Grids in Asia Pacific

What has been done? and what hasn't?

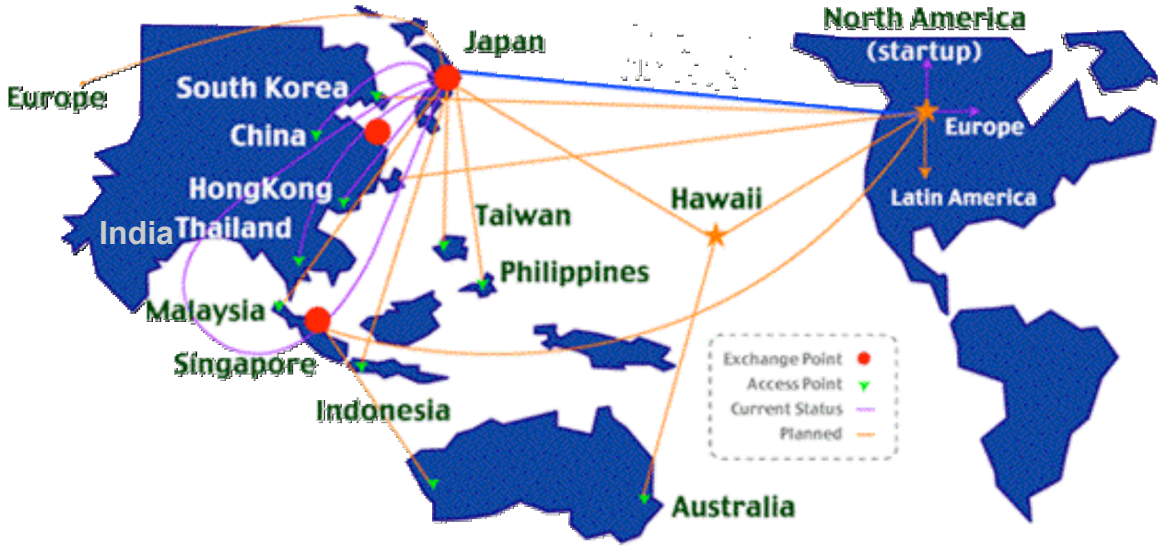
Yoshio Tanaka

Grid Technology Research Center,

AIST, Japan

ApGrid: Asia Pacific Partnership for Grid Computing

- **Open community for Grid researchers in Asia Pacific**

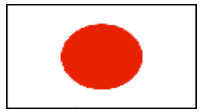


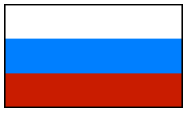
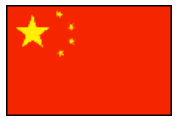
- **ApGrid can be**
  - A meeting point for all Asia-Pacific Grid researchers
  - A communication channel to the GGF, and other grid communities (e.g. TeraGrid, UK-eScience, EUGrid, etc.)
  - A pool for finding in

15 countries, 49 organizations

ApGrid Testbed – unique features –

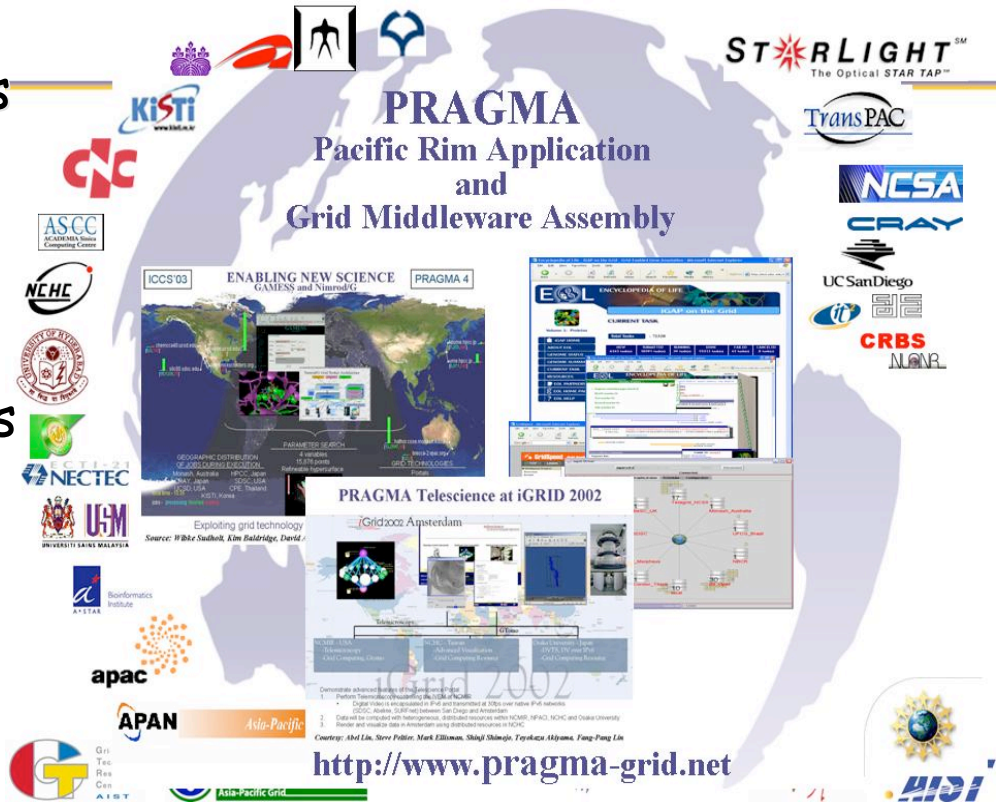
- **Truly (naturally) multi national/political/institutional VO beyond boundaries**
  - Not an application-dedicated testbed – general platform
  - Diversity of languages, culture, policy, interests, ...
- **Grid BYO – Grass roots approach**
  - Each institution contributes his resources for sharing
  - Not a single source funded for the development
- **Physical resources**
  - Most contributed resources are small-scale clusters
  - Networking is there, however the bandwidth is not enough
- **We can**
  - have experiences on running international VO
  - verify the feasibility of application to grid development



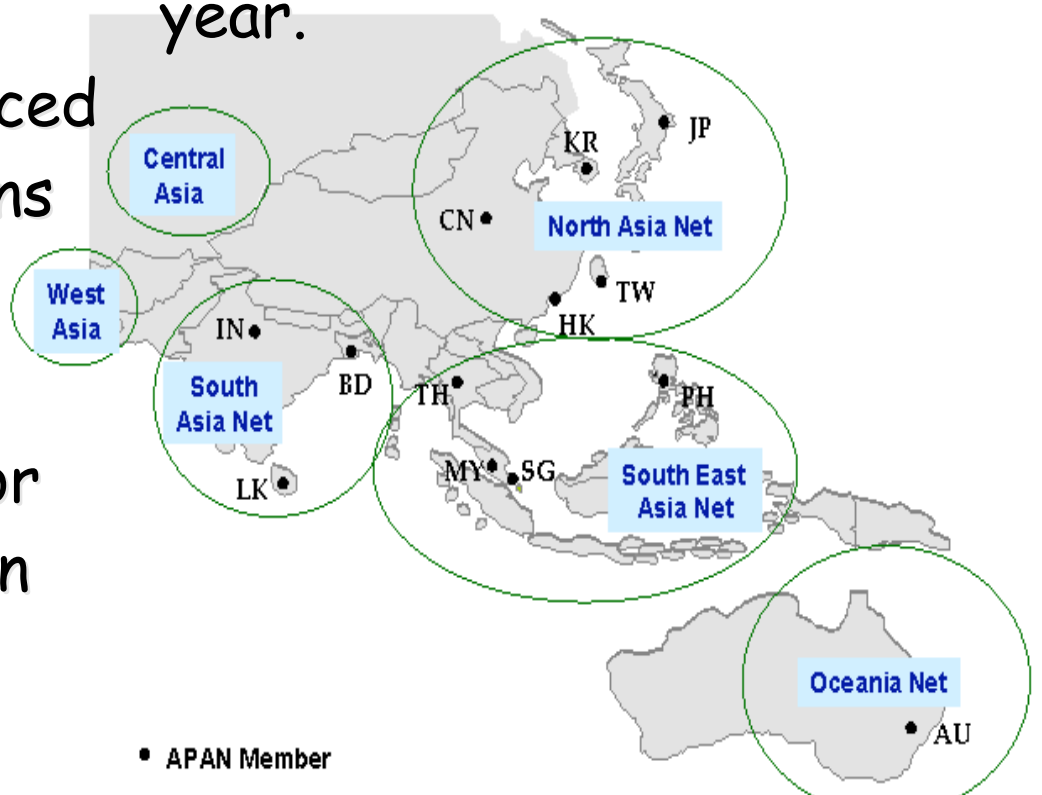
PRAGMA Pacific Rim Application and Grid Middleware Assembly

- NSF-funded project lead by UCSD/SDSC.
- 1st workshop was held in March 2002.
- Establish sustained collaborations and advance the use of the Grid technologies for applications.
- Expected outcomes:
  - Advance scientific applications
  - Increase productive and effective use of the grid by researchers and scientists in the Pacific Rim
  - Increase interoperability of grid middleware in Pacific Rim and throughout the world
- Tightly collaborating with ApGrid.
- Having workshops 2~3 times a year.



APAN Asia-Pacific Advanced Network

- Non-profit international consortium established on 3 June 1997.
- Being a high-performance network for R&D on advanced next generation applications and services.
- Providing an advanced networking environment for the research and education community in Asia-Pacific.
- Promoting global collaboration
- Tightly collaborating with TransPAC.
- Having meetings twice a year.



ApGrid/PRAGMA Testbed

- Architecture, technology
  - Based on GT2
    - Allow multiple CAs
    - Build MDS Tree
  - Grid middleware/tools from Asia Pacific
    - Ninf-G (GridRPC programming)
    - Nimrod-G (parametric modeling system)
    - SCMSWeb (resource monitoring)
    - Grid Data Farm (Grid File System), etc.
- Status
  - 26 organizations (10 countries)
  - 27 clusters (1674 CPUs)

Users, Applications and Experiences

- **Users**
  - Participants of both/either ApGrid and/or PRAGMA
- **Applications**
  - Scientific Computing
    - Quantum Chemistry, Molecular Energy Calculations, Astronomy, Climate Simulation, Molecular Biology, Structural Biology, Ecology and Environment, SARS Grid, Neuroscience, Tele Science, ...
- **Experiences**
  - Successful resource sharing between more than 25sites in the application level.
  - Lessons Learned
    - We have to pay much efforts for initiation
      - Installation of GT2/JobManager, CA, firewall, etc.
    - Difficulties caused by the bottom-up approach
      - Resources are not dedicated
      - Incompatibility between different version of software
    - Performance problems
      - MDS, etc.
    - Instability of resources

An experiment using Grid-enabled climate simulation

- **Climate simulation is used as a test application to evaluate progress of resource sharing between institutions**
- **Climate simulation is developed using GridRPC system called Ninf-G**
  - Ninf-G is developed using Globus C and Java API.
  - Layered in the same level of MPICH-G2
- **We can confirm achievements of**
  - Globus-level resource sharing
    - Globus is correctly installed
    - Queuing systems/job-manager work well
    - Mutual authentication based on GSI
  - Application-level resource sharing
    - Network configuration of the cluster (note that most clusters use private IP addresses for backend nodes)

Necessary steps for the deployment

- **Apply my account to each site**
- **Asked to add my entry to grid-mapfile**
- **Test globusrun**
  - authentication
    - Is my CA trusted? Do I trust your CA?
    - Is my entry in grid-mapfile?
  - DNS lookup
    - reverse lookup is used for server authentication
  - firewall / TCP Wrapper
    - Can I connect to the Globus gatekeeper?
    - Can the globus jobmanager connect to my machine?
  - jobmanager
    - Is the queuing system (eg. pbs, sge) installed correctly?
    - Does jobmanager script work as expected?
- **In case of TeraGrid**
  - Obtained my user certificate from TeraGrid CA (NCSA CA)
  - Asked ITECH and KISTI to trust NCSA CA
  - It was not feasible to ask TeraGrid to trust AIST GTRC CA

Necessary steps for the deployment (cont'd)

- **Install Ninf-G2**
  - Two frequently occurred problems due to inappropriate installation of GT2 SDK
  - Ninf-G2 links SDK libraries for GRAM, DATA, and INFO
    - All SDK libraries are expected to be built from source bundles
    - All SDK libraries are expected to be built using the same flavor
  - If SDK is built from binary bundles, we need to ask rebuild SDK from source bundles
  - Sample build flavor in GT2 manual:
    - GRAM and DATA: gcc32dbg
    - INFO: gcc32dbgpthr
  - Need to ask additional installation of INFO SDK with gcc32dbg
    - For GT2: simply run gpt-build for INFO SDK bundle with gcc32dbg
    - For GT3.0: need to modify install-gt3 script
    - For GT3.2: no smart solution ☹

In case of TeraGrid

Necessary steps for the deployment (cont'd)

- **Install and test the climate simulation**
  - Can Ninf-G server program connect to the client?
  - If private IP addresses are used for backend nodes, NAT must be available
  - Does the application run correctly?

Lessons Learned

- **We have to pay much efforts for initiation**
  - Troubles in installation of GT2/PBS/jobmanger-pbs
    - Firewall?
    - Possible reasons for authentication error
      - CA Cert is not placed in /etc/grid-security
      - My entry is not appeared in grid-mapfile
      - Reverse lookup failed or the returned hostname does not match to the hostname in hostcert
        - Add host entries in /etc/hosts in our resources
      - Time is not synchronized between a client and a server
        - Change configuration of the system to refer a NTP server
      - CRL is expired
    - Possible reasons for errors in PBS
      - failed in rsh/ssh from/to server to/from backend nodes
        - .rhosts, ssh key, mismatch of hostname
      - pbs_rcp was located in NFS mounted (nosuid) volume

Lessons Learned (cont'd)

- **MDS is not scalable and still unstable**
  - Performance is not good
    - GIIS lookup takes several ten seconds ~ minutes
  - Some parameters in grid-info-slapd.conf such as sizelimit, timeout, cahcettl, etc., should be set to appropriate values depends on your environment (number of registered objects, network performance between GRISes and GIISes, etc.).

Lessons Learned (cont'd)

- **Difficulties caused by the grass-roots approach.**
  - It is not easy to keep the GT2 version coherent between sites.
    - Some sites are still using GT2.0 since their application is built on the old version of CoG.
    - It is not easy to catch up frequent version up of the Globus Toolkit.
  - Most resources are not dedicated to the ApGrid Testbed. (though this is a common problem for Grids)
    - There may be busy resources
    - Need grid level scheduler, fancy Grid reservation system?

Lessons Learned (cont'd)

- **Requirements depend on the middleware and application**
  - For Ninf-G2
    - GT2 SDK must be built from source bundles
    - GT2 SDK must be built using the same flavor
    - NAT must be available for private IP address nodes
  - For other Ninf-G2 applications
    - Intel Fortran Compiler is necessary
    - GAMESS or Gaussian is necessary
  - In addition to the common software stack, we need to install software according to the requirements by the application

Lessons Learned (cont'd)

- **The testbed was unstable**
  - Unstable / poor network
  - System maintenance (incl. version up of software) without notification
    - realized when the application would fail.
    - it worked well yesterday, but I'm not sure whether it works today
  - Grid middleware should take care of this instability.

Observations

- **Still being a "grass roots" organization**
  - Less administrative formality
    - cf. PRAGMA, APAN, APEC/TEL, etc.
    - Difficulty in establishing collaboration with others
  - Unclear membership rules
    - Join/leave, membership levels
    - Rights/Obligations
  - Vague mission, but already collected (potentially) large computing resources

Observations (cont'd)

- **Duplication of efforts on "similar" activities**
  - Organization-wise
    - APAN - participation by country
    - PRAGMA – most organizations are overlapped
  - Operation-wise
    - ApGrid testbed vs PRAGMA-resource
      - may cause confusion
      - technically, the same approach
    - Multi-grid federation
  - Network-wise
    - Primary APAN – TransPAC

Observations (cont'd)

- **Two severe problems**
  - Lack of human resources
    - Few person must take care of everything
      - Cluster management, GT2 installation, software development, research...
    - Nobody is dedicated on this efforts (ApGrid)
  - Have interests, but less motivation for doing hard work
    - Everybody is interested in Grid, but Grid is still difficult to use
      - How can I develop Grid-enabled application?
      - Is it easy to make my application Grid-enabled?

Summary of current status

- **Difficulties are caused by not technical problems but sociological/political problems**
- **Each site has its own policy**
  - account management
  - firewalls
  - trusted CAs
  - ...
- **Differences in interests**
  - Application, middleware, networking, etc.
- **Differences in culture, language, etc.**
  - Human interaction is very important

Summary of current status (cont'd)

- **What has been done?**
  - Resource sharing between 26 sites (1674cpus are used by Ninf-G application)
  - Use GT2 as a common software
- **What hasn't?**
  - Formalize "how to use the Grid Testbed"
    - I could use, but it is difficult for others
      - I was given an account at each site by personal communication
    - Provide documentation
  - Keep the testbed stable
  - Build an environment for helping users
    - helpdesk
    - training

What's the role of ApGrid?

- **Deploy middleware on all institutions to support easy implementation of resource sharing and development/execution of Grid applications**
  - Define minimum requirements for middleware
  - Define recommended configuration of the middleware
- **Keep the testbed as stable as possible**
- **Define and authorize security policy in Asia Pacific.**
- **Collaboration with PRAGMA and APAN is definitely necessary for tackling with these**

Towards a production Grid

- **Grid technology is still too difficult to use.**
- **In order to solve this problem, ApGrid, PRAGMA, and APAN will collaborate on the following activities:**
  - publishing Grid documents
  - running the production grid
  - support for Applications
  - organizing meetings
  - building knowledge center
  - running Asia Pacific Grid PMA

Summary

- **It is tough work to make resources available for applications**
  - many steps
- **It is tough to keep the testbed stable**
- **Many issues to be solved toward a production Grid**
  - Technical
    - local and global scheduler
    - dedication / reservation / co-allocation
  - Political
    - CA policy
    - How can I get an account on your site?
  - Both
    - Coordination of middlewares
- **Tight collaboration between ApGrid, PRAGMA, and APAN is definitely necessary for running production Grid in Asia-Pacific**