



TWGrid: the Grid & e-Science Global Infrastructure of Taiwan

Eric Yen

ASGC, Taiwan

Asia Federation Session, EGEE 2006

27 Sep 2006

Academia Sinica Grid Computing



Outline

- Expectation to Asia Federation
- Only Collaborate and Share can make us Evolve smoothly and Successful
- ASGC Experiences and Services
- Actions



Asia Federation (I)

- Objectives

- Coordinate the Development Regional e-Science Infrastructure and Applications under WLCG/EGEE Framework
- Collaborate for larger scale and higher performance researches by Sharing of Power and Knowledge

- Strategy

- Based on the Consortium-based regional e-Science alliance: AU+CN+IN+JP+KR+ML+PK+SG+TH+VN +TWGrid
- Application-oriented Services and Development
- Visioning the future by sharing, and Learning by doing



Asia Federation (II)

- Challenges

- Using e-Science infrastructure and services for routine works
- Ensure the high quality services for all user communities
- ease-of-use and reduce the O&M cost
- Standardization and Interoperability
- Prepare for a sustainable long-term framework



TWGrid Introduction

- Consortium Initiated and hosted by ASGC in 2002
- Objectives
 - Applications to join the Global e-Infrastructure
 - Providing Asia Pacific Regional Operation Services
 - Fostering e-Science Applications collaboratively in AP
 - Dissemination & Outreach
 - Taiwan Grid/e-Science portal
 - Providing the access point to the services and demonstrate the activities and achievements
 - Integration of Grid Resources of Taiwan
 - VO of general Grid applications in Taiwan



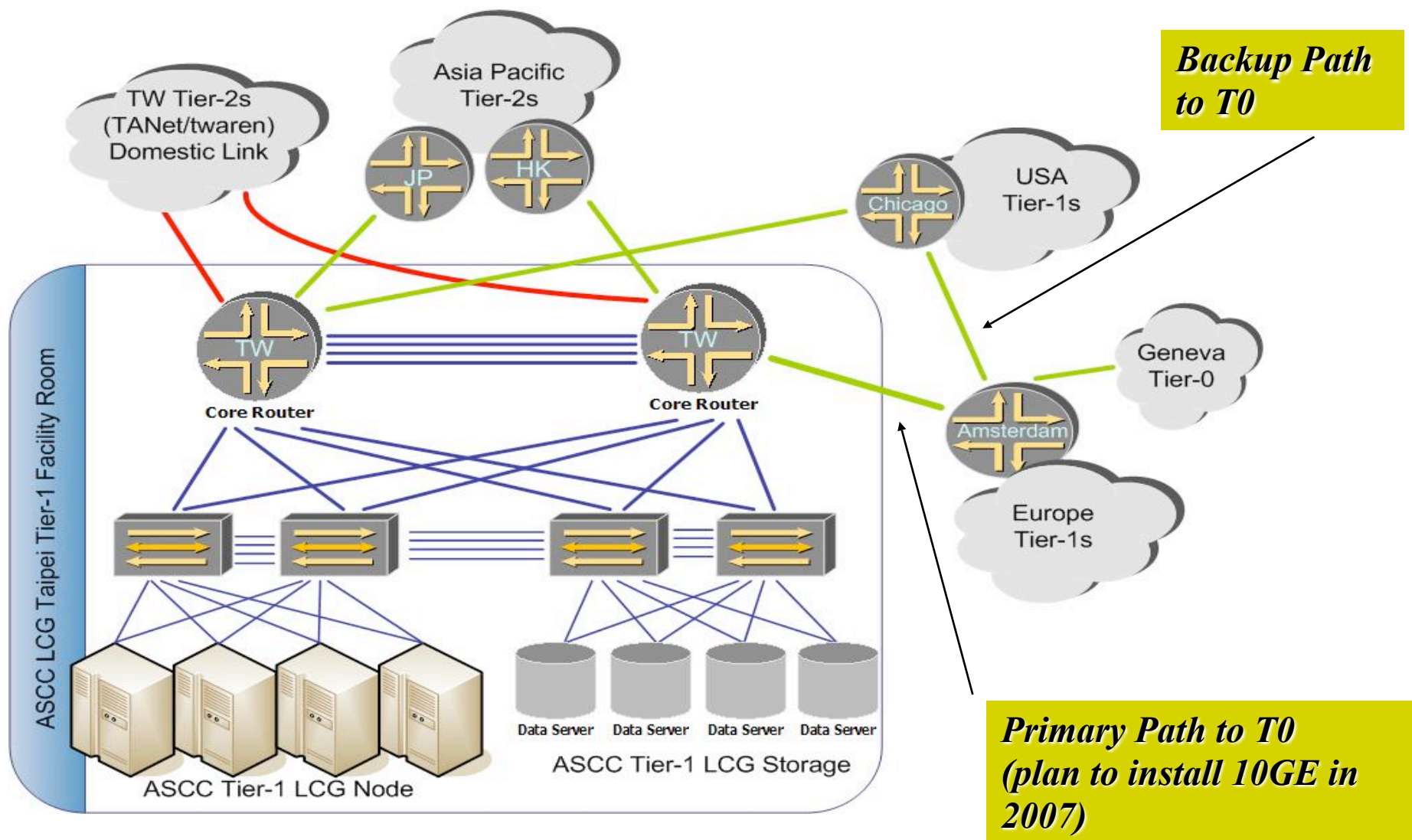


Potential Contributions to the World Wide e-Science/Grid

- Extend the global e-Science infrastructure to AP region
- Reduce the complexity of infrastructure interoperation
- Facilitate the worldwide collaboration by linking the people, data, CPU, instruments globally
- Bridge the digital divide
- Advance essential collaborations of e-Science applications
- Advance the quality of services and applications of worldwide e-Science

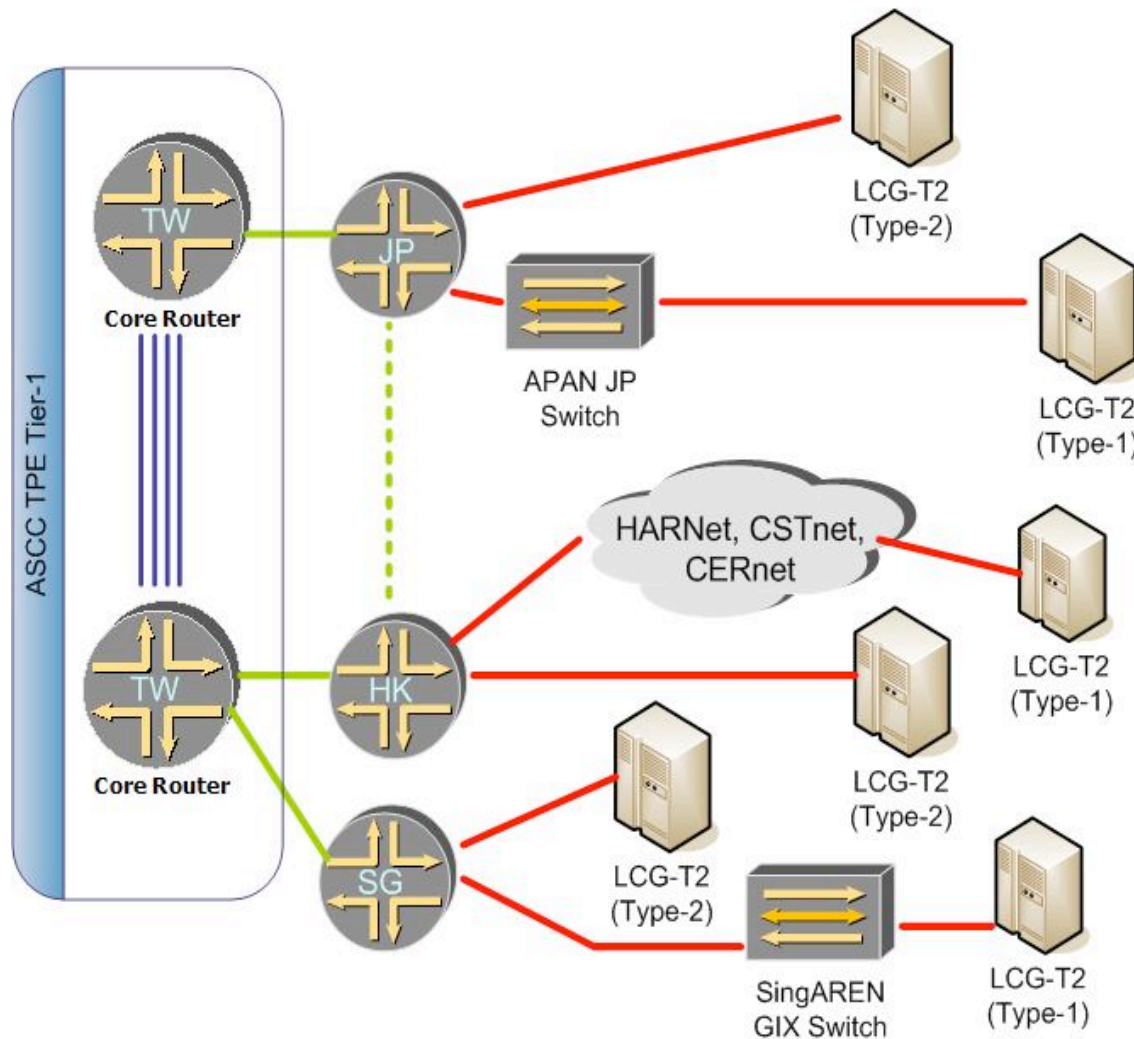


Plan for Taiwan Tier-1 Network



AP Regional LCG Network

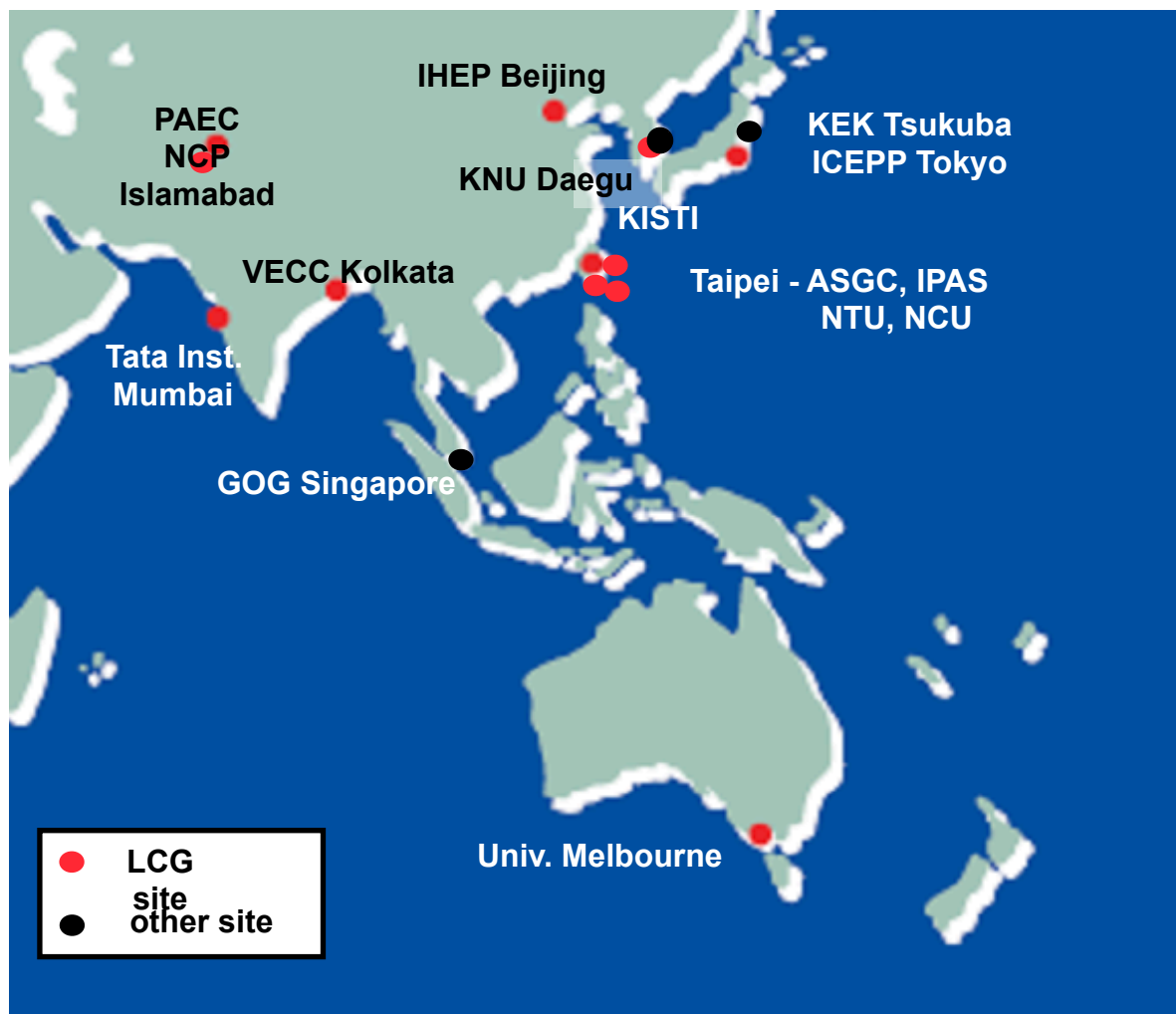
**Just Build a New Link
between TW-AU, reduce
RTT from 380 to 138 ms via
Singapore!**



- *Solid lines between routers (circle) and switches (box) and networks are already exist.*
- *Solid lines between T2 and routers / switches / networks are already exist and/or proposed.*
- *Dashed line are currently planned by ASnet and will be installed in 2006/7.*
- *Type-2 is “direct-connect”*
Type-1 is passing through 3rd party facility or 3rd party network



EGEE Asia-Pacific Federation



- 12 LCG sites and 3 EGEE sites in Asia Pacific
- Academia Sinica Grid Computing Centre (ASGC) is acting as the coordinator, also
 - the WLCG Tier-1 Centre
 - WLCG/EGEE Operation Centre in Asia Pacific Region
- Potential Sites
 - Thailand, Malaysia, Australia, New Zealand

■ AP Federation now shares the e-Infrastructure with WLCG

Academia Sinica Grid Computing



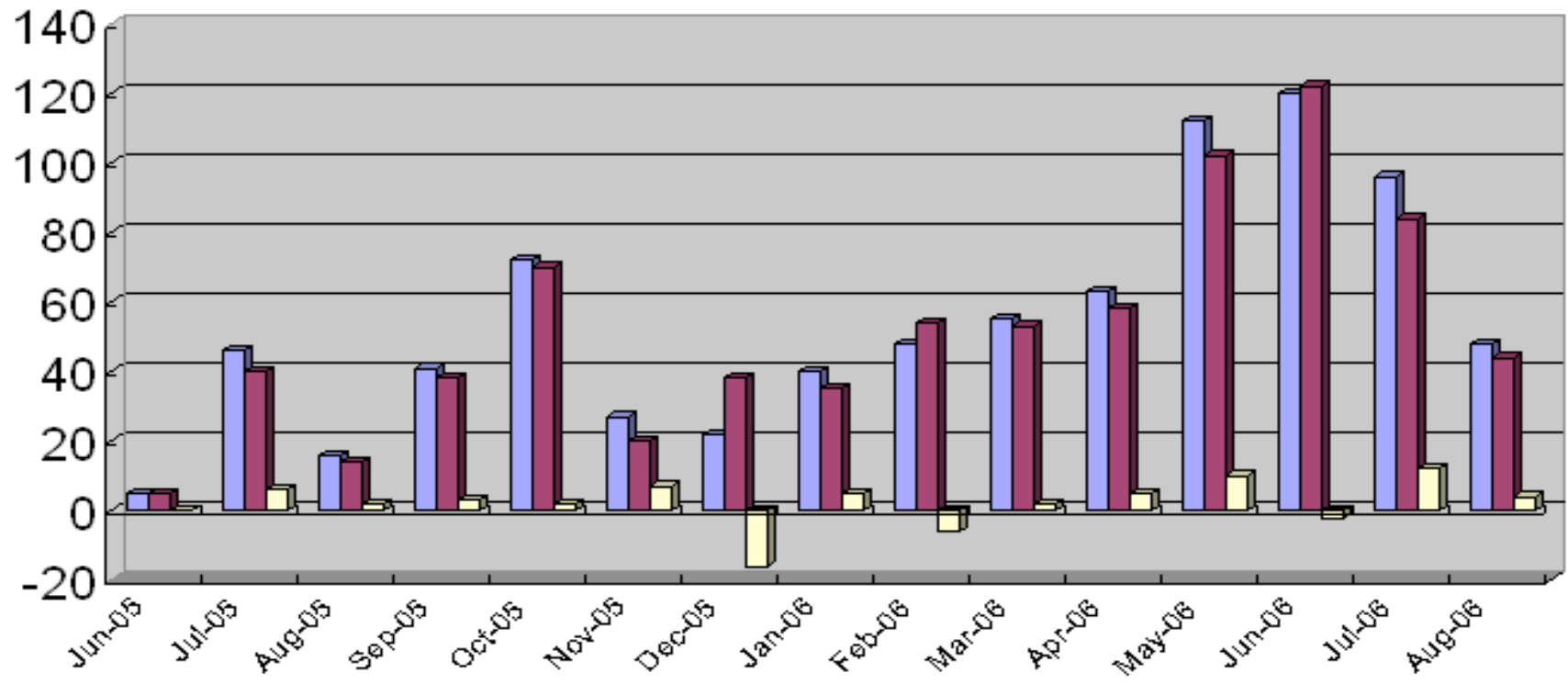
WLCG/EGEE Asia Pacific Services by Taiwan

- Production CA Services: production service from July 2003
- AP CIC/ROC: 17 sites 7 countries, > 800 CPUs
- VO Infrastructure Support: APeSci and TWGrid
- WLCG/EGEE Site Registration and Certification
- Middleware and Operation Support
- User Support: APROC Portal (www.twgrid.org/aproc)
- MW and technology development
- Application Development
- Education and Training
- Promotion and Outreach
- Scientific Linux Mirroring and Services



APROC Tickets Statistics

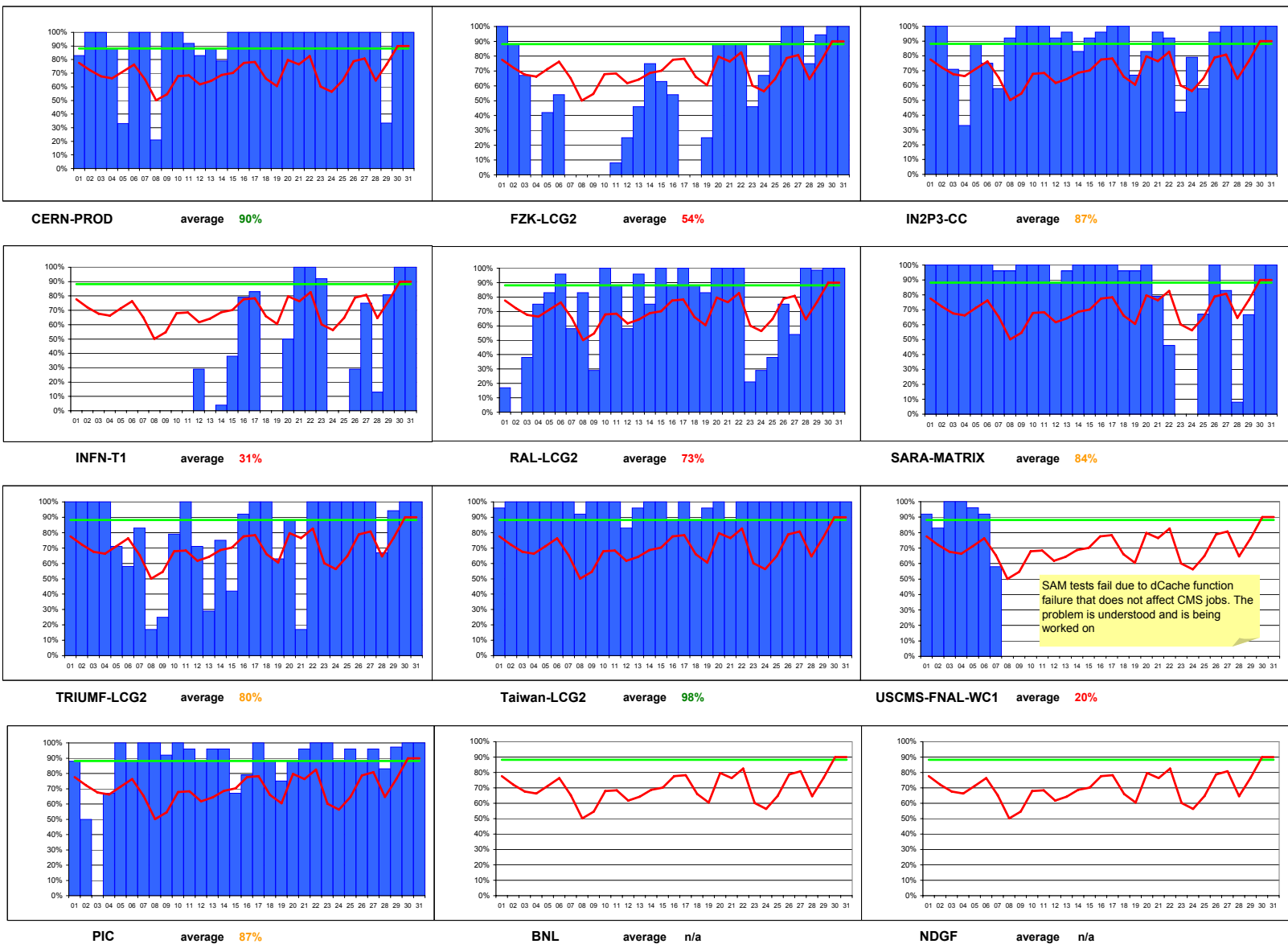
	Statistic (Tot/Ave)
Open tickets	34
Close tickets	777/51
Total tickets	811/54



Availability of WLCG Tier-1 Sites + CERN July 2006

29 July CERN power-off 08:00-24:00 - SAM failed through 31 July.
Other sites that were active before and after power-off assumed to be fully available throughout power-off and SAM failure. Availability on 29 July before 08:00 adjusted pro-rata.

Data from SAM monitoring. Site availability as agreed in WLCG MB on 07mar06 and 04apr06. Service considered unavailable from time of first failure to time of next success. Scheduled interruptions are considered as unavailable.
average (all sites) — 70% — target 88%
site average colour coding: < 90% of target ≥ 90% of target ≥ target





WLCG Availability of CERN & T1s

- Target is to have ≥ 8 sites reaching 88% availability by end of Sep. 2006.
- After that, full WLCG services should be in operation.

WLCG Availability of CERN & T1s

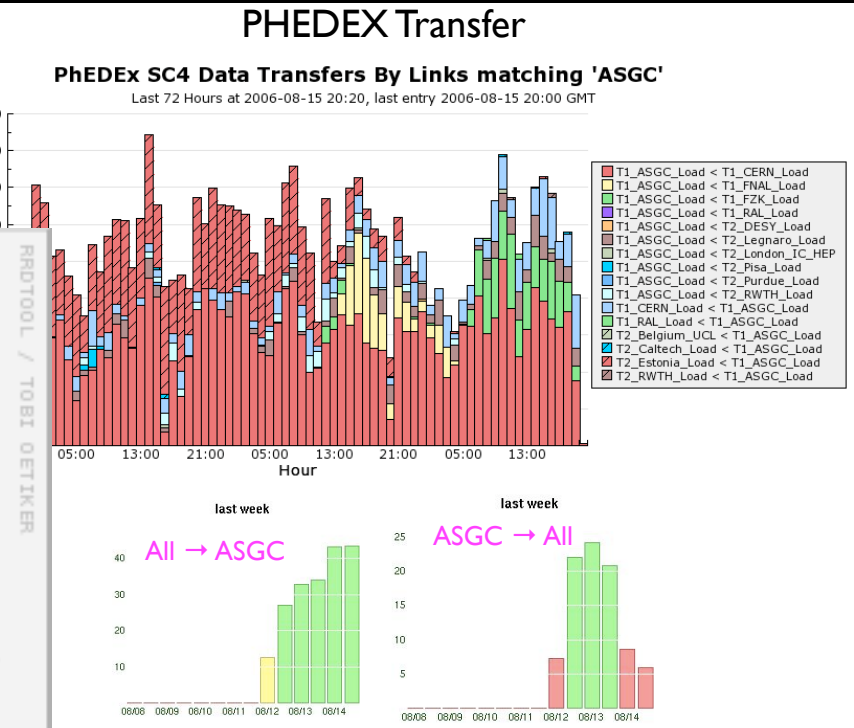
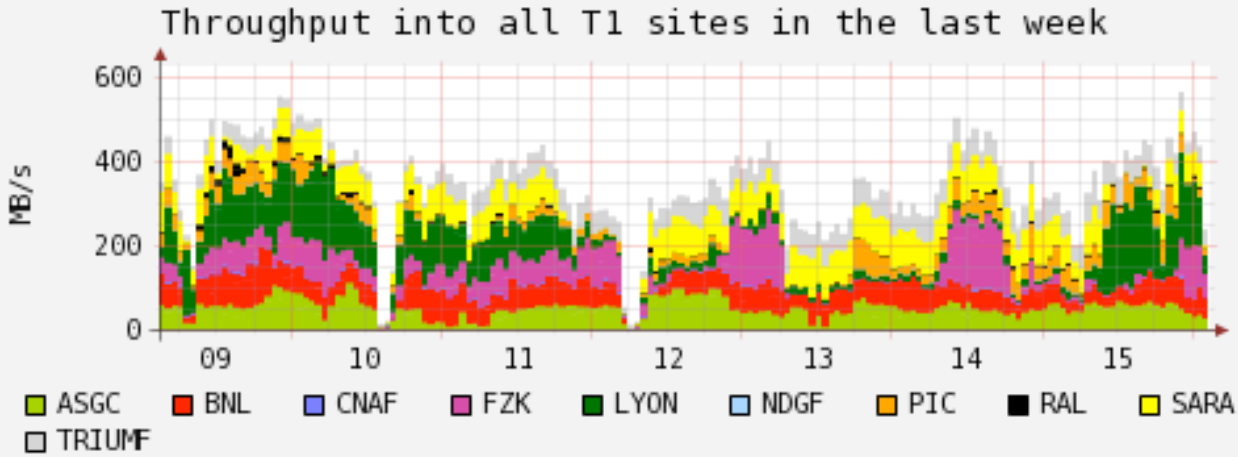
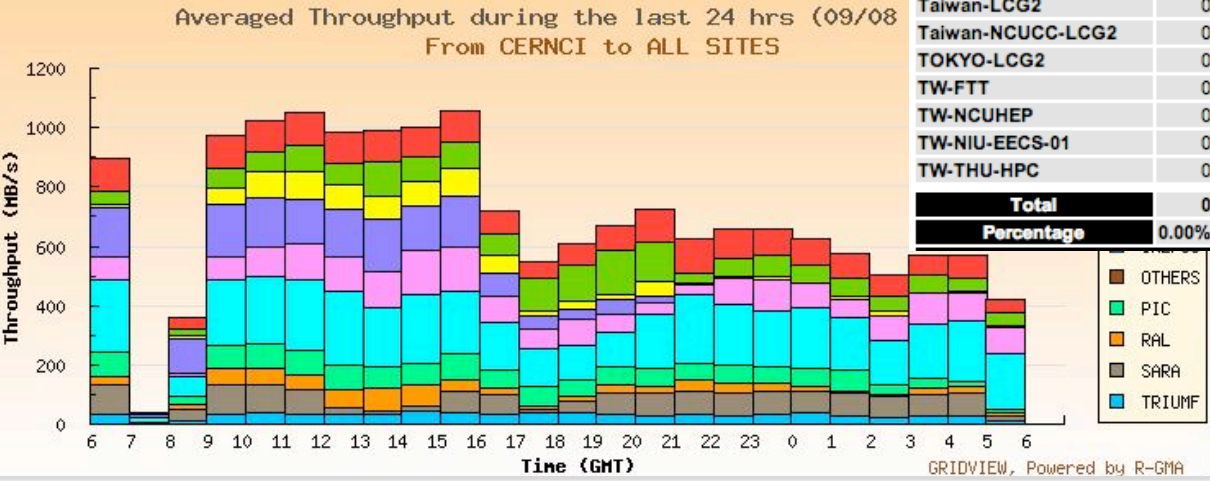
- Target is to have ≥ 8 sites reaching 88% availability by end of Sep. 2006.
- After that, full WLCG services should be in operation.

	CERN- PROD	FZK- LCG2	IN2P3- CC	INFN-T1	RAL- LCG2	SARA- MATRIX	TRIUMF- LCG2	Taiwan- LCG2	USCMS- FNAL- WC1	PIC	average - all Tier-1s
May'06	89%	85%	83%	89%	68%	58%	77%	87%	68%	61%	77%
June'06	92%	15%	89%	62%	76%	49%	88%	75%	64%	88%	70%
July'06	90%	54%	87%	31%	73%	84%	80%	98%	20%	87%	70%
average	90%	53%	86%	61%	72%	64%	81%	87%	50%	78%	72%



WLCG Services in Taiwan

Normalised CPU time [units 1K.Si2K.Hours] by SITE and VO														
SITE	alice	apdg	atlas	atlasprd	belle	blomed	cms	dteam	lhcb	na48	ops	twgrid	Total	%
Australia-UNIMELB-LCG2	0	0	3,748	0	0	0	0	2	0	0	2	0	3,752	0.67%
GOG-Singapore	0	0	4,795	0	0	1,277	3,734	7	0	0	0	0	9,813	1.76%
INDIACMS-TIFR	0	0	0	0	0	0	56,363	74	0	0	0	0	56,437	10.10%
JP-KEK-CRC-01	0	0	0	0	1,453	0	0	31	0	0	0	0	1,484	0.27%
JP-KEK-CRC-02	0	0	737	0	0	0	0	0	0	0	0	0	737	0.13%
LCG_KNU	0	0	1,259	0	0	0	10,386	12	586	0	1	0	12,244	2.19%
NCP-LCG2	0	0	6,628	0	0	0	11,148	8	1,722	0	0	0	19,506	3.49%
PAKGRID-LCG2	0	0	1,318	0	0	217	3,355	4	255	0	0	0	5,149	0.92%
Taiwan-IPAS-LCG2	0	0	10,670	1,165	0	0	0	20	0	0	0	0	11,855	2.12%
Taiwan-LCG2	0	0	242,140	0	0	25,595	140,613	33	0	0	2	153	408,536	73.13%
Taiwan-NCUCC-LCG2	0	0	1,945	0	0	0	1,476	1	131	0	0	0	3,553	0.64%
TOKYO-LCG2	0	0	13,758	0	0	0	0	2	0	0	0	0	13,760	2.46%
TW-FTT	0	0	126	0	0	0	3,400	7	0	0	0	0	3,533	0.63%
TW-NCUHEP	0	0	0	0	0	0	8,301	9	0	0	0	0	8,310	1.49%
TW-NIU-EECS-01	0	0	0	0	0	0	0	1	0	0	0	0	1	0.00%
TW-THU-HPC	0	0	0	0	0	0	0	4	0	0	0	0	4	0.00%
Total	0	0	287,124	1,165	1,453	27,089	238,776	215	2,694	0	5	153	558,674	
Percentage	0.00%	0.00%	51.39%	0.21%	0.26%	4.85%	42.74%	0.04%	0.48%	0.00%	0.00%	0.03%		





T1-T2 FTS: Performance Results

- Regional Centers:
 - Tokyo-LCG2: 48 MB/sec
 - KEK-LCG2 10 MB/sec
 - Australia-UNIMELB-LCG2: 3.23 MB/sec (16MBps by Iperf)
 - KNU 37 MB/sec
 - BEIJING-LCG2: 16 MB/sec
- Domestic:
 - IPAS: 100+ MB/sec
 - TW-NIU-EECS-01: 4.3 MB/sec
 - FTT: ~35 MB/sec
 - NCU-HEP: 40 MB/s
- SRM or storage issues:
 - PAKGRID, TIFR, NCP

Dissemination & Outreach

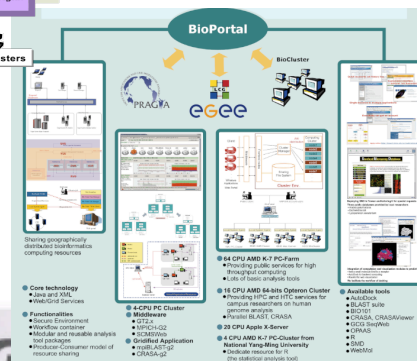
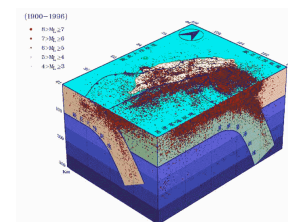
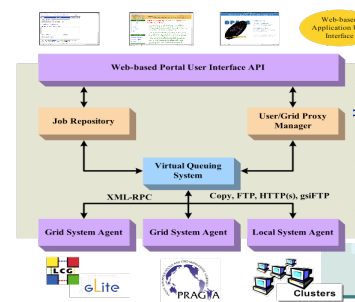
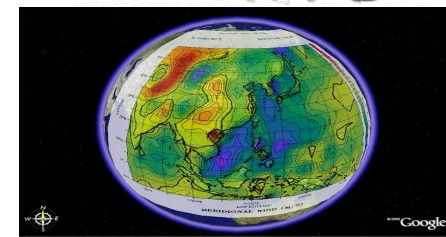
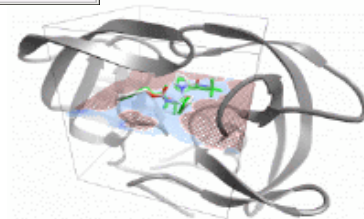
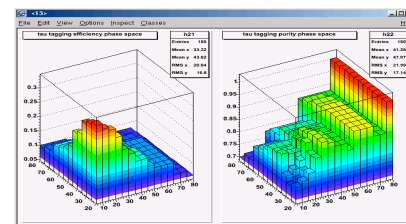
- International Symposium on Grid Computing from 2002
- TWGRID Web Portal
- Grid Tutorial, Workshop & User Training: > 700 participants in past 10 events
- Publication
- Grid Café / Chinese (<http://gridcafe.web.cern.ch/gridcafe/>)

Event	Date	Attendant	Venue
China Grid LCG Training	16-18 May 2004	40	Beijing, China
ISGC 2004 Tutorial	26 July 2004	50	AS, Taiwan
Grid Workshop	16-18 Aug. 2004	50	Shang-Dong, China
NTHU	22-23 Dec. 2004	110	Shin-Chu, Taiwan
NCKU	9-10 Mar. 2005	80	Tainan, Taiwan
ISGC 2005 Tutorial	25 Apr. 2005	80	AS, Taiwan
Tung-Hai Univ.	June 2005	100	Tai-chung, Taiwan
EGEE Workshop	Aug. 2005	80	20th APAN, Taiwan
EGEE Administrator Workshop	Mar. 2006	40	AS, Taiwan
EGEE Tutorial and ISGC	1 May, 2006	73	AS, Taiwan



e-Science Applications in Taiwan

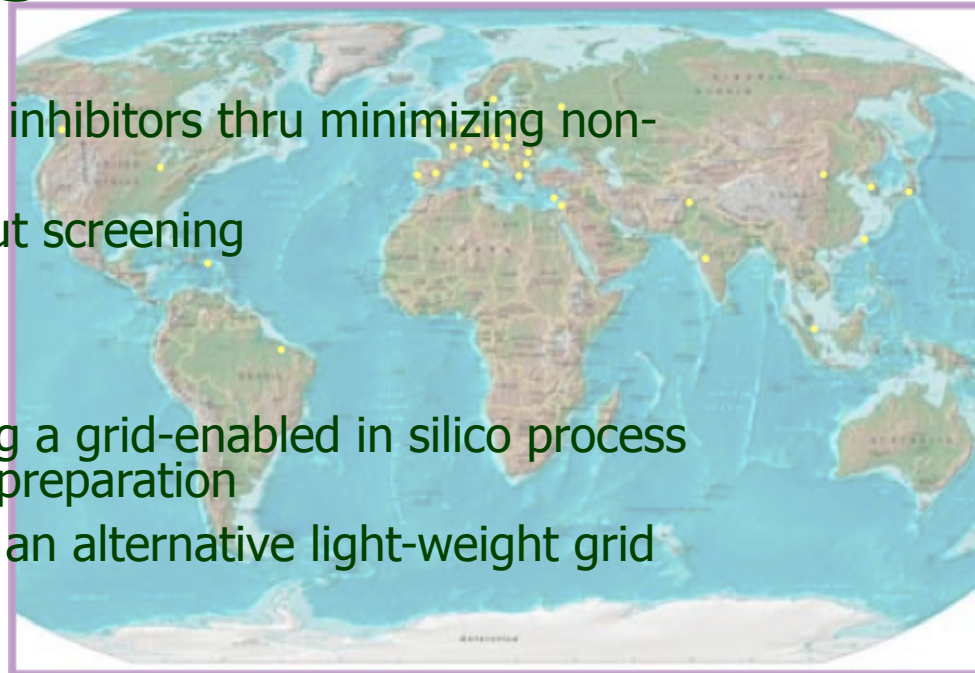
- High Energy Physics: WLCG
- Bioinformatics: mpiBLAST-g2
- Biomedicine: Distributing AutoDock tasks on the Grid using DIANE
- Digital Archive: Data Grid for Digital Archive Long-term preservation
- Atmospheric Science
- Geoscience: GeoGrid for data management and hazards mitigation
- Ecology Research and Monitoring: EcoGrid
- BioPortal
- Biodiversity: TaiBIF/GBIF
- e-Science Application Framework Development





EGEE Biomed DC II – Large Scale Virtual Screening of Drug Design on the Grid

- **Biomedical goal**
 - accelerating the discovery of novel potent inhibitors thru minimizing non-productive trial-and-error approaches
 - improving the efficiency of high throughput screening
 - **Grid goal**
 - aspect of massive throughput: reproducing a grid-enabled in silico process (exercised in DC I) with a shorter time of preparation
 - aspect of interactive feedback: evaluating an alternative light-weight grid application framework (DIANE)
 - **Grid Resources:**
 - AuverGrid, BioinfoGrid, EGEE-II, Embrace, & TWGrid
 - **Problem Size:** around 300 K compounds from ZINC database and a chemical combinatorial library, need ~ 137 CPU years in 4 weeks
- ⇒ a world-wide infrastructure providing over than 5,000 CPUs



Distributed Data Management & Long-term Preservation of NDAP

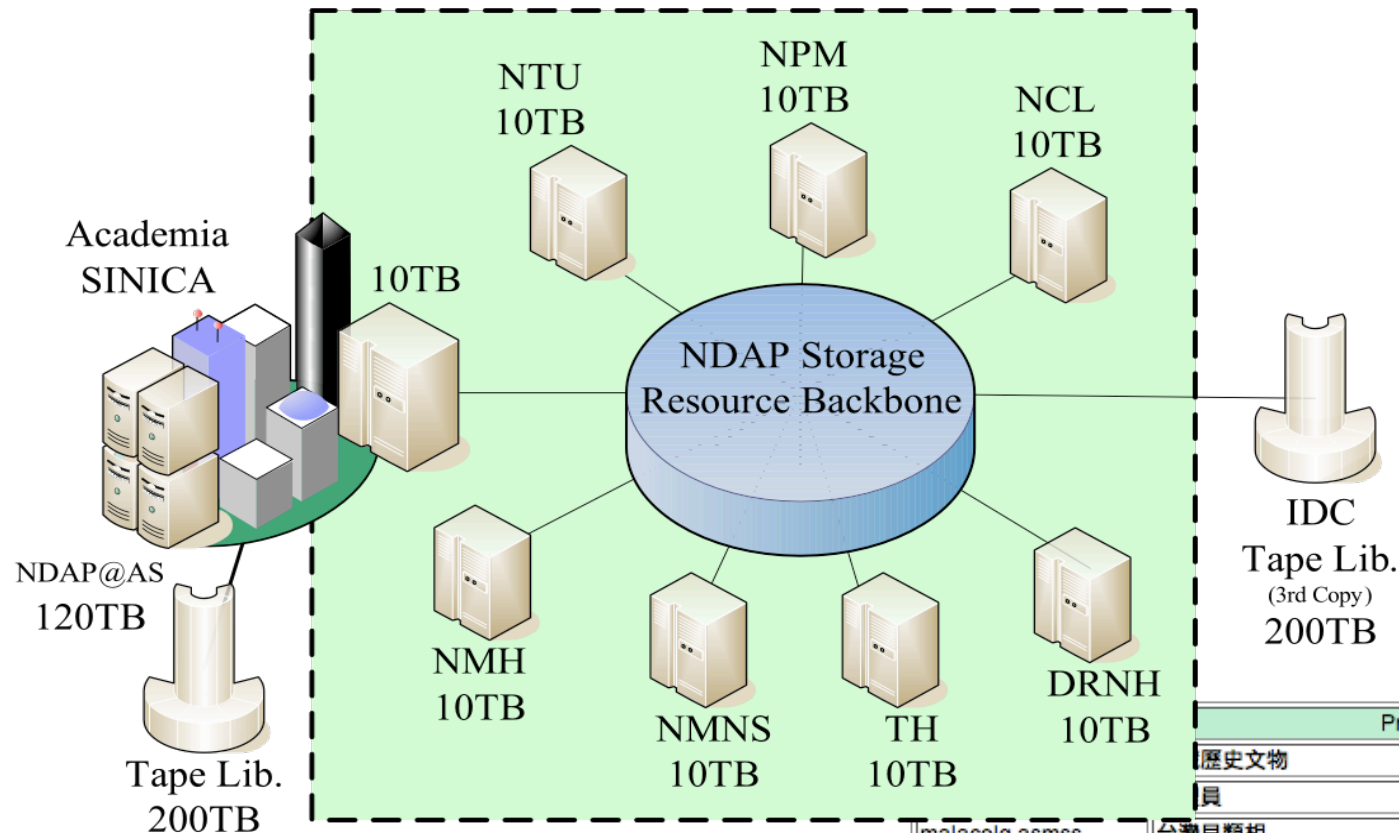
- Long-term Preservation
 - Automatic remote replication with 3 copies in different sites
 - Effective migration based on metadata
 - not just the digitized contents were archived, but also their metadata, methods/procedures, standard format, and management information
 - Separation of data representation and presentation
- Secure Access
- Reduce the total cost of management
- Data Management Framework could be shared for content-based applications, e.g., federation etc.
- Sustainable Operation and Services

Architecture

- NDAP LTP Infrastructure
 - Storage Resource Network: = 8 Storage Resource Centre + IDC + TWAREN/GSN
 - Middleware
 - replication, fail-over, uniform namespace, metadata catalog federation
 - Security Infrastructure
- Operation and Management
- Interface to Applications
 - Discovery
 - Management



SRB-based Data Grid System Architecture for NDAP



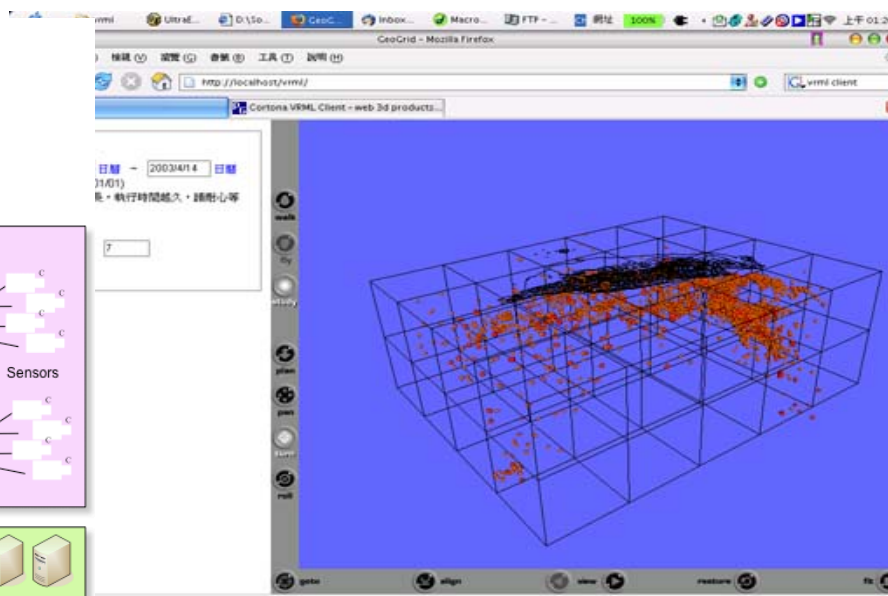
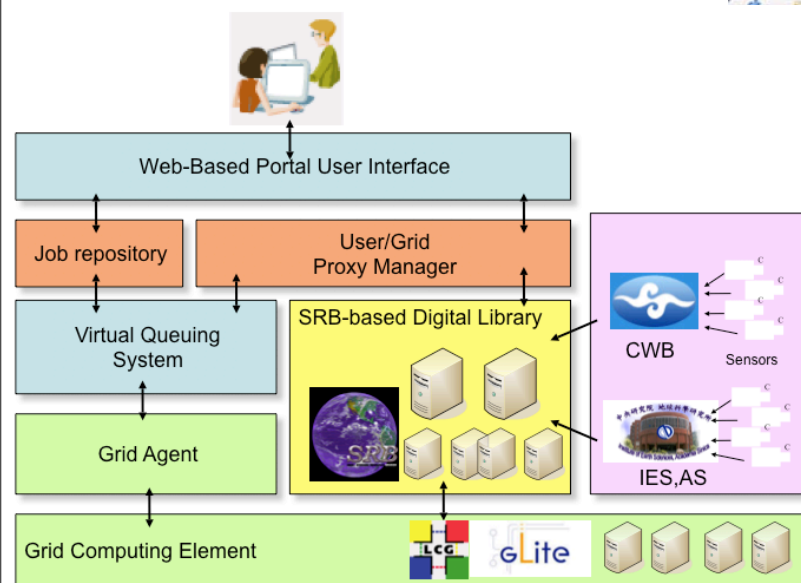
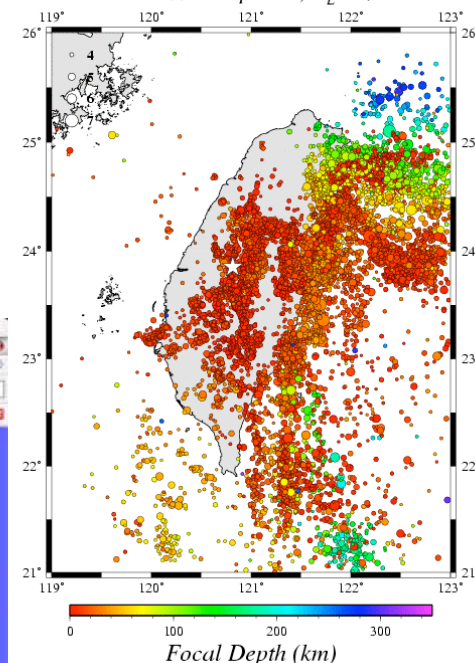
Project	Total Files	Total Size (Byte)
歷史文物	110,243	9,427,834,020,738
員	11,085	981,332,124
malacolg.asmss	22,127	71,677,099,489
gis.asmss	77,734	1,152,358,082,818
la.asmss	1	7,049,563
daal.asmss	473,106	670,011,927,564
fishdb.asmss	32,070	4,199,317,364
lthda.asmss	121,346	168,408,646,949
muchwood.asmss	32,542	1,640,317,250,276
archives.asmss	640,222	21,824,130,034,178
twnative.asmss	601,715	1,516,242,052,811
Total	2,122,191	36,476,166,813,874

Grid for Taiwan Seismology

- Data Management of 地震觀測網與統計資料
- Data Portal of Taiwan Earthquake Center
- Analysis by Grid for Seismology
- Services

Taiwan Seismicity

Jan. 1990 ~ Sep. 2003, $M_L \geq 3.5$

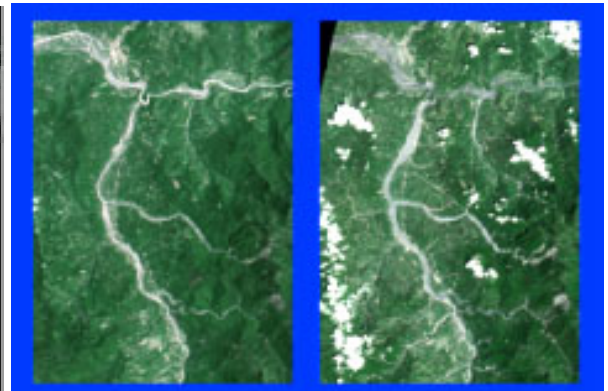
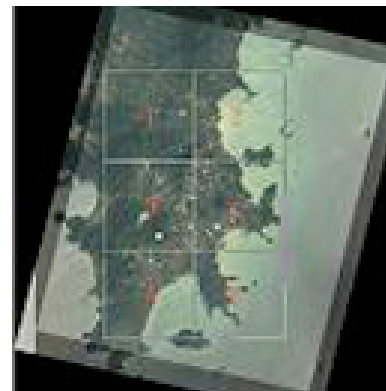




Taiwan GeoGrid

- Applications

- Grid for Geoscience, Earth Science and Environmental Research and Applications
- Land Use and Natural Resources Plan/Management
- Hazards Mitigation
 - Typhoon
 - Earthquake
 - Flood
 - Coast line changes
 - Landslide/Debris flow

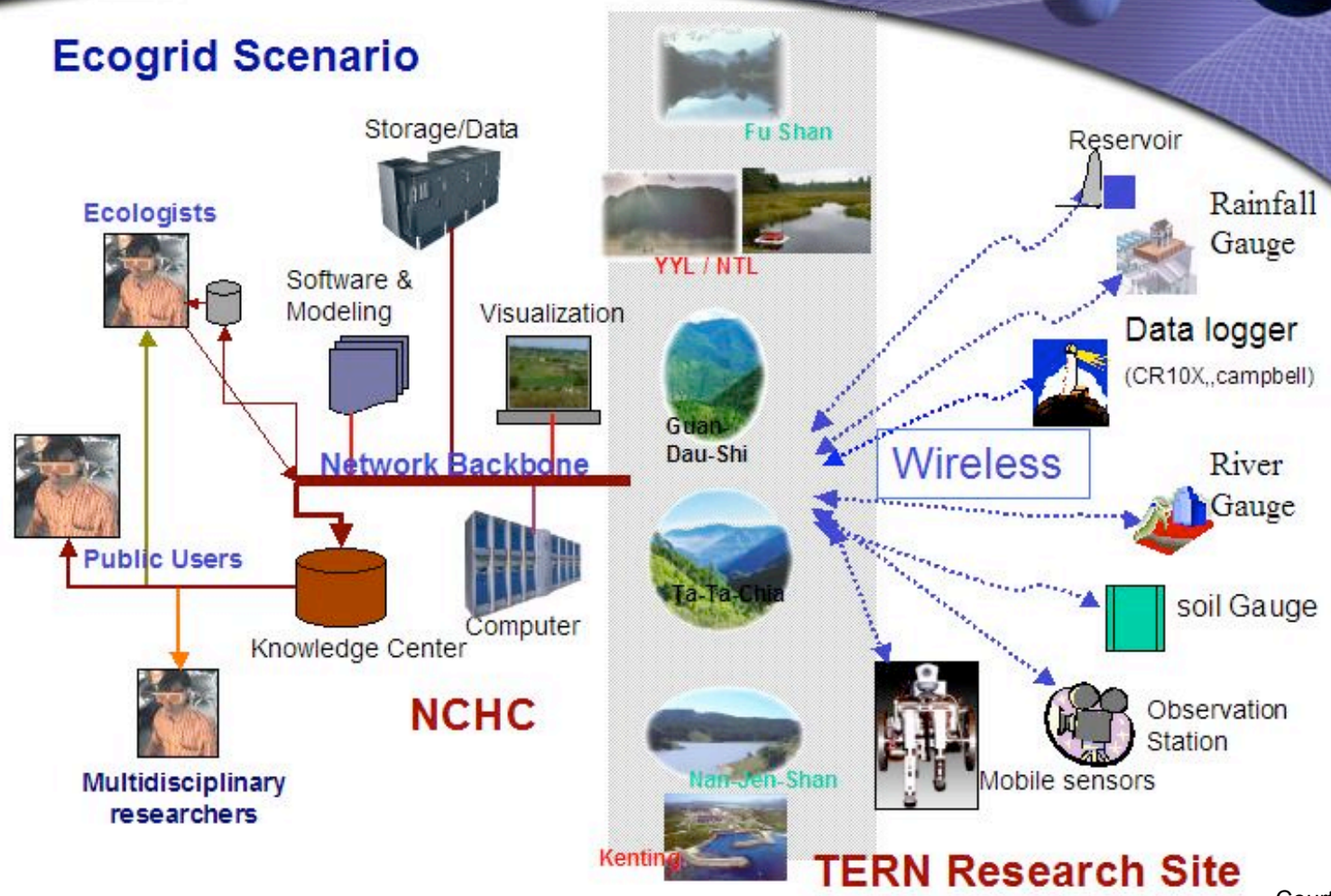


- On-the-fly overlay of base maps and thematic maps,

- from distributed data sources (of variant resolution, types, and time) based on Grid Data Management
- WebGIS/Goole Earth based UI
- Integration of Applications with Grid

EcoGrid

Ecogrid Scenario



Courtesy from NCHC



ASGC e-Science Application Focus

- Grid Portal
 - common data sharing environment as a one stop shop to search for and access data from difference administrative domains on heterogeneous system in a UNIFORM way
- Content Analysis & Management
 - Metadata model, Content management framework, data federation
- Security Framework
 - PKI based authentication, authorization, accounting and encryption
- SRM Interface with Storage Resource Broker (collaborate with SDSC)
 - pure distributed data management system, integration to Grid infrastructure, development of SRB-SRM
- Long-Term Preservation (LTP) & Data Curation
 - Persistent archive, Mass storage technology, sustainable operation/business model



SRM-SRB Development

- Objectives

- Middleware Persistence: Integrate SRB into the e-Science infrastructure (gLite+OSG) of Taiwan
- Interoperation:

- Approach & Focus

- Use Case Collection and Analysis
- Make use of the current SRM implementations as the code base
 - e.g., CERN Castor SRM, DESY/FNAL dCache SRM, LCG DPM, LBNL DRM, JLab SRM, etc.
- Evaluate how authentication works in both SRM and SRB
- Evaluate the similar services of gLite



Industrial Program

- NSC-Quanta Collaboration

- To help Quanta Blade System have best performance for HPC and Grid Computing
- Quanta is the largest Notebook manufacturer in the world
- Participants: AS, NTU, NCTS, NTHU, NCHC
- Scientific Research Disciplines: Material Science, Nano-Technology, Computational Chemistry, Bioinformatics, Engineering, etc.
- Performance Tuning, Grid Benchmarking

- ASGC-Microsoft

- To integrate Microsoft CCS into gLite middleware
 - To take advantage of Microsoft CCS computing power as the computing resources in gLite.
- The scientific application porting plan
 - BLAST (Basic Local Alignment Search Tool)



EGEE Activities

- Joined EGEE as a non-funded member from 2004
 - SA1: Support and Operation Management
 - CIC/ROC
 - SA3: Integration, Testing & Certification
 - NA2: Dissemination & Outreach
 - NA3: User Training & Induction
 - NA4: Application Identification & Support
 - High Energy Physics
 - Biomedical: HealthGrid
 - Digital Archive
 - JRA1: Middleware Re-engineering
 - gLite testbed
 - gLite pre-production site



Deployment Status in Asia Pacific

- APROC EGEE Sites
 - 17 production sites
 - 7 site deployed and certified this year
- 3 under deployment and certification
 - Korea: KISTI, KONKUK
 - Mongolia: (MAS IPT) Mongolian Academy of Sciences
 - Taiwan: NTCU
- ** Remove Australia as potential site
 - They are already certified.



APROC Introduction



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services
- Supports EGEE sites in **Asia Pacific** since April 2005



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services
- Supports EGEE sites in **Asia Pacific** since April 2005



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services
- Supports EGEE sites in **Asia Pacific** since April 2005
- Operates ASGCCA that provide certificates for AP EGEE sites without domestic CA.



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services
- Supports EGEE sites in **Asia Pacific** since April 2005
- Operates ASGCCA that provide certificates for AP EGEE sites without domestic CA.



APROC Introduction

- APROC Mission
 - Provide deployment support facilitating Grid **expansion**
 - Maximize the **availability** of Grid services
- Supports EGEE sites in **Asia Pacific** since April 2005
- Operates ASGCCA that provide certificates for AP EGEE sites without domestic CA.
- EGEE Operations
 - CIC-on-duty: EGEE global operations
 - Monitoring tool development: GStat and GGUS Search
 - VO services: VOMS and LFC
 - TPM: Front line user support (Q4 2006)



APROC Services: Site Deployment



APROC Services: Site Deployment

- Deployment consulting
 - Site architecture
 - Hardware requirements
- Middleware installation support
 - Configuration
 - Troubleshooting
- Site certification
 - Functionality testing
 - Official EGEE infrastructure registration



APROC Services: Continuous Support

Asia Pacific ROC
http://www.twgrid.org/aproc

Home

Products

Support

Online-Support

Contact

Logout

New Ticket

MyTickets

CompanyTickets

Search

FAQ-Area

Preferences

Welcome min tsai (minhong.tsai@gmail.com)
11/19/2005 18:21:43

Zoom Ticket#: 2005060910000025

Age: 162 days 21 hours
Created: 06/09/2005 20:44:43

[->>>customer (webrequest#) 06/09/2005 20:44:43

[->>>system (email-external) 06/09/2005 20:44:43

[->>>system (email-notification-ext) 06/09/2005 20:45:51

[->>>agent (email-external) 06/09/2005 21:24:22

State: closed successful
Priority: 3 normal
Queue: CIC/ROC
CustomerID: minhong.tsai@gmail[.]

From: min tsai <minhong.tsai@gmail.com>

To: CIC/ROC

Subject: I need some help

Attachment:

problem details...

Subject: I need some help

Text: problem details...

[OTRS]

Min-Hong Tsai (min17502@gate.sinica.edu.tw) Sat Nov 19 18:20:42 2005

Logout

QueueView

Phone-Ticket

Email-Ticket

Search

Preferences

Customer

Bulk-Action

Calendar

FieldManager

WebMail

FAQ

Admin

New message (0)

Locked Tickets (1)

[Queue: CIC/ROC]

Tickets shown: 1-5 - Page: 1 - Tickets available: 5 - All tickets: 6

Queues: My Queues (0) - CIC/ROC (12) - CIC/ROC - closed (158) - HPG - closed (39) - PRAGMA (2) - PRAGMA - closed (13) - CASTOR (4) - Tiet-1 (3)

[Ticket#: 2005101810000036]

Age: 32 days 4 hours

Created: 11/15/2005 11:55:00

Lock - Zoom - History - Priority - Note - Close

From: APROC Trouble Ticketing System <trs@roc.grid.sinica.edu.tw>

To: Niranjan <niranjan@tiffr.res.in>

Cc: ASGC-ROC <roc@lists.grid.sinica.edu.tw>

Subject: Re: [Ticket#2005101810000036] [Roc] Re: fail to cr file into your classic SE

Dear Niranjan,

we also report this prob to CIC, as we can solve this during the past three weeks.

but would like to proceed if you can help switching off the wn030 from DNS.

Thanks.

BR,

J

Thank you for your request.

APROC Trouble Ticketing System <trs@roc.grid.sinica.edu.tw> wrote:

> Dear Niranjan,

>

> sorry for the late,

[...]

State: open

Priority: 3 normal

Queue: CIC/ROC

CustomerID:

Escalation in: none

Compose Answer (email):
• empty answer

Contact customer (phone):
• phone call

Change queue:
CIC/ROC Move



APROC Services: Continuous Support

- Operations Support
 - Monitoring
 - Diagnosis and troubleshooting
 - Problem tracking via OTRS ticketing system
- M/W release deployment support
 - Pre-Production site operations
 - Certification testbed
 - Supplementary release notes
- Security Coordination
 - Security release announcement and instructions
- Documentation: APROC Portal and wiki
 - <http://www.twgrid.org/aproc>
 - <http://list.grid.sinica.edu.tw/apwiki>
- Site communication and support channels
 - Phone
 - Email
 - OTRS Ticketing System
 - Monthly meeting with AsiaPacific sites

The screenshot shows the 'Asia Pacific ROC' web portal. The header includes the logo and the URL 'http://www.twgrid.org/aproc'. A navigation menu on the left lists 'Home', 'Products', 'Support', 'Online-Support', and 'Contact'. The main content area displays a ticket with ID '2005060910000025'. The ticket status is 'closed successful', created on '06/09/2005 20:44:43', and has a priority of '3 normal'. The subject is 'I need some help' and the text is 'problem details...'. The ticket is assigned to 'min tsai' with email 'minhong.tsai@gmail.com'.

The screenshot shows the OTRS (Open Source Ticketing System) interface. The top navigation bar includes links for 'Logout', 'QueueView', 'Phone-Ticket', 'Email-Ticket', 'Search', 'Preferences', 'Customer', 'Bulk-Action', 'Calendar', 'FieldManager', 'WebMail', 'FAQ', and 'Admin'. The main content area shows a list of tickets with columns for 'Queue', 'Ticket ID', 'Status', and 'Priority'. A specific ticket with ID '2005101810000036' is selected, showing its details. The ticket is assigned to 'min-hong.tsai' and has a status of 'open'. The subject is 'APROC Trouble Ticketing System <trs@roc.grid.sinica.edu.tw>'. The text of the ticket includes a message from 'Niranjan' and a response from 'min-hong.tsai'.



Application Startup

- Initial startup: APESCI VO
 - Provided for new communities to test and develop Grid applications
 - Acts as incubator VO for fast access to Grid resources
 - Centralized services already running
 - Resource Broker, LFC and VOMS services
- Next step: Production VO
 - Discuss with NA4 to join existing VO and collaborate
 - Create a new VO
 - APROC can also help host LFC and VOMS for the new VO



APROC Collaboration Areas

- Experienced sites can work with APROC
 - TPM User support tasks
 - Regional site administrator training
 - Regional monitoring tool development
 - Other EGEE Activities
- Regional Specific Application Development
- Workshop and Training
 - Site Administration
 - Application Development
 - ...
- Dissemination and Outreach

Requirements: User's viewpoint

- Find Data
 - Registries & Human communication
- Understand data
 - Metadata description, Standard / familiar formats & representations, Standard value systems & ontologies
- Data Access
 - Find how to interact with data resource
 - Obtain permission (authority)
 - Make connection
 - Make selection
- Move Data
 - In bulk or streamed (in increments)

Requirements: User's viewpoint 2

- Transform Data
 - To format, organisation & representation required for computation or integration
- Combine data
 - Standard DB operations + operations relevant to the application model
- Present results
 - To humans: data movement + transform for viewing
 - To application code: data movement + transform to the required format
 - To standard analysis tools, e.g. R
 - To standard visualisation tools, e.g. Spotfire

Coming Events

- EGEE Tutorial in APAN 2007 (Jan. 2007) in Philippine.
- ISGC 2007 (International Symposium on Grid Computing) will be held in Taiwan, Late March 2007.

<http://www2.twgrid.org/event/isgc2006/>





Summary

- TWGrid is serving as the bridge and facilitator to
 - facilitate the regional e-Science applications
 - establish production e-Infrastructure and integrate such as EGEE/OSG into Taiwan and Asia Pacific
- e-Science has to be application-driven. We are also anticipating more collaboration in the future
- Collaboration mechanism has to be setup ASAP. Then identify the requirements based on applications. And take actions !
- To get users fun and profit by collective intelligence and leveraging the e-Infrastructure.