27 – 29 Oct. 2007 @ CC-IN2P3, Lyon, France Day1 15:25 - 16:05 (40min)

Workshop KEK - CC-IN2P3

LCG update and plan at KEK

Go Iwai, KEK/CRC

Agenda

- Introduction
  - Grid at KEK, its role and stance, etc
  - Brief history and summary on Grid
  - Ops stats
  - Grid related services
  - Networking (SINET3)
- Particular Activities over the VO
  - Belle experiments
  - Accelerator Science over the SINET3
  - ILC experiments
  - Geant4 Medical Application
- NAREGI
  - Beta-1: Activities at KEK
  - Beta-2: Current status
- Summary

Introduction

- Grid at KEK, its role and stance, etc
- Brief history and summary on Grid
- Ops stats
- Grid related services
- Networking (SINET3)

KEK's role, stance, etc

- KEK-EGEE II Collaboration
  - Focusing on operation and management (SA1)
  - Many supports by ASGC, Asia-Pacific ROC
- Domestic support
  - KEK has a role to offer necessary assistance to university groups in Japan
    - They are interested in Grid, but most of them are not production site
      - They hesitate to participate in EGEE?
      - Mostly, graduate students in physics are the main human resource to support IT unfortunately
  - Support for their deployment, operation and monitoring is a role instead of ROC
- We are T2 center, but not for any LHC exp.
  - T2 for ATLAS: ICEPP, Univ. of Tokyo
  - T2 for ALICE: Hiroshima University
- Currently, our activities on Grid are focused on operation and management, and developing gridified application over the LCG infrastructure, but not for LHC experiments
  - Encourage users to use the Grid
  - Just circumstance only at KEK, not at other institutions

People on Grid at KEK, CRC

- 7 persons in total
  - Today's participants are underlined
- CA
  - T. Sasaki and Y. Iida
- VOMS
  - Y. Watase and G. Iwai
- Site Operation and Security
  - KEK-1
    - T. Sasaki, Y. Iida, Y. Watase and G. Iwai
  - KEK-2
    - T. Sasaki, Y. Watase, and G. Iwai
  - NAREGI
    - Y. Watase, and G. Iwai
- Deployment
  - Y. Watase, Y. Iida and G. Iwai
- Documentation
  - Y. Watase
- Networking
  - S. Suzuki, S. Yashiro and Y. Iida
- Application (SRB, Portal and some Gridified applications)
  - K. Murakami, Y. Iida and G. Iwai

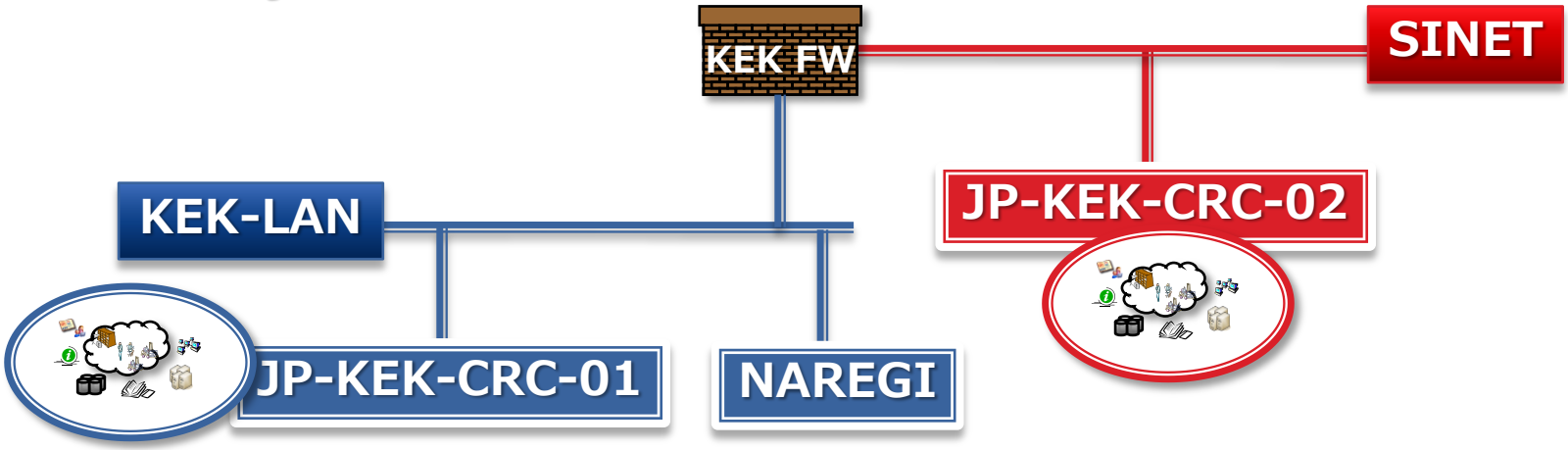
Brief History on Grid

2005	Q1 (Jan-Mar)	
	Q2 (Apr-Jun)	Testbed project was started with LCG-2.6 - Federation among a few institutions
	Q3 (Jul-Sep)	
	Q4 (Oct-Dec)	"*HEP Data Grid Workshop*" was held at KEK - KEK-1 system were introduced based on experience in workshop
2006	Q1	KEK Grid CA: accredited as production JP-KEK-CRC-01: approved as a certified site
	Q2	JP-KEK-CRC-02 (LCG-2.7): approved as a certified site NAREGI beta1 released
	Q3	1st KEK-IN2P3 workshop at IN2P3 KEK-1 and KEK-2 was upgraded to gLite-3.0
	Q4	
2007	Q1	2nd KEK-IN2P3 workshop at KEK
	Q2	
	Q3	NAREGI beta2 released
	Q4	3rd KEK-IN2P3 workshop at IN2P3

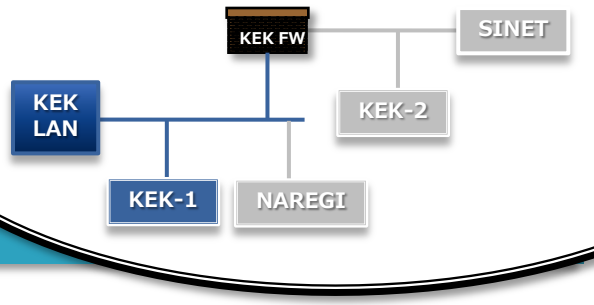
Brief Summary of Grid Deployment

- 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 and evaluation, what is lack for our usage?

- Accepted VOs are
  - belle
  - ppj
  - ilc, calice
  - g4med
  - dteam, ops
  - apdg, ail (have been gone)



KEK-1: JP-KEK-CRC-01



- SL30X/gLite-3.0
  - still not migration to SL4
- CPU: 14 x P4
- Storage: ~1.5TB

Service Name

dgXX.cc.kek.jp
130.87.208.xxx

KEK FW

KEK-LAN

130.87.208.0/22

FTS	RB	CE	SE	MON	WMS/LB	UI	BDII/PX	LFC
dg08 168	dg09 169	dg10 170	dg11 171	dg12 172	dg13 173	dg14 174	dg15 175	dg16 176

192.168.1.0/24

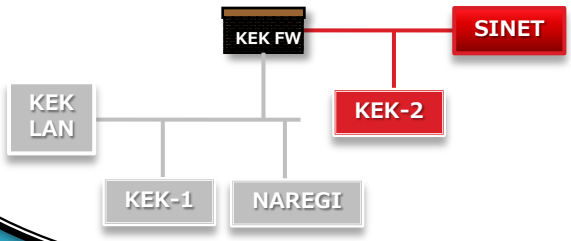
dg17 177	dg18 178	dg19 179	dg20 180
VOBOX	SE	WEB	DB

WNs

wn001 101 | wn002 102 | wn003 104

wn001 - wn015 (192.168.1.xxx)

KEK-2: JP-KEK-CRC-02



- SL30X/gLite-3.0
  - still not migration to SL4
- CPU: 48 x Opteron252
- Storage: ~2.0TB
  - Only disk in use, HPSS exclusive

Service Name

rlsXX.cc.kek.jp
202.13.197.xxx

L3 SW

SINET

202.13.197.0/24

CE	CE	RB	BDII	LFC	SE		MON
rls01 101	rls02 102	rls03 103	rls04 104	rls05 105	rls06 106	rls07 107	rls08 108

10.32.2.0/24

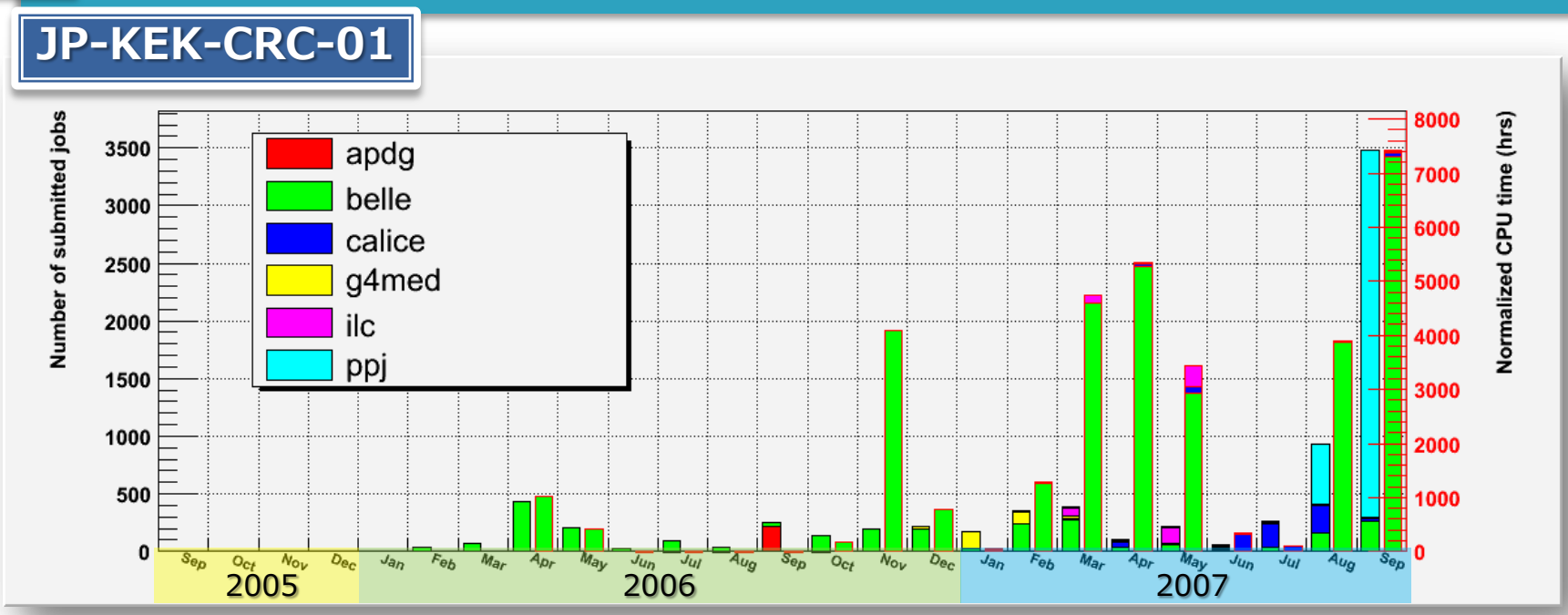
WNs

rlc01 51 | rlc02 52 | rlc03 53 | rlc04 54 | rlc05 55

rlc01 - rlc24 (10.32.2.xxx)

* 2 processors on each

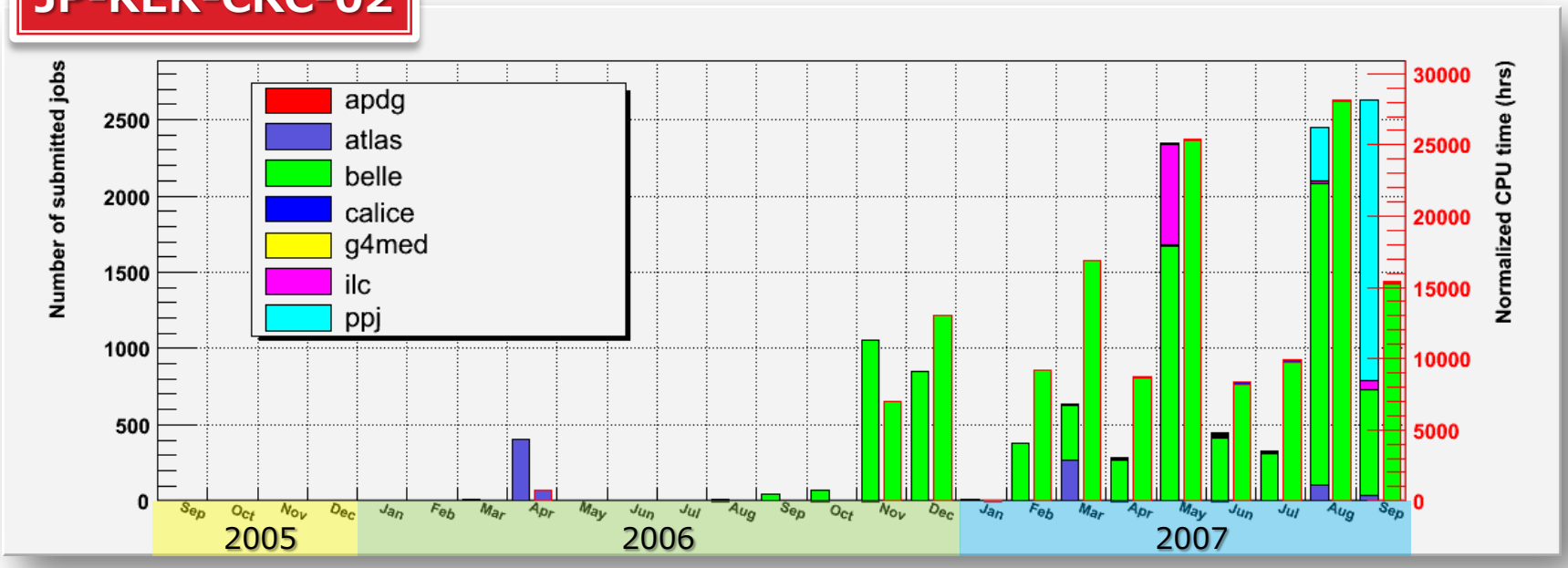
Ops Stats at KEK-1 last 2yrs



- Number of Jobs (w/o ops, dteam): 7,677
- 33,140 CPU time normalized by 1kSI2K (hrs*kSI2K)

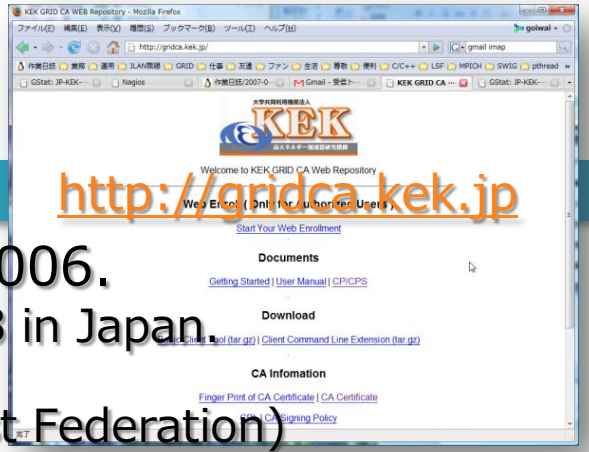
Ops Stats at KEK-2 last 2yrs

JP-KEK-CRC-02



- Number of Jobs: (w/o ops, dteam) 11,951
- 142,960 CPU time normalized by 1kSI2K (hrs*kSI2K)

KEK Grid CA



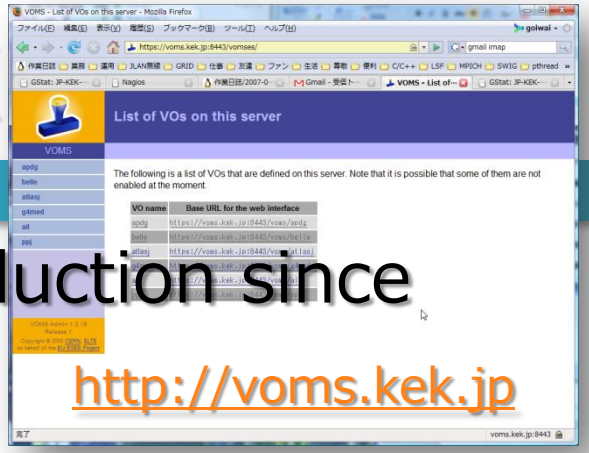
http://gridca.kek.jp

- 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
  - is accredited by IGTF (International Grid Trust Federation)
    - The IGTF consists of APGridPMA (AP), EUGridPMA (EU) and TAGPMA (US).
  - is recognized by LCG also.
- KEK Grid CA has been audited by Yoshio Tanaka (chair of APGridPMA), AIST on May 2007 and passed
  - Audited NAREGI CA also for cross checking on July 2007

KEK Grid CA: Statistics of Issued Certificates

	JFY2006 Apr 2006 - Mar 2007	JFY2007 Apr 2007 -
Globus Client Certificate (Personal cert.)	68	119
Globus Server Certificate (Host cert.)	139	238
Web Server Certificate	4	0

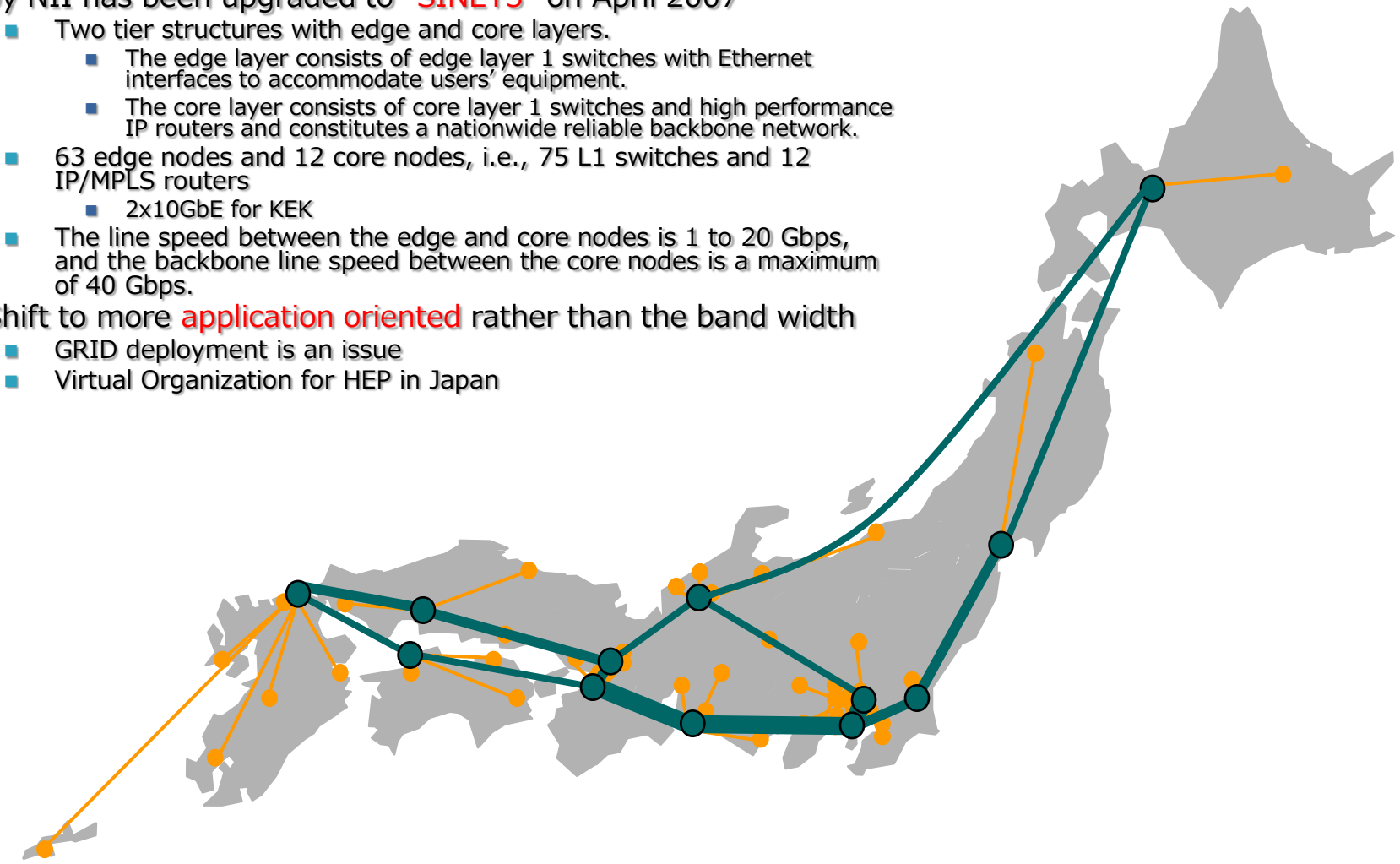
VOMS operated at KEK



- 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
  - 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

Networking on Grid

- The production R&E networks, "SuperSINET/SINET" operated by NII has been upgraded to "SINET3" on April 2007
  - Two tier structures with edge and core layers.
    - The edge layer consists of edge layer 1 switches with Ethernet interfaces to accommodate users' equipment.
    - The core layer consists of core layer 1 switches and high performance IP routers and constitutes a nationwide reliable backbone network.
  - 63 edge nodes and 12 core nodes, i.e., 75 L1 switches and 12 IP/MPLS routers
    - 2x10GbE for KEK
  - The line speed between the edge and core nodes is 1 to 20 Gbps, and the backbone line speed between the core nodes is a maximum of 40 Gbps.
- Shift to more application oriented rather than the band width
  - GRID deployment is an issue
  - Virtual Organization for HEP in Japan



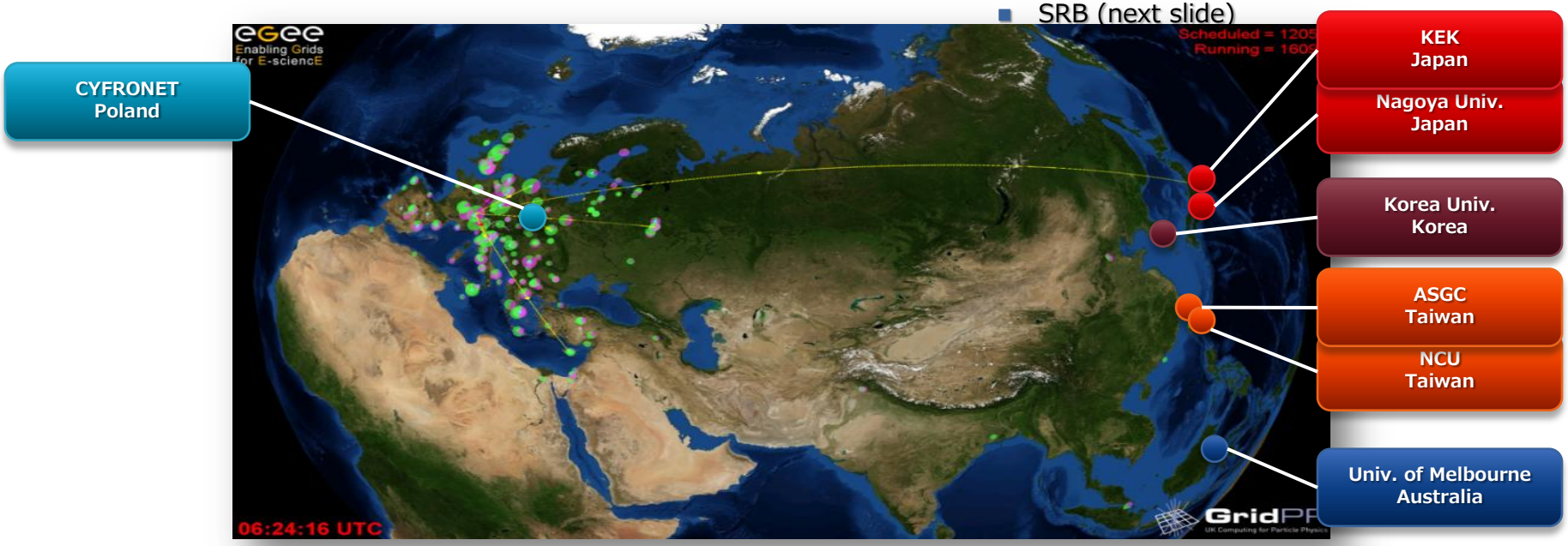
Particular Activities over the VO

- Belle experiments
- Accelerator Science over the SINET3
- ILC experiments
- Geant4 Medical Application

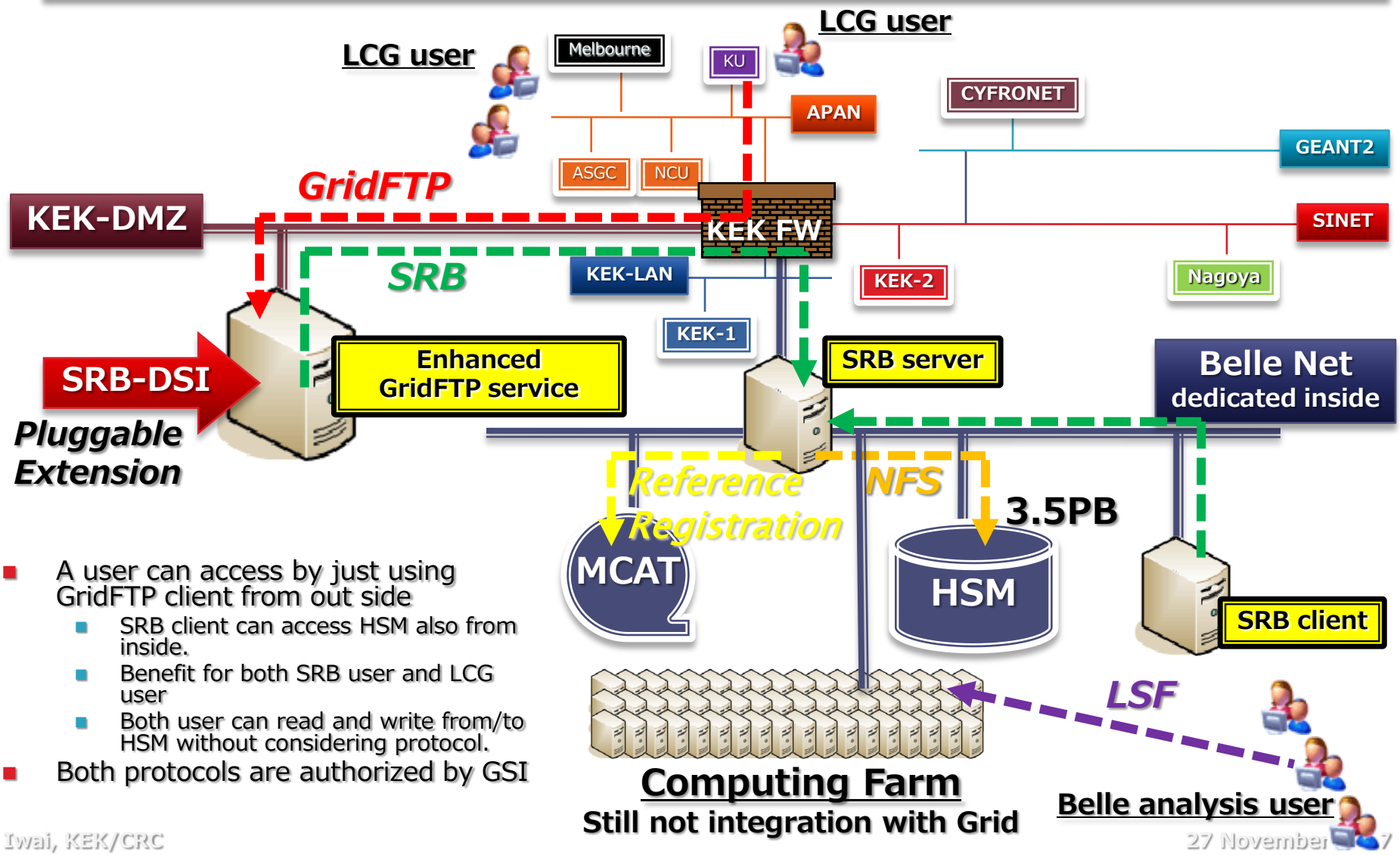
BELLE: The VO for the Belle Exp.

- Belle VO is federated among 5 countries, 7 institutes, 10 sites.
  - Nagoya University, University of Melbourne, ASGC, NCU, CYFRONET, Korea University
    - KEK-1/2 are also
- VOMS is supported by KEK
  - http://voms.kek.jp/

- Past Activities
  - Federation among sites
  - 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
  - **Finding a way to access existing data**
    - SRB (next slide)



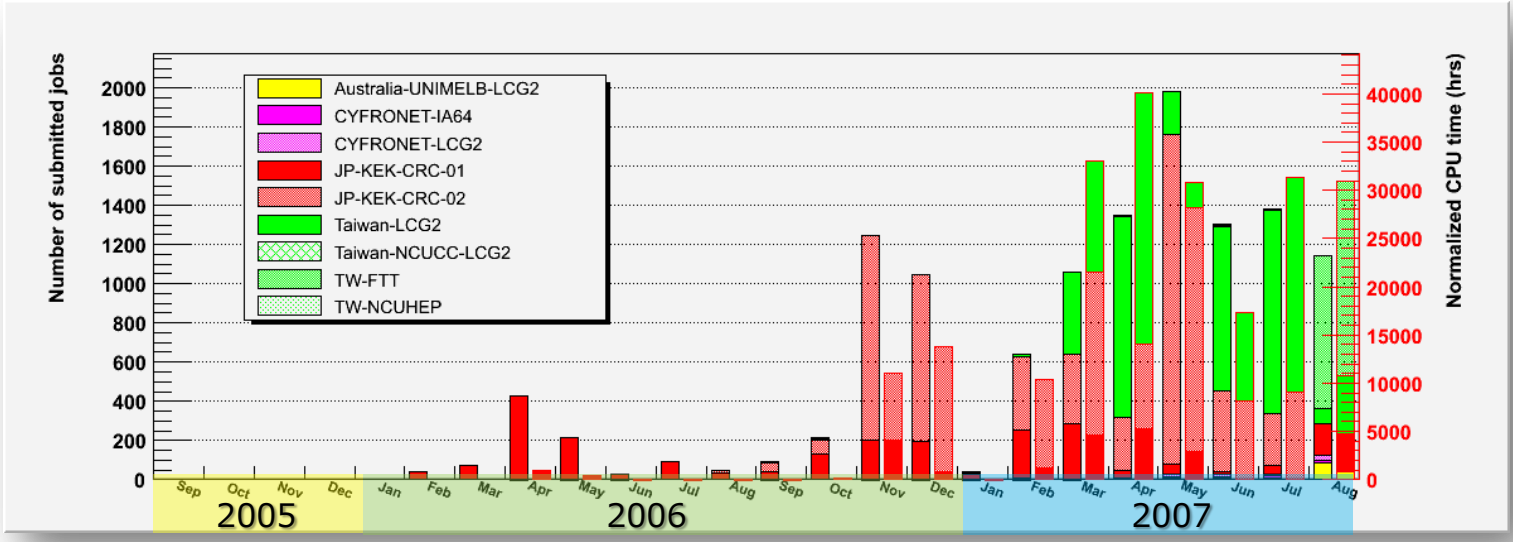
The choice and design by using SRB-DSI



- A user can access by just using GridFTP client from out side
  - 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

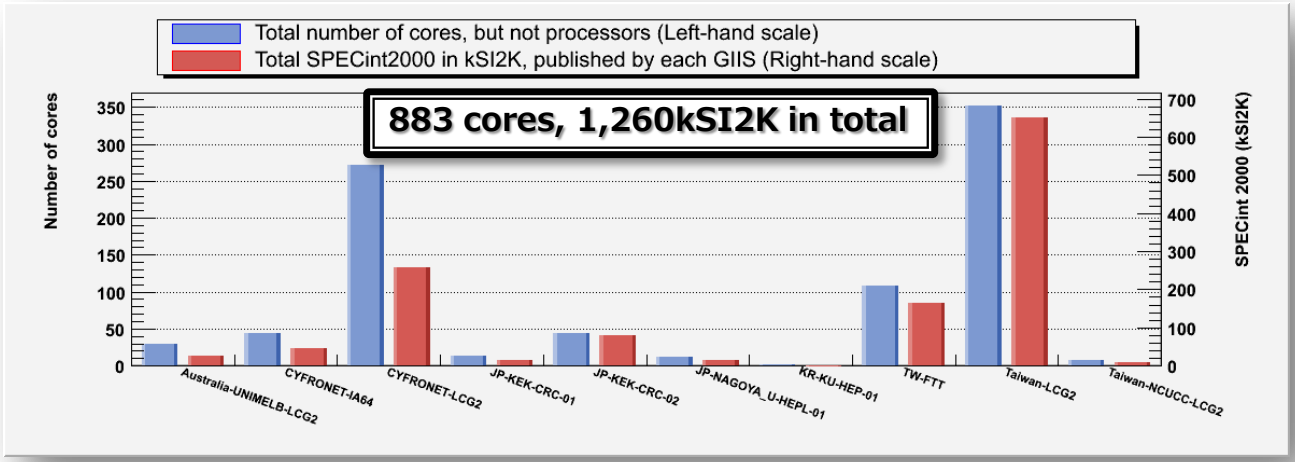
Ops Stats over the Belle VO last 2yrs

- Number of Jobs: 12,417
  - 7,706 of 12,417 has been processed at KEK-1/2
- 220,809 CPU time normalized by 1kSI2K (hrs*kSI2K)
  - 122,032 of 220,809 has been used at KEK-1/2

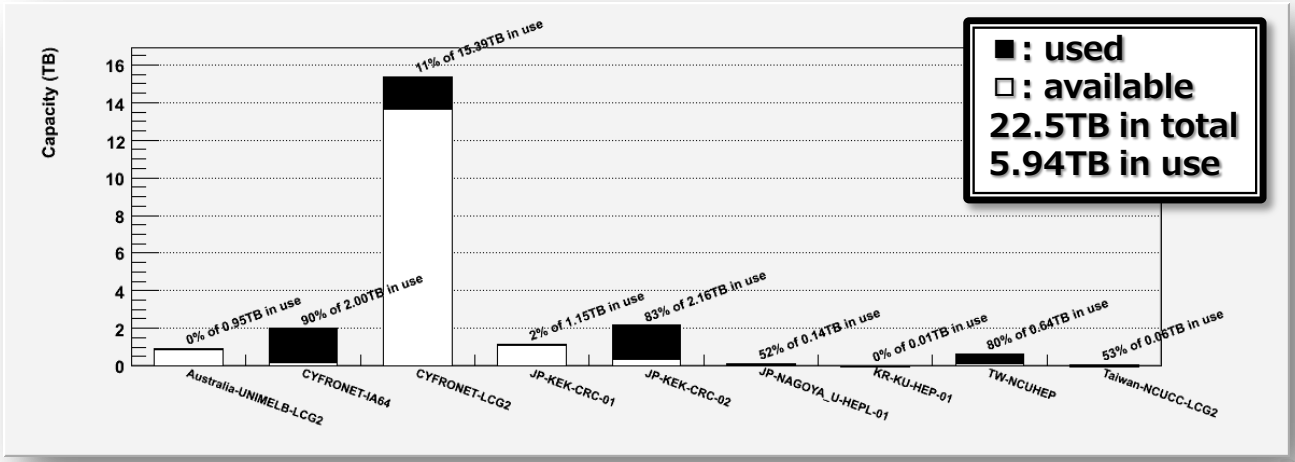


Computing Resources over the Belle VO

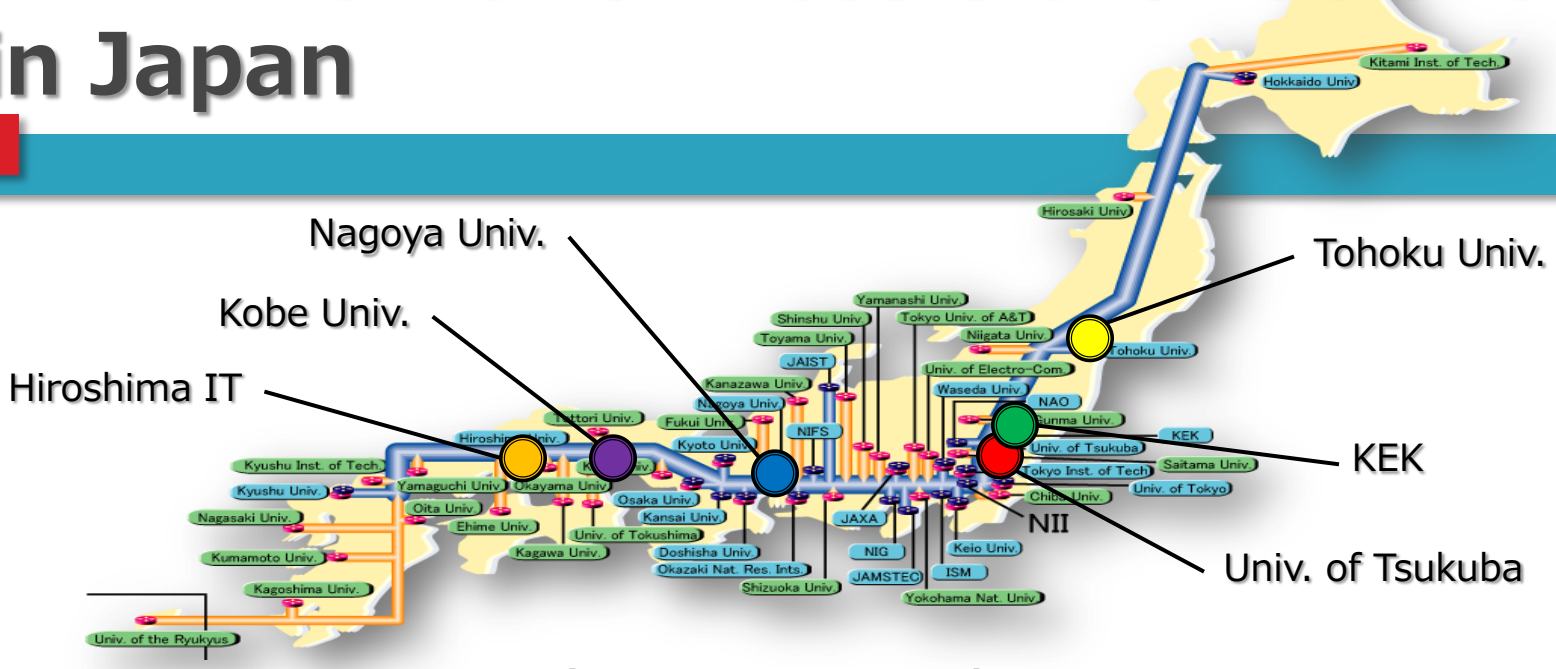
CPU resources over the belle VO as a function of sites



Storage capacities over the belle VO as a function of sites



PPJ: The VO for Accelerator Science in Japan



Computing Resources over the PPJ VO

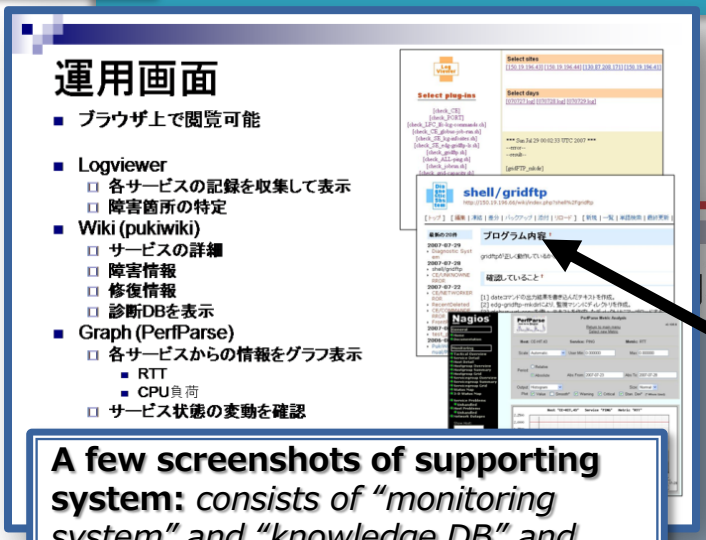
	Tohoku U	KEK	Tsukuba U	Nagoya U	Kobe U	HIT	Total
CPU (kSI2K)	0.68	91	5.1	8.3	8.5	1.2	115
SE (GB)	150	2,676	65	150	68	36	3,145

- Federated among major university groups and KEK only 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

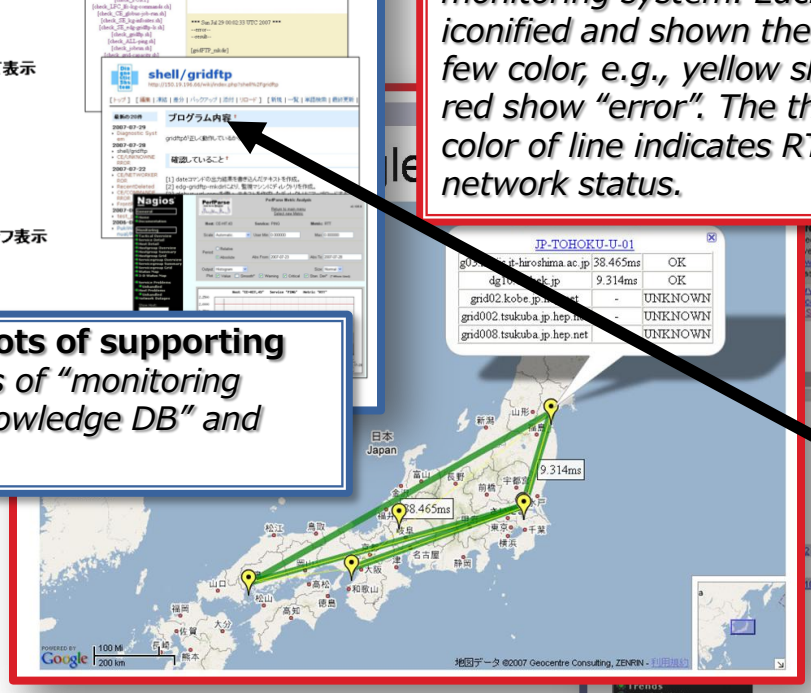
Mineo Fujii, Hiroshima-IT

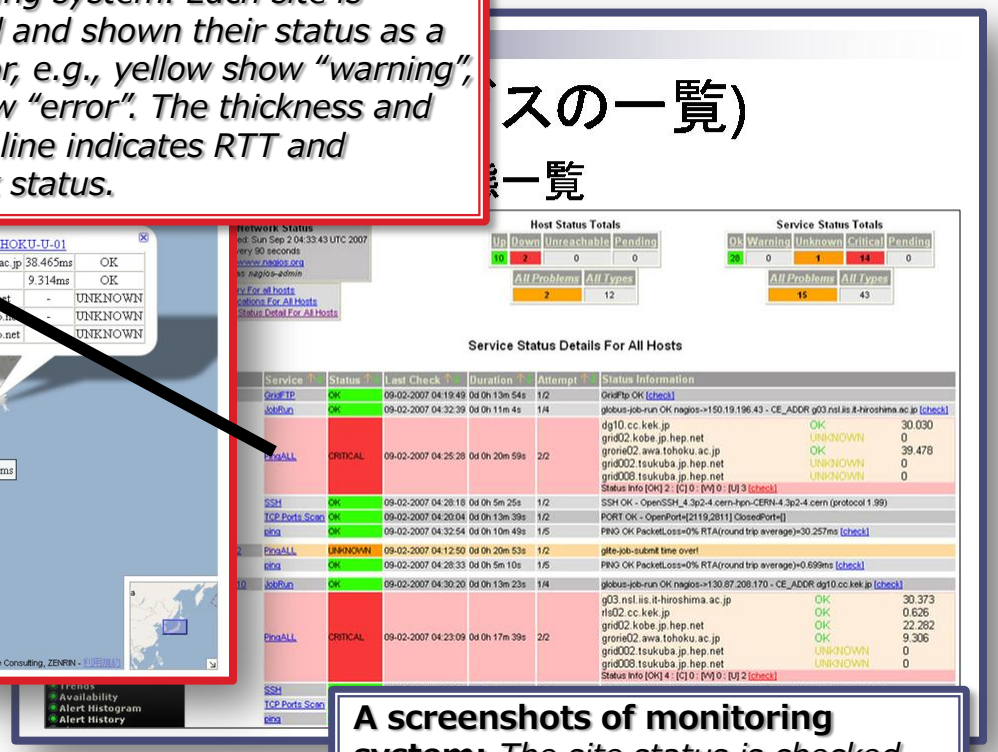
The Supporting System

A screenshot of summary view: *some of results are summarized and shown on the map as a part of monitoring system. 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.*



A few screenshots of supporting system: *consists of "monitoring system" and "knowledge DB" and "FAQ by wiki"*





Monitoring Portal

- The monitoring system based on nagios and wiki has been developed over the PPJ.
  - To support their operation at the university.
- The monitoring portal creates a link automatically based on knowledgebase and navigates administrators to appropriate troubleshooting page on wiki.

A screenshots of 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.*

ILC/CALICE: The VO for Linear Collider Exp.

- ILC/CALICE supported by DESY has bean ready at KEK
  - Since end of 2006
- Initial goal is achieved.
  - Triangle 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)

Ops Stats Last 2yrs

ILC

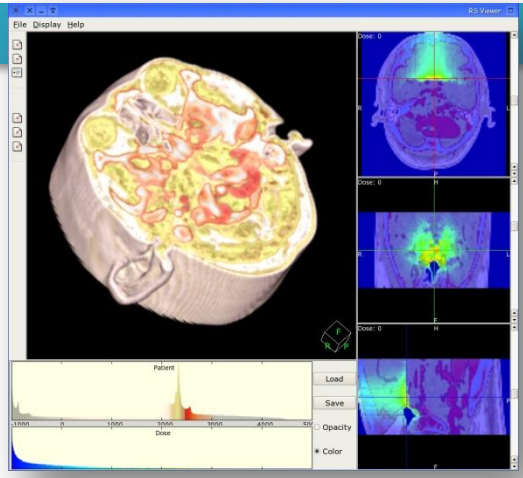
- Number of Jobs: 150,269
  - 955 of 150,269 has been processed at KEK-1/2
- 323,251 CPU time normalized by 1kSI2K (hrs*kSI2K)
  - 569 of 323,251 has been used at KEK-1/2

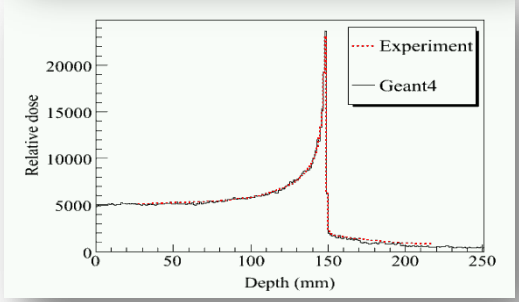
CALICE

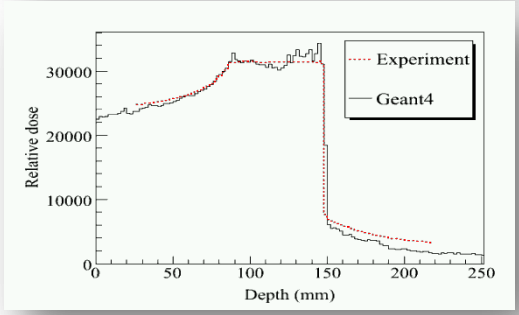
- Number of Jobs: 145,776
  - 579 of 145,776 has been processed at KEK-1/2
- 338,531 CPU time normalized by 1kSI2K (hrs*kSI2K)
  - 1,061 of 338,531 has been used at KEK-1/2

G4MED: Geant4 Medical Application

- We provide the framework and software toolkit for simulation in radiotherapy
- This software has following functions
  - CT/MRI data conversion to the Geant4 simulation
  - Simulation load sharing and data sharing with data grid technology
  - Visualization
  - Interactivity
  - Web based interface
    - Hospital is strongly protected by FW
  - etc

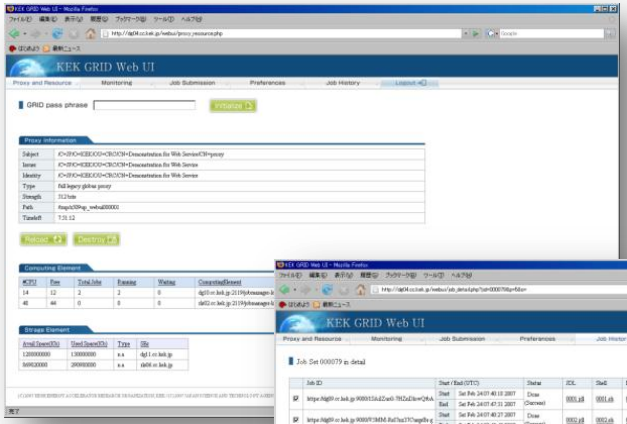




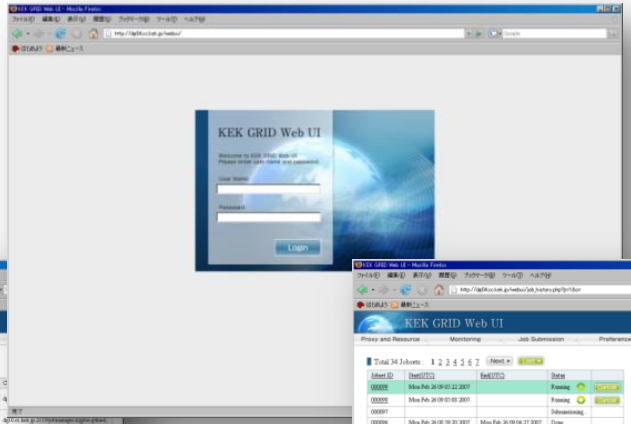


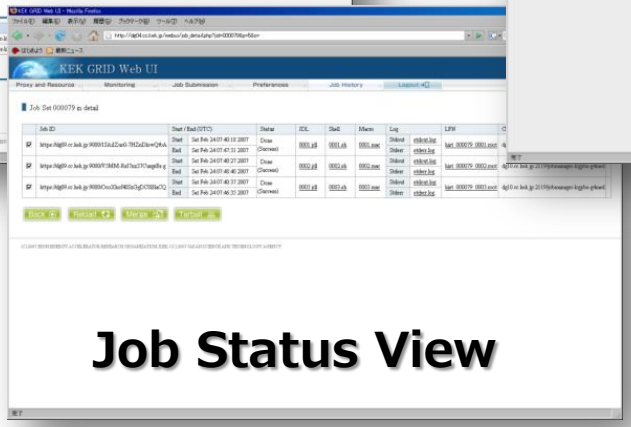
Screenshots of Web Interface

Proxy cert Creation

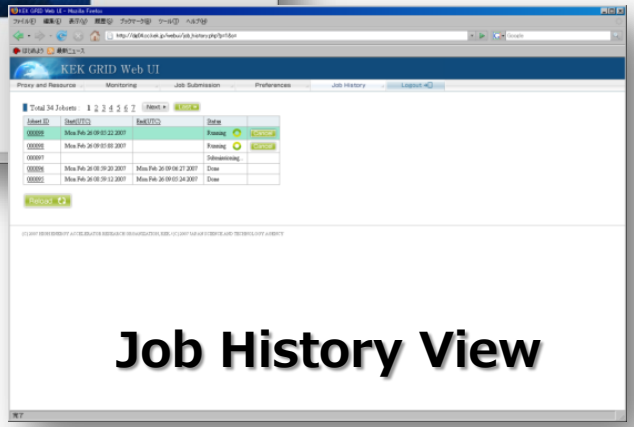


Login view





Job Status View



Job History View

- Web Interface has been developed as the first practical implementation.
- Job control and getting/viewing results are available only with web browser without any rich client.
- Keep the developing to improve the usability and to integrate into other components.

High priority issues in LCG

- CE and WN
  - Migration to SL4 (WN, lcg-CE)
    - Higher performance for the application
  - Queue settings
    - Always occupied by jobs, sometime jobs from ops are expired
  - Migration to LSF
    - Sharing resources with local user and grid user
    - Most of batch system at KEK are using LSF
    - Most of staffs have many knowhow
    - Maui is a bit poor
- SE integration
  - Currently only disk in use
    - A bit poor for us
    - We want larger scale storage, e.g. TAPE system
  - We are using DPM/SRM as a head node of SE
    - Established to access HSM by using SRB-DSI, hopefully HPSS by using HPSS-DSI soon
      - We can access to both SRB-DSI and HPSS-DSI via GridFTP, but not use LFC
  - Keep the contact with application team at ASGC
- Networking/Security
  - Always tradeoff relationship between convenience and security
  - Always important subject how to manage easily and quickly with security assurance
- More robust and higher performance services
  - Using VM
  - Redundant design
    - How? round robin DNS?

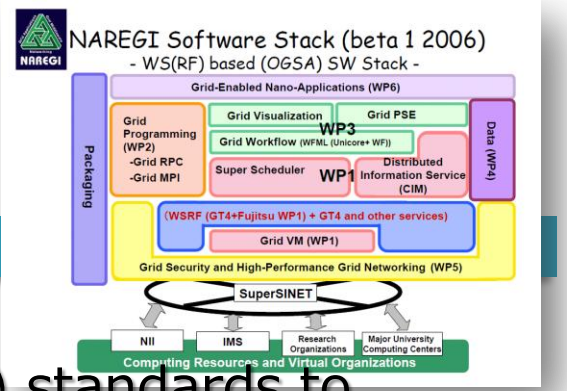
NAREGI

- Beta-1: Activities at KEK
- Beta-2: Current status

Introduction to NAREGI

- NAREGI: NAtional REsearch Grid Initiative
  - Foundation: 2003-2007 10 billion Yen for 5 years (Initial plan)
    - extended to 2009
  - Host institute: National Institute of Informatics (NII)
  - Core collaborators:
    - 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 Architecture

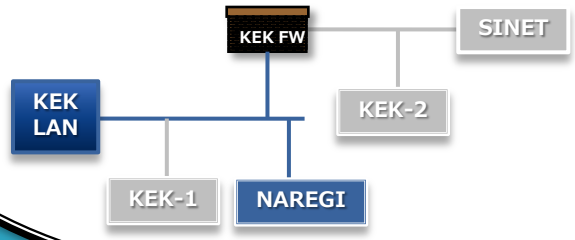


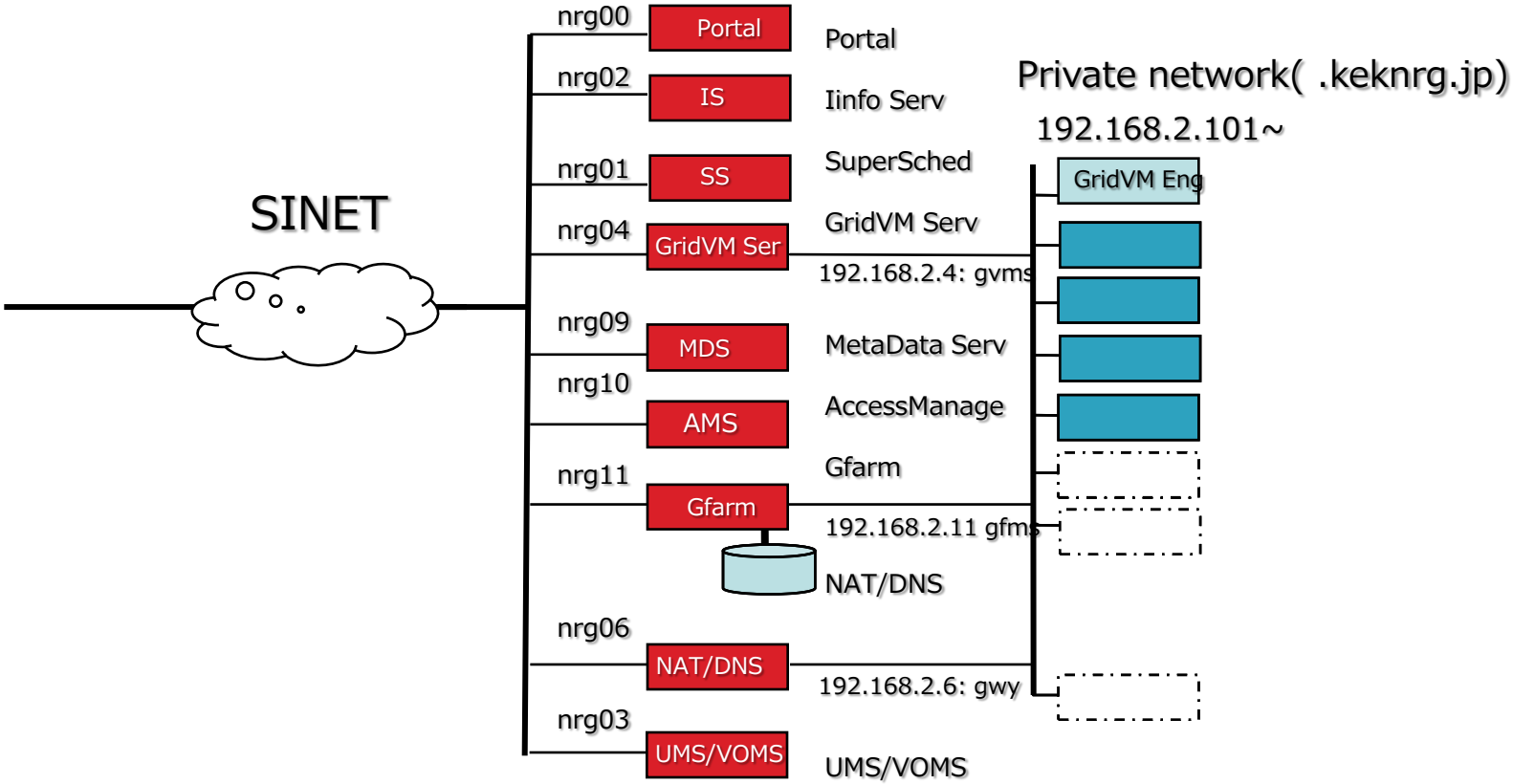
- Complete set of Grid middleware
- Advanced implementation of the OGF (GGF) standards to contribute for Grid world
- OGSA-WSRF(WebServiceResoureFramework) compliant architecture based on Globus ToolKits 4.0 (GT2 in gLite)
  - provides Web service interface components for various services
  - not only resource brokering. resource reservation, co-allocation, and co-scheduling
  - GGF JDSL 1.0 for job submission (JDL in gLite)
  - Resource Information based on CIM standard schema (GLUE schema in gLite)
- Security
  - adopt VOMS/EGEE
- Information service
  - Service interface: OGSA-DAI (Data Access & Integration)
- Adopt many standards
- NAREGI beta-1 released in May 2006

NAREGI-beta1 at KEK

- Testbed: 9 server nodes & 5 compute nodes
- Middleware installation
  - NAREGI-beta 1.0.1: Jun. 2006 - Nov. 2006
    - Manual installation for all the steps
    - Confirmed functionalities of Information Service, PSE (Problem Solving Environment), WFT (GUI Workflow Tool), GVS (Grid Visualization System)
  - NAREGI-beta 1.0.2: Feb 2007
    - DG comprehensive installation manual was released in Jan. 2007
- Site federation test Sep. 2007
  - KEK-NAREGI/NII
- Job Tests with Application software
  - Data analysis of ion scattering experiment
  - Geant4 simulation with MPI
  - Belle event simulation

Configuration of NAREGI Testbed at KEK





KEK Site

Data storage with Gfarm Test

- Data Grid part is consist of Gfarm
  - data files are stored in the multiple disk servers under the Gfarm file system
  - Input and output data are stage-in and stage-out to the GFarm storage.
  - Gfarm client installed in the Engine nodes (Worker nodes) can get access the data file through program read/write directly with no change in the application program (Belle event simulation).

Next Steps for us

- NAREGI beta-2 released in Oct. 2007
  - Under installation now
    - Mostly done, but DG part still not yet
  - Features
    - "**Easy**" installation by apt-rpm
      - Greatly improved
        - Beta-1: 6 months
        - Beta-2: 2 hours
        - DG part is exclusive
    - Interoperation with EGEE/gLite
      - Job submission, Data exchange, Information exchange
    - Various useful features for application
      - GridMPI: MPI jobs linking over sites
      - Burst job submission
      - GridFTP interface to the Gfarm files
        - Perhaps, expect to work with SRM
- Release of NAREGI ver1.0 in Apr-May 2008

Summary (1/2)

- 2 systems are in operation as production.
  - Main usage:
    - KEK-1: R&D
    - KEK-2: high quality services
- NAREGI beta1 is installed and tested.
  - being replaced with beta-2 very soon (this month)
- Other services related on Grid
  - KEK Grid CA
  - VOMS
- 19,628 jobs have been processed, 176,100 CPU hours have been used in KEK last 2 years.

SITE	# of cores	SPEC (kSI2K)	Storage (TB)	# of jobs	CPU (kSI2K*hrs)
KEK-1	14	15	0.26/1.41 (18%)	7,677	33,140
KEK-2	44	80	0.57/2.03 (28%)	11,951	142,960

Summary (2/2)

- BELLE VO
  - The GridFTP server plugged SRB-DSI is in operation.
    - Succeed to establish access method of existing data by using GridFTP client from outside.
  - Installation of belle library remotely, MC production is available at each site.
- ILC/CALICE VO
  - Triangle file sharing/transfer and job submission each among DESY, IN2P3 and KEK over the VO.
  - FW
    - Shigeo (Yashiro-san) talk about this issue on Thursday
- The VO for the Accelerator Science in Japan (ppj)
  - Installed and federated among 6 institutes, non-production sites mostly.
  - Developing supporting system.
- Geant4 Medical Application VO
  - Web Interface has been developed.
  - Job control and getting/viewing results are available only with web browser.
- NAREGI beta-2 has been released
  - We are installing and testing now

VO	# of cores	SPEC (kSI2K)	Storage (TB)	# of jobs	CPU (kSI2K*hrs)
BELLE	883	1,260	5.94/22.5 (26%)	12,417	220,809
ILC	32,793	35,384	12.6/68.4 (18%)	150,269	323,251
CALICE	13,469	15,140	15.6/204 (7.6%)	145,776	338,531

Items to be continued

- "*KEK new Grid system*" given by Koichi Murakami on Wednesday
  - CE and WN
    - Migration to LSF
      - Sharing resources with local user and grid user
      - Most of batch system at KEK are using LSF
      - Most of staffs have many knowhow
  - SE integration
    - We are using SRM as head node of SE
      - Established to access HSM by using SRB-DSI, hopefully HPSS by using HPSS-DSI soon
        - We can access to both SRB-DSI and HPSS-DSI via GridFTP, but not use LFC
  - More robust and higher performance services
    - Using VM
    - Redundant design

- "*Discussion on security issue on GRID system and network operation*" given by Shigeo Yashiro on Thursday
  - Networking/Security
    - We are planning to change the network configuration logically for more flexible operation

Acknowledge

- 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.)

Many thanks for your attention