

Enhancing APAN Efforts to Support Grid Activities in Asia-Pacific Region

Kento Aida, Tokyo Institute of Technology

Peter Arzberger, University of California, San Diego

Yoshio Tanaka, Satoshi Sekiguchi, National Institute of Advanced Industrial Science and Technology

Jysoo Lee, Korea Institute of Science and Technology Information

Seishi Ninomiya, National Agricultural Research Center

Chris Robb, John Hicks, James Williams, Indiana University

Putchong Uthayopas, Kasetsart University

1 Introduction

Several groups related to APAN are involved in grid activities either partially or fully. Considering the importance and perspective of the promising concept and technologies of the grid, we may need to consider how to approach it as APAN and how to coordinate these groups within the APAN framework. The APAN Grid Activities Committee was created in August 2003 in order to discuss APAN's efforts with respect to grid activities.

This position paper summarizes discussions in the Grid Activities Committee and presents the current status and the future role of grid activities in APAN. The focuses of the paper are: (1) to clarify activities/efforts that grid communities wish APAN to do and those that groups in APAN wish grid communities to do, (2) to clarify the role of APAN to support grid activities in Asia-Pacific, i.e. coordination between APAN and grid communities outside of APAN, and coordination among groups (areas, working groups and etc.) in APAN.

2 Activities/efforts of APAN and grid communities

This section describes grid activities inside and outside of APAN, and clarifies activities/efforts that grid communities wish APAN to do and those that groups in APAN wish grid communities to do.

2.1 Activities/efforts that grid communities wish APAN to do

Several grid communities in Asia-Pacific have strong relation to activities in APAN. The following subsections introduce overviews of these communities and describe what they wish APAN to do to support the grid activities.

2.1.1 PRAGMA

The Pacific Rim Application and Grid Middleware Assembly (PRAGMA) is an open, institution-based organization, founded in 2002, to establish sustained collaborations and to advance the use of grid technologies in applications among a community of investigators working with leading institutions around the Pacific Rim. PRAGMA was founded based on the following premises: the conduct of science is global and more examples arise that point to the challenges that must be faced internationally; the grid promises to revolutionize science as much as networking has done to our daily activities; and the grid is yet too difficult to use by most researchers. Finally, PRAGMA recognizes that constructing and using the grid to promote e-Science is inherently a global, collaborative undertaking. No one institution or economic entity has all of the talent or all of the resources to do this. Yet, each needs to participate in building the future scientific, social, and economic global infrastructure.

PRAGMA accomplishes its mission primarily by conducting joint projects that develop grid middleware to advance applications and by sharing resources to create a testbed, and addressing scheduling and allocation issues across institutional and international boundaries. In addition, PRAGMA is committed to disseminating the results of its efforts to the broader community and to work with regional and international groups to enhance the overall grid infrastructure and to promote global collaboration.

PRAGMA's activities to date highlight what can be attained by working together across political and disciplinary boundaries, with a common focus and shared principles underlying the collaborations. These activities illustrate the concept that the grid "brings remote resources (observational equipment, computers, data, and people) together to one's local work environment." The examples range from controlling a microscope in order to understand cell processes in the brain to monitoring the environment in national parks, distributed computations that can lead to insights into drug discovery to moving files essential to high-energy physics experiments, and from conducting a global structural genomic experiment to rapidly deploying technology to assist the world in fighting the outbreak of SARS.

In August 2003, prior to the 16th APAN meeting in Busan, PRAGMA and APAN signed a memorandum of understanding (MOU) that formalize a working partnership. Experience and expertise in networking infrastructure and research tied with grid technology, driven by applications, is key to the partnership between APAN and PRAGMA. The MOU that underlies this new partnership envisions these groups working together to promote and develop a common set of grid applications using the expertise and resources of both groups. Target areas for initial close interaction include: a) Natural Resources and b) Applications of APAN.

In particular, the MOU called for PRAGMA and APAN to work together to facilitate and support the creation of computational grids in and among the Pacific Rim institutions by

- Providing each other with input into networking needs by applications and networking directions;
- Hosting joint workshops and meetings in the Pacific Rim;
- Developing and promoting joint-training programs.

Ultimately, PRAGMA's success will be measured by the applications running on the grid, facilitated by PRAGMA members. Thus, involving APAN and its members in these application projects is a key need of PRAGMA. It is anticipated that the grid efforts (including peer to peer efforts) will push the networking demand both in capability (large pipes) as well as capacity (many pipes). Thus, having key points of contact to work on specific projects, even by giving advice, would be very constructive.

Two concrete ideas for discussion with APAN as PRAGMA moves forward with aspects of its testbed.

1. Make the grid usable on a routine basis for applications. What is APAN's role in this? One component is to have connectivity and a networking monitoring site that can be checked for connectivity.
2. As PRAGMA looks towards establishing a persistent infrastructure (testbed), what is the role of APAN?

A third example would be the involvement of application groups in APAN in grid activities, such as the efforts currently underway with the Natural Resources group.

Finally, APAN can play a role in continuing the dialogs and collaborations between ApGrid and PRAGMA, also with various regional networking activities, such as TransPAC.

2.1.2 ApGrid

The Asia Pacific Partnership for Grid Computing (ApGrid) is an open community encouraging collaboration between academia, industry, and government for the research and development of grid technologies. Its objectives are to develop a partnership among Asia Pacific communities to

- Build an international grid testbed,
- Provide a venue for sharing and exchanging ideas and information,
- Provide a venue for helping to initiate new projects,
- Collaborate and build on each other's work,
- Encourage application communities and assist them in using our technologies along with other technologies,
- Become an interface to global grid efforts such as the Global Grid Forum.

One of the most important objectives of the ApGrid is to build an international grid testbed that can be used for the evaluation of middleware as well as for running large-scale applications. The ApGrid testbed will not be simply an application-driven testbed but will be a truly multinational, multilanguage, multi-institutional virtual organization. The following are some issues to be considered in developing the ApGrid testbed as a production grid.

- How to encourage Asia Pacific grid researchers to participate in the ApGrid.
- How to settle administrative issues such as membership rules, policy management, and funding.
- How to develop the testbed.

These issues were discussed at the past ApGrid Core Meetings, and the ApGrid has drafted "ApGrid Testbed Development Notes" that describe security services, information services,

administrative issues, guidelines for contributing resources, and guidelines for use. These notes should be the basis of an “Asia Pacific Grid Middleware Deployment Guide” proposing a set of standard grid middleware to facilitate the implementation of resource sharing in the Asia Pacific region. That document is expected to be referred to by Asia Pacific organizations that have been or are going to participate in grid communities like the ApGrid and PRAGMA. Forty-one organizations from 15 countries were participating in the ApGrid by January 2004, when the ApGrid Testbed already comprised about 1000 CPUs contributed by more than 20 organizations.

Obviously, collaboration between grid communities and network communities is essential for both grid and network development. The APAN has supported the ApGrid by improving its network configuration, such as the routing between organizations in the Asia Pacific region. The ApGrid in turn, provides a guideline for building a grid infrastructure that can be used for various large-scale applications developed by APAN member organizations. Furthermore, grid middleware and applications can be considered network applications on network. That is, the ApGrid will provide applications that can be used for developing new network technologies and evaluating their performance.

The ApGrid expectation of APAN is:

1. Recommending guidelines for grid testbed development, maintenance, and utilization.
 - These guidelines will help provide a grid environment accessible to all Asia Pacific organizations.
 - The “Asia Pacific Grid Middleware Development Guide” written by ApGrid (Yoshio Tanaka, AIST) is expected to be approved as a recommendation.
2. Running, in collaboration with grid communities, a Grid Operation Center (GOC) that maintains a stable and safe grid testbed.
 - Experience in running the Network Operations Center would be valuable in running a GOC.
3. Using APAN members’ technologies to develop an evolutionary IT infrastructure in the Asia Pacific region.
4. Providing technical support.
 - Provide information about the network (status etc.)
 - Provide operational help (network optimization etc.)
5. Encouraging APAN application developers and users to use the testbed.

2.1.3 APEC-TEL ApGrid

Asia-Pacific Economic Cooperation (APEC) is the premier forum for facilitating economic growth, cooperation, trade and investment in the Asia-Pacific region. APEC was established in 1989 to further enhance economic growth and prosperity for the region and to strengthen the Asia-Pacific community. The Telecommunication and Information Working Group (TEL) is committed to improving the telecommunication and information infrastructure in the region and to facilitating effective cooperation, free trade and investment and sustainable development.

In an effort to contribute in narrowing down the technological gap in the science and technologies on telecommunications sector and to attain sustainable growth and balanced

development in the APEC region, the Korean government proposed the “Asia Pacific Grid Implementation (APGrid¹)” project, which was endorsed during the 24th APEC-TEL meeting in 2001. The project was designed to respond directly to the needs established by the APEC Leaders and Ministers to increase cooperation in infrastructure activities within APEC member economies.

The APEC-TEL APGrid project has the following objectives

1. Sharing the results of Grid researches carried out by the APEC member economies, including K*Grid through official channel.
2. Supporting the APEC member economies to initiate their own national Grid projects and forming collaboration among the Grid projects of the APEC member economies.
3. Advancing collaborative Grid application research environments for such as, Bio-technology, Nano-technology, and Environmental-technology in the region.
4. Reducing the digital divide that limits many economies, rural communities and geographic regions from fully exploiting the benefits of the internet and enhancing the utilization and collaboration between R&D networks, such as APII testbed and TEIN.

The APEC-TEL concerns much about governance issues related to the deployment of major Grid infrastructure in the APEC region. APEC-TEL can act as an official channel to establish policies regarding the implementation of Asia Pacific broadband networks, and the cooperation between governments is eventually necessary as technologies become mature. APEC-TEL can also act as an official channel for international collaborations of network related activities.

For the APEC-TEL ApGrid project, two Grid testbeds have been constructed so far: cluster and supercomputer Grid testbeds. The cluster Grid testbed consists of 2 Linux clusters with the total of 80 nodes, which is based on Globus 2.4. The supercomputer Grid testbed consists of 2 IBM p690s (from KISTI and iHPC in Singapore), which also is based Globus 2.4. The need for constructing and using the APEC-TEL ApGrid testbed is common to many countries in the APEC region, and it is intended to extend the testbeds by working with other efforts in the region. An advisory committee for the APEC-TEL ApGrid project has been constructed. The committee has currently seven members from seven countries---Korea, Japan, Australia, Canada, Singapore, Thailand, and China. To ensure that softwares and other results of Grid activities can be transferred to developing APEC member economies, APEC-TEL Grid Workshop was proposed and endorsed at the 29th APEC-TEL meeting in 2004. Working closely with other member economies such as Japan, Korea will host two workshops at the 30th and 32nd APEC-TEL meetings. The APEC supported Grid workshop would help to identify key issues of governmental roles and the creation of international Grid partnerships of infrastructure development. The proposed workshops will bring together governmental representatives and researchers to foster and share experiences related to the development of Grid technology.

¹ There are two grid projects called “ApGrid” in this paper, the Asia Pacific Partnership for Grid Computing described in the section 2.1.2 and the Asia Pacific Grid Implementation project presented in this section. This paper refers to the former project as “ApGrid” and refers to the latter project as “APEC-TEL ApGrid”.

To achieve the sustainable growth and balanced development in the APEC region, the APEC-TEL would like to work together with the other Grid activities. To do so, APEC-TEL expects that APAN can assist APEC member economies by doing the followings.

1. Supporting the high speed bandwidth and QoS of the APAN network
2. Providing the next generation network technologies developed by APAN
3. Helping the Grid projects of the APEC member economies to make progress
4. Providing technical supports of the necessary infrastructure, especially R&D network for carrying out Grid projects in the region.

2.2 Activities/efforts that APAN wish grid communities to do

Several groups related to APAN are partially or fully involved in grid activities. The following sections introduce grid activities in these groups and describe what they wish grid communities outside of APAN to do.

2.2.1 Natural Resource Area

Grid technology is surely promising in many activities of Natural Resource Area such as bioinformatics, agriculture, earth observation and earth system science. We are now at an initial stage to apply this technology to our field, performing a few projects related to the grid. A project to develop a data broker called Metbroker to provide clients applications consistent access to heterogeneous weather databases is a typical example. This project is now partially involved in the PRAGMA activity by providing the accesses to the weather stations under PRAGMA. The idea of Metbroker is now extending to other types of databases such as DEM and maps. We are also planning to provide a framework to handle several types of satellite images taken by different sensors of various satellites with a consistent interface to clients. Another tool we are trying to develop is a grid based localization system with which a single application can provide a user interface in several different languages without any modification.

In general, we need to combine several distinct data sets simultaneously to perform research activities in Natural Resource Area. These data sets or databases are widely distributed in the Internet and even databases with the same type of contents (e.g. weather data) are heterogeneous in their data formats, database management systems (DBMS) and user interfaces. Therefore, we used to collect these data sets one by one to perform researches. This process was extremely inefficient and constrained our activities. Regarding this issue, data grid technologies are now dramatically improving the research environment of Natural Resource Area by providing a sophisticated framework harnessed with meta-databases, agents and brokers. The MetBroker project mentioned above has convinced us of it, though we still need to standardize the framework as an internationally acceptable one.

The computational grid is also required especially for such fields as bioinformatics and earth observation where we need truly high performance computational power for huge scale genetic data, dynamic weather modeling and satellite image analysis. In some of the Asia Pacific countries,

supercomputers are not affordable and the computational grid is also a reasonable solutions.

Recently, the sensor grid is becoming a reality. Huge scale earth observations have been mainly based on satellite observations. However, such observations become more valuable when combined with ground data. The sensor grid can provide high resolution ground data suitably combined together with satellite images.

The grid in Natural Resource Area is still at an initial stage and we can foresee an unlimited number of its possible applications. We believe that grid technologies are constantly advancing and Natural Resource Area would always like to catch up with the latest one, keeping in touch with grid proper people. We should also mention that the activities in Natural Resource Area are partially or even fully involved in eScience through the grid.

2.2.2 Network Technology Area

Network is essential to construct a grid infrastructure. There are examples of collaborations between networking communities and grid communities, e.g. collaborations between APAN Tokyo NOC and ApGrid for the SC2003 demonstration, and the iVDGL projects below.

Based at Indiana University, the international Virtual Data Grid Laboratory (iVDGL) Grid Operations Center (iGOC) serves as a single point of contact and for iVDGL information and operations. The iGOC maintains watch over the hardware and software resources that comprise the iVDGL and ensures the smooth operation of the Grid2003 project. The Grid2003 project is the next phase of the iVDGL laboratory. Grid2003 is a grid used by physics and astronomy experiments to analyze data and consists of more than 25 sites with over 2500 CPUs.

The iGOC uses both proactive and reactive methods to identify issues affecting Grid2003 performance. Grid monitoring tools, including Ganglia and MonaLisa, are used by iGOC operators to watch over grid operations. The operators at the iGOC are available 24/7 and most problems are handled on or before the next business day.

The iGOC also maintains critical iVDGL services used by the Grid2003 project such as the Virtual Organization Membership Service (VOMS) that provides access to grid resources. Historically, each institution provided its own way for users to access its resources. The VO model allows grid members to access resources at multiple institutions. Grid2003 makes extensive use of VOMS to allow its users access to its resources. The iGOC also provides: Replica Location Services providing access to virtual data, the Globus Index Information Service which provides an index of available iVDGL resources, the Ganglia and MonaLisa monitoring tools, and the iVDGL web content at www.ivdgl.org.

2.2.3 Application Technology Area

Many activities in Application Technology Area have been tightly related to grid technology. It has the Grid working group, which plays the interface between grid communities and APAN. Also, application technology discussed in related groups, or peer-to-peer, education and bioinformatics, includes topics related to grid technology.

The Grid working group organized sessions/workshops for grid technology in the past APAN

meetings in order to introduce grid activities in APAN member countries and to disseminate standard grid technology to APAN member countries. Through these events, many activities about grid middleware, applications and testbeds in APAN member countries are reported. The reported activities include:

- Grid middleware, such as Ninf, Ninf-G and Peterpan
- Grid applications, such as earth monitoring, agriculture, bioinformatics, computational fluid dynamics, operations research, visual processing, virtual reality, Data Grid and Access Grid.
- Grid testbeds and national/international grid projects, such as ApGrid, PRAGMA, N*Grid, APEC-TEL ApGrid and ThaiGrid.

The Grid working group supported grid projects outside of APAN to help them utilize APAN network infrastructure for their grid applications. An example of the support is demonstrations at SC2003. The Grid working group helped the exhibitors, the Grid Datafarm team (the winner of the SC2003 Bandwidth Challenge Distributed Infrastructure Award), the ApGrid team and the BioGrid team (the winner of the SC2003 Bandwidth Challenge Application Award), to set up routing configurations for running their applications. Also, the Grid working group assisted in formalizing the partnership between APAN and PRAGMA.

Peer-to-peer technology shares many technical issues with grid technology. The P2P working group started to discuss challenges and opportunities that come from synergy between peer-to-peer technology and grid technology. The joint session between peer-to-peer technology and grid technology has been held at the APAN Honolulu meeting in January 2004. For education, rapid progress of Internet technology makes changes in educational methods. Distance learning using the Internet makes learning a borderless experience where teachers/students geographically distributed share a virtual class room. The Education working group has activities in this field. Access Grid technology is one of enabling technologies to develop borderless educational system or testbed on the Internet.

APBioNet was the first application project in APAN. The User Community Area, which is the former organization of the current Application Technology Area, contributed to run the project since 1998. APBioNet is currently an affiliate member of APAN, and a lot of collaborative projects between APBioNet and APAN are running. These include the BioMirror project, the BioDataGrid framework, the demonstration of grid computing at SC2003 (the winner of the SC2003 Geographically Most Distributed Award), and lectures/conferences on BioGrid technology. APBioNet is also planning/running grid-related projects in the Asia Pacific region, such as development of bioinformatics applications bundled with grid/cluster computing software (APBioBox), and setting up the bioinformatics educational grid using the APBioBox and Access Grid technology.

Although there are a lot of activities for grid applications in APAN, most of application developers do not have enough knowledge or know-how about grid applications and testbeds, that is, programming tools for grid applications, and know-how to set up, operate and use grid testbeds. Thus, it is hard for application developers to develop/run their applications on real grid testbeds. Even for some developers who have the chance to develop/run applications on real grid testbeds,

there may need to be a lot of tedious work to implement/run their applications on ad hoc and unstable testbeds. Furthermore, in many cases, the testbeds have no inter-operability because of the lack of standard settings for grid softwares. In order to support these grid application developers, standard grid middleware, programming tools and a stable grid testbed are needed. The testbed should stably provide information resources, such as computers, networks, databases, instruments or people, and should have inter-operability with grid testbeds operated by other domains. Also, documents that provide information about standard grid technology would help application developers in APAN.

3 Role of APAN to support grid activities

This section describes how APAN should collaborate with grid communities and proposes the organization to support grid activities in APAN.

3.1 Collaboration with grid communities

Grid technology is still too difficult for application researchers/developers/users to use. Although many application groups in APAN have been expecting to run their applications on the grid, they have not yet been able to use the grid as common infrastructure to run their applications on a routine basis. In order to solve this problem, this position paper recommends APAN to collaborate with grid communities on the following activities:

1. Running the production grid

The production grid is necessary to make grid technology usable for application communities. The production grid should be easy to use for application researchers/developers/users and be stably operated by an operational team, namely the Grid Operation Center (GOC). The role of the GOC includes monitoring and providing information about networking/computing resources, trouble shooting, policy management and etc.

Collaboration between APAN and the grid communities is necessary to run the production grid. ApGrid has a lot of experience to run grid testbeds. At the end of 2004, more than 1000 CPUs distributed over 20 organizations are available on the ApGrid testbed. They would be able to provide guidelines to set up the production grid. The current Internet infrastructure is managed by expert teams for operational issues, such as Network operation centers (NOCs). Networking groups in APAN have a lot of experience to run the NOC, and some groups in APAN have experience to support to utilize network infrastructures for grid applications, e.g. demonstrations at SC2003 and iGOC at Indiana University. Their experience would significantly help to run GOC.

2. Implementation of applications on the production grid

Research and development of applications themselves are performed by application groups, e.g.

Application Technology Area and Natural Resource Area. However, they would need supports from experts of grid technology to implement their applications on the production grid, because many application developers do not have sufficient know-how to implement applications on the grid. Collaboration between these application groups in APAN and grid communities, particularly PRAGMA, is necessary to encourage implementation of their applications on the production grid.

3. Publishing documents

In order to encourage the above activities, dissemination of useful information about grid technology is important. Publishing documents about grid technology from APAN would accelerate grid activities in the Asia-Pacific region. The topics of the documents would include:

- Guidelines to set up and operate a grid testbed
- Manuals about standard grid software, e.g. middleware and programming tools
- Introduction of grid technology

4. Organizing sessions on research, experiences or training at APAN meetings

To solve grid related issues in APAN and to further engage other APAN members, hosting meetings at APAN meeting will provide face-to-face interactions between members of the grid community and between grid and application specialists. Moreover, this would be an ideal forum for some aspects of training to use or establish grids.

5. Building resources and knowledge center

Archiving knowledge obtained by the above activities is necessary. Building the resource and knowledge center, which capture the expertise available and make it ready to be used by APAN member, would be of benefit to both APAN and grid communities. The archived information should include documents, training materials, media, news etc.

3.2 Proposal of the new organization

In order to fully support grid activities described in 3.1, APAN needs to have an expert group for the grid activities. This group will be representative of APAN for grid activities and will coordinate APAN and grid communities in order to play the role described in 3.1. Currently, the Grid working group partially plays above missions. However, its activities are limited because of the organizational structure in APAN. Currently the Grid working group is organizationally under Application Technology Area. However, grid activities extend beyond the Application Technology Area to various areas in APAN, such as, Natural Resource Area, Network Technology Area and the e-Science group.

This position paper proposes to create the new committee, which coordinates groups involved in grid activities in APAN (such as those mentioned in this paper and others) and plays an interface role between grid-related groups in APAN and the grid communities outside of APAN, as the successor of the current Grid working group. More concretely, the role of the proposed committee is:

1. To collaborate with grid communities in running the production grid in the Asia Pacific region.
2. To collaborate with grid communities in supporting application groups of APAN to implement their applications on the production grid
3. To publish documents about grid technology
4. To plan sessions/workshops/training activities related to grid technology in APAN meetings
5. To archive knowledge obtained by above activities

3.3 Benefit of the new organization

The newly established organization can play an important role that benefit APAN in many ways such as:

1. Helping to create a clear focus and direction for grid activities under APAN
2. Establishing a clear APAN grid related contact point for other grid communities, in particular those mentioned above, e.g. ApGrid, PRAGMA, APEC-TEL and others
3. Encouraging APAN application communities, e.g. Natural Resource Area, Application Technology Area, APBioNet and the e-science group, to run their applications on the grid
4. Utilizing APAN network infrastructure for grid activities
5. Reducing the duplicated and uncoordinated works related to grid technology under APAN

Acknowledgements

The authors would like to thank Rajkumar Buyya and Tan Tin Wee for helpful comments.