Cyber Science Infrastructure
And
NAREGI Grid Middleware

November 13, 2007

Kenichi Miura, Ph.D.
Center for Grid Research and Development
National Institute of Informatics
Tokyo, Japan
Outline

1. National Research Grid Initiative (NAREGI)

2. Cyber Science Infrastructure (CSI)
National Research Grid Initiative (NAREGI) Project: Overview

- Originally started as an R&D project funded by MEXT (FY2003-FY2007)
  2 B Yen (~17M$) budget in FY2003

- Collaboration of National Labs. Universities and Industry in the R&D activities (IT and Nano-science Apps.)

- Project redirected as a part of the Next Generation Supercomputer Development Project (FY2006-…..)

MEXT: Ministry of Education, Culture, Sports, Science and Technology
National Research Grid Initiative (NAREGI) Project: Initial 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 (100+Tflop/s expected by 2007) can be practically utilized by the nano-science research community over the Super SINET (now, SINET3).

(3) To Participate in International collaboration/Interoperability (U.S., Europe, Asian Pacific) ➔ GIN

(4) To Contribute to Standardization Activities, e.g., OGF
NAREGI Software Stack

Grid-Enabled Nano-Applications
- Grid Programming
  - Grid RPC
  - Grid MPI
- Grid Visualization
- Grid Workflow
- Grid PSE
- Super Scheduler
- Distributed Information Service

(Globus, Condor, UNICORE → OGSA)
- Grid VM

High-Performance & Secure Grid Networking

SINET3

Computing Resources
- NII
- IMS
- Research Organizations
- etc
An Example of NAREGI Workflow
Adaptation of Nano-science Applications to Grid Environment

GridMPI

IMS

RISM

Solvent Distribution Analysis

Grid Middleware

Data Transformation between Different Meshes

Electronic Structure in Solutions

(NINET3)

NII

Electronic Structure Analysis

Grid Middleware

Fragment Molecular Orbital method

Reference Interaction Site Model
Virtual Organization and Resources

Decoupling VOs and Resource Providers (Centers)

Research Org (RO)1

VOs & Users

Resource Providers

Policy

• VO-R01

• VO-APL1

• VO-RO1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM

Policy

• VO-R01

• VO-APL1

• VO-APL2

• GridVM
Collaborations in Data Grid Area

- High Energy Physics (GIN)
  - KEK
  - EGEE (e.g., CCIN2P3)

- Astronomy
  - National Astronomical Observatory (Virtual Observatory)

- Bio-informatics
  - BioGrid Project
## Roadmap of NAREGI Grid Middleware

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Ver. (Internal)</td>
<td>Development and Integration of □ Ver. Middleware</td>
<td>Evaluation on NAREGI Wide-area Testbed</td>
<td>Development of OGSA-based Middleware</td>
<td>Verification &amp; Evaluation Of Ver. 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **UNICORE-based R&D Framework**
- **OGSA/WSRF-based R&D Framework**
- **Utilization of NAREGI NII-IMS Testbed**
- **Utilization of NAREGI-Wide Area Testbed**
Highlights of NAREGI β release (2005-6)

1. Resource and Execution Management
   • GT4/WSRF based OGSA-EMS incarnation
     Job Management, Brokering, Reservation based co-allocation, Monitoring, Accounting
   • Network traffic measurement and control

2. Security
   • Production-quality CA
   • VOMS/MyProxy based identity/security/monitoring/accounting

3. Data Grid
   • WSRF based grid-wide data sharing with Gfarm

4. Grid Ready Programming Libraries
   • Standards compliant GridMPI (MPI-2) and GridRPC
   • Bridge tools for different type applications in a concurrent job

5. User Tools
   • Web based Portal
   • Workflow tool w/NAREGI-WFML
   • WS based application contents and deployment service
   • Large-Scale Interactive Grid Visualization
NAREGI Version 1.0

Operability, Robustness, Maintainability

- Being 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 sweep type jobs
- Improvement in maintainability
  - More systematic logging using Information Service (IS)
- Easier installation procedure
  - apt-rpm
  - VM
Science Grid NAREGI
- Version. 1.0 Architecture -
It has 63 edge nodes and 12 core nodes (75 layer-1 switches and 12 IP routers).
It deploys Japan’s first 40 Gbps lines between Tokyo, Nagoya, and Osaka.
The backbone links form three loops to enable quick service recovery against network failures and the efficient use of the network bandwidth.
Academic Networking of SINET (Asia) and Partners
NAREGI Testbed

Connectivity Demo At SC’07

SINET3 (10Gbps MPLS)

- TiTech Tsubame/Campus Grid
- Osaka Univ. CyberiMedia Ctr.
- Kyushu Univ. Small Test App Clusters
- Kyoto Univ. Small Test App Clusters
- AIST SuperCluster
- AIST Small Test App Clusters
- KEK Small Test App Clusters
- ISSP Small Test App Clusters
- Tohoku Univ. Small Test App Clusters
- Center for GRID R&D (NII) ~5 Tflops
- Computational Nano-science Center (IMS) ~10 Tflops

~3000 CPUs
~17 Tflops
Computer System for Grid Software Infrastructure R & D
Center for Grid Research and Development (5 Tflop/s, 700GB)

File Server
(PRIMEPOWER 900 + ETERNUS3000 + ETERNUS LT160)

File Server
PRIMEPOWER 900
ETERNUS3000
ETERNUS LT160

1node / 8CPU
(SPARC64V1.3GHz)

Memory
16GB

Storage
10TB

Back-up
Max.36.4TB

SMP type Compute Server
(PRIMEPOWER HPC2500)

1node (UNIX, SPARC64V1.3GHz/64CPU)

Memory
128GB
Storage
441GB

SMP type Compute Server
(SGI Altix3700)

1node (Itanium2 1.3GHz/32CPU)

Memory
32GB
Storage
180GB

SMP type Compute Server
(IBM pSeries690)

1node (Power4 1.3GHz/32CPU)

Memory
64GB
Storage
480GB

Intra NW

Intra NW-A

Intra NW-B

L3 SW
1Gbps

L3 SW
1Gbps

Ext. NW

Ext. NW

High Perf. Distributed-memory type
Compute Server (PRIMERGY RX200)
128 CPUs (Xeon, 3.06GHz)+Control Node
Memory 130GB
Storage 9.4TB

High Perf. Distributed-memory Type
Compute Server (PRIMEPOWER RX200)
128 CPUs (Xeon, 3.06GHz)+Control Node
InfiniBand 4X (8Gbps)
Memory 65GB
Storage 4.7TB

Distributed-memory type
Compute Server (Express 5800)
128 CPUs (Xeon, 2.8GHz)+Control Node
GbE (1Gbps)
Memory 65GB
Storage 4.7TB

Distributed-memory type
Compute Server (Express 5800)
128 CPUs (Xeon, 2.8GHz)+Control Node
GbE (1Gbps)
Memory 65GB
Storage 4.7TB

Distributed-memory type
Compute Server (HPC LinuxNetworx)
128 CPUs (Xeon, 2.8GHz)+Control Node
GbE (1Gbps)
Memory 65GB
Storage 4.7TB

SINET 3

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB

Memory 65GB
Storage 4.7TB
Outline

1. National Research Grid Initiative (NAREGI)

2. Cyber Science Infrastructure (CSI)
Cyber Science Infrastructure: background

- A new information infrastructure is needed in order to boost today’s advanced scientific research.
  - **Integrated information resources and system**
    - Supercomputer and high-performance computing
    - Software
    - Databases and digital contents such as e-journals
    - “Human” and research processes themselves
  - **U.S.A: Cyber-Infrastructure (CI)**
  - **Europe: EU e-Infrastructure (EGEE, DEISA, …)**

- **Break-through in research methodology is required in various fields such as nano-Science/technology, bioinformatics/life sciences,…**
  - the key to industry/academia cooperation:
    - from ‘Science’ to ‘Intellectual Production’

A new comprehensive framework of information infrastructure in Japan

Cyber Science Infrastructure
Architecture of Cyber-Science Infrastructure

Cyber Science Infrastructure

Virtual Research Organization over CSI
- Construction of new software and databases
- Development of Human resources and skills
Development of Scholarly digital contents and institutional repositories
High-performance Computing among university super-computer centers and research institutions

Middlewares
- Deployment of Research Grid Middleware for collaborative utilization of supercomputers
- University PKI Initiative for secure and reliable computing infrastructure

Networking
- High Speed Backbone Network: SINET3

Contributions to Industry and Society

International contributions
Structure of Cyber Science Infrastructure and Role of Grid Operation Center (GOC)

Center for Grid Research and Development

- Deployment & Operations of Middleware
- Tech. Support
- Operations of CA
- VO Users Admin.
- Users Training
- Feedbacks to R&D Group

GOC (Grid Operation Center)

- R&D and Operational Collaboration
- R&D/Support to Operations

Cyber Science Infrastructure

- Academic Contents Service
- Planning/Operation
- WG for Grid Middleware
- WG for Inter-university PKI
- WG for Networking

Center for Grid Research and Development

- Planning/Operation
- NAREGI Middleware

International Collaboration

- International Collaboration
- Academic Contents Service
- Planning/Operation
- WG for Grid Middleware
- WG for Inter-university PKI
- WG for Networking

e-Science Community

- Industry Project VOs
- Research Project VOs
- Univ./National Supercomputing Center VOs
- Domain-specific Research Organization VO (MS, AIST, KEK, NAO etc)

UPKI System

- Networking Infrastructure (SINET3)

National Institute of Informatics

Planning/Operation
Cyber Science Infrastructure

Cyber Science Infrastructure Plan Toward Petascale Computing

- National Leadership System (Next-generation Supercomputer)
- Grid
  - Grid Interoperability and International Collaboration
  - EGEE
  - Virtual grid
- National Infrastructure (Institute, University)
- Grid
  - Grid Interoperability and International Collaboration
  - EGEE
  - Virtual grid
- University/interuniversity research institutes VO
- Industry-university joint research VO
- Project VO
- INFRA (Infrastructure for Certification, etc.)
- VRE (Virtual Research Environment for various fields)

Development and Application of Advanced High-performance Supercomputer Project

NAREGI Grid Middleware
  - URKI (Inter-University Public Key Infrastructure)
  - National Institute of Informatics
Future Expansion of NAREGI Grid

- National Supercomputer Grid (Tokyo, Kyoto, Nagoya…)
- Domain-specific Research Organizations (IMS, KEK, NAOJ…)
- Interoperability (GIN, EGEE, Teragrid etc.)
- Laboratory-level PC Clusters
- Departmental Computing Resources
- Domain-specific Research Communities

Petascale Computing Environment

NAREGI Grid Middleware
Summary

• NAREGI Grid middleware will enable seamless federation of heterogeneous computational resources.

• Simulation in nano-science/technology applications over Grid is being tested, including participation from industry.

• NAREGI Grid Middleware has been adopted as one of the important components in the new Japanese Cyber Science Infrastructure Framework.

• NAREGI is planned to provide the access and computational and data/sharing infrastructure for the Next Generation Supercomputer System.
Thank you!

http://www.naregi.org