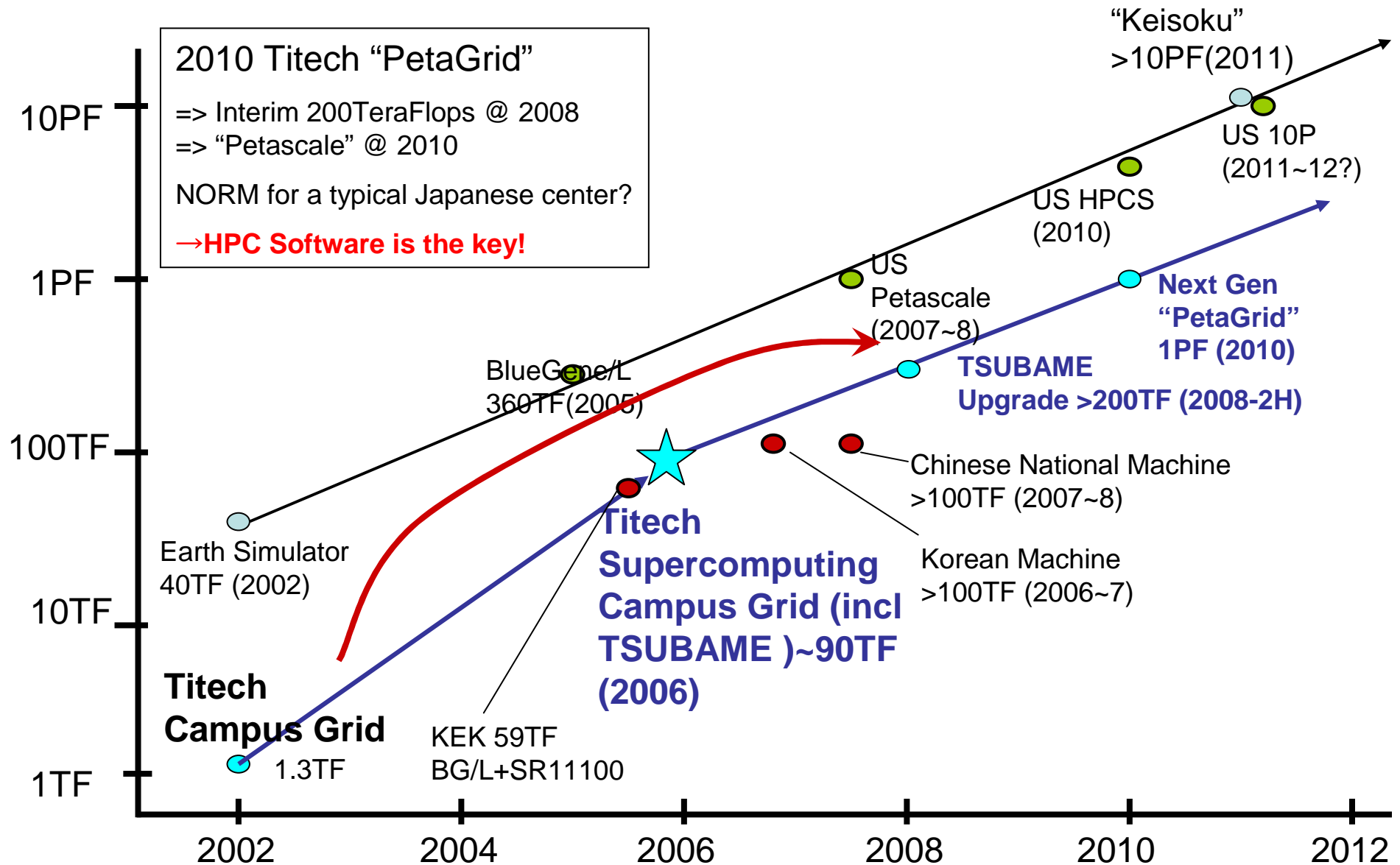
- ~13,000 CPUs, 90 TeraFlops, ~26 TeraBytes Mem, ~1.1 Petabytes Disk
- CPU Cores: x86: TSUBAME (~10600), Campus Grid Cluster (~1000), COE-LKR cluster (~260), WinCCS (~300) + ClearSpeed CSX600 (720 Chips)



# The Titech TSUBAME Production Supercomputing Cluster



# Scaling Towards Petaflops...







# Plan

1. Introduction
2. **Des projets nippons :**
  - Naregi, *slides de Kenichi Miura (NII)*,
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - **Next-Gen, *slides de Tamaki Watanabe (Riken)***
3. Le projet CNRS-JST NEGST
4. Conclusion

# Development & Application of Next-Generation Supercomputer Project by MEXT

FY2006: 3,547Million yen / FY2007: 7,736Million 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

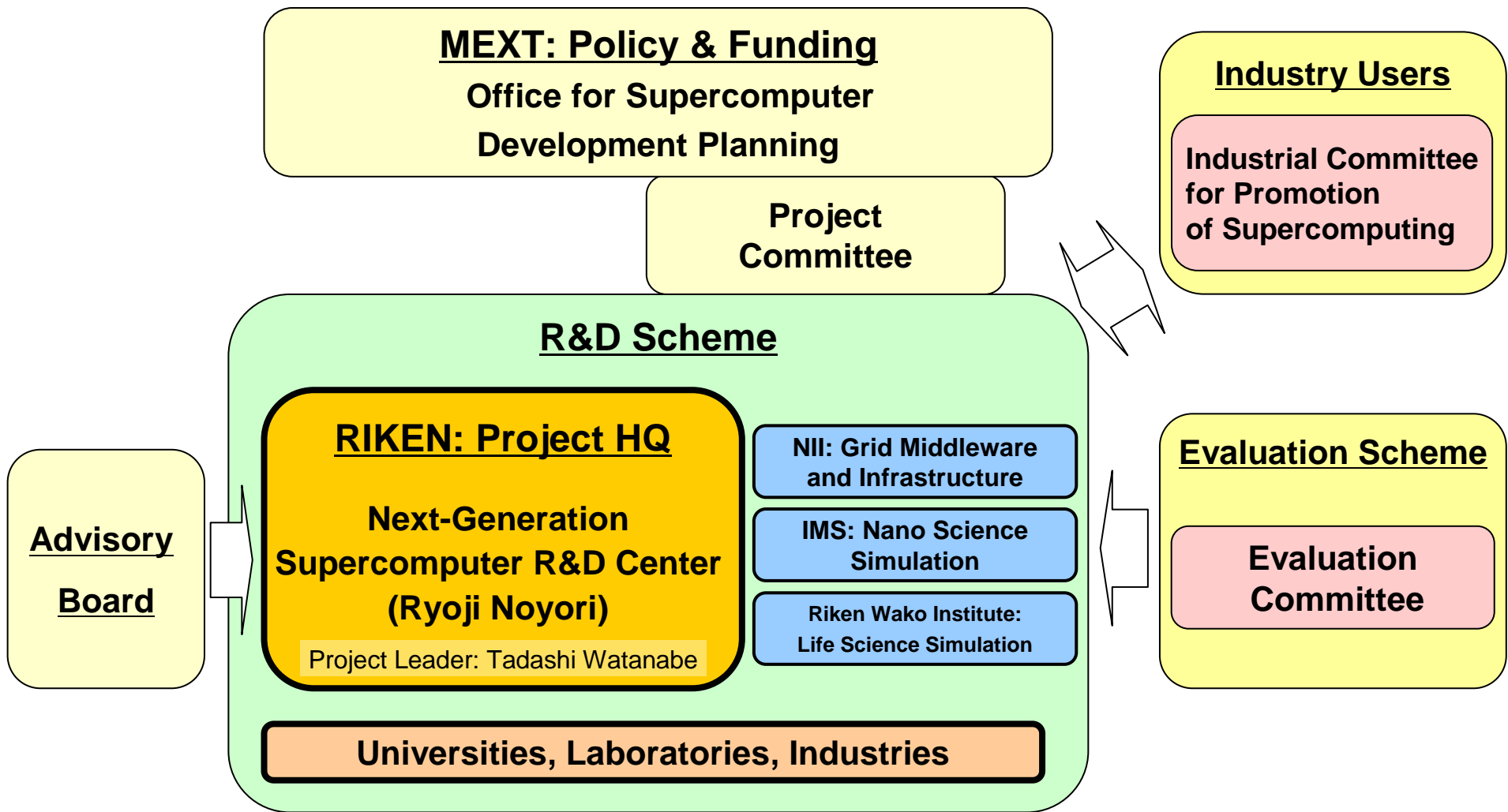
- Integrated development of computer and software
- Establishment of nationwide academic-industrial collaborative structure, with RIKEN as the project headquarters
- A new law has been introduced for the framework of usage and administration

# Application Programs

21 application programs have been identified as candidates to be used for architecture evaluation

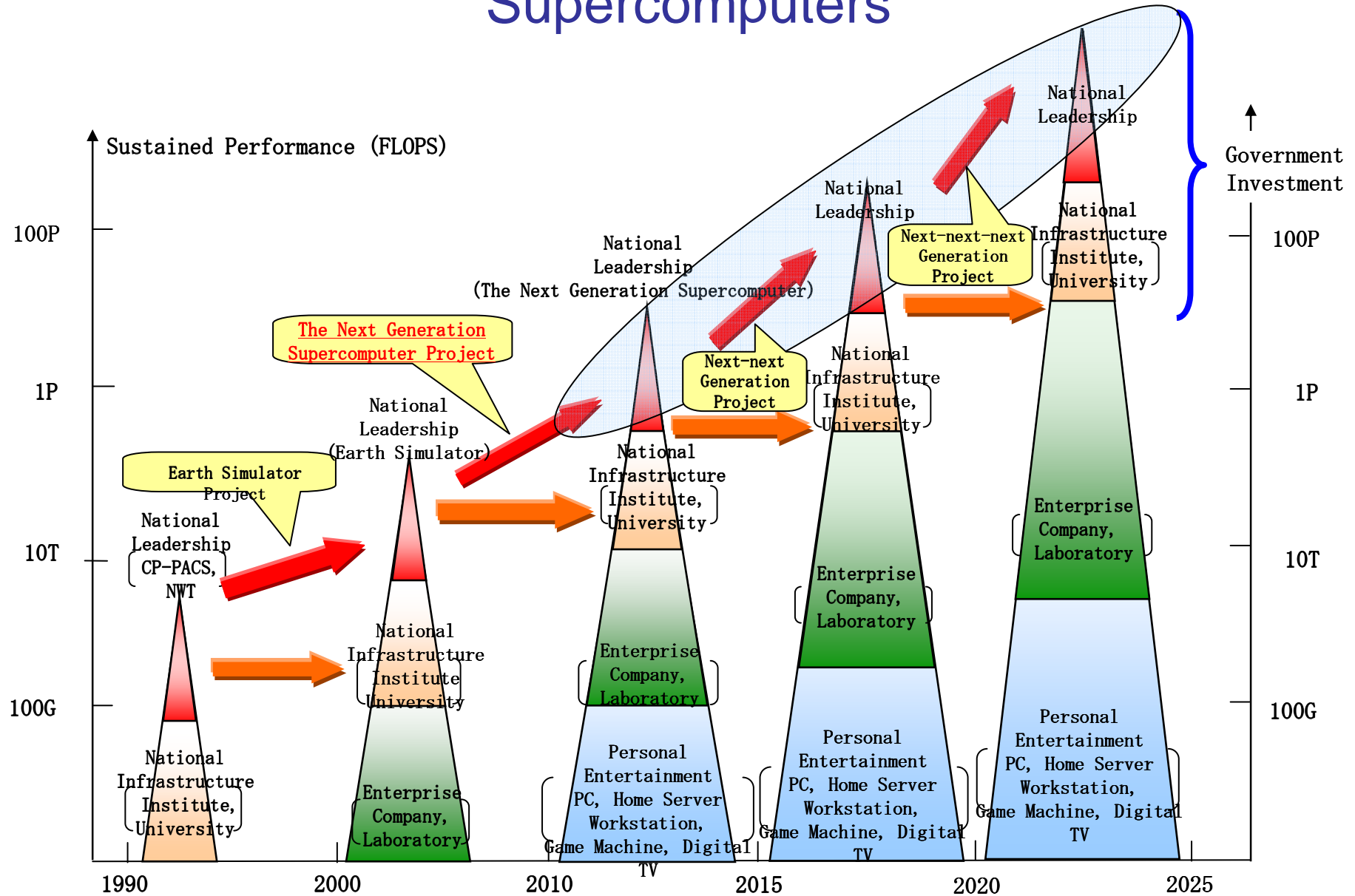
- Nanoscience (6), e.g., MD, Fragment MO, First Principle MD (physical space formulation), RISM (Reference Interaction Site Model), OCTA
- Life Science (6), e.g., MD (Protein), Density Function (Protein), Folding, Blood Flow, Gene network
- Climate/Geoscience (3)
- Earth Quake Wave Propagation, Hi-res. Atmospheric General Circulation Model, Hi-res. Ocean GCM
- Physical Science (2)
- Galaxy Formation (hydro + gravitational N-body), QCD
- Engineering (4), e.g., Structural Analysis by FEM, Non-steady Flow by FDM, Compressible Flow,
- Large Eddy Simulation

# Current Project Organization



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

# MEXT's Vision for Continuous Development of Supercomputers



# Summary

- The 3rd Science and Technology Basic Plan started in April 2006.
- The Next Generation Supercomputer Project (FY2006-2012, ~1B\$) is one of the high priority projects, aimed at the Petascale computation in the key application areas.
- NAREGI Grid middleware will enable seamless federation of heterogeneous computational resources.
- Computations in Nano-science/technology applications over Grid is to be promoted, including participation from industry.
- NAREGI will provide the access and computational infrastructure for the Next Generation Supercomputer System.
- NAREGI Grid Middleware is to be adopted as one of the important components in the new Japanese Cyber Science Infrastructure Framework.



# Plan

1. Introduction
2. Des projets nippons :
  - Naregi, *slides de Kenichi Miura (NII)*,
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. **Le projet CNRS-JST NEGST**
4. Conclusion

# **Projet CNRS-JST NEGST (NExT Grid Systems and Techniques) 2006-2007-2008**

## **International Collaboration and Promotion on Interoperability and advanced technologies of Grid**

The objective of this project is to promote the collaborations of Japan and France on grid computing technology on three important research issues:

- 1) Instant Grid and virtualization of grid computing resources,
- 2) Grid Metrics, and
- 3) Grid Interoperability and Applications.

An international testbed including the French Grid5000 project and the Japanese NAREGI project would be used to demonstrate and validate systems, software and applications.



# Instant Grid and virtualization of grid computing resources

The complexity of the existing Grid middleware and systems raises the need of discovering and developing original mechanisms and interfaces for users and administrators

- Experiments using several middleware such as OmniRPC, Xtremweb, CONDOR and YML, on Japanese-French Platforms and on TSUBAME (LIFL, PRISM, INRIA, Tsukuba-U, Titech,..).
- Data persistence : JuxMen and G-farm Grid File System (Univ. of Tsukuba, IRISA, AIST...)
- Redundant Interconnections : RI2N/UDP and PVC (Paris-Sud Univ, Univ. of Tsukuba,...)
- Kadeploy/Grid5000 and Naregi (LIG, INRIA, NII,...)
- Grid middleware DIET-Ninf fully GridRPC compliant (LIP, AIST,...)
- YML and OmniStorage: Performance improvement by Data Management Layer on a Grid RPC System, for a YML backend (LIFL, PRISM, INRIA, Univ. of Tsukuba,...)
- .../.....

# Grid Metrics

## **Grid is related to methodology and metrics for accurate measurement methods and simulation, virtualization tools**

- A Gnet box has been installed in the GRID5000 testbed at Lyon in order to realize several experiments on transport protocols and grid traffic measurement (LIP, INRIA, AIST,....)
- We are discussing with people in NAREGI around some of the approaches developed within the GRID-TLSE project (IRIT, NII,...) :
  - The concept of scenario: this is a high-level description of the sequence of operations required for answering an user-request with the framework of expertise in sparse linear algebra. The execution of scenarios generates a sequence of workflows that are dynamic since one workflow may depend upon the results of the previous one.
  - Semantic description of computational services : it is used for high-level service trading i.e. identifying all services capable of performing some algorithmic task (e.g. finding all services capable of solving a symmetric linear systems using a direct method)
- Others joint research topics :
  - Network measurement, emulation, protocol benchmarking and scheduling, lead by AIST,
  - Network resource control, traffic measurements, protocol benchmarking, lead by ENS Lyon,
  - Grid interoperability in HEP, lead by KEK

# Grid Interoperability and Applications

This issue concerns the interoperability between existing Grid middleware and deployments. How to connect and exchange resources securely and efficiently between several Grids? How to run application on Interoperable Grids.

- Grid interoperability of software using a large scientific dataset (IN2P3, KEK,...).
  - Phase one: installation of the NAREGI middleware at IN2P3, the EGEE/LCG middleware being already deployed. KEK and French side will assess what is necessary to achieve interoperable Datagrid,
  - Phase two: In full collaboration with KEK, the French side will take an active part in developing the software necessary to achieve interoperability, according to the work plan established during phase one. Part or all the applications considered by the Japanese side will be also deployed at IN2P3 in order to assess the interoperability on the same test cases.
- Matrix computation on several Japanese-French platforms (LIFL, PRISM, IRIT, Tuskuba, Titech, Todai,.....)
- Large Scale Optimization problems (LIFL, Titech, Todai, ...)
- .....



# Plan

1. Introduction
2. Des projets nippons :
  - Naregi, *slides de Kenichi Miura (NII)*,
  - Titech Supercomputing GRID, *slides de Satoshi Matsuoka (Titech)*
  - Next-Gen, *slides de Tamaki Watanabe (Riken)*
3. Le projet CNRS-JST NEGST
4. **Conclusion**

# Conclusion

- Le déploiement de Grilles au Japon est fait dans le cadre d'un écosystème global pour le calcul haute performance,
- Une planification globale sur plusieurs années existe,
- Des liens étroits existent entre les organismes de recherches, les universités et les constructeurs informatiques,
- Plusieurs des résultats obtenus (middleware, langage, *workflow*, tolérance aux pannes, répartition de tâche, ordonnancement, ..., seront repris en « *Petascale HPC* »
- Des collaborations existent entre la France et le Japon dans le domaine depuis plusieurs années et devraient se développer,
- Les résultats du projet NEGST sont attendus en 2008,
- Projet d'un LIA CNRS franco-japonais en STIC, dont l'un des thèmes devrait être « Grille/HPC »

# Liens

Projet CNRS-JST NEGST

<http://www2.lifl.fr/MAP/negst/>

GRID-Petascale HPC

<http://www.bulletins-electroniques.com/actualites/51795.htm>

<http://www2.lifl.fr/MAP/paap/>

# Grid Visualization: Visualization of Distributed Massive Data

