GRID activities and future plan at KEK

Takashi Sasaki
KEK Computing Research Center
Basic strategy

• GRID deployment and operations for projects at KEK
  – Belle, J-PARC, ILC and so on

• LHC regional center
  – Univ. of Tokyo is operating tier-2, but no tier-1 in Japan
  – Preparing for future tier-1 level GRID center operation at KEK toward ILC
    • Depends on the decision of the society
    • Technical assessments and staff training

• University support
  – Help for deployment and operations of LCG at universities outside of WLCG

• NAREGI
  – Our human resource is very limited and we want to depend on the operation infrastructure of NAGEI if possible
  – Would help university people and non-HEP users
INTRODUCTION
High Energy Physics in Japan

- Major High Energy Activities in Japan
  - Terminated in the last FY
    - K2K Experiment at Kamiokande and KEK
  - Active Experiments
    - Belle Experiment at KEK
    - KamLAND at Kamioka
    - CDF at FermiLab/USA
  - Under construction
    - J-PARC
      - T2K Experiment at Tokai and Kamioka
    - ATLAS and ALICE at LHC
  - Future Plan
    - SuperB Factory
    - International Linear Collider (ILC)
  - HEPnet-J:
    - KEK provides the network facility, NEPnet-J, on the SINET/SuperSINET (NII).
**J - PARC (Tokai)**

*J-PARC = Japan Proton Accelerator Research Complex*

Joint Project between KEK and JAEA

- **Linac (330m)**
- **3 GeV Synchrotron (25 Hz, 1MW)**
- **50 GeV Synchrotron (0.75 MW)**
- **Materials and Life Science Experimental Facility**
- **Hadron Beam Facility**
- **Neutrino to Super-Kamiokande**
Introduction to KEK Super-Kamiokande Neutrino Experimental Facility

J-PARC (T2K Experiment)

295 km West

Super-Kamiokande
KEK (Tsukuba site)
**KEKB $e^+ e^-$ Collider**

**Belle Experiment**
13 countries, 57 institutes, ~400 collaborators

**Observation of CPV in the $B$ meson system**

$B^0 \rightarrow J/\psi K_S$

$B^0 \rightarrow J/\psi K_S$
LCG
Brief Summary of GRID Deployment

- 2 sites are in operation
  - Deployed in different network logically
- JP-KEK-CRC-01
  - Since Nov 2005
  - Usage: experimental use and R&D, but production in LCG framework
- JP-KEK-CRC-02
  - Since Jan 2006
  - More stable services based on experience at KEK-1
- NAREGI
  - Using NAREGI beta1 released on May 2006.
  - Testing, evaluation, requirement assessment
- Accepted VOs are
  - belle
  - ppj
  - ilc, calice
  - g4med
  - dteam, ops
  - naokek (The same VOMS is used for NAREGI)
  - apdg
VOMS operated at KEK

- VOMS has been serviced in production since Sep 2006.
  - Tested from Nov 2005.
- VOMS support the VO for
  - **BELLE**: Belle Experiments (belle only registered in CIC)
    - The biggest target for us
  - **PPJ**: Accelerator Science in Japan
  - **G4MED**: Geant4 Medical Application for Radiotherapy
  - **NAOKEK**: Interpretational VO between National Astronomical Observatory of Japan and KEK
  - **APDG**: The R&D of Data Grid among Asia-Pacific region
  - **ATLASJ**: The ATLAS Experiment only for Japanese Group
  - **AIL**: Associated International Laboratory between KEK and France

http://voms.kek.jp
BELLE: The VO for the Belle Exp.

- Belle VO is federated among 5 countries, 7 institutes, 10 sites.
  - Nagoya University, University of Melbourne, ASGC, NCU, CYFRONET, Korea University and KEK
- VOMS is provided by KEK
  - [http://voms.kek.jp/](http://voms.kek.jp/)

- Past Activities
  - Federation was established
  - Library installation
  - Submitting MC production job for more realistic use
  - Long-term jobs, MC is taken ~1 week usually
  - Functional tests and performance tests over the VO
  - Interface to existing peta-bytes of data
Integration of existing storage using SRB-DSI

- A user can access by just using GridFTP client from outside
  - SRB client can access HSM also from inside.
  - Benefit for both SRB user and LCG user
    - Both user can read and write from/to HSM without considering protocol.
- Both protocols are authorized by GSI

Computing Farm
Still not integration with Grid
ILC/CALICE: The VO for Linear Collider Exp.

- ILC/CALICE VO is supported at KEK
  - Since end of 2006
- File sharing/transfer among DESY, IN2P3 and KEK over the VO
- ILC
  - Number of cores: 32,793
  - SPEC: 35,384 kSI2K
  - Storage: 68.4TB (12.6TB in use)
  - Members: 69 (4 from Japan)
- Calice
  - Number of cores: 13,469
  - SPEC: 15,140 kSI2K
  - Storage: 203TB (15.6TB in use)
  - Members: 52 (3 from Japan)
- KEK offer small resource
Operation statistics in Last 2yrs

**ILC**
- Number of Jobs: 150,269
  - 955 of 150,269 has been processed at KEK-1/2
- 323,251 CPU time normalized by 1kSI2K (hrs*kSI2K)
  - 569 of 323,251 has been used at KEK-1/2

**CALICE**
- Number of Jobs: 145,776
  - 579 of 145,776 has been processed at KEK-1/2
- 338,531 CPU time normalized by 1kSI2K (hrs*kSI2K)
  - 1,061 of 338,531 has been used at KEK-1/2
To do lists relating LCG

- CE and WN
  - Migration to SL4 (WN, lcg-CE)
  - Queue settings
    - Always occupied by jobs, sometime jobs from ops are expired
  - Job scheduler
    - LSF is the standard job scheduler at KEK
      - Necessary to share resources with local user and grid user
    - Maui is not satisfactory
- SE integration
  - HSM integration
    - Currently only disk in use
    - HPSS-DSI became ready and under testing
  - We are using DPM/SRM as a head node of SE
  - Keep the contact with application team at ASGC
    - SRB-SRB is desirable
- Networking/Security
  - KEK has very tight network security policy (aka, not GRID friendly)
    - Always tradeoff relationship between convenience and security
    - Always important subject how to manage easily and quickly with security assurance
  - Network bandwidth/throughput
    - Higher throughput between CC-IN2P3 and KEK/CRC is requested
- More robust and higher performance services
  - Using VM
  - Redundant design
Prototype of GOC

- Federated among major university groups and KEK in Japan.
  - Tohoku-U (KAMLAND, ILC)
  - Tsukuba-U (CDF)
  - Nagoya-U (BELLE, ATLAS)
  - Kobe-U (ILC, ATLAS)
  - Hiroshima-IT (ATLAS, Computing Science)
- We have a common VO, but do NOT depend on scientific projects.
  - To test each site.
- KEK assists their operation over the this VO
  - same motivation with ops VO

KEK behaves as the GOC
- Remote installation
- Monitoring
- Software updates

<table>
<thead>
<tr>
<th></th>
<th>Tohoku U</th>
<th>KEK</th>
<th>Tsukuba U</th>
<th>Nagoya U</th>
<th>Kobe U</th>
<th>HIT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (kSI2K)</td>
<td>0.68</td>
<td>91</td>
<td>5.1</td>
<td>8.3</td>
<td>8.5</td>
<td>1.2</td>
<td>115</td>
</tr>
<tr>
<td>SE (GB)</td>
<td>150</td>
<td>2,676</td>
<td>65</td>
<td>150</td>
<td>68</td>
<td>36</td>
<td>3,145</td>
</tr>
</tbody>
</table>
Monitoring System for our GOC

Monitoring Portal
- The monitoring system based on nagios and wiki
- The monitoring portal creates a link automatically based on knowledgebase and navigates administrators to appropriate troubleshooting page on wiki.

Summary view: Each site is iconified and shown their status as a few color, e.g., yellow show “warning”, red show “error”. The thickness and color of line indicates RTT and network status.

Support system: consists of “monitoring system” and “knowledge DB” and “FAQ by wiki”

Monitoring system: The site status is checked by a few simple jobs or commands, and is listed here. Link to FAQ is generated as to error description.

Strongly inspired ASGC NAGIOS Monitoring Service maintained by Joanna Huang, APROC
NAREGI
• NAREGI: NAAtional REsearch Grid Initiative  
  – Foundation: 2003-2007 10 billion Yen for 5 years  
  – Host institute: National Institute of Infomatics(NII)  
  – Core collaborations: IMS(molecular science), AIST(industrial app.), TIT, Osaka, Hitachi, Fujitsu, NEC  

• Mission:  
  – R&D of the Grid middleware for research and industrial application toward the advanced infrastructure  
  – Primary target application is nano technology for innovative and intelligent materials production.  
    • More focused in the computing grid for linking supercomputer centers for coupled simulation of multi-scale physics  
    • Support heterogeneous computer architectures (vector & super parallel & clusters)  
    • Data grid part were integrated in 2005
NAREGI β2 status

• Released in Oct. 2007, revised in Jan. 2008
• Additional features to β1 version
  – Auto installation with apt-rpm
  – Provides script for operation (Full Server start/stop/restart/status)
  – Full VOMS integration
    • Resource Sharing with multi VO’s
    • Data access and data space management with the VOMS
  – Provides Gfarm-DSI: GridFTP interface of the DataGrid
  – Improved performance
  – Interoperation with gLite/EGGE (not released yet, internal evaluation process)
NAREGI-β at KEK

• Testbed: 9 server nodes + 5 compute nodes
• Middleware installation
    • Manual installation for all the steps
    • Confirmed functionalities of Information Service, PSE, WFT, GVS
  – NAREGI-beta 1.0.2 : Feb 2007
    • DG comprehensive installation manual was released in Jan. 2007
  – NAREGI-beta 2.0.0: Oct. 2007  apt-rpm installation
  – NAREGI-beta 2.0.1: Dec. 2007

• Site federation test
Testbed for NAO-KEK federation
March 2008

User Browsers

NAO Site

Portal
SS
IS-NAS
IS-CDAS
GridVM Ser
MDS
AMS
Gfarm
NAT/DNS
UMS/VOMS

Comput nodes

SINET3

FireWall

KEK Site

Private network

192.168.2.101~

GridVM Eng

SuperSched
Infosys-NAS
Infosys-CDAS
GridVM Serv
MDS
AMS
Gfarm
Gfarm Storage
NAT/DNS
UMS/VOMS

DataGrid part

Test VO: naokek

VOMS/KEK

User Browsers

Compute nodes

User Browsers
Federation Test with NAO
(National Astronomical Observatory)

• **Aim:** Evaluation of application environments of NAREGI
• **Test Applications**
  – NAO: JVO(Japanese Virtual Observatory) applications
  – KEK: HEP Data Analysis, ex Belle simulation
    Geant4 MPI simulation
• **Status:**
  – NAO installed NAREGI β2 in the testbed Feb. 2008
    (DataGrid part is not yet installed.)
  – Test VO: naokek hosted by KEK VOMS server/gLite
  – Simple Job submission and retrieve were successfully tested in the end of March
  – Remote data file staging-in/-out has been confirmed.
  – Astro application job has been submitted to KEK site and retrieved the result to post-process for visualization. Apr. 2008
Federation Test with NAO-KEK

- Setup Astro Libraries at KEK site
- Job submission to KEK with Work Flow Tool (WFT) at the NAO Portal
- Input data are transferred from NAO and Output data are staged-out to NAO portal

- Output data was processed with vis. software as shown in the right picture.

SUBARU Telescope in Hawaii

Input Data: (2.7 GB)
10 CCD mosaic images
160MB x 17

Process: 10 Hours
- sensor calibration
- adjust deformation
- positioning
- mosaicing
- summing 17 frames

Visualization

50,000 objects identified in this frame.
Test of data storage with Gfarm

- **DataGrid part is consist of Gfarm distributed file system**
  - data files are stored in the multiple disk servers under the Gfarm file system software

- **Input and output data are staged-in and staged-out to the Gfarm storage.**
  - confirmed within KEK site and between KEK and NAO site

- **GridFTP interface supported in the β2**
  - data file transfer tested between gLite site and NAREGI site
  - File name space is the same between TURL and Program IO(using Gfarm-Fuse) as shown in the next slide.

- **Data file access from application program**
  - Gfarm client installed in the Engine nodes (Worker nodes) with FUSE
  - User can mount the Gfarm file system in the job by user privilege
  - Get access the data file through program read/write directly with no change in the application program (Belle event simulation), as if local files.
  - IO speed is several times slower than the local disk file IO. Tolerable depending on the data IO portion in the application program.
DataGrid: Gfarm File access

- Data Staging
- File transfer
  - Grid FTP
- Program IO
  - with Gfarm Fuse
  - Data analysis program can read/write data files without any modification

Transfer URL through Gfarm-DSI:
gsiftp://nrg11.cc.kek.jp/gfarm/user01/data1-22.dat

Program access name space for VO:
/gfarm/user01/data1-22.dat
Test Applications

- **Data Analysis program**: Carbon Ion Scattering in the water measured with Emulsion at 150 MeV, 300MeV
  - Data analysis program (written in ruby)
  - Input data was in the Gfarm and analyzed data were stored in the Gfarm files and also transferred to the SRB storage with Grid FTP
  - Typical elapsed time of a job is about 2 hours.

- **Geant4 Simulation with MPI**
  - Parallel processing Geant4 simulation with GridMPI of NAREGI has been test on the b1.

- **Belle event Simulation**
  - Full simulation softwares with libraries and database are installled and tested successfully in KEK site.
  - Plan to inter-operate with gLite/EGEE belle VO

- **SUBARU telescope image enhancement**
Data Grids Installation at KEK

2007.2.9

Naregi-kek

gLite/EGEE

gLite/ CRC-01

gLite/ CRC-02

SRB-DSI

SRB server

SRB files

SRB MCAT

CPUs

Local files

Grid files

HPSS
Future Plan on NARGI at KEK

- Migration to the production version
  - Release of NAREGI v. 1.0 will be in May 2008
- We will cooperate with Grid Operation Center in National Institute of Informatics
  - Planned to be started the operation JFY2008
- Multi site federation test with full specification will be done
- KEK leads improvement of the middleware in the application domain
OTHER ACTIVITIES
**KEK Grid CA**

- KEK Grid CA has been started since Jan 2006.
  - 75 CAs are in production all over the world, 3 in Japan.
    - AIST, NAREGI, and KEK
  - accredited as an IGTF (International Grid Trust Federation) compliant CA

- KEK Grid CA has been audited by Yoshio Tanaka (chair of APGridPMA), AIST on May 2007

---

**KEK Grid CA: Statistics of Issued Certificates**

<table>
<thead>
<tr>
<th>Certificate Type</th>
<th>JFY2006 Apr 2006 - Mar 2007</th>
<th>JFY2007 Apr 2007 -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Globus Client Certificate (Personal cert.)</td>
<td>68</td>
<td>119</td>
</tr>
<tr>
<td>Globus Server Certificate (Host cert.)</td>
<td>139</td>
<td>238</td>
</tr>
<tr>
<td>Web Server Certificate</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

[http://gridca.kek.jp](http://gridca.kek.jp)
RNS

• Middleware independent file catalogue is strongly desirable to operate multi-middleware and share data
  – Robustness and scalability are issue
• RNS: Resource Naming Service is standardized at OGF already
  – Two independent implementations are going on
    • U. of Tsukuba
    • University of Virginia
• We have requested NAREGI to support RNS
iRODs

- We are waiting for iRODs becomes matured enough to replace SRB
- Small contribution to the development
  - [https://www.irods.org/index.php/Performance](https://www.irods.org/index.php/Performance)
    - People at iRODs workshop had seen
    - Collaboration with Adil Hasan (RAL, UK) and Jean-Yves Nief (CC-IN2P3)
GRID interoperability

• gLite and NAREGI interoperability is our great concern
  – Discussion and collaboration with the NAREGI team
    • GIN
    • Data GRID
  – Collaboration with CC-IN2P3, Lyon
    • NAREGI has been installed at CC-IN2p3
SUMMARY AND ACKNOWLEDGEMENTS
Summary

• Many kinds of GRID middleware have been deployed and in operation
  – Related R&D is also going on
• Future strategy for the regional center will be decided by the Japanese HEP society
  – Super B, T2K, ILC, LHC and etc.
• In our experience, human network is much more important than computer network. We thank Simon and his staff for their efforts to provide us the opportunity every year.
  • We thank ASGC for their great support on the LCG operation
Acknowledgement I

- **Daily operation**
  - All members of APROC (ASGC)
  - All of ROCs
  - K. Ishikawa, M. Matsui (ISE Co., Ltd)

- **Belle virtual organization**
  - K. Inami, M. Kaga (Nagoya Univ.)
  - P. Lason (CYFRONET)
  - J. Shih, M. Tsai (ASGC)
  - M. Rosa, G. Moloney (Univ. of Melbourne)
  - S. Lee (Korea University)

- **ILC/Calice virtual organization**
  - R. Poeschl (LAL)
  - A. Miyamoto (KEK)
  - People in DESY-IT

- **Accelerator Science**
  - M. Fujii, Y. Nagasaka (Hiroshima-IT)
  - J. Ebihara, K. Yamada (Soum Co., Ltd)
  - Y. Takeuchi (Univ. of Tsukuba)
  - K. Kawagoe (Kobe University)
  - T. Nagamine (Tohoku Univ.)
Acknowledgements II

• Some part of works are partly supported by CSI Program of NII

Research Organization of Information and Systems

National Institute of Informatics