PROCEEDINGS OF SCIENCE

Grids and Clouds Activities in Asia-Pacific

Simon C. Lin
Academia Sinica Grid Computing Centre (ASGC)
Taipei, Taiwan
E-mail: Simon.Lin@twgrid.org

Eric Yen
Academia Sinica Grid Computing Centre (ASGC)
Taipei, Taiwan
MIS Department, National Cheng-Chi University, Taiwan
E-mail: eric.yen@twgrid.org

The International Symposium on Grids and Clouds (ISGC) 2012
Academia Sinica, Taipei, Taiwan
February 26 – March 2, 2012

1.Australia

Following the bankruptcy of a major Australian fiber optics company in 2006, the Australia's Academic and Research Network (AARNet) partnered up with another large company to secure a large portion of this fiber network, which resulted in the establishment of an extensive fiber network between the major cities in Australia that could be run at whichever speed they wished to. This marked the starting point of the AARNet 10 Gbps backbone.

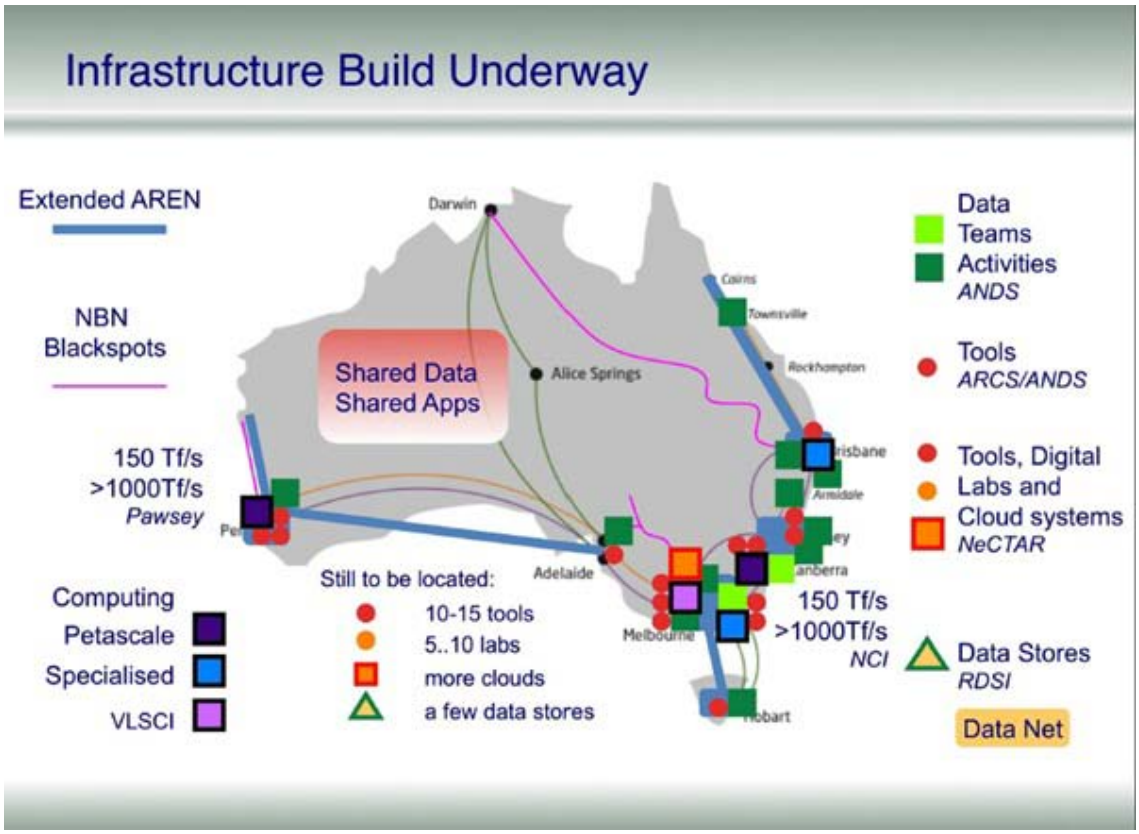

Fig. 1: Current infrastructure projects

Recently, Australia has focused on extending this research network to a significant extent (cf. Fig. 1). As part of the Australia's and New Zealand's joint bid between to host the Square Kilometer Array, one of the most ambitious radio astronomy project to date, preparations to establish a light path between the two countries and optimize the existing network are currently underway.

Australia has also seen the rise of several specialized facilities in high performance computing, such as the Victorian Life Sciences Computation Initiative (VLSCI), which is a state-funded project at the University of Melbourne aimed at providing the life sciences with substantial amounts of computing power. There are also a number of other facilities emerging across the country, in addition to an expansion of the existing network in an attempt to reach some of the black spots, which were not previously serviced.

With a large number of projects underway, Australia intends to achieve a substantial expansion of the current infrastructure by 2013. There are plans to have teams working on metadata in several locations as well as software tools, virtual laboratories to provide examples of eResearch workflows, and some cloud fabric to underpin progress, in addition to the expansion of the specialized and general high performance computing facilities. Australia is also currently funding a nationwide data storage fabric, which is looking to have four or five sites with a proposed capacity of 20 PB per site. By 2013, the newly built infrastructure will be able to facilitate research and provide researchers with a substantial amount of computational power for their projects. Overall, the country's infrastructure has seen a sizable amount of investment, with funding for individual projects ranging from 22 to 80 million Australian dollars, which will put it in a good position to provide a solid infrastructure to researchers.

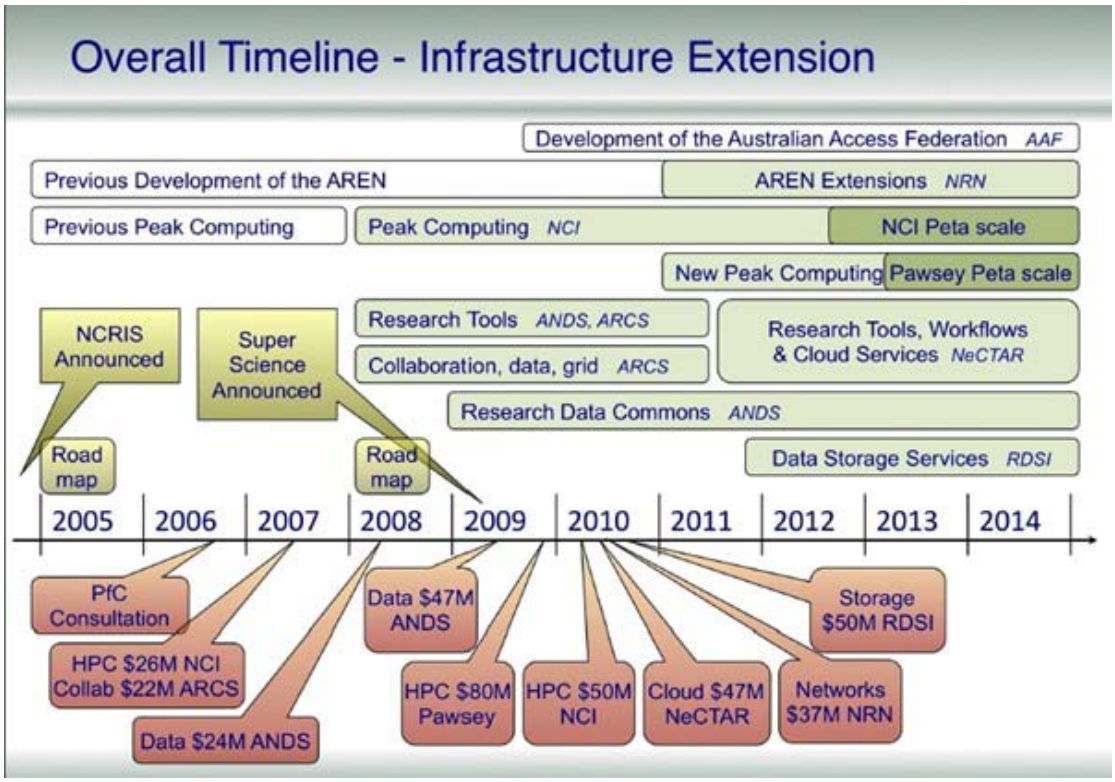

Fig. 2: Timeline of the ongoing infrastructure extension

In 2011, the eResearch landscape in Australia underwent a major change, after the government opted to make some structural adjustments (cf. Fig. 2), which resulted in the creation of the National eResearch Collaboration Tools and Resources (NeCTAR) project that is led by the University of Melbourne. NeCTAR is a 47 million dollar, Australian Government project, conducted as part of the Super Science initiative and financed by the Education Investment Fund. The objective of NeCTAR is to facilitate the use of tools and applications to work remotely and collaboratively, since research is usually conducted in cooperation with other partners.

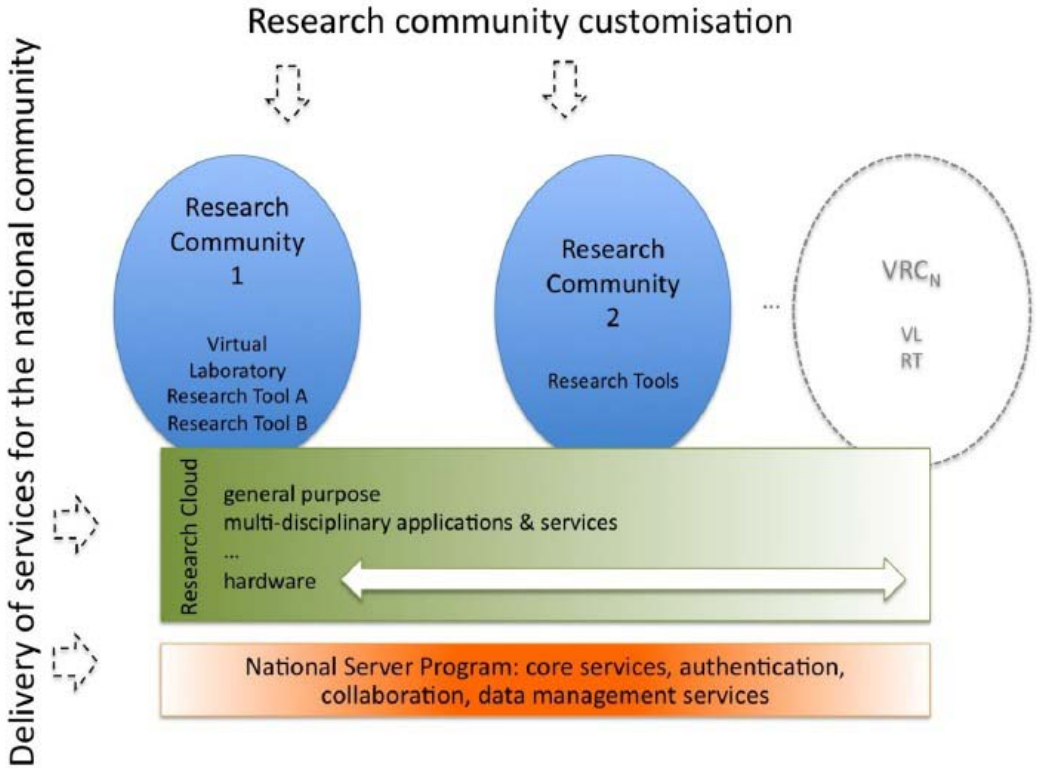

Fig. 3: Organizational structure of NeCTAR

NeCTAR is devided into four programs (cf. Fig.3). Realizing the need for highly reliable infrastructure for hosting core eResearch services, such as authentication and access service, which need to be available around the clock, NeCTAR is currently funding the National Server Program, which provides a robust national server infrastructure. As the central part of the infrastructure, the Research Cloud acts as a service research cloud, designed for any researcher in any discipline at any institution, which is widely expected to be free to use. The NeCTAR Research Cloud went live in February 2012 with a total of 4176 cores and several hundred terabytes of object storage. The aim of the Research Cloud is to break down some of the barriers of entry that can be usually observed in traditional eResearch infrastructure. To help spread its use, NeCTAR is also funding the Virtual Laboratories program, designed to create exemplars for eResearch by supporting and funding communities that are already engaged in conducting high quality eResearch. The fourth pillar of NeCTAR is the Research Tools Program, which sponsors software development projects that help fill existing gaps and have a transformative impact in their respective research areas.

Following a Request For Proposal (RFP) process, NeCTAR has received substantial response from universities and research groups during the first round, including 21 Virtual Laboratory, 48 eResearch Tool, and 6 Research Cloud Node proposals. As of March 2012, the process has reached its final contract negotiation stage for the first round, with a second round scheduled to start the following month. The first results from the first round of projects are expected within the next 12 months.

Other current projects include: (1) the Australian National Data Service (ANDS), a mature metadata project with ongoing workshops aimed at increasing the reuse of research data by educating researchers on how to annotate their data or information with metadata so that it can be shared; (2) the Australian Urban Research Infrastructure Network (AURIN), which was created to improve the understanding of urban resource use and management through the development of multifaceted research infrastructure; (3) the Research Data Storage Infrastructure (RDSI), which currently has proposals under evaluation to procure a substantial amount of data storage; and (4) the Pawsey Centre that is currently under construction and has the primary goal of hosting new supercomputing facilities and expertise to support Square Kilometre Array (SKA) research and other high-end science.

2.China

In the recent years, great efforts have been made toward the development of e-Science in China. As China has entered the second year of its twelfth five-year plan in 2012, the government is committed to supporting the sustainable development of e-Science.

Launched in 2002, the China National Grid (CNGrid) is the national grid initiative supported by the Ministry of Science and Technology and currently the largest grid system in China. The objective of the initiative is to develop PFlop supercomputing systems and integrate these supercomputers into a grid system. On top of this grid system, efforts are also being made to develop grid computing and high-performance computing applications in selected disciplines.

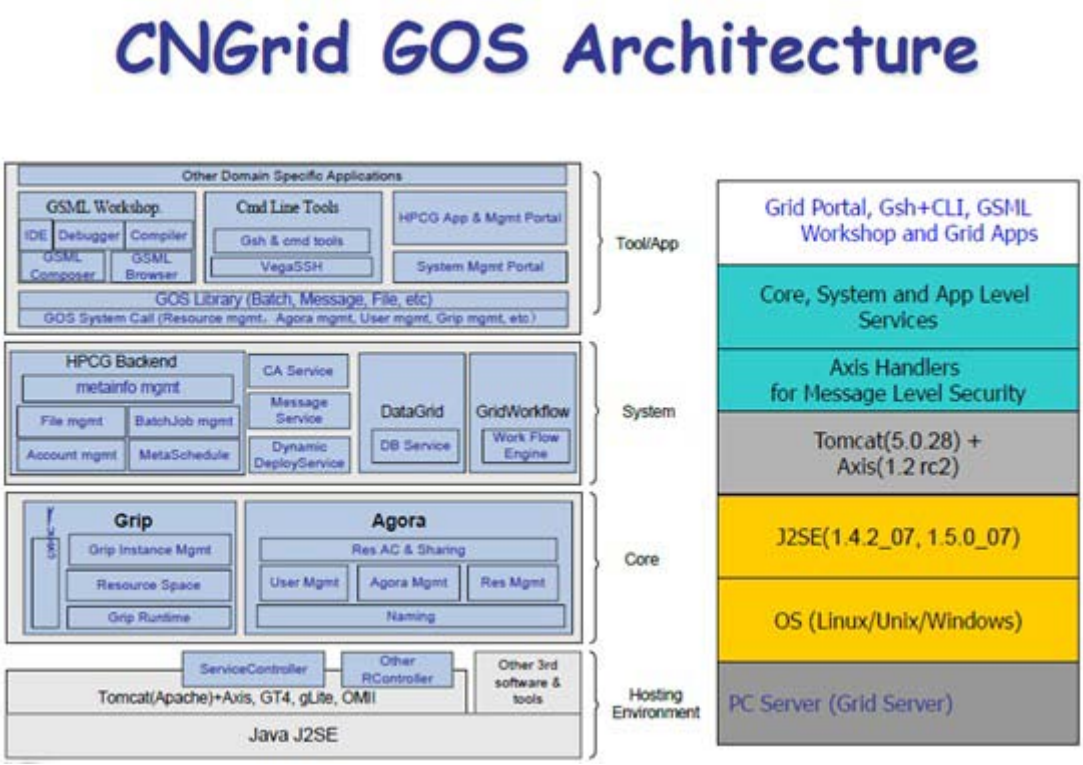

Fig. 4: Architecture of the CNGrid Grid Operation System (GOS)

The figure above illustrates the structure of the CNGrid middleware, also known as the Grid Operation System (GOS). Grip and Agora form the core of the system, which are responsible for the resources and security management. On top of them are three service systems, namely the High Performance Computing Gateway (HPCG), the data grid and the workflow system. Over the past few years, the CNGrid GOS has been deployed at 10 sites and on several application grids. It supports heterogeneous high-performance computing and multiple platforms, such as UNIX, Linux and Windows. CNGrid is connected through the public network with only the HTTP port enabled. Users can access grid resources through various clients, such as a web browser, special client or GSML client. CNGrid currently has 11 sites across China, in addition to a few larger

supercomputer centers and several computer centers operated by universities and research institutes.

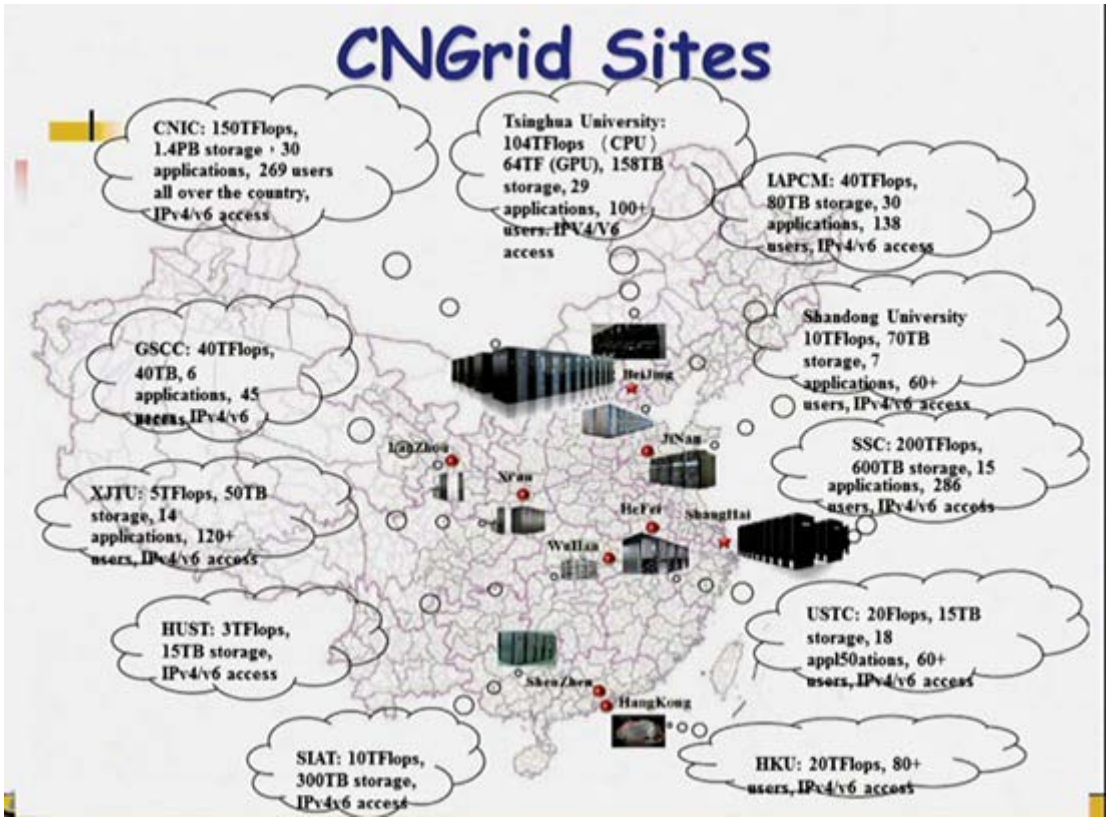

Fig. 5: Distribution of CNGrid sites across China

The 11 sites provide more than 430 TFlops of processing power and 2200 TB of storage, with three PFlop-scale sites to be integrated into CNGrid soon. CNGrid currently supports about 230 services to approximately 1.500 users with the major users either coming from the industry or the scientific computing community at research institutes and universities. Examples of current applications supported by CNGrid include bird flu drug screening, car safety design analysis, tunnel construction simulations, Monte Carlo simulations, computational physics, aircraft design aerodynamics, sand storm weather forecasting, and magnetic hydrodynamics.

Another grid system is the so-called ChinaGrid, which is funded by the Chinese Ministry of Education and based on the China Education and Research Network (CERNET). It is also a pilot grid application supported by CNGrid. The first phase of ChinaGrid entails the establishment of a large-scale high-performance grid environment, which has resulted in the creation of a grid system covering more than 20 key universities throughout the country. The grid encompasses computing capabilities of up to 16 TB and a storage capacity of up to 180 TB. It also already supports the IPv6 network environment and can be deployed on CERNET 2, the IPv6 network in China.

Current applications include the Bioinformatics Grid, the Image Processing Grid, the CFD Grid and the Massive Information Processing Grid. Another important application

is the Course On-Line Grid, which provides free educational VoD services on 22 servers of ChinaGrid that can be accessed by university students across China. The video resources include more than 300 college courses from 20 key universities and many disciplines. The total length is more than 8000 hours and the average daily hits amount to 5,000.

In 2011, the second phase of ChinaGrid was launched. Its goals are to (1) expand the scale of ChinaGrid; (2) build high-performance grid centers; (3) increase the number of campus grids; (4) advance the key technologies of ChinaGrid; (5) enforce the development of the second phase of CGSP; (6) deploy more applications on ChinaGrid; (7) establish ChinaGrid management organizations.

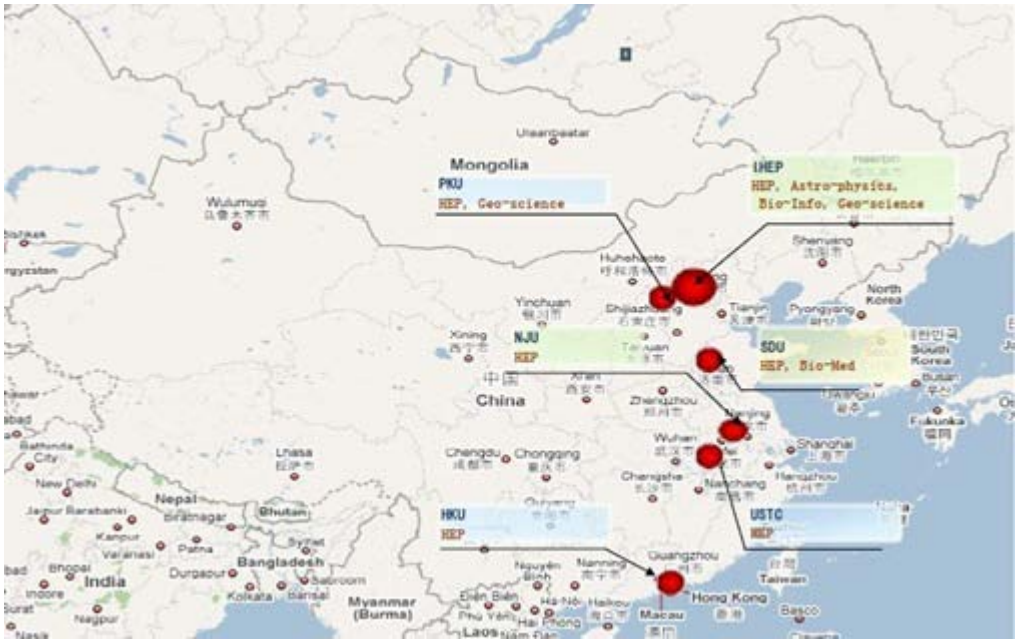

Fig. 6: Map of High Energy Physics (HEP) Grid in China

The High Energy Physics (HEP) Grid in China (cf. Fig. 6) is coordinated by the Institute of High Energy Physics (IHEP) of the Chinese Academy of Sciences (CAS). IHEP is also a WLCG Tier-2 site with the support of CAS. The IHEP grid site currently operates 112 nodes (about 900 CPU cores) and 640 TB of disk storage. It runs gLite middleware, but plans to move on to EMI in the future. International links have been established with ORIENT/TEIN3 in Europe and GLORIAD in North America. In 2011, more than 1 PB of data from the LHC experiment was exchanged.

Over the past few years, IHEP has been able to ensure very stable operations of the site, which is reflected by a high rate of availability and reliability. About a millon jobs were completed for ATLAS in 2011, while the number of completed CMS jobs amount to 570,000.

CNGrid and ChinaGrid are currently financed by the government as part of the twelfth five-year plan. However, funding for the time beyond has not yet been secured. The operation of the WLCG HEP site is going smoothly due to the support from CAS. Plans

to integrate cloud computing, volunteer computing and desktop grids at IHEP are currently being developed to build a larger and comprehensive computing system for high energy physics.

3.India

The National Knowledge Network (NKN) is a pan-India network, providing a unified high speed network backbone for all knowledge related institutions in the country. Major advantages of NKN are its wide geographical coverage (cf. Fig. 7), high scalability, high capacity, and high reliability.

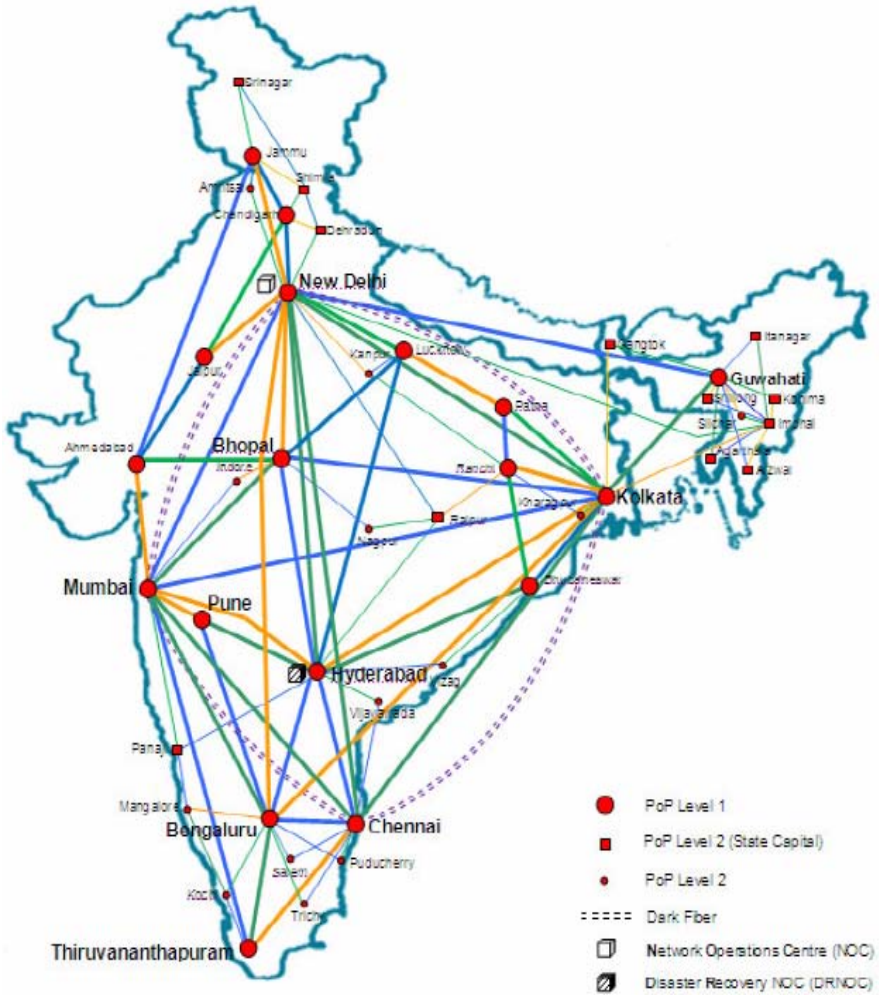

Fig. 7: Map of the National Knowledge Network (NKN)

The Global Access to Resources Using Distributed Architecture (GARUDA) grid is India's national grid computing initiative and brings together academic, scientific and research communities for developing their data and computationally intensive applications in collaboration with 60 institutions. International partner projects include EGI and caBIG. All resources and approximately 600 institutions across India are connected through a 10 Gbps high speed network with plans to increase the participating institutes to 1,500 by 2015.

GARUDA currently uses three 4-teraflop high-performance computing clusters located at Bangalore, Chennai, and Hyderabad with fourteen of the partner institutions also

contributing resources. The total computing power has reached more than 6200 CPUs, which is equivalent to 70 teraflops, whereas the storage space is 220 TB. The supercomputers PARAM Yuva at Pune and PARAM Padma at Bangalore are also available if users require more resources.

GARUDA has also teamed up with a large number of partners to collaborate on research and engineering technologies, architectures, standards, and applications as well as to contribute to the aggregation of GARUDA resources. It currently maintains partnerships with more than 50 research and academic institutions in 17 cities, 8 centers of C-DAC, and is connected to more than 20 additional labs with plans to connect 40 more research laboratories in India by 2015. It currently supports eleven virtual user communities and intends to increase the number to 20 by the end of 2013.

One of the applications currently running on the GARUDA grid infrastructure is the OSDD Chemoinformatics Community, which has about 400 users and was initiated by the Council of Scientific and Industrial Research (CSIR) in India. It operates under the umbrella of the Open Source for Drug Discovery (OSDD) project, which has more than 5300 registered users from more than 130 countries. Bringing together computer science and chemistry to make progress in the field of drug discovery, the Chemoinformatics Community uses GARUDA for data mining and analysis as well as high-throughput virtual screening. Depending on the data intensity of the submitted jobs, users are either connected through NKN, a high-speed network, or the Internet.

The Protein Structure Prediction (PSP) project, an in-house developed, grid-enabled code is an application in bioinformatics. Concurrent PSP jobs are processed by splitting the protein molecule into multiple overlapping parts. This Devide-and-Construct approach is used to achieve a reduction in complexity, allow for the possibility of concurrency, and enable the processing of larger molecules.

Another major application is the Collaborative Class Room (CCR) project, which uses grid resources like storage, distributed computing power to support high quality video transmission and simulations based on virtual laboratories. Working professionals, students, teachers and other stakeholders of educational institutes are expected to benefit from having access to distributed computational resources and storage capacities. This project is of particular significance to India, since academic institutions located in remote areas of the country do not have sufficient access to educational resources. GSRM-based data storage is used for maintaining course repositories, while indexing of course material is based on key words, after which the material can be downloaded and used for lectures throughout India.

Climate modeling is another major research area that has prompted scientists to implement the Seasonal Forecast Model (SFM) on GARUDA to make ensemble forecasts, after several simulations were performed using different models to assess the accuracy of these models. A framework has been developed to make ensemble forecasts using the existing grid middleware services. It uses GRSM for data storage and supports a resolution of 40 km to 40 km.

Another recent project is the simulation of the safety of nuclear reactors aimed at studying the heat extraction from nuclear reactors. This is particularly important, since an increase in heat extraction makes the generation of electric power less expensive. However, the generation of electric power has to meet certain safety standards, which includes the ability to keep the temperature below the specified limit. Due to an increase in available resources, CDAC and ZN were able to reduce the simulation time to mere three days despite an initial estimated time of 24 days. A paper is now being written for publication.

Other areas of collaboration include: (1) the Indian Cancer Grid; (2) protein folding analysis using the caGrid workflow, transport and security technology; (3) caTissue; and (4) the establishment of a biobanking system based on caTissue at the Tata Memorial Center and Hospital in Mumbai.

As for the interoperability with international grids, current priorities include the integration of the technical components of GARUDA and EGI, such as gLite and Globus as well as the customization of the Gridway meta-scheduler so that applications can be run across both infrastructures.

4.Indonesia

e-Science activities in Indonesia began in 2006 when the Institut Teknologi Bandung (ITB) together with University of Indonesia, University of Gajah Mada and Institut Teknologi Sepuluh Nopember developed the Indonesia Higher Education Network (INHERENT), the first National NREN in Indonesia. A few years later, ITB joined the EUAsiaGrid project with the goal of assisting regional integration to the wider grid infrastructure and promoting e-Science applications.

Despite the limited availability of resources in Indonesia, several universities have initiated e-Science activities for several disciplines. ITB currently focuses on areas such as weather forecasting, desktop grid initiatives, and the establishment of a digital library.

The Institut Teknologi Bandung (ITB) is currently home to the only meteorological institute in Indonesia, where weather forecasting remains one of the major research areas in e-Science. In 2011, the resolution of the data input was changed from 1° with an interval of 6 hours to 0.5° with an interval of 3 hours, which translates into a larger download volume with roughly the same latency. Experimental weather forecast extensions use the tropical belt as a prediction domain, predict the distribution of volcanic ash, and assimilate radar data for local high-resolution predictions.

The Weather and Climate Prediction Laboratory (WCPL) at ITB uses computational resources to conduct research on quick forecasting of volcanic ash dispersion up to 24 hours after a volcanic eruption. Computational resources are used to enable rapid processing of available satellite data in near real time, calculate the trajectory of the volcanic ash using the PUFF model and estimate the height of the volcanic plume. These calculations are of particular importance to ensure an accurate and prompt implementation of hazard mitigation measures, such as the determination of the evacuation area in case of a volcanic eruption.

The prediction for the tropical belt domain is another research area of the WCLP. Its objectives are to provide downscaled weather prediction for the whole tropical area with a horizontal grid resolution of 27 km that is suitable for middle-range (up to one week) weather forecasting and to establish a common platform for collaborative research in tropical weather forecasting using the WRF model. There are plans to increase the grid resolution to 9 km using grid resources in the future. However, these simulations generally require a significant amount of computer memory and CPU cores.

Efforts are also being made to develop a prototype system of radar data assimilation for local weather forecasting, which requires large bandwidth for the data transfer in near real time. Local predictions use a grid resolution of 3 km, which requires a large amount of data storage and processing power.

The weather forecast for the Java region is available to the general public at http://weather.meteo.itb.ac.id. The WCPL is planning to extend its weather forecasts

to cover the whole of Indonesia and increase the grid resolution to three or nine kilometers, depending on the availability of computational resources.

Last year, the Desktop Grid Initiative has organized an event on desktop grids with a focus on group discussions, seminars and mini workshops. The objectives of the Desktop Grid Initiative is to foster knowledge on desktop grids, establish a cooperation between ITB and IDGF, and form a grid computing community in Indonesia. Desktop grids are of particular importance to Indonesia due to the lack of high performance computing resources. Participants of the seminar included delegates from BMKG (Meteorological, Climatological and Geophysical Agency), LIPI (Indonesian Institute of Sciences), and the Ministries of Education as well as Communication and Informatics.

The Digital Library is a management system for library resources in a digital format, which includes text, video and audio data. INHERENT is used to build the nationwide Indonesia Digital Library Network with other universities across Indonesia.

ITB will continue its efforts to extend grid technology to other universities in Indonesia, raise awareness of the grid infrastructure and its applications, enable the access to grid resources, and encourage interested individuals to help propel the development of the grid infrastructure and grid applications. Activities that are currently being planned will focus on the development of grid technology, grid system administrators and users as well as desktop grids. Other plans include the increase of the resolution of WCPL experiments to 3 km. As for the Digital Library, preparations to adapt GDL applications to run on grids, which can assist in the indexing process, and storage, are currently underway.

5.Japan

REsources liNKage for E-scIence (RENKEI) is a research and development project on the development of new middleware technologies for e-Science. "RENKEI" means "federation" in Japanese. The goal of the project is to develop middleware to utilize resources, such as files, databases, and applications, in different operational environments, such as local machines and grids with different middleware. The standards for interoperability among multiple grid infrastructures are currently discussed in the Open Grid Forum (OGF), but still leave much room for improvement.

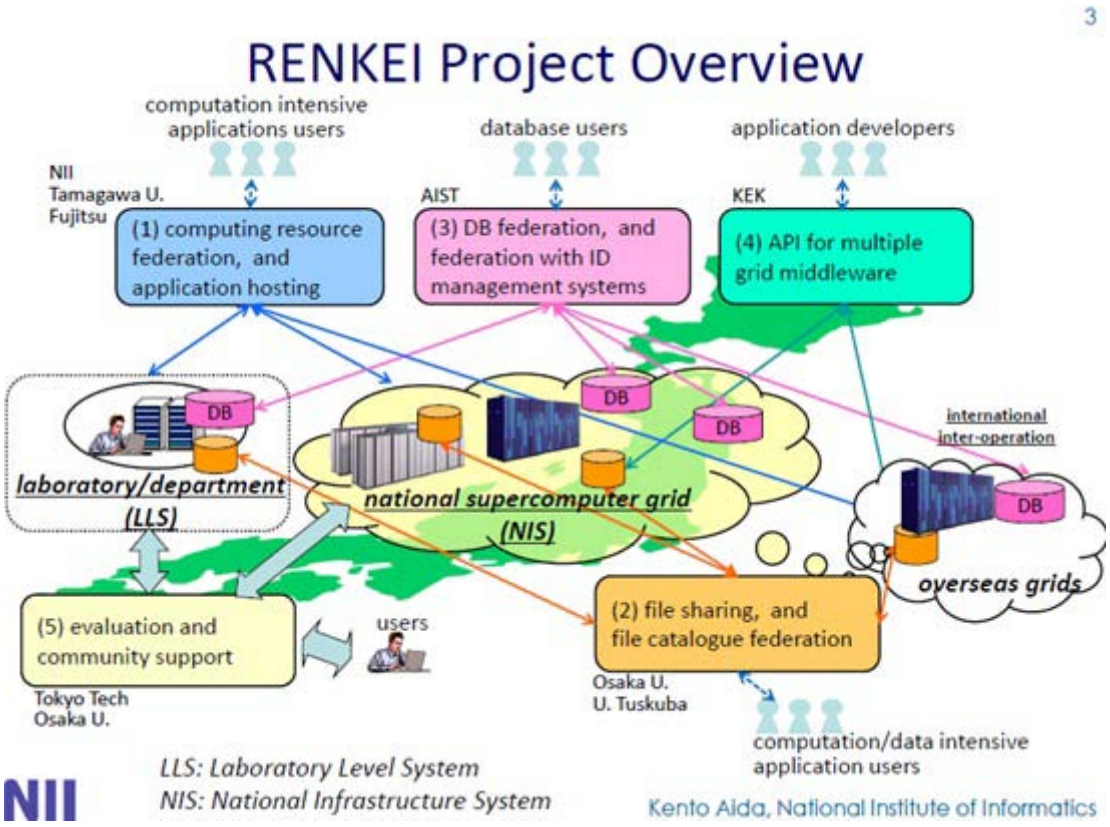

Fig. 8: Overview of the RENKEI project

RENKEI currently has four groups for middleware development (cf. Fig. 8), each of which focuses on different resources. The first group focuses on computing resource federation and application hosting with the goal to develop middleware that allows users to submit jobs to resources using different middleware, whereas the second group specializes in file sharing and file catalogue federation aimed at enabling unified access to files and file catalogue systems through resources using different middleware. The third group focuses on database federation and federation with ID management systems. The fourth one is dedicated to the development of application programming interfaces for different grid middleware. A fifth group is responsible for the evaluation and the community support. It is also in charge of building the testbed for this project.

The testbed is currently operated by a number of research institutions in Japan as well as several partner institutions in Europe and the US (cf. Fig. 9). One problem was that some institutions offered supercomputer resources in production level service, but deploying the research results on such supercomputers has proven a very difficult task. This resulted in the deployment of an appliance service named RENKEI-POP at some sites, making it easy for large quantities of research data in the supercomputers of various universities and research organizations to be utilized at another location. Data can be retrieved and transmitted between nodes. This technology is being developed using software resources like the Globus Toolkit and Gfarm, and a SINET3 layer 3 VPN is used to network the nodes together.

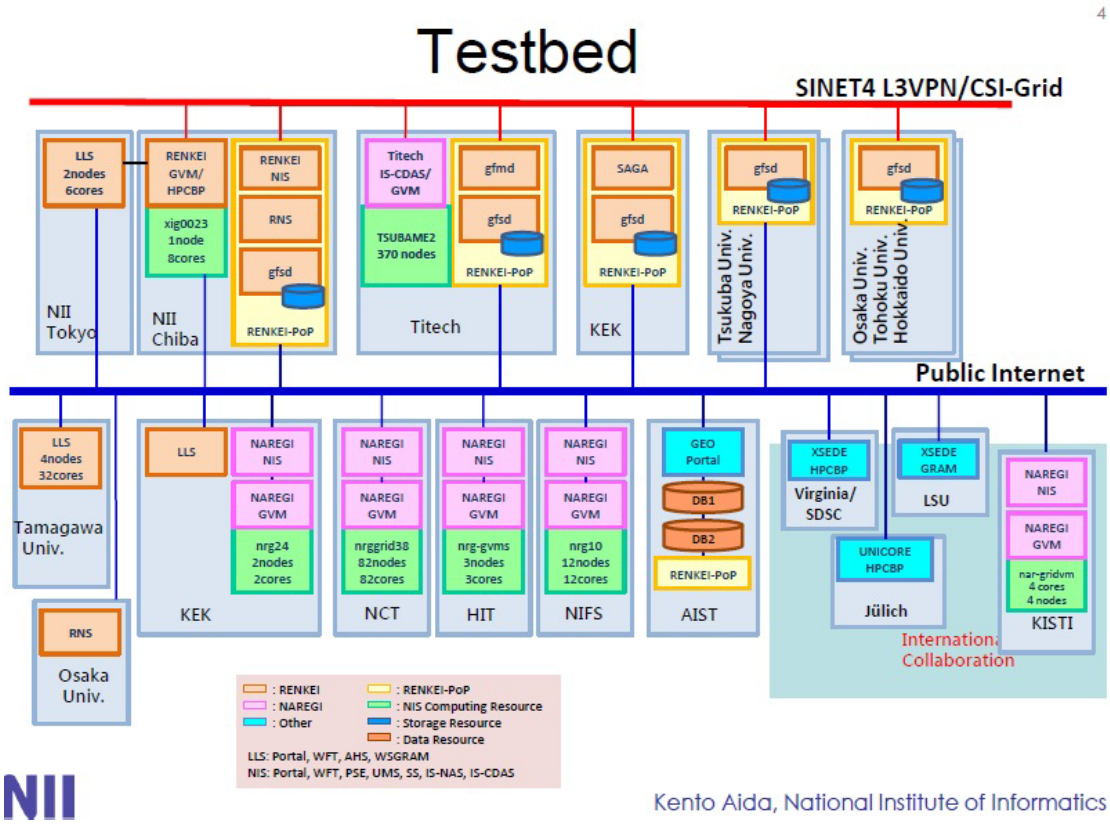

Fig. 9: Composition of the RENKEI testbed

The above-mentioned testbed was used for a number of research purposes, one of which focuses on computing resource federation and file sharing. Two software kits were developed, one of which is a workflow system aimed at enabling job submission in a single work to resources using different middleware. The other one is a job scheduler based on OGF standards, such as HPCBC, which also features interoperability with different middleware. As for file sharing, RENKEI developed a distributed file system named Gfarm, which grants users at remote sites access to the file sharing system.

The federation of file catalogues is another major research area. File catalogues are used to manage the physical location of files, so users can query the file catalogues to determine the location of the requested file. However, since there are numerous different file catalogue systems, users cannot query multiple file catalogues in a single query. The

development of a new file catalogue system, which is based on RNS, an OSF standard, aims at standardizing OSF interfaces and offering users a single interface that can be used to access various catalogues at different locations.

Another focus is database federation. The goal is to allow for a unified, easy access to a large variety of databases with different implementations, such as RDB, XML, Web, RDF, and search engines. RENKEI has been collaborating on a breath gas analysis project with medical scientists at the University of Vienna, Austria to enable access to several databases with varying implementations. This project has succeeded in providing a distributed, heterogeneous database environment to support a collaborative workspace. The only distributed database server is currently located at the National Institute of Advanced Industrial Science and Technology (AIST) in Japan, but there are plans to establish additional sites in Europe.

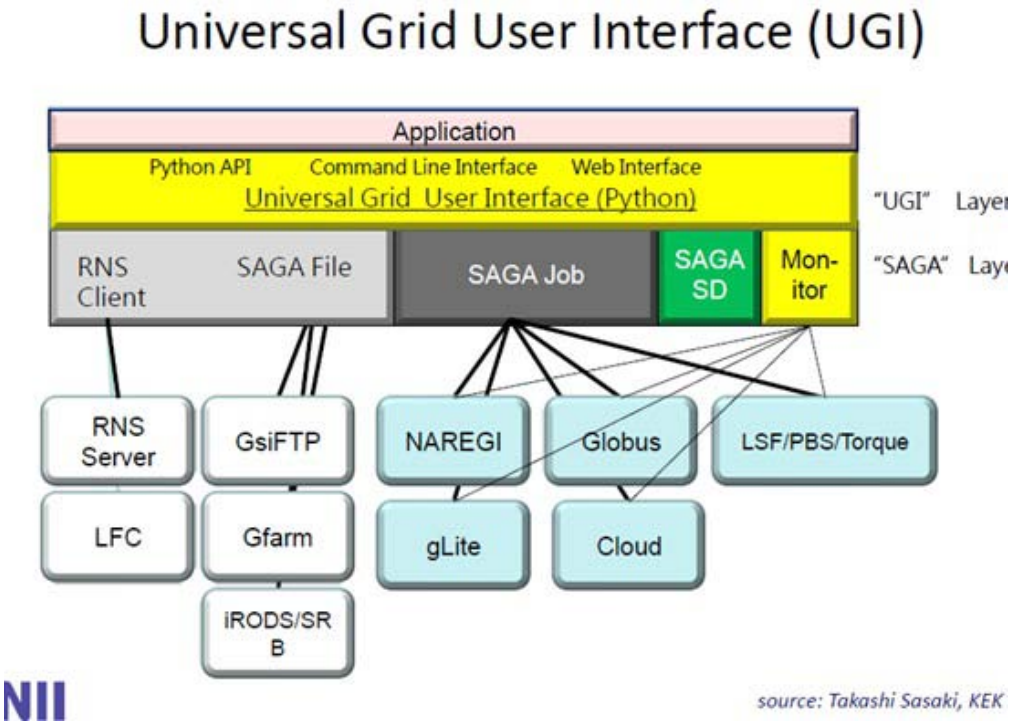

Fig. 10: Universal Grid User Interface (UGI)

RENKEI is also developing an application programming interface (API) for multiple grid infrastructures. This is being developed using the Universal Grid User Interface (UGI), which allows developers to write code in Python, which provides the application interface for the SAGA interface, the OGF standard. In other words, developers can develop application programs than can run on various kinds of grid middleware without additional modifications. Two examples of applications that have been implemented using UGI are the Particle Therapy Simulation project and the Human Cell Model project

RENKEI-POP is the appliance used for middleware deployment and evaluation within the RENKEI project. It offers a VM hosting service, allowing application developers to

deploy their applications on VMs running on the supercomputer sites, in addition to a high-performance shared storage service, based on Gfarm 2 and a 10 Gbps network connection. Authentication is based on GSI. A project is currently running space plasma simulations on a supercomputer at the University of Nagoya, Japan, and uses RENKEI-POP to share and visualize the simulation results after transferring the data. These results can then be accessed from anywhere. In order to share and analyze genome data on the supercomputer TSUBAME 2.0, RENKEI-POP has also been deployed at the Tokyo Institute of Technology and the National Institute of Genetics, Japan.

Some results of the RENKEI project will also be used for the High Performance Computing Infrastructure (HPCI), a national project promoted by Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan. It aims at establishing a distributed computing infrastructure for high-performance computing, which includes the integrated operation of the K computer with other supercomputer centers, enabling seamless access to the K computer from supercomputers and user machines. Large-scale storage systems will be shared among the supercomputers. This is also the first government project that focuses on production-level distributed computing infrastructures with the launch of production-level operations scheduled for November 2012.

6.Malaysia

Launched in March 2005, the Malaysian Research & Education Network (MYREN) is aimed at enabling high-speed dedicated network connectivity to enhance and uplift the education sector while exploring the research capacity and capability of the Malaysian education and research sectors. MYREN offers a dedicated high-speed network for research and education and is connected to other similar RENs. It is not affected by fluctuations in the number of users and their behavior like the Internet, since it does not provide access to commercial services, such as Facebook, Twitter, online games, streaming videos and radios. The current applications cover a wide range of research areas, such as computational science and engineering, e-Culture (Digital Culture and Heritage), and office automation.

In April 2010, its successor MYREN 2, which is characterized by its advanced network architecture and greater autonomy, had a soft launch. MYREN also offers a range of cloud services under the umbrella of MYRENCLOUD, a private community cloud that relies on the availability of MYREN 2 as a dedicated private network, single entity policy and governance to subscribe and the long-term availability of the infrastructure.

During ISGC 2012, the Malaysian Identity Federation and Access Management (MyIFAM) was inducted as a member to the Asia Pacific Grid Policy Management Authority (APGridPMA) and International Grid Trust Federation and became the Certification Authority (CA) for Malaysia. As of March 2012, Malaysia has seen an increase of certified grid sites from two to six sites, with plans to raise the number to 20 sites, which equals the current number of universities in Malaysia. This will become part of the distributed computing infrastructure in Malaysia.

Proposed activities for MYREN 2 include the On Demand Infrastructure for Research and Education Collaborative Infrastructure (OnDIRECT), which will further enhance international collaboration and will offer video collaboration services for virtual meetings between individuals or groups at 158 concurrent sites, VideoBridge, the uCAST video sharing portal and on-demand rich media for learning. Important goals of MYRENCLOUD will be to ensure secure data services, including a data fabric system for storing research data, a data transfer service, a database hosting service for collaborative research efforts, and the implementation of 256-bit encryption. Other objectives include the provision of virtual servers, identity and security services, and web collaboration services. Preparations to establish OnVM and OnDisk are currently underway, which includes the installation of 1,000 virtual servers with dual Quadcore processors and 8 GB RAM of memory and will offer 50 GB of disk storage to 10,000 researchers in the first phase respectively. MYREN 2 will also facilitate access to more research and education entities in Europe, the U.S., China, Japan, Korea and the Asia-Pacific region.

Launched in 2008, the Biruni Grid Center is a project led by Associate Professor Dr. Suhaimi Napis, Director of the InfoComm Development Centre (iDEC) at the University Putra Malaysia. Its work has focused on the development of the technical

infrastructure that serves as the grid foundation for scientific research. Deployed applications concentrate on bioinformatics, cheminformatics and medical informatics. Research topics such as next-generation sequencing and genome comparison illustrate the transition from computationally intensive to RAM intensive analysis, as a single run produces approximately 7.5 GB of raw data and 200 GB of sequence data, which makes the use of MYREN necessary to move the data. Other applications cover graphic/image screen rendering and CFD using Open-FOAM.

MYREN 2 will continue to promote its educational purpose in its regional outreach efforts to countries such as Laos, Cambodia and Myanmar. A further extension of MYREN 2 to all ministries, hospitals, libraries and community centers is also intended.

7.Mongolia

The ICT development of Mongolia has advanced rapidly over the last five years. The telecommunications network of over thirteen thousand kilometers of fiber optic cable covers all of Ulaanbaatar, the capital city of Mongolia, and connects all *aimag* (or province) centers and some 151 *sum* (or county) centers . The number of mobile phone subscribers has reached over 2.3 millions. Internet is widely used in all sectors of economic and social development of Mongolia.

Currently, the three major operators, Railcom, Gemnet, and Mobicom, provide wholesale Internet connection services in Mongolia. The total connection bandwidth is 12 Gbps.

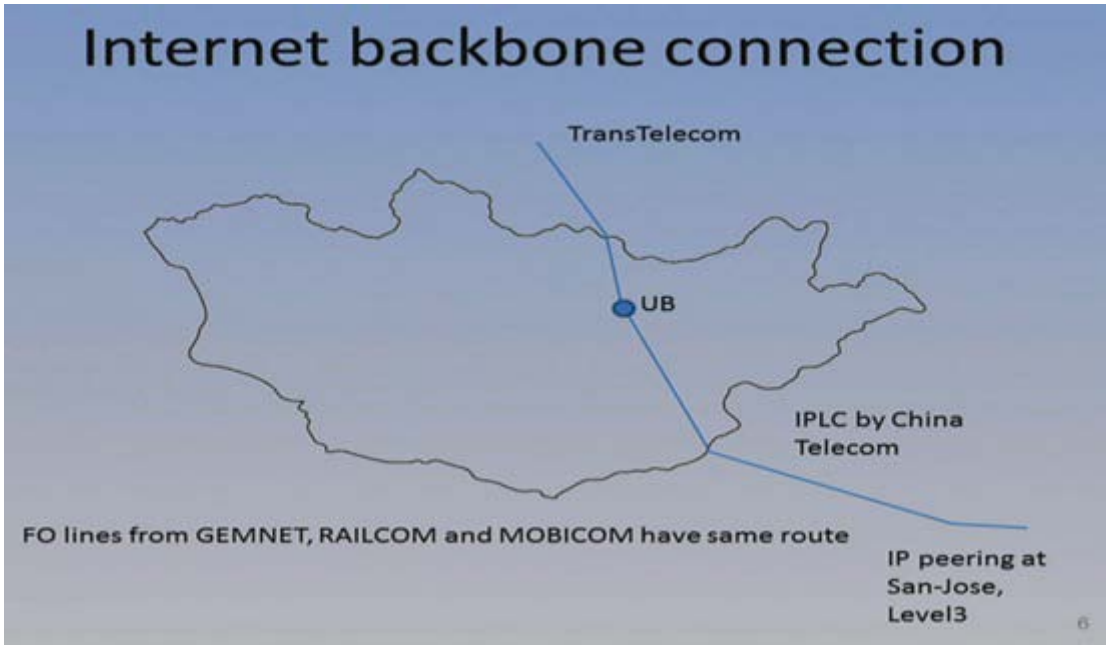

Fig. 11: Map of the current Internet backbone connection

In Mongolia, research activities are being carried out at different universities and various institutes of the Mongolian Academy of Sciences (MAS). The MAS consists of 22 research institutes and centers, which are located at six different locations within a distance of up to 17 km.

Unlike different universities in Mongolia, the institutes and centres of the Mongolian Academy of Sciences have collected a multitude of information on history, archaeology, paleontology, geography, botany, biology, natural resources, language and literature of Mongolia, which are the result of the long-term research activities and field investigation. In order to grant public access to this information, it is vitally important to create an integrated computer database and disseminate the information via the Internet and electronic publishing.

Since 1996, the Institute of Informatics has been providing the Mongolian Academy of

Sciences with electronic mailing and Internet service. Over 600 computers are connected to the current network (speed of 15 Mbps, Fiber backbone).
In recent years, research of nano-technology, biotechnology and producing natural medicine is developing rapidly in Mongolian universities and science institutes. Mongolian physicists have been successfully participating in LHC research.

Mongolia has only recently begun to promote e-Science. However, the use of cloud computing has become an urgent necessity, as Mongolia is located in an active seismic zone of Central Asia. In recent years, research in nanotechnology, biotechnology, the production of natural medicine has been developing rapidly at Mongolian universities and research institutions. Two major impediments are Mongolia's economic situation and a shortage of well-trained professionals, which makes it necessary for Mongolian scientists to participate in international research projects.

A consortium named "Mongolian Grid" was founded under the guidance of Mongolia's Nuclear Energy Agency, which consists of the following organizations: (1) the Institute of Informatics, MAS; (2) the School of Mathematics and Computer Science at the National University of Mongolia; (3) the Institute of Telecommunications and Information Technology at the Mongolian University of Science and Technology.

The 11th Science Council of Asia (SCA) Conference was held on 4-6 July 2011 in Mongolia. In 2011, the Mongolian Academy of Sciences presided over the Science Council of Asia. At the request of the Mongolian Academy of Sciences, Dr. Simon Lin, Director of the Academia Sinica Grid Computing Centre (ASGC), organized "Developing e-Science in Asia", one section of the conference. During the conference, the Memorandum of Understanding for Research Cooperation between the e-Science Working Group of Mongolian Academy of Sciences and the Academia Sinica Grid Computing Centre was signed. The memorandum started showing its first results very shortly after its signing. For instance, a workshop on e-Science applications and technology was organized on 3-5 October 2011 on Dr. Lin's initiative. The workshop attracted the attention of MAS institutes, various universities and scholars as well as ISP operator companies, with a total of more than 90 participants. Three delegates from the Mongolian Academy of Sciences attended the TELDAP International Conference 2012 and ISGC 2012 in Taipei.

In 2011, the Institute of Informatics became a member of the Pacific Neighborhood Consortium (PNC). Two delegates from the Mongolian Academy of Sciences attended the Annual Conference and Joint Meetings of the Pacific Neighborhood Consortium, which were held in Bangkok, Thailand, on 19-21 October 2011.
As a member of the International Council of Science (ICSU), the Mongolian Academy of Sciences also sent delegates to the conference "Data-Intensive Science and Discovery – CODATA 45 Years On" held on 30 October 2011 in Beijing, China and expressed its willingness to become a member of this organization.

The Mongolian side successfully organized the "Workshop on Open Knowledge and Data Environment for Innovative Research, Education and Society (OKDE) in Asia" for the International Council of Science on 4-6 July 2011 in Ulaanbaatar.

The Mongolian Academy of Sciences also collaborates with the Information Technology Laboratory of the Joint Institute for Nuclear Research in Dubna, Russia, on the establishment of an educational grid system. Two teachers from the National University of Mongolia and Technology are being trained on-site as system administers for the educational grid, which is based on gLite 3.2, while six students are scheduled to participate in the 2012 summer school.

A three-year joint project on the establishment of an integrated information and technology network of Mongolia's research institutions is still ongoing in cooperation with the Irkutsk Centre of the Siberian Branch of the Russian Academy of Sciences. The project is aimed at the creation of an integrated information and technology infrastructure to support interdisciplinary research at the institutes and centers of the Mongolian Academy of Sciences as well as to enhance openness and access to scientific information of the Mongolian Academy of Sciences for the global community. The Irkutsk Centre and the MAS bear the costs of the project using a ratio of 70:30. There are also plans to lease a high-speed communicational channel to link the Irkutsk Centre with the Institute of Informatics later this year. The Irkutsk Centre will bear the expenses of the lease.

At present, metadata is being added to the information resources of the Mongolian Academy of Sciences using international standards for metadata, such as Dublin Core, FGDC-STD-001-1998, and ISO-19115/19119/19110, while extensive testing and configuration of the comprehensive catalogue are also currently being carried out.

8.The Philippines

Established in 1987, the Advanced Science and Technology Institute (ASTI) serves as the research and development institute on ICT and microelectronics of the Department of Science and Technology in the Philippines. It operates and manages the National Research Network and National Grid Infrastructure.

From January 2008 to June 2011, the Philippine e-Science Grid (PSciGrid) was funded by the Department of Science and Technology. As of March 2012, preparations to submit a proposal for a second phase are underway. The program had two objectives: (1) to establish a national e-Science infrastructure in the Philippines, enabling collaborative research among local educational and research institutions; (2) to provide seamless access to high-performance computing resources and applications to life science and Earth sciences.

The reasons for selecting these focus areas is mainly based on the geographical location of the Philippines. The Southeast Asian country consists of about 7,000 islands, which are aligned along the so-called Pacific Ring of Fire, a horseshoe-shaped area characterized by its large number of earthquakes and volcanic eruptions. The area is also frequently struck by typhoons (cf. Fig. 12). Life sciences are a major research area due to the Philippines' status as a mega biodiversity hotspot. The Philippines are home to more than 52,000 species with more than half found nowhere else in the world. Many of these species are endangered and threatened by environmental destruction. To help conserve the biodiversity, the government established two leading research institutions, the National Biotech Institute in 1979 and the Philippine Genome Center in 2009, both of which collaborate with ASTI.

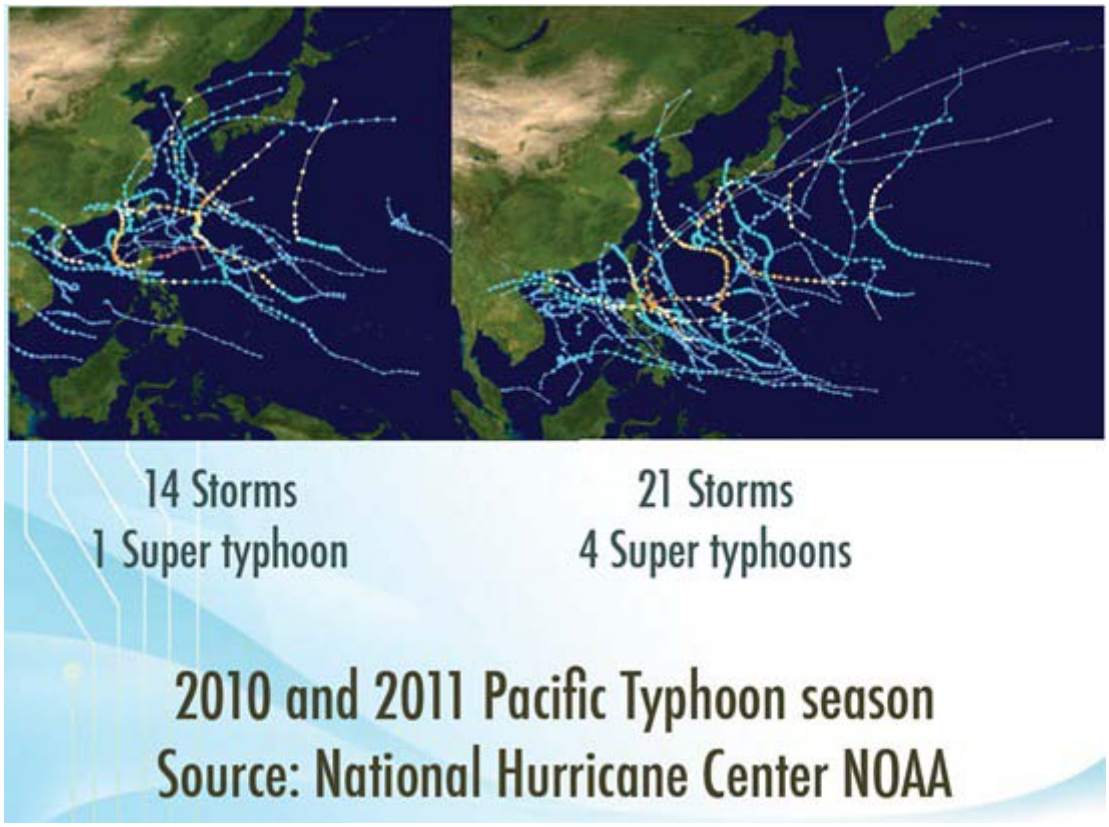

Fig. 12: Map of the 2010 and 2011 Pacific typhoon season

In 2010, El Niño, a warm oceanic phase, resulted in an increase of the surface temperature of the tropical eastern Pacific Ocean and high air surface pressure in the western Pacific, which was also the caused a record low in the number of typhoons that year. El Niño was followed by a corresponding cold phase, La Niña, in 2011, which triggered a decrease of the surface temperature of the eastern Pacific and low air surface pressure in the western Pacific. The extremes of this climate pattern's oscillations cause extreme weather, such as floods and droughts, in many regions of the world, among which developing countries dependent on agriculture and fishing are the most affected. This usually results in a substantial number of fatalities and several billion dollars in damages.

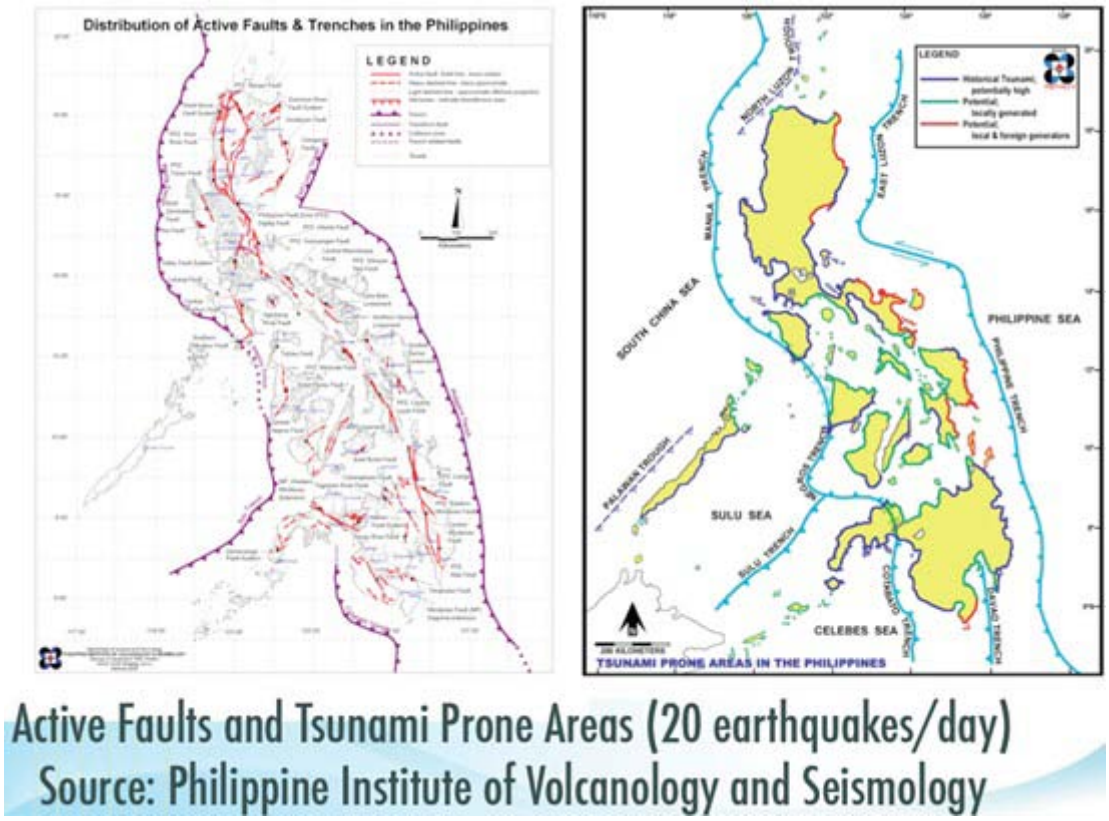

Fig. 13: Active faults and tsunami-prone areas

The Philippines are located between two tectonic plates, the Eurasian and the Philippine Plates (cf. Fig. 13). This causes a large number of earthquakes across the country. On average, seismologists detect 20 earthquakes per day, while a magnitude 5.0 event occurs once per month. This makes the area very susceptible to tsunamis from the Pacific Ocean and the South China Sea.

The map below (cf. Fig. 14) illustrates the active volcanoes in Southeast Asia. The count of active volcanoes located on the Philippines tends to vary between different institutions depending on their definition, but is generally estimated to range between 21 to 25. However, only 5 active volcanoes are currently being monitored due to resource constraints.

ASTI has developed three tsunami warning stations, 80 automated weather stations, and field monitoring systems that can be deployed flexibly.

Fig. 14: Map of active volcanoes in Southeast Asia

There are two projects under the PSciGrid program: (1) boosting Grid computing using reconfigurable hardware technology; and (2) boosting social and technological capabilities for bioinformatics research. The PSciGrid network consists of three sites, ASTI as the central site, the Ateneo de Manila University, and the University of the Philippines. The physical link for the network, which also connects the sites to the international research community and other partner sites in the Philippines, is provided by the Philippine Research, Education and Government Information Network (PREGINET).

ASTI's high performance computing (HPC) facility consists of 51 nodes, 408 cores, 51 computing nodes with 300GB/500 GB of disk space and 16GB/24GB of RAM per node in computing capacity and 30 TB in storage capacity. There used to be 6 computing clusters using gLite middleware and ROCKS 5.2.2 for the operation system: the bioinformatics cluster (8 nodes), the meteorology cluster (6 nodes), the cluster sandbox (2 nodes), the EUAsiaGrid and EGEE collaboration cluster (EGEE certified production cluster) (7 nodes), virtual cluster 1 (16 nodes), and virtual cluster 2 (13 nodes). The installed applications include Bioinformatics (BioRoll, Progeniq BioBoost, AutoDock), Seismology (SPECFEM3D), Meteorology (WRF , RegCM, MM5), and Oceanography (SeaDAS). Since last year, these clusters have been integrated into ASTI's private cloud that is based on Open Nebula 3 to merge the physical infrastructures into one. As of March 2012, 17 nodes are already in operation, while the integration of the rest of the clusters is scheduled to be completed later this year.

As for middleware and the Cluster operating system, the PSciGrid currently uses Rocks 5.2.2 and gLite. Installed applications include Bioinformatics (BioRoll, Progeniq BioBoost, AutoDock), Seismology (SPECFEM3D), Meteorology (WRF , RegCM, MM5), and Oceanography (SeaDAS). To facilitate access to the resources, a job submission portal for PSciGrid was developed and made available on the Internet.

Several trainings were conducted, including the SPECFEM 3D in August 2010, which is a workshop in collaboration with ASGC for senior programmers on in-depth usage of SPECFEM 3D application for seismic wave propagation simulation. Other activities include the Philippine Grid Computing Forum in 2008, 2009 and 2011, training on UPLB Biotech Data Warehouse Portal in February 2011, and an one-day talk in the Philippine Genome Center.

PSciGrid is a contributing member of EUAsiaGrid and an institutional member of the Pacific Rim and Grid Middleware Assembly (PRAGMA). It also has international ties with PANDA Grid, the Asia-Pacific Advanced Network (APAN), and the Trans-Eurasia Information Network 3 (TEIN3).

Current users include: UP Marine Science Institute, UPLB Biotech, International Rice Research Institute, PAGASA (Weather Bureau), UP National Institute of Physics, UP Computer Science Department, PHIVOLCS (Volcanology and Seismology), Manila Observatory, Philippine Genome Center, and the Energy Development Corporation.

There are plans to develop other FPGA applications, expand visualization as well as computational and data capacity, which is currently impeded by a limitation of physical space. This year will also the integration of PSciGrid with the Philippine Government Data Center and the program proposal for phase 2. Meanwhile, PSciGrid will continue to support, advocate and promote grid technology to local universities and research communities.

9.South Korea

Since signing the WLCG MoU in 2007, the Korea Institute of Science and Technology Information (KISTI) has served as an official Tier2 site for ALICE experiment, offering production-level grid services to the ALICE Grid. It is funded by the Ministry of Education, Science And Technology. In 2011, KISTI was able to maintain a high level of reliability and availability (cf. Fig. 15) with two minor exceptions in August and October due to a monitoring problem and site reorganization, respectively. As of March 2012, KISTI has been processing nearly 35,000 jobs per month on average for the past six months, which equals a 0.62% contribution to ALICE computing in terms of total job execution.

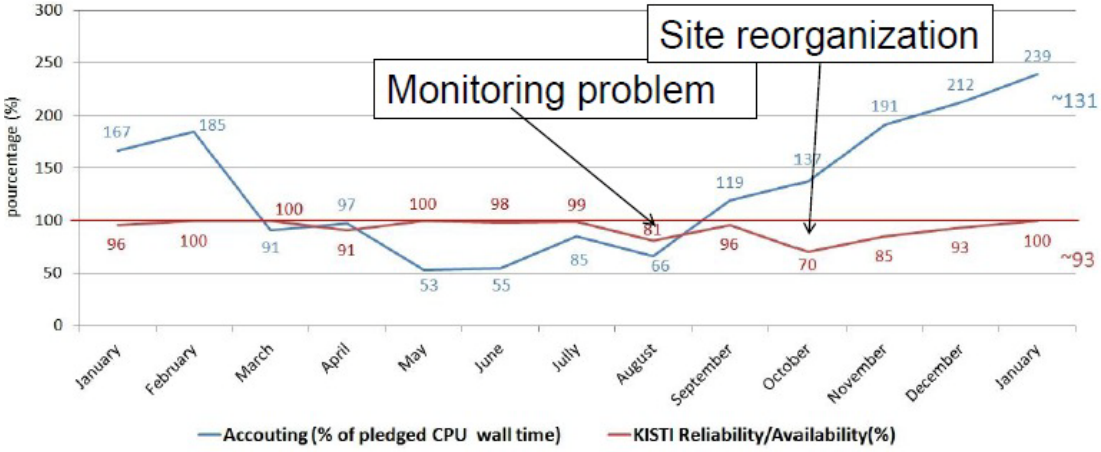

Fig. 15: Accountability and reliability of KISTI CREAM

The France-Asia Virtual Organization (VO) was successfully deployed in October 2011 and is now fully functional at three sites, namely KISTI, CC-IN2P3 from France and KEK from Japan (cf. Fig 16), providing a total of 25,000 CPU cores and 8 TB of disk storage. The migration of FKPPL VO to the France-Asia VO was preceded by the joining of KEK in Japan. The goal of the newly deployed France-Asia VO is to share computing resources to make them accessible to multidisciplinary scientific projects and scientists of other Asian countries. As of March 2012, 106 users have joined France-Asia VO as members. IHEP-Beijing and Vietnam have expressed their interest in having access to it for the deployment of applications in astronomy and drug discovery, respectively. IHEP-Beijing is currently also considering the establishment of a grid site at IHEP to contribute to the France-Asia VO.

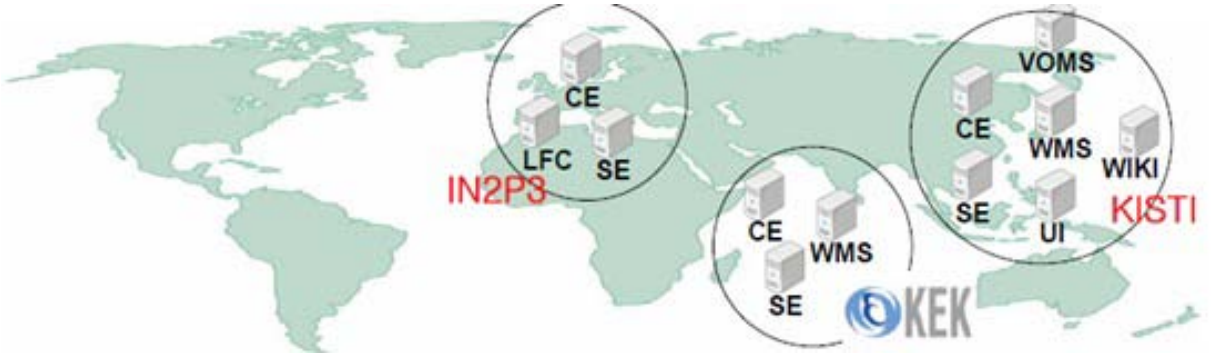

Fig. 16: Map of current France-Asia VO sites

Applications currently running on the France-Asia VO include: (1) in-silico drug discovery in collaboration with Prof. Doman Kim from Chonnam National University, which entails a large-scale deployment of docking simulations, with hundreds of thousands of potential chemical compounds against viruses; (2) Geant4 applications for medical physics in collaboration with Dr. Se-Byeong Lee from the National Cancer Center in Korea aimed at performing computationally intensive simulations relevant to cancer treatment planning; and (3) two-color Quantum ChromoDynamics (QCD) simulations in theoretical physics in collaboration with Prof. Seyong Kim from Sejong University, which includes several hundreds or thousands of QCD jobs to be run on the grid, with each job taking about 10 days. In 2011, two papers were published on in-silico drug discovery in the Bioorganic & Medicinal Chemistry Letters (BMCL) and another one on the QCD simulations in Computer Physics Communications (CPC).

A workshop named “2011 Geant4, GATE and Grid Tutorial for Medical Applications” was held in Seoul in October with 45 participants from 15 institutes in Korea and 7 lecturers from France, the US, Japan and Korea.

The Korea-Japan NAREGI Federation has been established recently, after the KISTI NAREGI site was set up in January 2012. RENKEI is a testbed for e-Science Infrastructure based on NAREGI middleware to ensure seamless usage between local and remote grid and coud resources. This collaborative effort is also aimed at allowing Korean scientists access to KEK resources and the Tsubame 3 supercomputer.

Launched in 2007 and funded internally by KISTI, the Partnership and Leadership for Nationwide Supercomputing Infrastructure (PLSI) is a consortium of 14 high-performance computing centers in Korea. Its goal is to establish a distributed HPC computing environment for world-class computational science research. The target is to reach 400 Tflops by linking 140 HPC centers across Korea. Despite some recent cuts in its funding, PLSI has already established around 100 Tflops of computing capacity by combining 18 computing resources at 10 partner sites through dedicated high-performance networks.

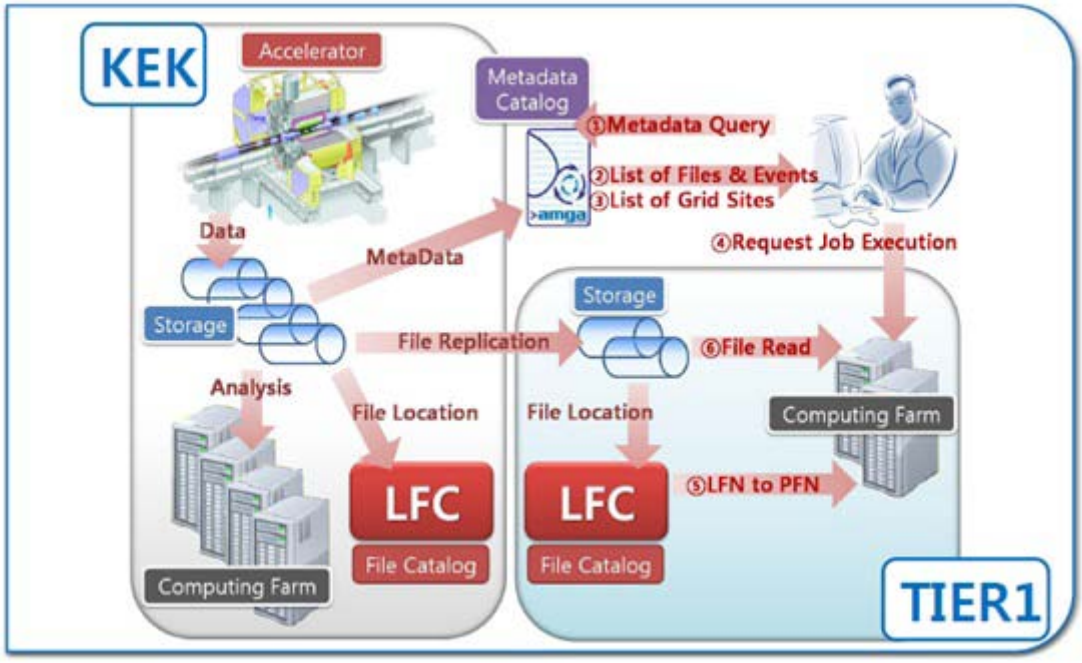

Fig. 17: The use of AMGA as Belle II metadata service

AMGA is an EMI metadata catalogue software that has been developed at KISTI. As such, AMGA provides access to metadata for files distributed on the grid. AMGA version 2.1.2 was released as part of the EMI-1 distribution in April 2011, with version 2.3 scheduled to be released as part of EMI-2 in April 2012, which features publishing service information in the BDII, enabled Nagios monitoring, and support for SL6 and Debian. To promote AMGA to a wider audience, a training session was held in Lyon, France, during the EGI Technical Forum last year. AMGA has been adapted for use in the BELLE II experiment with KISTI being a driving force behind these adaptions. KISTI is leading the Data Handling Working Group as part of the Belle II computing activities and is also in charge of data registration and the design of the metadata schema for Belle II. Preparations to set up and operate the Belle II metadata service at KEK are currently underway.

10.Taiwan

The Academia Sinica Grid Computing Center (ASGC) has been the WLCG Tier-1 Center since 2005. As such, it has been conducting and fostering e-Science collaborations as well as the development and operations of the infrastructure in the Asia-Pacific region. ASCG has also been participating in the global grid-based distributed computing infrastructure and e-Science collaborations to advance human knowledge. By joining WCG and working closely with the scientific communities, ASGC explores topics of human sustainable development through e-Science. Another objective is the development of a new generation of the distributed computing infrastructure and related technologies.

As part of its overall strategy, ASGC focuses on e-Science development with an emphasis on life science, earth sciences, environmental change, social sciences and high energy physics in an effort to support advanced research and international collaborations conducted through the distributed computing infrastructure. The e-Infrastructure development, another important core area, is designed to maximize availability, performance, and operation automation while minimizing operation costs, which includes integrating new technology to allow the e-Infrastructure to keep evolving, such as merging distributed clouds with grids. This concept also incorporates the support of flexible computing models by integrating service grid, desktop grids, and high-performance computing as well as ensuring intelligent operations through intelligent monitoring. Dissemination and outreach also forms an integral part of the ASGC's overall strategy for a broader promotion and implementation of e-Science concepts. Collaborations with regional and international projects, such as WLCG, EGI, EMI, DEGISCO, CHAIN, and WeNMR, are fundamental drivers for the continuous development of e-Science.

As of January 2012, ASGC dedicates 4,688 CPU cores, 4,400 TB of disk storage, and 4,000 TB of tape libraries to the worldwide grid infrastructure, including WLCG, EUAsiaGrid, EGI, and e-Science applications. For WLCG alone, this translates into the handling of 33,300 jobs per day or about 1 million jobs per month on average. High-performance computing, which includes applications in earth science, environmental change and HPC, takes up another 5,000 CPU cores and 120 TB of disk storage. 1,984 CPU cores and 1,110 TB of disk storage are allocated as a separated pool of resources for flexible on-demand use. All resources are integrated and managed by the grid with cloud technology being integrated into the worldwide grid to provide flexible on-demand resources and services across sites. Last year, ASGC upgraded the networks within its data center to 10 Gb Ethernet connections, while some clusters are connected through InfiniBand, DDR and QDR, so the seamless integration between the 10 Gb Ethernet connection and the other links has become a top priority.

Within the last two years, ASGC has transmitted more than 13 PB of inbound and outbound traffic for WLCG. Due to the high reliability and availability of networks in Europe, the U.S. and the Asia-Pacific region, ASGC has been able to reach the maximum international traffic speed of 11 Gbps (cf. Fig. 18).

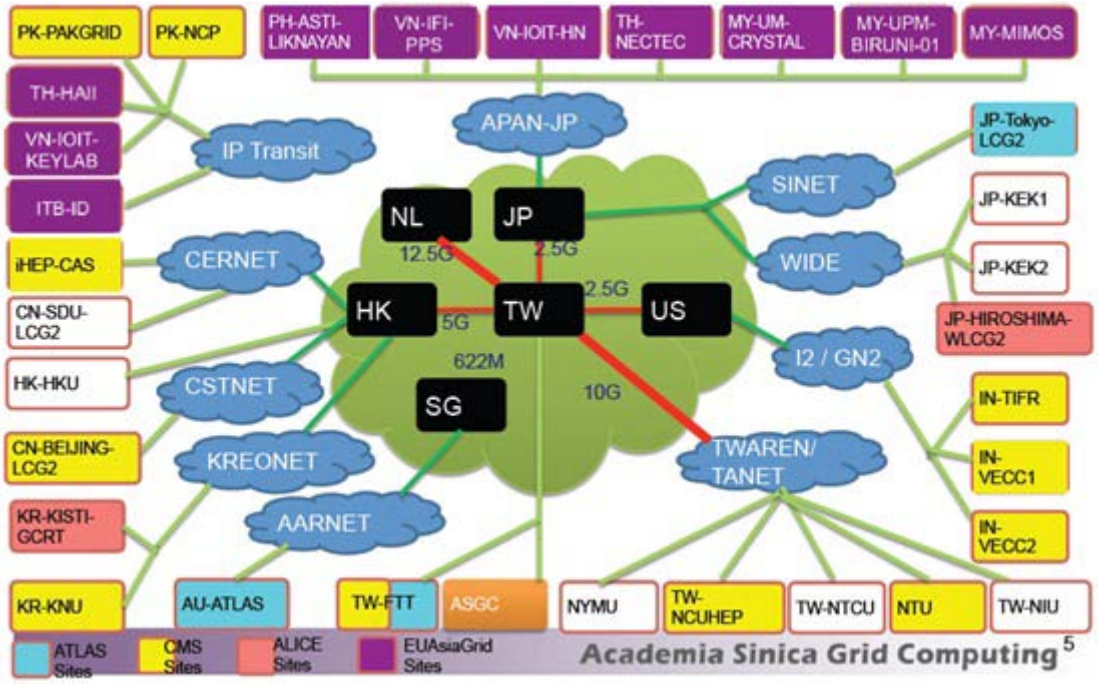

Fig. 18: e-Science networking in the Asia-Pacific region

System optimization has been an ongoing priority for ASGC to maximize performance, reduce costs, improve its energy saving properties, and enhance early warning and automation systems. In terms of the storage system and data management, ASGC deployed a higher density disk array with a large bandwidth right after the installation of the new 10 GbE network, tuned the Castor and DPM performance, merged the ATLAS storage class with DPM, as well as evaluates and deploys distributed file systems. The optimization of the computing system, particularly between various work nodes as well as between work nodes and storage servers, is still ongoing. In early February 2012, the CMS access to ASGC's Castor systems reached a new peak speed of 4 GB/s on five storage servers after their reconfiguration, while access to the ATLAS tape libraries reached a speed of 1.6 GB/s.

ASGC has also focused on ensuring intelligent operations through monitoring. The purpose is to gather information on the system, the services, and the applications to identify the threshold range of resource efficiency and potential limitations of the respective system. It also enables the establishment of early warning mechanisms to prevent the potential failure of critical services. This includes the monitoring of the data center regarding power consumption to save energy, and the tracking of the storage system to ensure optimal performance, data transmission and networking, job execution and other grid services.

In terms of data capacity and data throughput, ASGC's data center is now the largest in Taiwan. ASGC also has a backup power system and a power generator at its disposal. However, these Uninterruptible Power Supply (UPS) systems only protects the most

critical services, such as networking, the databases, and data services from potential data loss.

The deployment of new technologies aims to improve the e-Science infrastructure and its services. As part of its efforts to promote distributed cloud services, ASGC has developed a service-oriented science environment by integrating cloud and virtualization technologies into the grid-based distributed computing infrastructure, thus enabling the automation of data analysis, management and research life cycle and allowing for resource integration with other clouds. ASGC's target user communities in life science, earth sciences, environmental change research, social sciences and high-energy physics, generally make use of the distributed computing infrastructure due to their need for complex modeling and simulations, large-scale data analysis, flexible on-demand resources, and resource federation. Another important task is the performance evaluation aimed at identifying bottlenecks and right metrics for software and tools, such as Autodock 3, Blast and MPI applications. Future plans include the optimization of the use of information systems, the optimization of the file system, the integration of the pilot job, data management, VM & appliance repository, trust framework, etc.

The e-Science applications that are currently in development at ASGC are aimed at providing reliable services for daily usage within research communities and ensuring long-term sustainability.

In life science, ASGC has been working together with several partners in the Asia-Pacific region on cheminformatics for drug discovery in cases such as the avian flu and Dengue fever for several years. It is also part of a joint project with a local research group in large-scale genome sequencing using Blast and HMMR. About half of the genome annotation pipeline has already been ported to the EUAsiaGrid. Large-scale implementation of BLAST and HMMR has increased the performance by sixty times, cutting down the usual running time from one month at a 20 node computer farm to 0.5 days with 100 workers on the grid.

In the process of drug discovery, molecular docking, or virtual screening, is costly and time consuming. ASGC developed the Grid-Enabled Virtual Screening Service (GVSS) portal (http://gvss2.twgrid.org/) using Autodock as the docking engine, allowing users to easily, swiftly, and accurately simulate the ligand-protein docking. GVSS portal provides services from the pre-process of preparing ligand and protein structure parameter file to the docking simulation and the post-process of result analysis (cf. Fig. 19). All that users need is a certificate authentication and a browser to access the GVSS portal. As for computing resources, ASGC team has recently incorporated the Cloud resources on top of the existing Service Grid and Desktop Grid resources in the GVSS portal. Users can choose which resource to use on a docking simulation.

On the GVSS portal site, users can apply protein PDBQS/PDBQT structure file and then set up the simulation parameter step by step on the GVSS portal site. Currently there are about 700 proteins from the Protein Data Bank (PDB) in its system, presented in 3D. As for the ligand preparation, users can upload the small molecule or build one in the compound library. There is no need to remember the chemical formula—the

structure of compounds is searchable in the system; there are 10 million compounds in our database for users to search from. The docking simulation results are downloadable, and users can see online visualization of the conformation of each result. The analytical tools include the histogram showing the docking energy distribution of all docking results, the ranking table, and the 2D and 3D principle component analysis (PCA) (cf. Fig. 20).

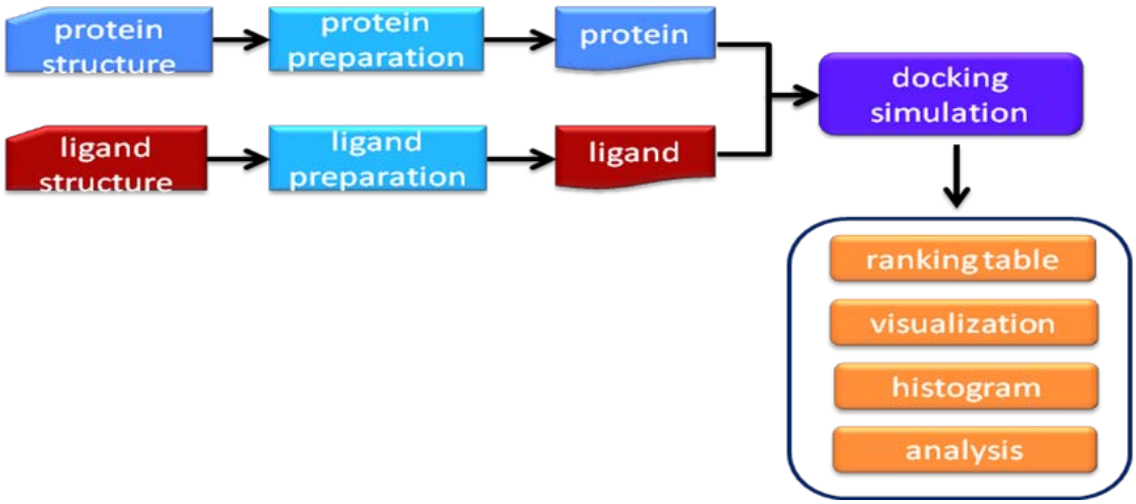

Fig. 19 GVSS Workflow

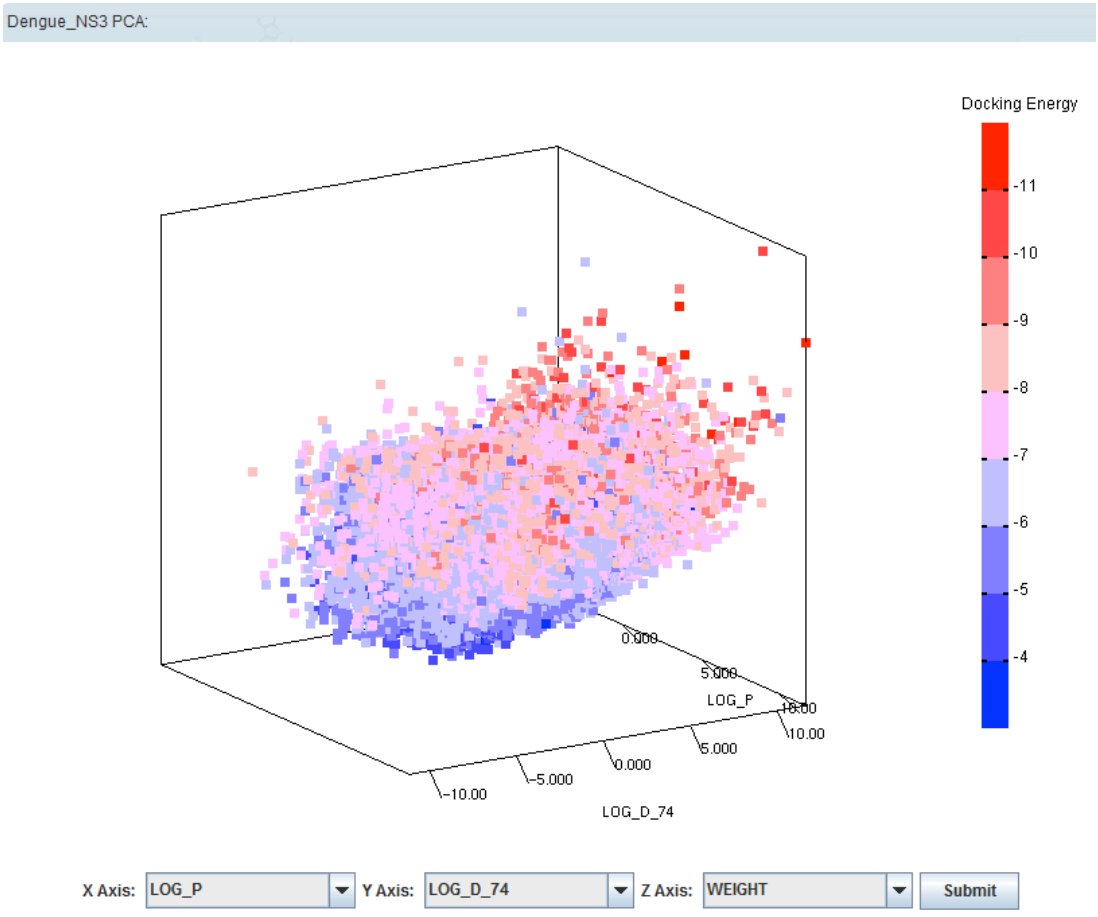

Fig. 20 GVSS PCA in 3D

In December 2011, ASGC, along with Academia Sinica's Institutes of Biological Chemistry and Physics, joined WeNMR, a worldwide e-Infrastructure project for NMR and structural biology. A detailed program has been defined for contributing to all activities: (1) Networking, dissemination and outreach, which includes translating documentation and publication abstracts into Chinese and covering major NMR, biophysics and structural biology events in the Asia-Pacific region; (2) Service activities with the aim to allocate and operate ASGC resources for WeNMR VO, provide dedicated queues for MD simulations with multi-threaded GROMACS as well as offer training material and user support in Chinese; (3) joint research activities aimed at developing protocols for joint usage of SAXS and NMR with docking and enhancing the working capacity of the SAXS beamline at the National Synchrotron Radiation Research Centre (NSRRC) in Taiwan.

In earth sciences, ASGC has integrated the sensor network, forward seismic wave simulations and the seismic data center. Intensity maps, shake movies, and tsunami simulations that could help improve disaster mitigation have also been generated and performed. For instance, shake movies visualize, in the form of animation, the ground motion of seismic events after they occur. They provide information such as where the strongest shake occurs so that rescue resources can be directed to where they are needed the most. Shake movies are produced according to calculations performed based on earthquake models as well as the geologic structure. However, the production process is computing resource intensive; it takes a few hours to create one frame on a large computing cluster. With the volunteer computing resources, the time required to generate the shake movies shortens significantly, providing valuable information to rescuers swiftly once earthquakes occur.

In environmental change research, ASGC has been working on weather simulations for regional or global weather events, such as typhoons, but also for prevailing issues, such as long-term dust transportation, aerosol dispersion, and biomass burning. Particularly, regarding the weather simulations, a gLite-based WRF (gWRF) program was developed by our team, porting the Weather Research and Forecasting Model (WRF) to the gLite grid infrastructure. *The gWRF program is currently implemented on the EUAsia and GISELA Grid, and its* features include:

- Running WRF solver (MPI) on gLite grid infrastructure through a package of bash scripts
- Providing basic job submission and monitoring function
- Using GridFTP for input preparation and output retrieve
- Generating real-time Job log and result outputs

There are four parts in gWRF workflow (cf. Fig 21):

- User Interface (UI): gLite UI is a suite of clients and APIs that users and applications can use to access the gLite services from both the LCG and gLite software stack

- Workload Management System (WMS): gLite WMS is responsible for tasks distribution and management across grid resources, it passes a job to an appropriate CE for execution
- Computing Element (CE): Computing nodes, usually a cluster, where a job gets executed
- GridFTP: A gLite GridFTP server, for job input/output transfer

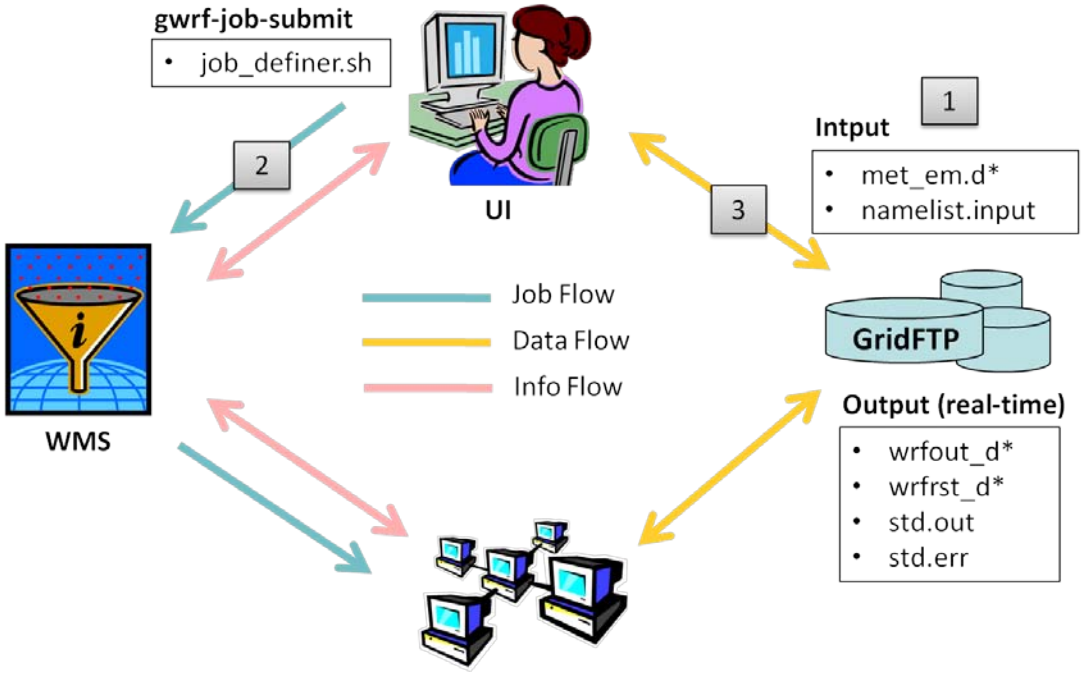

Fig. 21 gWRF Workflow

Efforts have been made in putting this program in the WebGUI so users can upload and manipulate their simulations through their browsers. This service is now available at this website-- https://gwrf.twgrid.org. Users need to log in with their certificates to use this service.

In addition to social sciences, ASGC also supports projects in high-energy physics, such as ATLAS, CMS, Belle and AMS.

Issues to be addressed in the future are the continuous improvement in reliability and the reduction of operation costs, including simplified and robust middleware, a sustainability model, and intelligent monitoring in collaboration with industrial partners, as well as the enhancement of resource efficiency through the introduction of new technologies. The advantages of DCI could also be further enhanced through the continuous development of data distribution and management as well as networking. ASGC will also place an emphasis on regional collaboration, which could help address common concerns, such as disaster mitigation.

11.Thailand

The Thai National Grid Project was a five-year project that ran from 2004 to 2009 and produced 192 nodes. With a budget of 3.7 million US dollars, its goal was to provide infrastructure for e-Science in Thailand. Focus area applications included life science, digital media, enterprise computing, computational science, engineering, and education. Resources were shared through gLite middleware. Similar to other projects in digital media and enterprise computing, the Thai National Grid Project was supervised by the Software Industry Promotion Agency (SIPA) and the Ministry of ICT. The project was run as a consortium, which means that all members managed their own budget, while the heads of the institutes constituted the committee. However, the project was discontinued after its first phase due to lack of funding for the second phase.

The EUAsiaGrid project was a 2 year project (EU-FP7 Coordination and Support Action) spanning from 2008 to 2010. In the end of the project, EUAsiaGrid has four European partners and eleven Asian partners, including National Electronics and Computer Technology Center (NECTEC) and Hydro and Agro Informatics Institute (HAII) from Thailand.

In September 2011, the consortium, including its founding members Chula University, Suranaree University, King Mongkut's University of Technology, Thonburi (KMUTT), the Hydro and Agro Informatics Institute (HAII), and the National Science and Technology Development Agency (NSTDA), signed the National e-Science Infrastructure Consortium MoU at Chulalongkorn University. It is open to new members with the intention of collaborating on the national infrastructure. Current application areas include high energy physics, climate change, water and energy resources and environment, computational science and engineering, and computer science and engineering.

In 2011, the NSTDA and HAII Cluster was set up with a total of 340 computing cores and 300 TB of data storage. In July 2012, KMUTT will connect their resources to the infrastructure, bringing up the number of cores to 706 CPU cores and 315 TB of data storage (cf. Fig 22). Within the next two years, National Electronics and Computer Technology Center (NECTEC), a statutory government organization under the NSTDA, and HAII intend to double their current capacity, depending on the availability of affordable hardware (cf. Fig. 23). The targeted output focuses on the strengthening of the human capacity, including 59 Master’s students, 24 Ph.D students and 45 undergraduates. Moreover, 59 projects covering the above-mentioned application areas will be serviced and 77 papers are to be published.

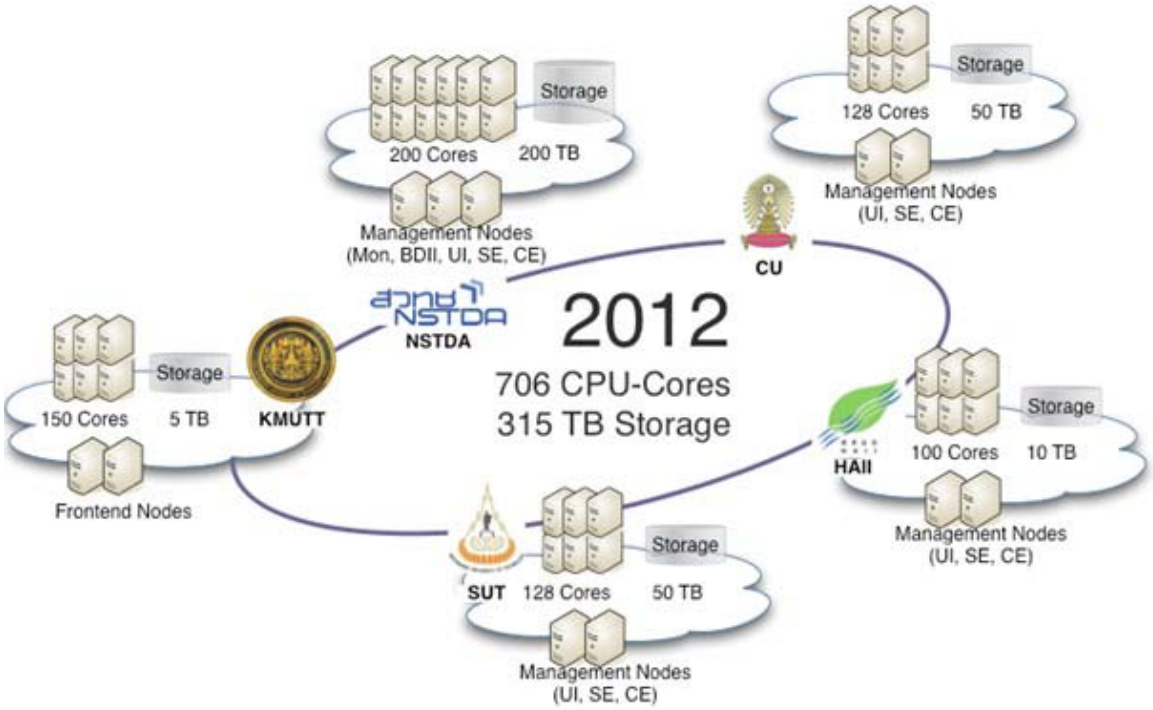

Fig. 22: Planned upgrades and expansions for 2012

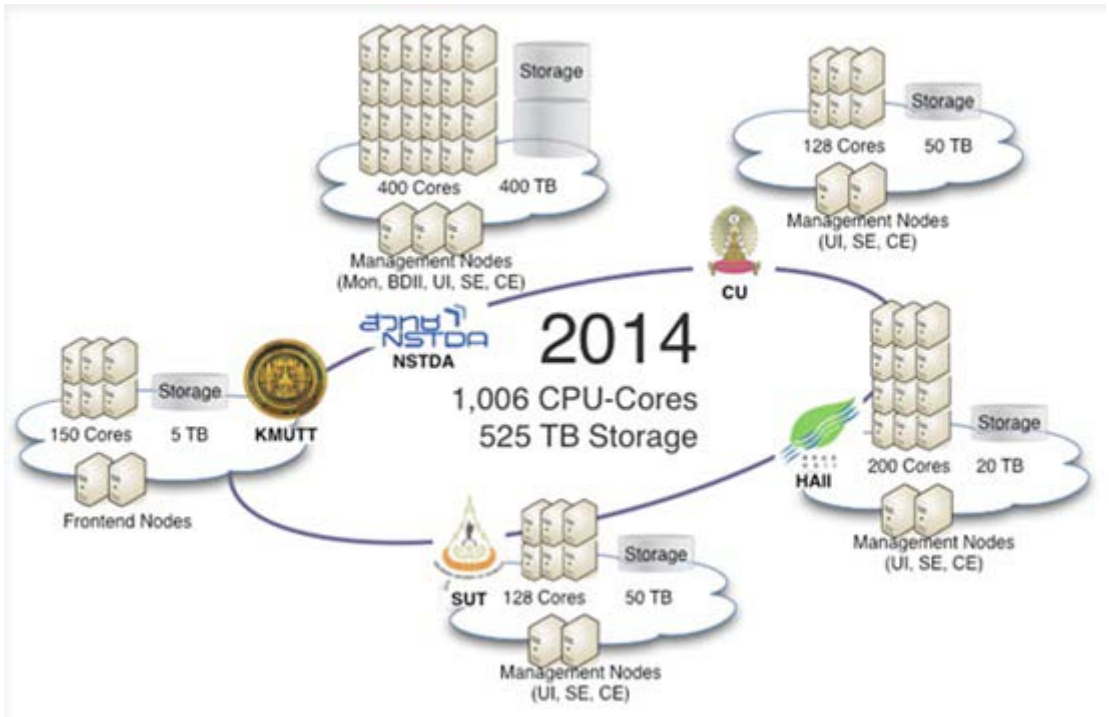

Fig. 23: Planned upgrades for 2014

Following the signing of the MoU, a workshop named “e-Science Infrastructure Consortium Seminar”, was held at Sofitel Centara Grand Bangkok, covering Thailand’s e-Science Infrastructure Consortium Project; windmap simulations; potential

collaborations with Thai universities though the National e-Science Infrastructure Consortium; strategy for a national science and engineering grid (a paper from NSF); cloud computing as a mechanism to deliver scientific and engineering computing services; Service Science, Management and Engineering (SSME) and successful projects that have been completed with universities and in collaboration with Janis E. Landry-Lane from IBM to transform business processes.

On January 18, 2012, 13 institutions signed the Thailand-CERN collaboration MoU, followed by a press conference, both of which were hosted by the Thailand Synchrotron Light Research Institute, MOST. Past and ongoing CERN-related activities include the CERN Summer Students Program & High School Teachers Program, post-graduate level research at CERN, and the CERN School Thailand 2010 (4-6 and 7-13 October 2010). Similar activities planned for 2012 include 2 postdoctoral visiting researchers at CERN, the first Thailand Experimental Particle Physics Novice Workshop at SUT in April 2012, the second CERN School Thailand 2012 at SUT in May 2012, the CERN Summer Students Program 2012 (2 university students, 8-week activities), and the CERN High School Teachers Program 2012 (2 teachers, 3-week activities).

National e-Science Infrastructure Consortium is an example of collaborative infrastructure development that ensures the right design and sustainability by user involvement and commitment. Current members are developing a procedure for accepting new members and expanding the consortium. Other issues to be addressed are governing, managing and operating procedures; the admission of additional members and the ongoing expansion of the current infrastructureas well as the ongoing participation in CERN related activities. A trend toward an increase of flood and disaster related applications is also possible.

12.Vietnam

A workshop named "International Workshop on Grid Computing Applications for Vietnam" was held in Hanoi on December 1-3, 2010. The research community concluded that grid applications have begun to take root in Vietnam due to the support of CNRS and the EUAsiaGrid project partners. A real need for applications has been identified in research areas, such as earth science, life science, physics, and chemistry, as several projects have been launched. However, Vietnam still needs to overcome several obstacles, such as a lack of experience, very limited research funding, very limited grid resources (CPU cores and storage capacity), limited bandwidth, unstable Internet connections, and lack of a National Grid Initiative (NGI) to coordinate efforts and resources.

Despite these difficulties, the Vietnam Research and Education Network (VinaREN) was re-established in June-2011, after links from VinaREN to IFI, IOIT and HUT had been suspended for 12 months. After its re-launch, all grid nodes were reinstalled, certified and re-integrated into the EUAsia VO. VinaREN is a TEIN3 member and hosts all Vietnamese EGI nodes (IFI, IOIT-VAST, HPC-HUT). No improvements to the processing power (3-5 CPU cores) and the data storage (733 GB HDD/node) have been made, which renders the network unsuitable for the use of real applications.

Current project include a performance comparison between the Wisdom Production Environment (WPE) and DIRAC, and next-generation sequencing, which makes the sequencing of genomes and transcriptomes much faster and cheaper than ever before by parallelizing nucleotide sequencing procedures, but poses a huge challenge in terms of data analysis and storage. Vietnamese scientists are also porting ePANAM, a phylogenetic pipeline dedicated to analyzing massive sequencing results that automatically affiliate sequences from SSU rRNA amplicons and builds phylogenetic trees of very large numbers of sequences, to the grid, which can help reduce the processing time to a significant degree.

Drug discovery on local biodiversity compounds is another important research area. In collaboration with INPC, IOIT, IFI, HCMUS, CNRS, HealthGrid, IRD, ASGC, and UPM, 45 compounds were tested and docking tests in a grid environment with DIRAC were successful. The compound database is also ready. Within the next 3 years, the scientists plan to continue the grid development and deployment.

In collaboration with IGP, IOIT, IFI, CNRS, IPGP, IRD and ASGC, efforts are also being made to establish a database of pre-calculated earthquake/tsunami scenarios. So far, 25 out of 1,000 planned simulations have been performed using ArcGIS software. It is planned to establish a portal with thousands of simulations that can be queried to get a clearer picture of the risks stemming from tsunamis. Since these simulations require a large amount of processing time, power and data storage, the international grid infrastructure will process the data, whereas private cloud systems will be used for data storage, so users can submit jobs and access the data through a web portal.

In collaboration with the Auckland University of Technology in New Zealand, the Grid and Cloud Computing Group at the University of Science in Ho Chi Minh City is currently building a web-based prototype of a private virtual cloud learning environment based on cloud computing technology. Their goal is to create various kinds of virtual labs in order to serve different coursework practices.

There are plans to establish a new government-funded institute named “Computational Science and Technology Institute” at Vietnam National University (VNU) with a proposal to be finalized and submitted to the government for approval by the end of 2012. Its objective will be to develop and gather a team of experts with strong research capacities as well as to build facilities with modern equipment and advanced computing technology in order to solve tasks addressing national interests and to meet the demand for computing applications in the fields of science and technology, economic, social and national security.

Sponsored by CNRS and IRD, the third school in "Advanced Computing and GRID Technology for Research" was held from October 24 to November 4, 2011), focusing on DIRAC middleware and StratusLab cloud. A workshop on grid and cloud computing for scientific applications was also held. 30 participants with grid certificates for the EUAsia VO attended the school and ran more than 5,000 jobs with DIRAC. After its successful installation, a DIRAC server is now in service in Hanoi, which allows access to resources in Vietnam, France (Biomed VO), and Korea (KISTI).