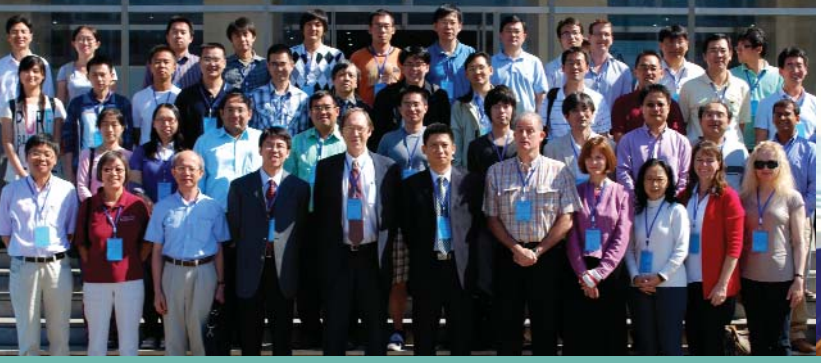
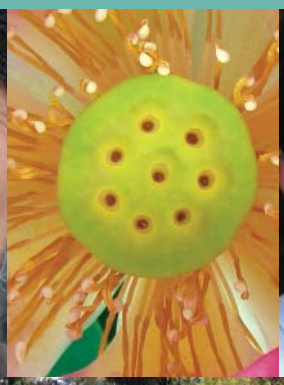
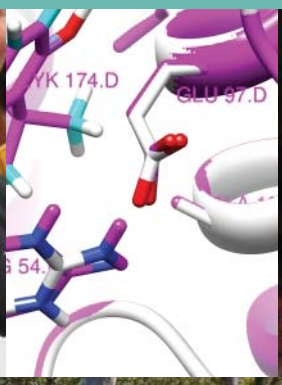


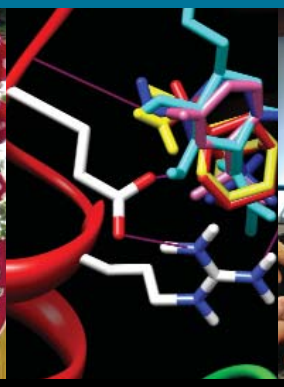

PACIFIC RIM APPLICATIONS AND GRID MIDDLEWARE ASSEMBLY








COLLABORATION OVERVIEW 2010-2011



PRAGMA

Ladder that leads to heaven (TianTi)—courtesy of Alicia Lee

INTRODUCTION

THE PACIFIC RIM APPLICATION AND GRID MIDDLEWARE ASSEMBLY, PRAGMA, was launched at its inaugural workshop in March 2002. The two components of its mission focus on both technologies and people: first, use applications to drive forward grid technology development, and second, build sustainable collaborations among the members. Since our inception, there have been many technological successes, among them are: the first international software, Ninf-G, integrated into the US National Middleware Initiative software stack; broad, international adoption of technologies, such as Rocks; and the testing, refining, and deployment of multiple software packages, including Duckling and eAirs. These technological achievements are made possible by the participants, our open collaborative framework, and the ability to attract new participants and grow collaborations.

PRAGMA continues to be a forum for EXCHANGE—OF INFORMATION, IDEAS, AND EXPERTISE. Members come to PRAGMA to learn what the trends are, where science, technology, and collaboration are heading, both in general, and at other institutions. Because of the nature of PRAGMA, we are structured to easily share technologies, test the codes of others, provide useful feedback, and then develop improved applications. Our exchanges of people (from the most senior researchers to students) create global scientists and engineers, namely, individuals who can work internationally and collaboratively in the shared development and testing of the tools needed for emerging technologies and scientific advancement.

PRAGMA has also seen tremendous GROWTH, not only in its membership (from 13 to more than 30 institutions), but also in helping member institutions across the globe compete successfully for new awards and funding. Equally important are the contributions towards the professional growth of many of our participants. A number have been promoted from the ranks of junior researchers to senior status, while others have advanced into positions as either group leaders or institutional directors.

"The framework of PRAGMA allows for new activities to arise and to create their own trajectory. The culture of PRAGMA allows for open sharing of ideas and has built trust between members."

Finally, PRAGMA is a FRAMEWORK FOR COLLABORATION, a vehicle that encourages new projects and groups to grow and develop. In some cases the groups stay tightly connected with PRAGMA, in others, the groups develop in parallel. We discuss several in this brochure.

In this eighth Collaborative Overview, we highlight a wide range of activities, including collaborations which

- have spanned several PRAGMA institutions,
- were driven by applications for real-world needs (volcanic ash circulation, lake modeling, and avian flu), and
- have advanced components of the technology (experimentation with clouds, virtualization of resources to be built once and used many times, and creating secure environments for virtual organizations).

We also highlight software contributions of individual institutions which were user-tested in the PRAGMA community (Duckling, Gfarm, and e-AIRS). In addition, we illuminate the benefits of PRAGMA to four institutions (NICT at Osaka University, NARL, JLU, and Monash). Finally, we highlight activities of our partner organizations (GLEON, CREON) and activities (PRIME, MURPA) that either started within PRAGMA or, are based on it.

The strength of a grassroots organization is the diversity, commitment, and active participation of its members. The framework of PRAGMA enables new activities and allows them to arise and create their own trajectories. The culture of PRAGMA fosters open sharing of ideas and has built trust between members. Sustaining these collaborations requires an ability to evolve with, as well as lead, the evolution of the technologies, to remain relevant to the members, and to continually engage new ideas and members. Over the next year we see new opportunities in technology that we are uniquely able to test and develop; new applications to drive forward those technologies; and a keen interest to broaden the impact of our work far beyond our membership, both in terms of software, as well as education and training.

In short, PRAGMA has produced collaborations wherein researchers are not just better together in a straight additive way, but multiplicatively better, in advancing technologies and staying on top, and ahead, of changes in the capabilities and needs of the global community, as well as preparing researchers for the future.

We look forward to continuing to building our collaborative framework and advancing technologies driven by societally important challenges.

p. 15

p. 26

p. 18

KEY COLLABORATIVE SUCCESSES

NG-TEPHRA: Simulating Ash Dispersal via Collaborations and Cloud Resources

FROM 1963-1965, COSTA RICA SURVIVED ONE OF THE MOST IMPORTANT EVENTS OF ITS MODERN VOLCANIC HISTORY[1]. The Irazú volcano, located northeast of the Great Metropolitan Area (GAM) of San José, erupted violently, dispersing large amounts of tephra (fragmental material of any composition and size produced by a volcanic eruption) and ash across surrounding areas. The GAM of San José received abundant ash falls, to the point of stressing health systems and the economy.

Livestock and agricultural production areas were heavily affected by tephra deposits due to the strong presence of sulphates in their composition. Across the country, acute medical conditions, such as asthma and pneumonia, were aggravated. Buildings near the vent (most had low, sloped roofs designed to withstand only the rigors of wind and rain) succumbed to excessive weight, leading to severe damage of major infrastructure. Limited emergency attention was available at the time, due to the scarcity of information regarding ash distribution and the magnitude across the affected areas.

At the time, Costa Rica had less than one million inhabitants across the entire country, approximately two million fewer than the current population of the GAM alone. Therefore, it is essential to have a system that facilitates analysis, characterization and prediction of the impact of volcanic events. It must provide timely results, be integrated with geographic information systems (GIS), and be applicable to other regions of interest where volcanic activity is considered possible and hazardous.[2, 3, 4]

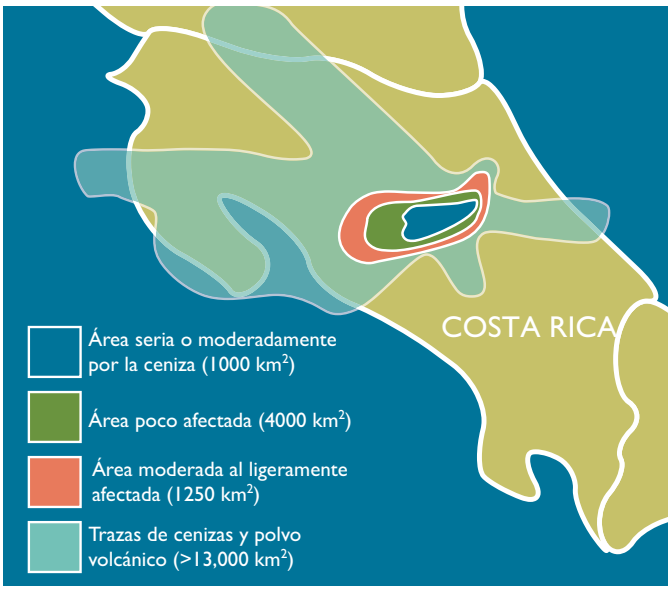


Fig. 1. Irazú volcano location and ash distribution during 1963-1965

Using PRAGMA's collaborative framework, researchers from the Costa Rica Institute of Technology, Monash University, and the National University of Costa Rica are collaborating to address this important problem. This collaboration was initiated by discussions at the PRAGMA 18 workshop (March 2010, San Diego, CA) with an aim to research and develop a volcanic ash dispersion simulation and spatial forecasting system that could be used not only for Irazú but also where needed across the Pacific Rim. The workshop facilitated rapid prototyping of a PRAGMA-Grid and Amazon EC2 cloud-based prototype driven by Nimrod/G. Recent work has focused on validation of NG-TEPHRA, a numerical simulation program implementing the Suzuki model, developed for the Irazú case. The current research group has extended and modified the original software[5], further parameterizing and making it suitable for large-scale distributed execution.

The model is based on a deterministic approximation of the advection-diffusion phenomenon coupled with classical ballistic motion for the tephra particle fall. It is a semi-analytical model given that some of the variables are not directly determined and thus left to the scientist to judge. The Irazú case requires variation of nine of these parameters, producing 4.41 million parameter combinations. Fortunately, a full sweep is not required (and would be impractical) due to the low probability of some parameter combinations. The effect of these is combined by means of a weighted average, with similar scenarios and results in a filtering of the parameter space to 217,125 valid combinations. As a by product of the work Nimrod/G's parameter generator was modified, saving significant time on this work and enhancing Nimrod's functionality.

Early results indicate good potential for this model and the execution framework[6]. Topics of ongoing research include visualization of model output and map generation suitable for urban planning.

[1] ALVARADO, G. Volcanology and petrology of Irazú volcano, Costa Rica. Doctoral dissertation. University of Kiel, Germany. 261p, 1993.

[2] HEIKEN, G., M. M. H. W.-S. W. Hazards and energy infrastructures. United Status. U.S. Department of Energy Code EH-33. Office of Risk Analysis and Technology., UR 95-1087, 1995.

[3] SMITH, K. Environmental hazards: assessing risk and reducing disaster. Routledge, London, UK, 1992.

[4] TOMLINSON, R. Thinking about GIS. ESRI Press, California, USA, 2003.

APPLICATIONS DRIVING TECHNOLOGIES

[5]CONNOR CH., E. A. Estimation of volcanic hazards from tephra fallout. Natural Hazards Review 2, 1 (2001), 33-42.

[6]NUNEZ S, Bethwaite B, Brenes J, Barrantes G, Castro J, Malavassi E, Abramson D. NG-TEPHRA: A Massively Parallel, Nimrod/G-enabled Volcanic Simulation in the Grid and the Cloud. IEEE e-Science 2010 (accepted)

PARTICIPATING RESEARCHERS: Blair Bethwaite, David Abramson, *Monash Univ.*; Santiago Núnez-Corrales, José Castro, *Costa Rica Institute of Technology, Costa Rica: Santiago*; Gustavo Barrantes, Eduardo Malavassi, *National University of Costa Rica*; José Brenes, *Citizen's Network of Meteorological Stations, Costa Rica*; Philip Papadopoulos, *SDSC/UCSD*

From Grid to Cloud - VM Image Replication: Providing Greater Opportunities for Avian Flu Researchers

TO MAKE THE PRAGMA GRID EASIER TO USE, THE RESOURCES WORKING GROUP HAS ADOPTED VIRTUALIZATION TECHNOLOGY AND IS WORKING ON GRID-TO-CLOUD MIGRATION. To achieve this, virtual machine (VM) image replication was a key issue that we needed to confront. Working collaboratively, three PRAGMA teams (SDSC, AIST, and NBCR) have successfully developed and tested a solution. The current result provides a key building block for paving the way to cloud computing in PRAGMA.

Unlike a homogeneous and specialized grid, where a software stack can be easily defined and uniformly provided, PRAGMA Grid is a heterogeneous grid with very diverse user applications and requirements. Providing all software required by all users at all sites would consume too much time from site administrators. In addition, software requirements for multiple applications often result in many conflictsthat are impossible to resolve. On the other hand, to have users install required software themselves at every site would be extremely inefficient and too much of a burden for users to handle. Cloud computing and virtualization offers a solution to this problem. A system customized for an application can be build once then run everywhere, with the proper virtualization platforms.

Since PRAGMA 17 (October 2009, Hanoi), the Resources Working Group has decided to migrate from grid to cloud through experimentation with virtualization technologies. Currently, there are three virtualization platforms in the works among PRAGMA Grid sites: Rocks-based VM by SDSC; WebOS-based VM by NCHC; PIAX-based VM by Osaka University. Our short-term goal is three-fold: setup VM hosting services at selected sites, build and share VM images for applications, and test interoperability among different virtualization platforms.

Currently, most applications executions in PRAGMA Grid are Globus-based and use batch-job schedulers. The configurations for Globus and job schedulers (SGE, PBS, etc.) are sensitive to host name and IP address changes, therefore a VM image migration among different VM hosting services may require more than trivial modifications. We need to find and resolve all issues.

As of PRAGMA 19 (September 2010, Changchun), five sites (SDSC, AIST, NBCR, NCHC, and Osaka University) have setup VM services. Among these, three are Rocks-based (SDSC, AIST, and NBCR).

In order to experiment with building and sharing VM images, our strategy is:

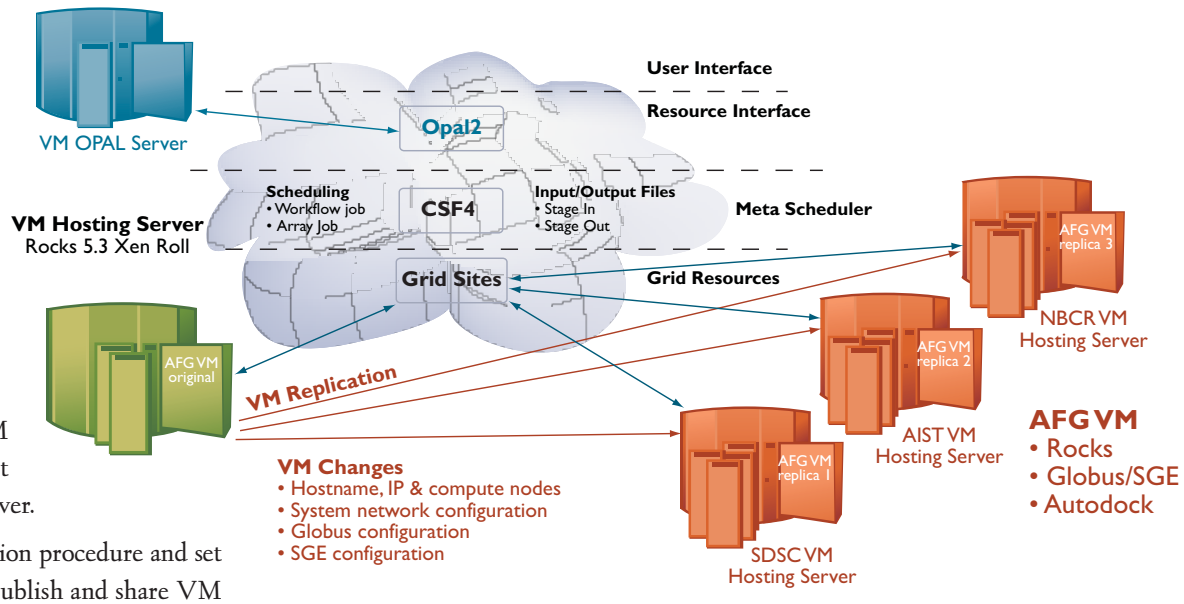
- Start with Rocks-based VM, and then add additional VM platforms.
- Start with manual procedures, and then automate.

We planned our work as follows:

1. Build a base-line VM (which includes Rocks cluster 5.3, Globus 4 and SGE 6.2) on a Rocks VM hosting server.
2. Replicate the base-line VM image on the same Rocks VM hosting server.
3. Use this baseline VM image and add software stack for Avian Flu Grid (AFG).
4. Replicate the AFG VM image on the same VM hosting server.

KEY COLLABORATIVE SUCCESSES: APPLICATIONS DRIVING TECHNOLOGIES

Fig. 1 PRAGMA Rocks Virtualization Experiment



5. Replicate the AFG VM image on a different Rocks VM hosting server.
6. Automate the replication procedure and set up Gfarm storage to publish and share VM images.
7. Build and replicate VM images for more applications.
8. Replicate VM images on different VM platforms.

For step 1, the SDSC team built a base-line VM on their Rocks VM hosting server and documented the work (available at goc.pragma-grid.net/wiki/index.php/Base-line). In step 2, the team replicated the base-line VM manually on the same VM hosting server and documented the procedures (available at goc.pragma-grid.net/wiki/index.php/VC-replication-1).

Step 3: SDSC invited the NBCR team (the Avian Flu Grid lead), to build an AFG software stack (AutoDock) on the replicated VM cluster and to test it with an AFG application run.

Step 4: the SDSC team replicated the AFG image on SDSC's Rocks VM hosting server and documented the work (available at goc.pragma-grid.net/wiki/index.php/VC-replication-3). The NBCR team tested the replicated VM cluster by running AFG applications.

Step 5: AIST and NBCR replicated the AFG VM image on both of their Rocks VM hosting servers, tested with an AFG application run and documented the procedures (available at goc.pragma-grid.net/wiki/index.php/VC-replication-2).

The results of these experiments have concretely proven that VM replication is a possible solution to achieve a "build once and run everywhere" environment. We are now working on automating this replication procedure and setting up Gfarm storage to publish and share VM images. Ultimately, we will be able to use the automated procedure to build, share, and replicate VM images for additional applications.

An interoperability experiment between Rocks-based VM and PIAX-based VM was initiated this summer as a PRIME undergraduate student project (please see Adrian Ho (PRIME 2010): "Implementation of a Multi-Site Virtual Cluster System With a P2P Overlay Network", in the Highlighted PRIME projects section for more information). Interoperability issues among different VM platforms will come into focus as we gain experience and have success with Rocks VM replications and services.

We believe that enabling virtualization in the PRAGMA Grid will allow us to use both public and private cloud platforms in the future. The cloud deployment will enable on-demand, scalable resources and applications provisioning. Taking advantage of cloud computing will enable us to speed up the delivery of applications as a finished product and will shorten their prototype-to-production chain.

PARTICIPATING RESEARCHERS: Philip Papadopoulos, Mason Katz, Cindy Zheng, *SDSC/UCSD*; Yoshio Tanaka, Akihiko Ota, *AIST*; Wilfred Li, Nadya Williams, Jane Ren, *NBCR/UCSD*

GEO Grid: Providing Secure Grid Resources for the Virtual GLEON Community

GEO GRID IS A PROJECT AT AIST, established not only to integrate various types of data regarding the earth systems in a way that allows for access, but also provides protection of special data as needed (e.g., licensing). GEO Grid is developing an infrastructure to integrate remote sensing and GIS data from various sources. One of the architectural features of GEO Grid is its focus on the concept of a Virtual Organization (VO), which integrates all relevant data and application services among a virtual community of distributed users.

In GEO Grid, a VO is associated with any kind of scientific activity and is created dynamically by integrating available services and resources according to the interests and requirements of the users. In this case, the system should have the ability to provide flexible management of VOs. To achieve this, the current implementation supports multiple VOs by using Grid Security Infrastructure (GSI) and VO Membership Service (VOMS), which allow fine-grain control of the use of the available resources to both the user's VO and the resource owners.

Despite the general acceptance of GSI and its use over the course of many years, GSI-based security systems are well-known to be difficult to use. For example, users must obtain their own public key infrastructure (PKI) certificates and manage their PKI private keys carefully. In order to access Grid resources, users have to generate proxy credentials from the PKI certificates and embed VOMS attributes to the proxy credentials. These are troublesome tasks for users who are unfamiliar with PKI and can be considered as one of the obstacles to wider use of Grid-based infrastructure. GEO Grid has developed security software called Tsukuba-GAMA that manages users' PKI certificates and private keys and generates proxy credentials with VOMS attributes on behalf of the user. Tsukuba-GAMA enables members to access Grid resources by a familiar authentication method such as username/password and OpenID.

Collaborating with the Resources Working Group, the PRAGMA GEO Working Group has been developing PRAGMA GEO Infrastructure which supports geosciences and practical applications based on huge datasets of satellite imagery and GIS data. Tsukuba-GAMA is deployed to manage appropriate access control for sensitive data (e.g., ASTER data provided by AIST and Formosat-2 data provided by NARL/NSPO, which has collaborated with PRAGMA to provide these data). PRAGMA GEO Infrastructure is not only currently used by application scientists in PRAGMA, but is ready for use by other scientific communities.

Launched in March 2005, GLEON, the Global Lake Ecological Observatory Network, is a network of people, institutions, programs and data, whose mission is to understand and predict responses of lake ecosystems to natural processes and human activities at regional, continental, and global scales (see Partners: Advancing Technologies, Training Researchers for the Future section for more details). During this past year, researchers from GLEON and PRAGMA have been working together to integrate sensor data from GLEON lakes with satellite data. PRAGMA GEO Infrastructure has been provided to GLEON members for their routine use since May 2010. Any GLEON member may have an account on PRAGMA GEO Infrastructure. S/he does not need to understand the details of the underlying security (e.g., PKI, GSI, and VOMS); they will be allowed to access the sensitive satellite data by a familiar username/password authentication method.

Monitoring the lake for basic properties is the fundamental objective of GLEON. Currently, single point, but high-accuracy, buoy data are collected; models then use these data as input to predict circulation of water and thus distribution of nutrients and phytoplankton. By linking data from sensors in lakes with data derived from various remote imaging approaches, researchers will be able to extrapolate the point measure—for example, temperature and fluorescence—to all parts of the lake, leading to more accurate modeling. This methodology is not limited to lake observation but is applicable to other fields such as carbon tracking and biodiversity. The PRAGMA-GLEON collaboration drives the development of tools and systems that will make PRAGMA outcomes practical for use by a wide range of scientific applications.

ADDITIONAL INFORMATION: GEO Grid: **www.geogrid.org**; Tsukuba-GAMA: **devel.geogrid.org/TsukubaGAMA**; PRAGMA GEO Infrastructure for GLEON: **goc.pragma-grid.net/wiki/index.php/GLEON_CREON_Information**

PARTICIPATING RESEARCHERS: Yoshio Tanaka, Ryosuke Nakamura, Naotaka Yamamoto, Sarawut Ninsawat, *AIST*; Franz Cheng, NARL; Bo Chen, National Space Agency (NSPO)/ NARL; Fang-Pang Lin, *NCHC*; Peter Arzberger, *UCSD*; Mathew Allan, *U. Waikato*

Members are the most important asset of PRAGMA, supporting individuals to participate in PRAGMA, hosting exchanges of researchers and students, organizing workshops, and contributing resources to the PRAGMA grid.

A list of active members together with their sponsors is provided at the end of the brochure. In this section we highlight three software tools developed by PRAGMA members and tested by others; and we highlight the benefits of PRAGMA membership to five institutions.

MEMBERS: CONTRIBUTIONS AND BENEFITS

DUCKLING: Collaboration Environment for e-Science

DUCKLING OPEN-SOURCE SOFTWARE IS A COLLABORATION ENVIRONMENT SUITE, developed by a team at the Computer Network Information Center (CNIC) of the Chinese Academy of Sciences—a PRAGMA member institution. PRAGMA served as the collaborative framework in the development cycles of DUCKLING, providing the collaboration with use cases and feedback, which ensured the software's usefulness, enhanced the software quality, and helped with its dissemination.

DUCKLING is a comprehensive resource sharing and collaboration platform created specifically for scientific research groups to improve their collaborations and workflow. All resources of a research group, including hardware, software, data, information and the researchers themselves, can be organized and integrated together to form an efficient and easy-to-use system, supporting and advancing new research activity models.

Duckling consists of core toolkits and a Duckling Application Integration Framework (DAIF) portal environment. The core toolkits include a Document Collaboration Tool (DCT), Virtual Organization Management Tool (VMT), User Management Tool (UMT), and a Collaboratory Library (CLB). Portlet Standard JSR168 and JSR286 are also implemented in the portal framework. Developers and users can build their collaboration environments at a glance and customize them as much as desired. Duckling helps organize collaboration behavior, which, in turn,enables greater resource sharing and innovation.

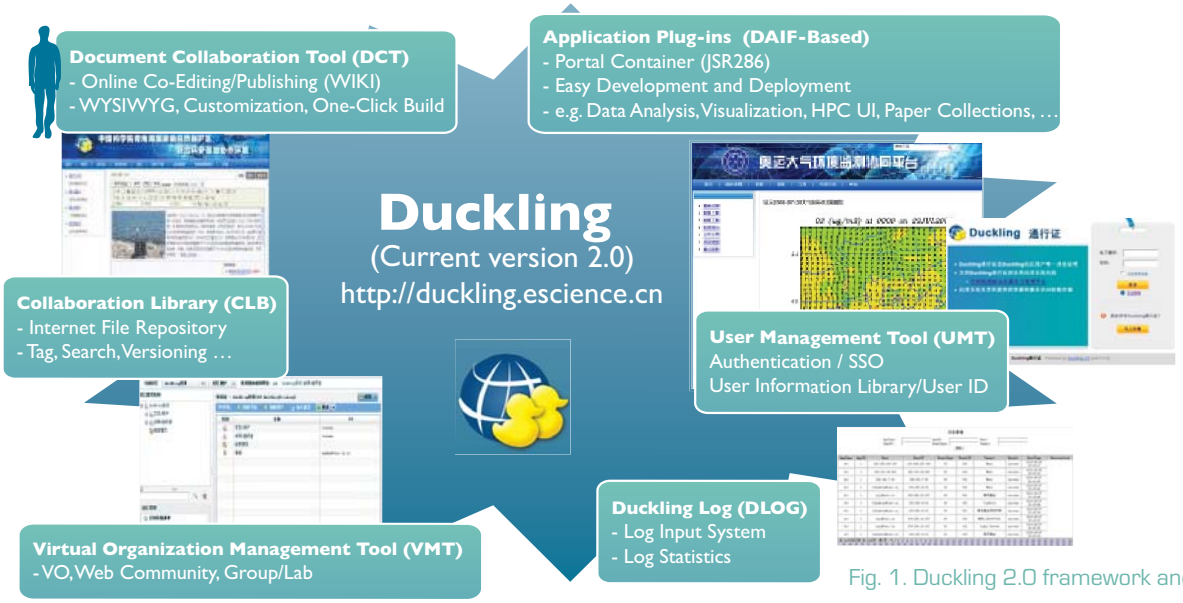


Fig. 1. Duckling 2.0 framework and core toolkits

PRAGMA 18 group photo—courtesy of Teri Simas

An application plug-ins set, based on DAIF, has been developed which further enriches Duckling as a collaboration environment, extending it for additional requirements. Some of the many plug-in-style applications in the set are Data Lumber Mill (DLM), Library Management Tool (LMT), and Conference Service Platform (CSP).

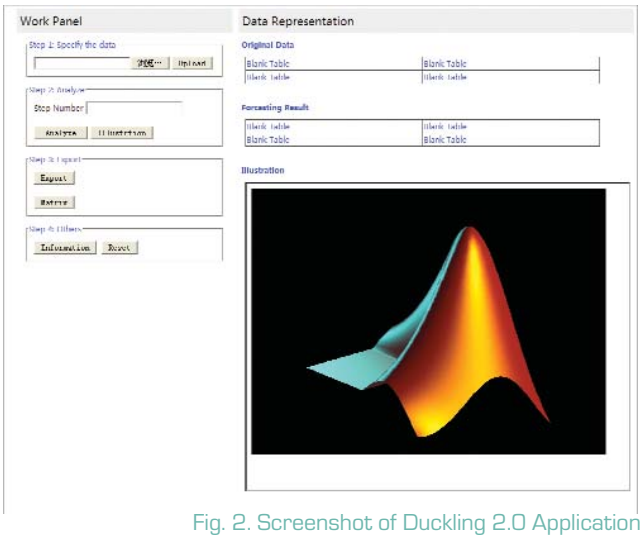


Fig. 2. Screenshot of Duckling 2.0 Application

The CNIC team worked with many of the PRAGMA teams to conduct use tests, as well as enhance DUCKLING capabilities, two of which are discussed here. 1) Working with UCSD and JLU teams, DUCKLING was used to set up two PRAGMA workshop websites which successfully handled international conference planning, arranging and organizing tasks. As a result, DUCKLING has become the standard tool for PRAGMA workshop websites. 2) CNIC and UCSD teams jointly developed a Real-time Biomedical Data Streaming platform based on DUCKLING and OPAL2 that extends scientists' desktops to biomedical grid/cloud computation resources and mass storage systems.[1] These successful use cases of DUCKLING have prompted more interest in collaborations from both inside and outside of the PRAGMA community. For example, the M*Grid development teamat Kookmin University has recently met with the CNIC team to plan work on integrating and testing DUCKLING with M*Grid.

CNIC released Duckling 2.0, their first open source software, in March, 2010; the source code is available as a free download from the SourceForge website. The hope is that Duckling can be widely tested and used with more PRAGMA partners, as well as researchers outside of PRAGMA.

[1] Kejun Dong, Sameer Tilak, Kai Nan, Cindy Zheng, Dong Xu, Jurgen Schulze, Peter Arzberger, Wilfred Li, "Real Time Biomedical Data Streaming Platform (RIMES) - a data intensive virtual environment", The 2010 International Workshop on HPC and Grid Applications (IWHGA2010).

ADDITIONAL INFORMATION: **en.wikipedia.org/wiki/Duckling_%28software%29**; **duckling.sourceforge.net**; **english.cnic.cas.cn/ns/es/201003/t20100318_51695.html**

PARTICIPATING RESEARCHERS: Kai Nan, Kejun Dong, Yihua Zheng, *CNIC*; Wilfred Li, Sameer Tilak, Cindy Zheng, Teri Simas, *SDSC/UCSD*

MEMBERS: CONTRIBUTIONS

Gfarm File System Supports PRAGMA Cloud Computing

THE GFARM FILE SYSTEM IS AN OPEN-SOURCE GRID FILE SYSTEM that has been developed by the University of Tsukuba (U. Tsukuba). Gfarm can be used as a global-distributed file system to share data and support distributed data-intensive computing. It federates local file systems on compute nodes to maximize distributed file I/O bandwidth and can store multiple file replicas—in any location—to avoid read-access concentration of hot files. Data-location aware process scheduling improves the file I/O performance of distributed data-intensive computing.

Gfarm has been an important middleware effort in PRAGMA Grid since 2004. It has been integrated with many grid middleware projects, such as the Community Scheduling Framework (CSF), Opal, and DataTurbine and tested with real applications, such as Avian Flu Grid and Geo Grid. Users can access files in Gfarm via an Opal portal and share input and output data, as well as application codes among institutes around the Pacific Rim in the PRAGMA Grid. Submitted jobs can access input and output data in the Gfarm system without any modification.

As PRAGMA Grid adopts virtualization technologies and migrates to cloud computing, the traditional need for a global file system to support data and computing needs will be expanded by the need for sharing virtual machine (VM) images which can be 10GB or larger each. Most PRAGMA sites do not have the capability for storing such big files. PRAGMA's Resources Working group has decided to set up Gfarm with large-capacity storage to support the migration from grid to cloud in PRAGMA.

The initial effort is a collaboration between teams at U. Tsukuba, SDSC, and AIST. SDSC has set up a Rocks VM for the Gfarm meta-server, is working on a 40TB storage system (which will be used for Gfarm storage), and will also provide support for these systems. All three teams will work together to implement and test the Gfarm setup among SDSC and AIST sites. The U. Tsukuba team is leading this collaborative effort.

The work of virtualization in PRAGMA Grid has already produced VM images ready to be shared. More VM images are in production. Once the Gfarm Storage System is established and in place, we will be testing it by publishing and sharing VM images using the setup. We are looking forward to discovering and resolving any issues that may arise, in order to advance the Gfarm system as an essential cloud-computing facility.

ADDITIONAL INFORMATION: Gfarm File System: **sourceforge.net/projects/gfarm**

PARTICIPATING RESEARCHERS: Osamu Tatabe, *U. Tsukuba;* Philip Papadopoulos, Mason Katz, Cindy Zheng, *SDSC/UCSD;* Yoshio Tanaka, Yusuke Tanimura, *AIST*

e-AIRS for Computational Fluid Dynamics

COMPUTATIONAL FLUID DYNAMICS (CFD) is one of the branches of fluid mechanics that uses numerical methods and algorithms to solve and analyze problems that involve fluid flows. This applied science seeks to understand thermofluidic movement and the relationshipbetween thermofluids and the forces acting on objects (including the properties of thermofluids which do or do not move). CFD is actively taught and researched in many academic departments, including mechanical engineering, aerospace engineering, ship building and marine engineering, civil engineering, architecture and building engineering, and others; it is utilized commercially in various sectors.

To apply and use the research findings related to CFD in higher education and maximize the usefulness of this research in teaching, it is valuable to know users' opinions about existing teaching methodologies. Based on the results of such a study, a new teaching environmentnamed e-Aerospace Integrated Research System (e-AIRS) was designed and developed as a pilot project for "Education-hub Development by converging high-end IT and computational Science for providing new Opportunities to the Next generation" (EDISON). EDISON can be considered a cyber-education environment, science gateway, or portal for the research and education of the computational science communities. The contents of e-AIRS are as described below.

- Develop various CFD analyzers in order to maximize the use of e-AIRS in teaching.

"There are always PRAGMA members working on interesting problems ready to push the boundaries of what we do."

- Enhance the ability of instructors in the numerical analysis used in multidisciplinary fluid dynamics.
- Develop web-based visualization technology to enhance understanding in multiple fields of fluid dynamics.
- Develop cyber-education portal for the effective support of education service.

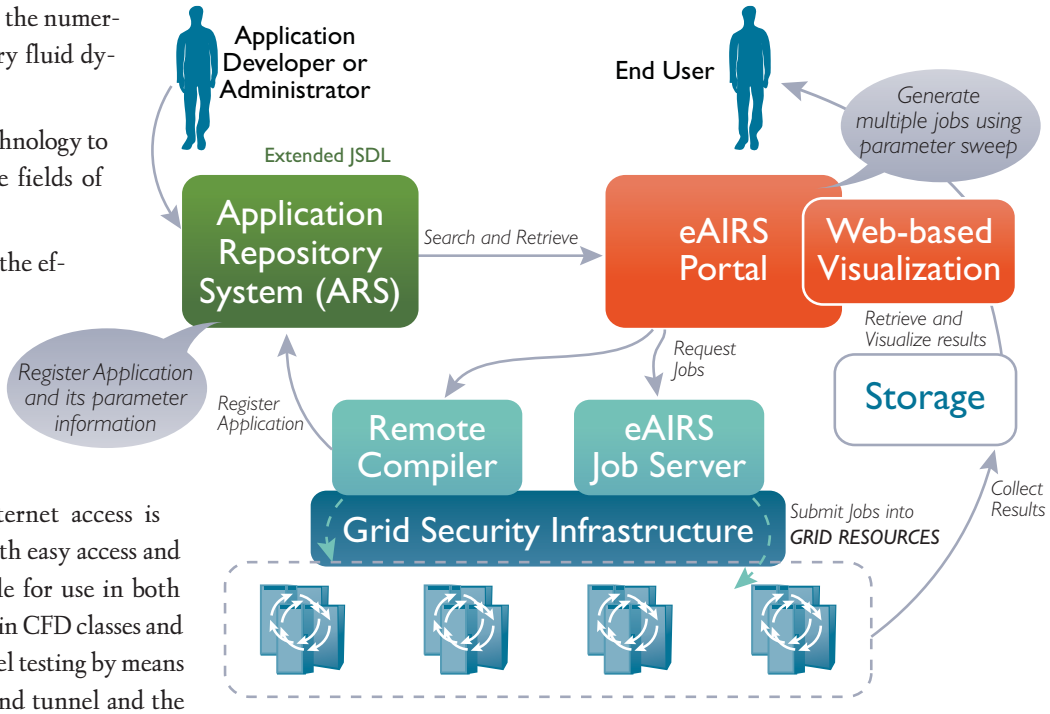


Fig. 1. Operations of e-AIRS

Figures 1 and 2 show how e-AIRS works and what the components are. One of the many advantages of e-Airs includes that users can continue on-line research anywhere internet access is available. This system was designed with easy access and ease of use in mind, making it suitable for use in both teaching and research. It has been used in CFD classes and students have learned about wind tunnel testing by means of video conferencing between the wind tunnel and the classrooms. Students have carried out numerical analyses and analyzed important physical phenomena with the tools in e-AIRS.

Experience with the web-portal based e-AIRS has satisfied students' requirements and boosted their interest in CDF over ordinary theory-based teaching methods. e-AIRS was used in practical experiments with 180 students across seven universities in 2008 and more than 400 students in nine universities in 2009, with the support of the PRAGMA testbed. In conjunction with the theory-based classes, term projects were given to students to assess their knowledge about physics. According to the follow-up questionnaire regarding e-AIRS, approximately 77% of the students said they were satisfied in terms of their ease in understanding the theories.

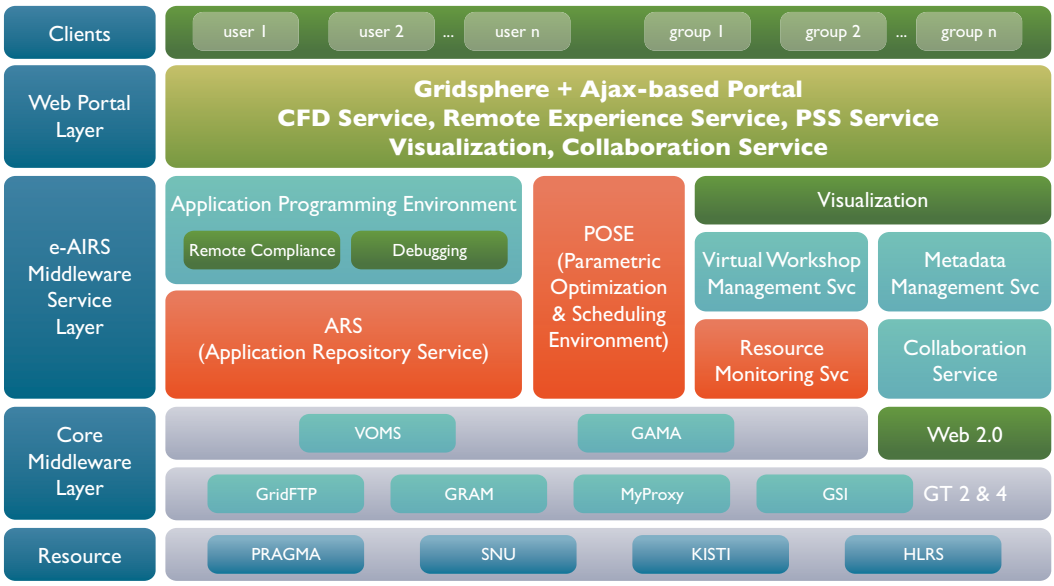


Fig. 2. Components of e-AIRS

ADDITIONAL INFORMATION: e-Aerospace Integrated Research System (e-AIRS): **eairs.kisti.re.kr;**

PARTICIPATING RESEARCHERS: Kum Won Cho (PI of e-AIRS), Jongbae Moon, Dukyun Nam, JongSuk Ruth Lee, *KISTI;* Chongam Kim, *Seoul National University*

Value of Participation in PRAGMA: Members choose to participate in PRAGMA for many reasons, and through participation, bring value to their own research activities. In this section, five institutions indicate how their participation has benefited their organization and its members.

MEMBERS: BENEFITS

Osaka University and the National Institute of Information and Communications Technology (NICT)

We have been developing technologies for cyberinfrastructure. By actively participating in a large community such as PRAGMA, we can contribute our technology to real-world use in science. For example, the OptIPlanet infrastructure becomes an example for future cyberinfrastructure with applications to telescience, geoscience and e-culture. Sensor networks are also foundations for cyberinfrastructure in environmental monitoring such as GLEON and urban monitoring. An educational component is mandatory for international collaborative science. PRIME is one of the key components of our lab in terms of Japanese students as well as international collaboration. Because we valued this PRAGMA effort involving education, we initiated a sister program, PRIUS, which sends our students to PRAGMA institutes. The large-scale international collaboration that happens in PRAGMA is unique. Therefore, for students especially, joining this international collaboration team is a real value.

National Applied Research Laboratory (NARL) and the National Center for High-performance Computing (NCHC)

The National Center for High-performance Computing (NCHC) has been a member of PRAGMA since 2002 and has continuously participated in PRAGMA workshops and associated activities. NCHC is one of the eleven national laboratories affiliated with the National Applied Research Laboratories (NARL). Learning from the PRAGMA collaboration, NCHC has built grid application platforms which have led to extensive involvement from NARL. For example, in addition to NCHC, both the NARL headquarters and the National Space Organization (NSPO) have been deeply involved with the GEO Working Group and have achieved the Aster/Formosat-2 image federation with PRAGMA colleagues at AIST. The National Center for Research on Earthquake Engineering (NCREE) participates in the Telescience Working Group, and through this PRAGMA alliance, NCREE strengthened its relationship with the US NSF Network for Earthquake Engineering Simulation (NEES) program. The Taiwan Ocean Research Institute (TORI)—following the path of the PRAGMA Telescience Working group—with technical support from NCHC, has deployed a coral reef observatory site at Orchid Island that is able to support marine biology research studies in Taiwan, as well as across the globe. In the near future, we anticipate more institutes under NARL with research interests in earth observations will take part in PRAGMA activities and collaborations.

College of Computer Science & Technology (CCST), Jilin University

Jilin University (JLU) is a leading national university under the direct jurisdiction of the Ministry of Education of China. The College of Computer Science & Technology (CCST) of JLU became a member of PRAGMA at the PRAGMA 9 workshop, held in 2006. Since then, PRAGMA has provided CCST with many opportunities to participate in international grid activities and build close collaborations with other PRAGMA members.

CCST joined both the Biosciences and Resource Working Groups of PRAGMA. Each group has a specific focus and projects that enable people to work together and make progress. Through working group activities, CCST is able to feel the pulse of technology development in the high-performance computing (HPC) arena and is able to better know how scientists in other areas use HPC to conduct their research.

CCST is involved in the Avian Flu project, an international project in the Biosciences Working Group, which is aimed at building a comprehensive grid platform to use simulations and in silico screening to support the fight against avian influenza virus. The Avian Flu project provided specific

usage scenarios for meta-scheduling; this pushed CCST to develop a number of useful features in meta-scheduler CSF4, including the following: array job, working flow, data staging in/out, and application aware scheduling.

CCST has established close relationships with the Opal team, Gfarm team, and other research groups in the PRAGMA community. Several CCST faculty members and graduate students have become short-term visiting researchers at UCSD. In turn, UCSD researchers have visited CCST on several occasions. In the last few years, CCST has co-published approximately ten papers with PRAGMA researchers in international journals and conferences, such as New Generation Computing, e-Science, Grid and Cloud Computing (GCC), Parallel and Distributed Processing with Applications (ISPA), Grid Computing and Applications (GCA), Grid and Distributed Computing (GDC), and others.

Above all, PRAGMA brings great value to CCST in terms of broadening international collaboration, engaging applications, publishing scientific papers, and cultivating student.

Monash University

Monash University has benefited from involvement in PRAGMA since its inception in 2002. The Monash e-Science and Grid Engineering (MeSsAGE) Lab specializes in the development of software tools and techniques for programming e-Science applications. e-Science is characterized by computationally intensive modeling and/or large data processing requirements. Due to the scale of computing required to support e-Science it is usually carried out across distributed environments such as grids and clouds, which requires high levels of collaboration between application scientists and distributed-computing specialists.

PRAGMA provides a unique incubator for this kind of collaborative science. Researchers in various disciplines from leading institutions provide a diversity of expertise, ripe with scientific opportunity. MeSsAGE Lab focuses on application development motivated by underlying scientific goals, rather than on technology for its own sake. This is why PRAGMA involvement is so important for us—there are always PRAGMA members working on interesting problems ready to push the boundaries of what we do.

This year, MeSsAGE Lab had the pleasure of hosting five students participating in the PRIME program, all of them working on projects related to e-Science. PRIME is another great PRAGMA opportunity for engagement, plus it has the added benefit of bringing some new friends to the Lab each year. At the PRAGMA 18 Workshop in March 2010, we identified and embarked on a collaboration with the Costa Rica Institute of Technology, modeling ash fallout from volcanic eruptions (please see NG-TEPHRA in the Key Collaborative Successes section). Early work has already been accepted for international publication, with full-scale modeling and a data visualization workflow to follow. We look forward to continuing and new PRAGMA collaborations in 2011.

Images: (above) View of Tokyo Tower from the bottom in the evening—courtesy of Kelly Lo; (below) Traditional Taiwanese Opera at the National Center for Traditional Arts in Yilan County, Taiwan—courtesy of Cato Chan

PARTNERS: ADVANCING TECHNOLOGIES, TRAINING RESEARCHERS FOR THE FUTURE

Environmental Observing

EARLY WORK BY PRAGMA'S TELESCIENCE WORKING GROUP HELPED CREATE THE ECOGRID IN TAIWAN. At the PRAGMA 5 workshop in 2003 (hosted by NCHC), we brought together ecological researchers with PRAGMA researchers. That meeting generated a plan to place a buoy in Yuan Yang Lake (YYL) in the eastern part of Taiwan. Between October 2003 and April 2004, through the concerted efforts of many groups and funding from several sources, a functioning buoy with sensors was placed on YYL and began near-real-time data streaming to the world. Furthermore, these data has formed the basis for several joint publications and student exchanges, with more to come. Through a series of workshops, two new organizations (described below) were developed based on these initial PRAGMA efforts.

Global Lake Ecological Observatory Network (GLEON) has grown dramatically

THERE ARE NOW MORE THAN 200 GLEON MEMBERS AROUND THE WORLD; it has an active graduate student group; proposals have been generated for growing the network; new research directions have been identified that only a global group such as this could tackle; and they continue to develop and deploy new technologies and link models to data to help understand the multiple impacts of changes in the environment on lakes. Like PRAGMA, GLEON has regular working meetings; the most recent ones took place in Torres, Brazil and Nanjing, China

As noted elsewhere in this Collaborative Overview, PRAGMA continues to work with GLEON to provide new technologies for GLEON researchers (GEOGrid, page 7). Another example is that PRAGMA researchers at UCSD, through additional funding from the Gordon and Betty Moore Foundation, have been working to test new technologies at the North Temperate Lakes Long-term Ecological Research (NTL LTER) site. At NTL, six buoys and one meteorological station are maintained; each is equipped with up to 20 sensors producing up to 45 separate data streams at a one- or ten-minute frequency. Traditionally, this data volume has been managed in many matrix-style tables, each described in the Ecological Metadata Language (EML) and accessed online by a query system based on the provided metadata. For developing an updated and more flexible information system, several technologies are currently being investigated and tested. A DataTurbine server is employed to stream data from the data logger files into a database with the Real-time Data Viewer being used for monitoring sensor health. The Kepler workflow processor is being explored to introduce quality control routines into this data stream in order to take advantage of the DataTurbine actor.

KLEON, the Korean Lake Ecological Observatory Network, described in the next section, grew out of PRAGMA introducing its member institution, KISTI, to lake researchers and the broadening awareness of the extent of the use and benefits of cyberinfrastructure in environmental observing networks.

ADDITIONAL INFORMATION: GLEON: **www.gleon.org**; GLEON interactive map: **gleon.org/DB_Lists/gleon_map.php**

Background Image: Racha Yai—courtesy Michael Nekrasov

PRAGMA provides a framework for collaboration which has frequently led to the creation of new activities, research groups, and organizations. In this section, we highlight the progress of two sets of activities: a) environmental observing networks for lakes and coral reefs, (GLEON and CREON respectively) and b) the undergraduate research experiences programs at UCSD (PRIME) and Monash University (MURPA) which involve and depend upon PRAGMA members and in the process, strengthen collaborations between participants. Looking forward, we hope to further strengthen ties between PRAGMA and these diverse groups to drive our development via applications and to engage the minds and enthusiasm of students and junior researchers.

Coral Reef Environmental Observatory Network (CREON)

THE FOCUS OF CREON HAS PRIMARILY BEEN ON SHARING EXPERTISE; developing, deploying, and enhancing different tools for use at several coral reef sites, notably Moorea Coral Reef, the Great Barrier Reef in Australia; and the Kenting and Orchid Island sites in Taiwan, and using data produced at those sites.

Above: Preparing the sensor deployment site at Racha Island—courtesy Michael Nekrasov

This year, we held two workshops at, or temporally near, PRAGMA workshops, one in San Diego in March 2010, the other on Orchid Island in September 2010, which resulted in welcoming researchers from Thailand into CREON (see the next section). At the most recent meeting on Orchid Island, hosted by NCHC, attendees were able to see the new installation by a team from the Taiwan Ocean Research Institute (TORI) and NCHC (see National Applied Research Laboratories (NARL) in the Value to PRAGMA section). At the Orchid Island meeting it was decided that equipment will be loaned to Racha Island researchers by researchers in Australia for a sampling season in order to gain experience and sufficient data to build the case for a permanent sensor at the Racha Island site.

Enlarging the CREON Network

UCSD undergraduate Michael Nekrasov (PRIME 2009) spent eight weeks at Walailak University working with researchers on developing a new system for near real-time monitoring of the coral reefs at Racha Island. Michael is working with Mullica Jaroensutasinee and Krisanadej Jaroensutasinee and their team of graduate students. Michael's project is part of a larger activity by CREON. Teams from UCSD, the Great Barrier Reef Ocean Observing System, and Walailak University are collaborating by sharing expertise and equipment to install new ocean sensors for measuring temperature, pressure, and conductivity in near real-time. The team plans to demonstrate the operational system at the next CREON meeting in Brisbane, on 10 December 2010.

ADDITIONAL INFORMATION: Coral Reef Environmental Observatory Network (CREON): **www.coralreefeon.org**; Great Barrier Reef Ocean Observing System (GBROOS): **www.imos.org.au/gbroos.html**

PARTICIPATING RESEARCHERS: Mullica Jaroensutasinee, Krisanadej Jaroensutasinee; *Walailak University;* Michael Nekrasov, Peter Shin, Tony Fountain, *UCSD;* Scott Bainbridge, Damien Eggeling, Geoff Page, *GBROOS*

In addition, NCHC hosted a meeting that allowed for researchers to discuss the loaning of equipment from GBROOS to Racha Islandresearchers, and had earlier hosted a PRAGMA Institute that first engages those researchers in CREON.

Korean Lake Ecological Observatory Network (KLEON)

KLEON'S GOAL IS TO ESTABLISH AN ECOLOGICAL OBSERVATORY NETWORK by sharing lake observation resources, managing observatory information, and supporting analysis tools.

To fully understand and predict changes in the ecosystems of lakes, rivers, and wetlands requires a lake ecosystem monitoring system to manage the various kinds and volume of observation data collected from many buoys and sampling sites. We are in the process of installing water condition sensors such as temperature, conductivity, pH, luminescent dissolved oxygen (LDO), total dissolved gas (TDG), and turbidity at selected Korean observatories (lakes: Soyang, Euiam, and Yeongrang; rivers: Soyang, Anyang, Han, and Gapyeong, wetland: Upo). Figure 1 shows the organization of KLEON.

KLEON provides graphs, xml format files, and CSV files on the web to check the sensor data, which is transmitted and available from the buoys in near real time. KLEON also manages the sampling data, such as water quality and plankton, with GLEON's Vega data model. Thirty-four kinds of sampling data types and a variety of plankton data are collected to check the water quality. These collected observations are used to analyze phenomena such as eutrophication, water turbidity, and the cause of fish kill.

KLEON allows for integration of data from various sensors, the observation site, lake, dam, floodgate, and weather. This is important for accurate analysis; for example, sampling data is affected by the conditions at the sampling site, heavy rain, construction, and sampling method of the sensors. The Sensor Model Language (SensorML) of the Open Geospatial Consortium (OGC) is used to describe the sensors' attributes including model number, sampling interval, and the error rate of the sensors. The information from the observation site, lake, and dam is handled with a data management service on the web. To access the floodgate and weather data, a graph and CSV files are also provided on the web interface. A web interface using Google Earth provides spatial information on the observatory such as the positions of sampling sites and the shapes of lakes.

It is our hope that KLEON resources will support the ecological research community, national policy decision making, and foster a better understanding of science. As an example, the results of lake ecosystem health assessments will lead to advances in environmental/ecology research by improving the accuracy of water quality models and updating teaching resources.

PARTICIPATING RESEARCHERS: Kum Won Cho (PI of KLEON), Bu-Young Ahn, Youngjin Jung, *KISTI*; Bomchul Kim, *Kangwon National University*

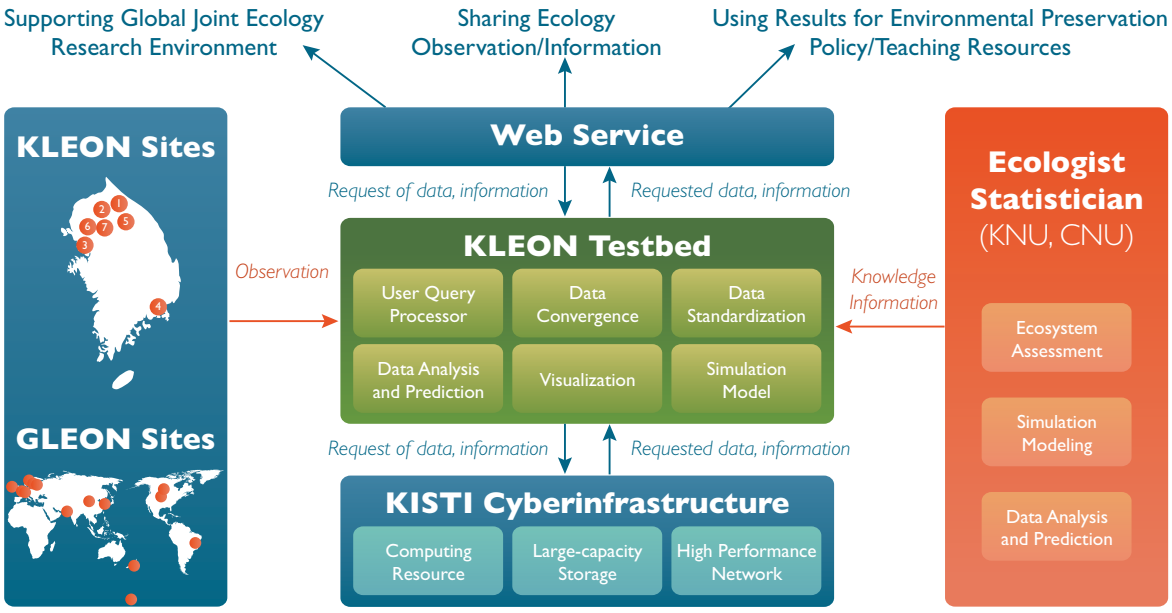
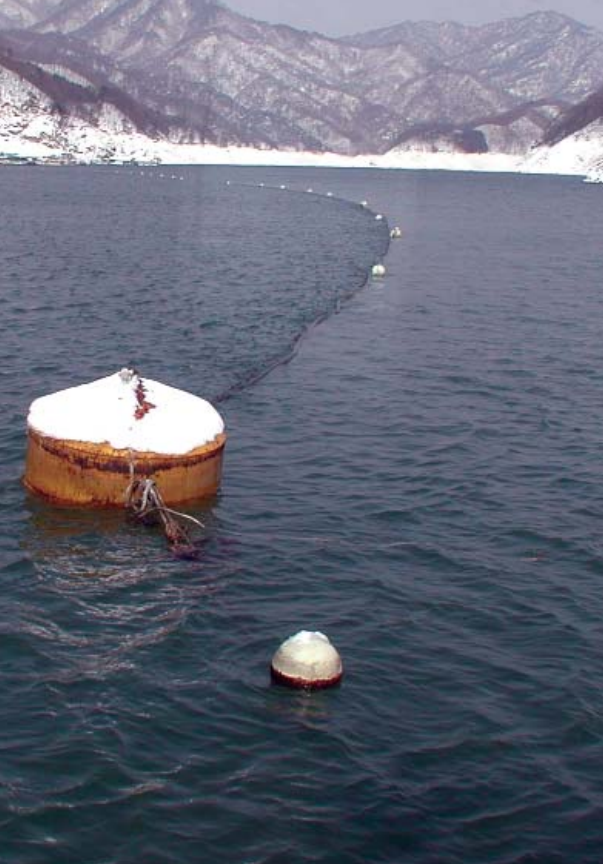


Fig. 1. Organization of KLEON

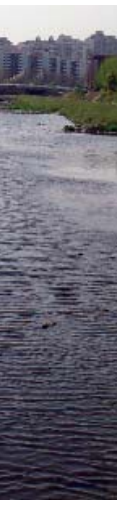



7

1. Soyang River—courtesy of Peter Arzberger
2. Lake Euiam
3. Looking downstream along the Anyangcheon—courtesy of Jdbarrass via Wikimedia Commons
4. Lake Yeongrang
5. Gapyeong River—courtesy of Erich Iseli via Wikimedia Commons
6. Han River—from Wikimedia Commons
7. Waterweed at Upo Wetland—courtesy of Jabon Oh via Wikimedia Commons


6


5




4

Target Place	Sensor Type	Data Collection Method	Analysis
1. Lake Soyang	Troll multiprobe (water temperature, DO, electrical conductivity)	Real-time data transmission CDMA modem	Eutrophication evaluation
1. Soyang River	Hydrolab (water temperature, electrical conductivity, turbidity)	Regular download	Turbid water evaluation
2. Lake Euiam	Combination sensor (water temperature, DO, electrical conductivity, turbidity)	Real-time data transmission CDMA modem	Eutrophication evaluation, turbid water evaluation
3. Anyang River	Troll multiprobe (water temperature, DO, pH, electrical conductivity)	Real-time data transmission CDMA modem	Analysis of causes of fish kill by oxygen deficit
4. Lake Yeongrang	MS5 Hydrolab (water temperature, DO, electrical conductivity)	Real-time data transmission CDMA modem	Analysis of causes of fish kill by oxygen deficit
5. Gapyeong River	Hydrolab (water temperature, electrical conductivity)	Regular download	Long-term ecological research, climate change
6. Han River	Hydrolab (water temperature, electrical conductivity)	Regular download	Long-term ecological research, climate change
7. Wetland Upo	Hydrolab (water temperature, DO, electrical conductivity)	Regular download	Long-term ecological research Eutrophication evaluation

Table 1. Sensor type, data collection method, and analysis

MONASH UNIVERSITY, MELBOURNE

COMPUTER NETWORK INFORMATION CENTER (CNIC), CHINESE ACADEMY OF SCIENCES, BEIJING

UNIVERSITY OF HYDERABAD, HYDERABAD

DOSHISHA UNIVERSITY, KYOTO

NATIONAL INSTITUTE FOR INFORMATION AND COMMUNICATIONS TECHNOLOGY (NICT), TOKYO

OSAKA UNIVERSITY, OSAKA

UNIVERSITI SAINS MALAYSIA, PENANG

UNIVERSITY OF AUCKLAND, AUCKLAND

NATIONAL CENTER FOR RESEARCH ON EARTHQUAKE ENGINEERING (NCREE), TAIPEI

NATIONAL TAIWAN UNIVERSITY (NTU), TAIPEI

TAIWAN FOREST RESEARCH INSTITUTE (TFRI), TAIPEI

Pacific Rim Experience for Undergraduates

THE PACIFIC RIM EXPERIENCE FOR UNDERGRADUATES (PRIME) program was created in 2004 to provide a project-based research internship program, combined with a cultural awareness experience for science and engineering undergraduates at UC San Diego. PRIME grew out of the PRAGMA collaborative framework and people network. PRIME's projects are based on PRAGMA collaborations, as well as additional collaborations between UCSD and PRAGMA researchers. 2010 was the seventh year of the program, sending 29 students to the 11 sites listed on the left. (For an overview of this year's students, please see **prime.ucsd.edu/PRIME2010_table.html**. To view the PRIME 2010 students' progress and final reports on their research and cultural interactions, please go to: **prime.ucsd.edu/student_collections2010.htm**.) To date, more than 125 students have participated in PRIME.

In the summaries to follow, students report on results from several of this year's PRIME projects. To fulfill the cultural aspect, PRIME has instituted a curriculum that starts students on a path of cultural awareness. We use the Intercultural Development Inventory (IDI) to measure the students' attitudes towards cultural differences. More about this index, as well as about PRIME and the PRIUS and MURPA programs, can be found in a soon-to-be-published article:

> PRIME: an integrated and sustainable undergraduate international research program. P Arzberger, G Wienhausen, D Abramson, J Galvin, S Date, FP Lin, K Nan, S Shimojo. *Advances in Engineering Education*. 2010, Vol 2, No 2 (**advances.asee.org/vol02/issue02/05.cfm**).

We are convinced that engaging students in research is essential to grow the next generation of scientists, and that challenging them with an international experience will prepare them to excel in the global workplace and participate in the forefront issues of science and society.

PARTNERS: ADVANCING TECHNOLOGIES, TRAINING RESEARCHERS FOR THE FUTURE

The 1300th Anniversary Celebration at Nara: Visual Technologies to Enhance the Cultural Heritage Experience

This year, three PRIME students, one at Osaka University and two at the National Institute for Information and Communications Technology (NICT), worked towards the same goal: showcase cultural heritage contents using modern technology at the exhibition of the 1300th Anniversary of Nara Heijo-kyo Capital celebrations in Nara, Japan.

Kevin Nguyen (PRIME 2010), hosted by NICT, improved upon the design of the natural user interface (NUI) and built a Fourier transform infrared spectroscopy (FTIR) multi-touch display table, an open-source tracking application and a photoviewer application which is compatible with the touch table, in collaboration with NICT researchers. He modified the application so that users are able to interact with cultural heritage images using the multi-touch user interface.

The other two students, Lex Lacson and Velu Ganapathy (hosted by Osaka University and NICT, respectively) collaborated together on an exhibition at the Nara celebrations. Lex Lacson designed and built a stable and easy-to-use interactive interface for real-time interaction between people and 3-D models. The system uses markerless facial motion capture (camera-based face tracking) to manipulate the perspective of 3-D models. Velu Ganapathy developed a program of 3-D reconstruction from stereo image pairs. The result of reconstruction is rendered through the virtual reality (VR) renderer, openCOVER, then displayed on a tiled-display wall (TDW). OpenCover's virtual camera viewpoint is controlled by the user through face-tracking algorithms developed by Lex's head tracking module. Both results were successfully shown during the demonstration of the interface at the 1300th anniversary of Nara, the old capitol of Japan.

ADDITIONAL INFORMATION: 1300th Anniversary of Nara Heijo-kyo Capital website: **www.1300.jp/foreign/english/kyuuseki/special/index.html**; Natural User Interfaces Forums: **nuigroup.com/forums**

PARTICIPATING RESEARCHERS: Kevin Nguyen, Lex Lacson, Velu Ganapathy, *PRIME 2010 Students, UCSD*; Shinji Shimojo, Taku Morinobu, Yoshinori Kobayashi, Masaki Chikama, *NICT*; Kiyochi Kiyokawa, Haruo Takemura, Susumu Date; *Osaka University*; Jurgen Schulze, Falko Kuester, *Calit2/UCSD*

Images (left to right): Village kids supported by the M. Venkatarangia Foundation sharing an iPad—courtesy of Keith Lee; PRIME students with Merbok locals—courtesy of Serena Chang; Old Summer Palace lotus—courtesy of Wendy Fong; "Eloquent Rain"—courtesy of Nita Chen

PARTNERS: ADVANCING TECHNOLOGIES, TRAINING RESEARCHERS FOR THE FUTURE

Implementation of a Multi-site Virtual Cluster System with a P2P Overlay Network

Grid computing environments harness the power of many clusters to facilitate large computational jobs. One caveat to grid computing environments is that they are heterogeneous—not all clusters on the grid are the same. For example, two clusters on the same grid can differ both in their hardware and software compilers and capabilities. This heterogeneity can cause problems when applications are deployed on the grid. One way to solve this problem is the use of virtual clusters. A virtual cluster is a group of virtual machines mounted on the resources of physical computers. Hardware virtualization hides the physical characteristics of a computing platform, instead presenting users with an abstract computing platform. Because each virtual machine can be held to the same standard, each virtual machine in the cluster will have the same set of hardware capabilities and software, therefore, the end user does not have to deal with heterogeneity. The goal of this project was to deploy a multi-site virtual cluster and analyze its performance.

A multi-site virtual cluster was successfully developed and deployed over two physical clusters located continents apart: Osaka University and UCSD (Figure 1). The virtual machines that comprised the virtual cluster were connected using a virtual network technology consisting of PIAX (peer-to-peer network [P2P] technology) and OpenVPN (open-source virtual private network [VPN] technology). The virtual machines were mounted using Rocks. Initial testing of the network revealed very high latency (500ms or greater) between the nodes of the virtual cluster across the two sites, as well as between nodes located at the same site. The high latency was due to the architecture of PIAX, which had issues recognizing when other nodes came up on the network. Subsequently, we tested two other virtual network technologies: N2N (a layer-two P2P VPN) and IPOP (IP over P2P, an open-source decentralized, distributed VPN application). Both had more reasonable latencies (in the 100-200ms range). Additional performance assessments are underway.

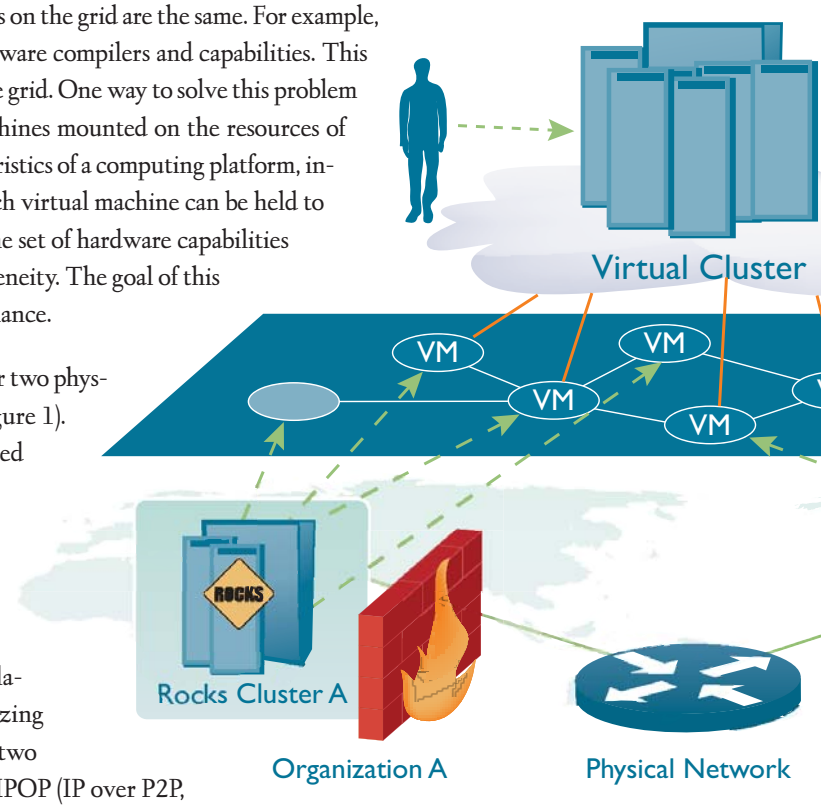


Fig. 1: Diagram outlining the multi-site virtual cluster established between Osaka University and UCSD.

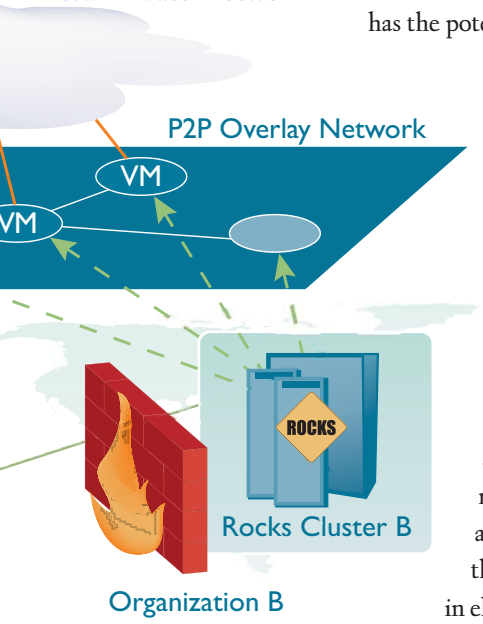
PARTICIPATING RESEARCHERS: Adrian Ho, *PRIME 2010 Student, UCSD*; Susumu Date, Kei Kokubo, Kohei Ichikawa, *Osaka University*; Jason H. Haga, *UCSD*

Application of Watermarking Techniques to Biomedical Research Data

The protection of intellectual property ownership in biomedical research data is rapidly becoming an important area of concern for cloud computing centers. One type of research data that needs to be secure are the results from virtual screenings performed by academic and industry researchers. Virtual screening is a computational method that analyzes the molecular interaction of a chemical compound with a protein target of interest and generates a large amount of plain-text (ASCII) data that describes the spatial orientation of the chemical and protein. These data are used to discover new chemical compounds that may inhibit the activity of a protein target of interest and may have significant intellectual property rights which need to be protected. We developed an algorithm from basic correlation watermarking that embeds and retrieves an encoded message within ASCII data and satisfies the requirements of security, imperceptibility, and robustness.

"We are convinced that engaging students in research is essential to grow the next generation of scientists and that challenging them with an international experience will prepare them to excel in the global workplace and participate in the forefront issues of science and society."

This algorithm first transforms a copyright message into ASCII code, then generates a seemingly nonspecific, pattern-independent, numerical watermark, which is then embedded by directly altering molecular positional information in the files (i.e., the x and y coordinates). The detection and strength (α) of the watermark within the data files was determined by statistical testing that correlates the original data with the watermarked data. The watermark was optimal with α=1.5, which had the lowest detection error probability and was the least visible in the data files. Because virtual screening data files must adhere to a specific, standard file format (known as mol2) in order to be visualized, other methods of watermarking (adding whitespace, additional significant figures, and line swapping) were easily attacked and thus, the watermark was vulnerable to easy removal. In conclusion, this new method of securing biomedical ASCII data is effective and has the potential to be used with other types of biomedical data that are plain-text based.



PARTICIPATING RESEARCHERS: Kai En Tang, *PRIME 2010 Student, UCSD*; Susumu Date, Toru Fujiwara, Maki Yoshida, *Osaka University*; Christopher Lau (PRIME 2009), Jason H. Haga, *UCSD*

Multiple Regression Analysis of Parameter Sensitivity in Shannon-Bers Ventricular Myocyte Model Using Nimrod

The electrophysiological computational models of cardiac ventricular myocytes are complex and involve numerous parameters and component interactions. Therefore, viable methods are required to assess model stability and to learn more about the interactions between various model parameters. In silico experiments are easy to conduct and manipulate and allow investigators to collect large-scale data sets, which enable statistical methods to be used to study model outcomes and qualitatively evaluate parameter sensitivity in relation to the outputs. In addition, statistical methods can correlate multiple variables and identify novel interactions that are oftentimes elusive in in vitro experiments. Sobie suggested that the use of partial least-square regression is an exploratory analysis tool in analysing parameter sensitivity in electrophysiological models (Biophysical J 2009). He demonstrated its potential as a viable method for characterizing and assessing computational models. PRIME 2010 student Yuan-Hung Lo compared two statistical methods of partial least-square regression versus fractional factorial design (Nimrod) by performing a multiple regression analysis of parameter sensitivity in a Shannon-Bers ventricular myocyte model (Biophysical J 2004). He found that these new methods for studying cardiac muscle cells are useful and may help in explaining important disease pathologies and offer insights into potential treatment options.

Eric A. Sobie, Parameter Sensitivity Analysis in Electrophysiological Models Using Multivariable Regression, *Biophysical Journal*, Vol. 96, No. 4, 18 February 2009, pp 1264-1274, ISSN 0006-3495, DOI: 10.1016/j.bpj.2008.10.056.

Thomas R. Shannon, Fei Wang, Jose Puglisi, Christopher Weber, Donald M. Bers, A Mathematical Treatment of Integrated Ca Dynamics within the Ventricular Myocyte, *Biophysical Journal*, Vol. 87, No. 5, November 2004, pp 3351-3371, ISSN 0006-3495, DOI: 10.1529/biophysj.104.047449.

PARTICIPATING RESEARCHERS: Yuan-Hung Lo, *PRIME 2010 Student, UCSD*; David Abramson, Tom Peachey, Blair Bethwaite, Neil Diamond, *Monash Univ.*; Anushka Michailova, *UCSD*

PARTNERS: ADVANCING TECHNOLOGIES, TRAINING RESEARCHERS FOR THE FUTURE

Collaborative Approaches to Extend Observing Systems: Integrating Real-time Audio and Video Data into Open Source Data Turbine and Integrating Event Detection

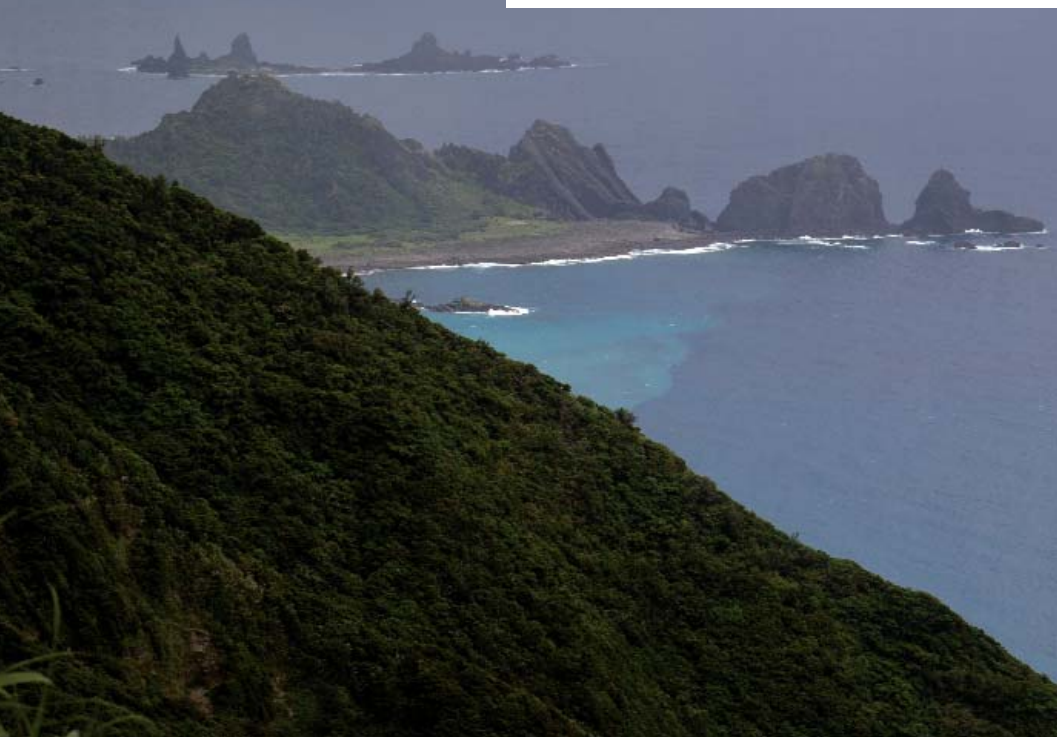
PRIME 2010 students Allison Johnston and Michael Perry, hosted by first time PRIME host—Taiwan Forestry Research Institute (TFRI), worked closely with TFRI scientists Chau Chin Lin and Sheng-Shan Lu to integrate real-time audio and video data into Open Source DataTurbine for use with a variety of applications including bee and wasp interaction and frog species identification. Their development efforts have enhanced the observing system infrastructure at the TFRI research sites and are being used to address important questions in biology and resource management. Specifically, one group of researchers currently using these new tools is interested in understanding the interactions of predatory wasps and honey bees, in another instance, the influence of changing environmental conditions on amphibian populations is being investigated. The system developed is currently deployed at two TFRI field sites.

Two other PRIME 2010 students, Raymond Wingfield and Keith Lee, working at the University of Hyderabad, India, were also involved with integration into DataTurbine. They worked closely with Arun Agarwal on real-time detection and modeling of tsunamis. For event detection, they integrated complex event stream processing technology (the open-source ESPER engine) with the DataTurbine middleware. The modeling aspect of the project involved the use of a cloud computing paradigm.

ADDITIONAL INFORMATION: Open Source DataTurbine: **www.dataturbine.org**

PARTICIPATING RESEARCHERS *(Real-time Audio and Video Data)*: Allison Johnston, Michael Perry, *PRIME 2010 Students, UCSD;* Chau Chin Lin, Sheng-Shan Lu, *Taiwan Forestry Research Institute (TFRI);* Tony Fountain, *SDSC/UCSD*

PARTICIPATING RESEARCHERS *(Integrating Event Detection)*: Raymond Wingfield, Keith Lee, *PRIME 2010 Students, UCSD;* Arun Agarwal, *U. Hyderabad, India;* Sameer Tilak, *SDSC/UCSD*



3-D Protein Modeling on a Grid Computing Environment

The dual-specificity phosphatase (DSP) family is a group of enzymes that is responsible for catalyzing the dephosphorylation of tyrosine/serine amino acid residues of other molecules. An inhibitor of DSPs has enormous potential as a tool for research and even potentiallyas a pharmaceutical molecule, however, among the DSP family, roughly half have unknown tertiary structure. Since a potential inhibitor must be tested for all of the members of the family, the only reasonable option is to computationally determine the 3-D structure via the resources of in silico studies and use this predicted structure in subsequent virtual screening experiments. This project deployed and tested a grid-enabled 3-D modeling program (MODELLER) on the PRAGMA Grid.

The homology simulator MODELLER was used because of its built-in task-based parallel support functions. Opal-OP was used to wrap the MODELLER application as a web service and Perl scripts were created to utilize the mpich parallel environment with the SGE scheduler used to execute the application and for data handling. During a simulation, MODELLER produces thousands of 3-D models across multiple clusters in parallel. The best models are determined and are subsequently refined to achieve a more accurate 3-D model. A DSP with known3-D structure, SSH-2, was used to test our grid-enabled modeling program. The total time to generate the predicted structure was six days. As shown in Figure 1, there is good qualitative agreement between the predicted and known 3-D structures. In summary, we have created a grid-enabled version of MODELLER that produces accurate 3-D representations of DSPs. Using this workflow to determine the unknown tertiary structure of members of the DSP family will help facilitate the search for effective DSP inhibitors in the future.

PARTICIPATING RESEARCHERS: Charles Xue, *PRIME 2010 Student, UCSD;* Susumu Date, *Osaka University;* Matthew K. Mui, PRIME 2009 Student, Jason H. Haga, *UCSD*

Images (left to right): TFRI meteorological station—courtesy of Michael Nekrasov; Dinner with TFRI scientists—courtesy of Michael Nekrasov; A young shepherd learning his father's trade along the roads to Hampi—courtesy of Keith Lee

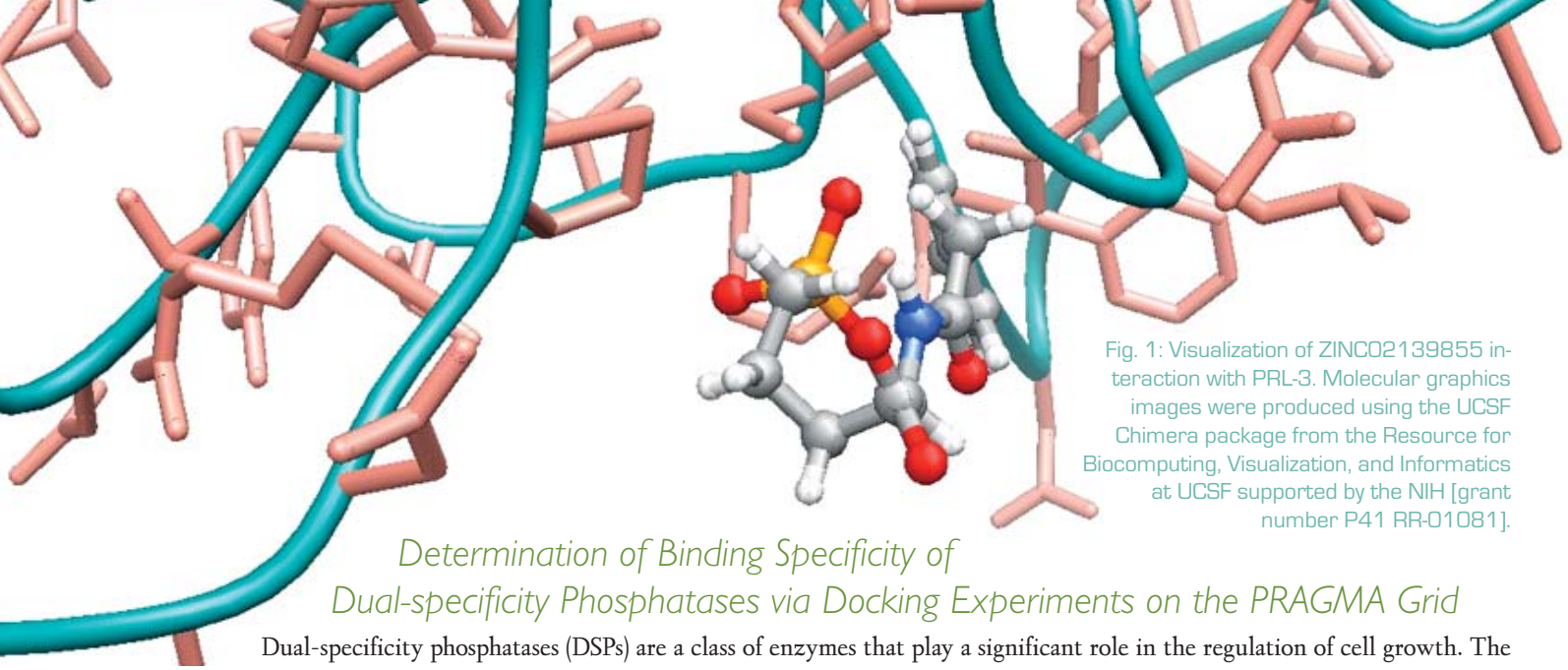


Fig. 1: Visualization of ZINC02139855 interaction with PRL-3. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at UCSF supported by the NIH [grant number P41 RR-01081].

Determination of Binding Specificity of Dual-specificity Phosphatases via Docking Experiments on the PRAGMA Grid

Dual-specificity phosphatases (DSPs) are a class of enzymes that play a significant role in the regulation of cell growth. The slingshot (SSH2) protein is a member of the DSP family; it is responsible for the regulation of cofilin activity, which plays a critical role in the development of Alzheimer's disease and cancer. Even though a list of high-potential inhibitors for SSH2 proteins has been identified in previous experiments, the binding specificity of these potential inhibitors has not been entirely defined, especially in light of the high homology between all DSP family members. Thus, the purpose of this project was to identify the binding specificity of DSP family members via docking experiments using our PRAGMA Grid framework. It is of great interest to identify the binding specificity of these potential inhibitors because it will give new insights into the molecular interactions of small molecule chemicals and various DSP family members. In this experiment, the binding specificity of DSP family members PRL-3 and MKP-6 was screened against ligands from the ZINC database using the molecular docking software, DOCK 6. Top binding ligands for both PRL-3 and MKP-6 were identified and ranking results were compared with SSH-2. According to consensus ranking, the three highest binding compounds for PRL-3 are as follows: ZINC02139855 (Figure 1), ZINC02888956, and ZINC02210369. The three highest-ranking compounds for MKP-6 were ZINC00310164, ZINC00074595, and ZINC02384698, according to the energy score alone. These results, in combination with those from the SSH-2 screening, showed that ZINC02561394 has the highest specificity towards SSH-2, followed by ZINC03906727 and ZINC04226651. Wet-lab experiments will be conducted to verify the efficacy and specificity of ZINC02561394.

PARTICIPATING RESEARCHERS: Kelly Lo, *PRIME 2010 Student, UCSD*; Susumu Date, *Osaka University*; Matthew K. Mui (PRIME 2009), Jason H. Haga, *UCSD*

PRIME Time for International Collaborative Influenza Research

The summer of 2010 became 'PRIME time' for five UC San Diego undergraduate students to conduct computer-aided drug discovery for influenza virus prevention and treatment at three different PRAGMA sites, CNIC in Beijing, USM in Penang, and NTU in Taipei. Thestudents conducted research on several topics including: predicting the pandemic potential of influenza virus mutations in the hemagglutinin (HA) proteins from humans and swine; the effect of sulfation on glycan receptor binding affinity to HA; virtual screening for novel inhibitors of viral membrane fusion to host endosomal membrane, targeting HA and the M2 proton channel. Working with researchers at the different PRAGMA member institutions, as well as mentors back at UCSD, the students experienced first hand the power of cloud computing and grasped complex workflows that support computer-aided drug discovery, using the NBCR CADD pipeline. The NBCR CADD pipeline provides a visually friendly workflow interface to access grid and cloud resources contributed by PRAGMA member institutions. The software stack utilizes the metascheduler CSF4 from JLU to distribute molecular dynamics simulations and virtual screening experiments using AutoDock 4. Moreover, the students benefited from interactions with host institution researchers and learned new algorithms and techniques, which they can apply to their own research projects now and in the future. The students contributed not only to extending cultural exchanges with the people at the participating sites, but also to future progress on influenza virus research, and helped promote international collaboration in the fight against pandemic infectious diseases. The work of two of the students (Cato Chan and Wendy Fong) is highlighted below.

PARTICIPATING RESEARCHERS:

- Identification of Residue Mutations that Increase the Binding Affinity of LSTc to HA RBD (Wendy Fong)
- Comparative Sequence Analysis of Circulating and Pandemic Flu Viruses in Swine and Humans (Alicia Lee)

Wendy Fong, Alicia Lee, *PRIME 2010 Students, UCSD*; Kai Nan, JianjunYu, *CNIC*; Wilfred Li, Jane Ren, *NBCR/UCSD*

- Modeling Glycan Diversity to better understand H1N1 Receptor Binding Specificity (Serena Chang)
- M2 proton channel and potential inhibitors (Qin Li)

Serena Chang, Qin Li, *PRIME 2010 Students, UCSD*; Habibah Wahab, Yusuf Muhammad, *USM*; Wilfred Li, Jane Ren, *NBCR/UCSD*

- Virtual Screening and Analysis on Membrane Fusion Inhibitors targeting Group 2 Hemagglutinin(Chan)

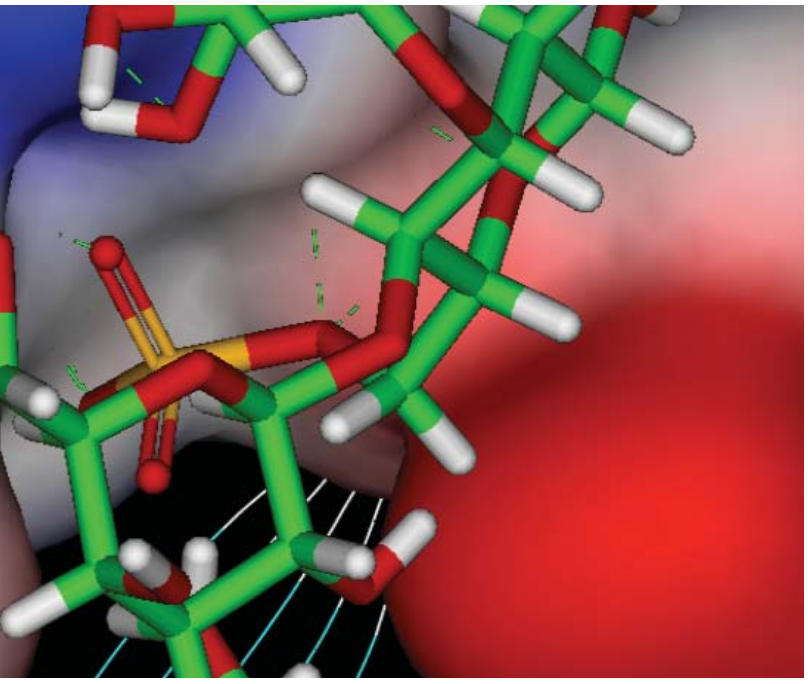
Cato Chan, *PRIME 2010 Student, UCSD*; Jung-Hsin Lin, Yu-Hsuan Chen, *NTU*; Wilfred Li, Jane Ren, *NBCR/UCSD*

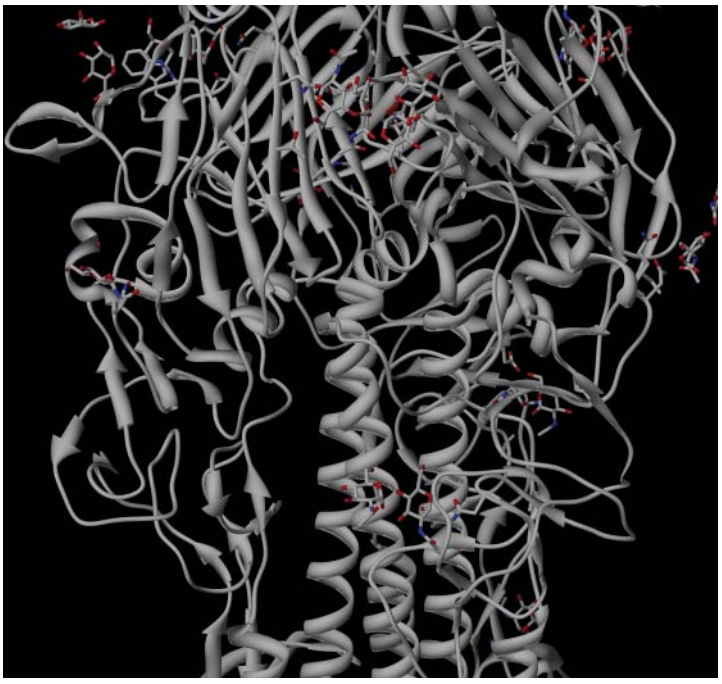
Virtual Screening and Analysis on Membrane Fusion Inhibitors targeting Group 2 Hemagglutinin

Hemagglutinin (HA) is a major membrane envelope surface protein for the influenza virus responsible for host cell attachment and and membrane fusion. Due to its importance in the viral life cycle, it has been the target of many experimental drugs including tert-butyl hydroquinone (TBHQ). TBHQ is a hydrophobic molecule capable of binding to conserved residues in the HA fusion domain which are located between two alpha helices of the HA trimer. The binding of TBHQ to the HA has been shown to prevent the conformational changes necessary for membrane fusion, thus interfering with viral capsid disassembly within the host.

An in silico approach was adopted to understand the HA-TBHQ interactions and to search for molecules with similar properties, which could be potentially more potent inhibitors. Control docking experiments with Protein Data Bank (PDB) structure 3EYM (H3 Hemagglutinin) [1] initially determined that Autodock Vina, Autodock 4 (with and without RESP charges), and Autodock2MMGBSA did not reproduce the binding conformation present in the co-crystal structure. Due to the poor results from the control experiments, the binding mode of TBHQ was investigated with a short (200 ps) molecular dynamics (MD) simulation using the Generalized Born Implicit Solvent Model. Analysis of this MD sim-

Images: (left) PRIME researchers found through computer simulation that sulfation of Sialyl Lewis X (SLe-X) increased its interaction with the H1N1 hemagglutinin (H1) protein. The SLe-X is shown is ball and stick representation within the H1 binding domain. Hydrogen bonds are shown as dashed green lines—courtesy of Serena Chang. (right) Protein structure of the H3 subtype that was mutated at the RBD—courtesy of Wendy Fong.





ulation suggests that key interactions discussed in an earlier paper[1] are not present in the simulated solution environment and may only be present under crystallizing conditions. Further analysis of the MD simulation and additional MD simulations will need to be conducted to confirm these findings.

Minimized PDB structure of H3 (Hemagglutinin) in complex with tert-butyl hydroquinone (TBHQ). Hemagglutinin is depicted blue and TBHQ is depicted as green. The original PDB file is 3EYM—courtesy of Cato Chan

To optimize the structure for docking, an iBelly minimization was conducted. Subsequent control dockings showed that the minimized structure was acceptable for use in virtual screening experiments with Autodock Vina. An Autodock Vina virtual screening of four drug banks, including FDA approved, Nutraceuticals, Small Molecule and NCIDS II, determined several potential drug candidates from a pool of 8198 compounds. The majority of these compounds were large flexible hydrophobic compounds with some polar functional groups. The control experiments with 3EYK (H14)[1] successfully validated the use of Autodock Vina for future virtual screening experiments.

[1]Russell, R.J. et al. Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion. Proceedings of the National Academy of Sciences 105, 17736-17741 (2008).

Identification of Residue Mutations That Increase the Binding Affinity of LSTc to HA RBD

As noted in previous studies, hemagglutinin (HA) is a major membrane envelope surface protein for the influenza virus, instrumental to cellular uptake and membrane fusion with the host. By focusing on mutating residue sets in the HA receptor binding domain (RBD) to artificially increase the binding affinity of a human glycan receptor analogue (LSTc) to the HA RBD, we can further understand the mechanisms behind glycan recognition and binding. The target residue positions are known to have high affinity for an avian glycan receptor analogue (LSTa). Hence, by mutating these amino acids to ones that favor LSTc-binding, not only are the sites for LSTa interaction weakened or blocked, but the interaction strength with LSTc is also increased. Most of the chosen residues also contained hydrogen bonds important for sialic acid contact, which is necessary for viral infection. When mutations were introduced to disrupt or preserve these hydrogen bonds based on amino acid structures, HA exhibited more favorable binding to LSTc.

The experiments started with single mutations, but reached up to quintuple mutations in order to evaluate the effects of combining residue changes that seemed to show preference for LSTc. AutoDock Vina was used to assess and compare the strength of the interaction (measured in kcal/mol) between the ligands and the mutated RBD. All of the potential mutations were first docked with the human glycan receptor analogue before they were docked with the avian glycan receptor analogue, in order to verify that there was, indeed, a greater affinity for LSTc over LSTa. The different outputs of the analogue were compared against the corresponding original structure isolated from the protein structure. The sialic acid portion of the ligand was used to see which docking pose was closest to the original, an indication that it might contribute to preferential LSTc-binding. The docking poses closest to the original were then rescored using AutoDock2MMGBSA to better compare the binding affinities.

These experiments can ultimately help us further understand the mechanisms behind glycan binding and species specificity switch in the influenza virus so that it may be prevented from infecting different species, in particular, humans.

Using Static Pushover Procedures in PISA4SB for Applications in the Taiwan School Building Seismic Retrofit Program

Taiwan is located in an earthquake-prone zone on the Pacific Rim. In 1999, the Chi-Chi earthquake destroyed nearly half of the school buildings in central Taiwan. This demonstrated the need to evaluate the seismic resistance of existing school buildings. The National Center for Research on Earthquake Engineering (NCREE) has developed a method for detailed seismic evaluation using the Platform for Inelastic Structural Analysis of School Buildings (PISA4SB) software platform. This project focused on the application of pushover procedures in PISA4SB for the evaluation of school buildings.

The school building investigated in this project was the Puying building at the Lujhou Junior High School in northern Taiwan. The building is comprised of reinforced concrete columns and beams, as well as brick windowsills and partitions. The building was analyzed under dead load, live load, and pushover load cases in a nonlinear static pushover analysis. Through the analysis, the building's capacity was found and compared to the seismic resistance requirements for the area. The seismic resistance requirement for a 475-year earthquake in the 1st Taipei District region in Taiwan was calculated to be AT = 0.24g. This means that the peak ground acceleration (Ap) for buildings in this location must be greater than 0.24g to have sufficient seismic resistance. From the pushover analysis with PISA4SB, the building was found to be lacking in seismic resistance in the x-direction with an Ap = 0.217. The results from PISA4SB were compared to an analysis previously performed by professional engineers using ETABS and found to be similar, indicating that additional reinforcement is required in the x-direction. These results demonstrate that PISA4SB can be a useful application for the analysis of school buildings.

ADDITIONAL INFORMATION: NCREE: **www.ncree.gov.tw/eng**

PARTICIPATING RESEARCHERS: Vicki Lai, *PRIME 2010 Student, UCSD;* KC Tsai, MC Chuang, Edward Liao, Tzu-Kang (William) Lin, *NCREE;* Lelli Van Den Einde, *UCSD*

Fig. 1. Puying Building, Lujhou Junior High School, northern Taiwan

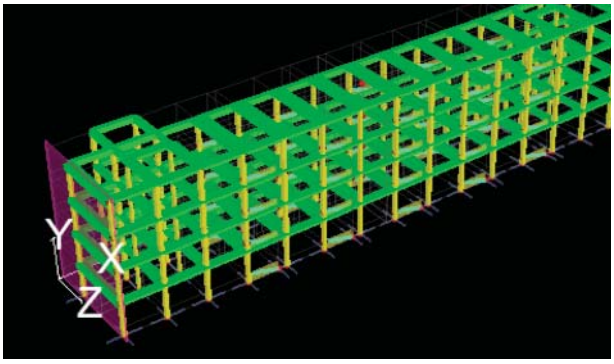


Fig. 2. Modeled Building in PUSHX Direction

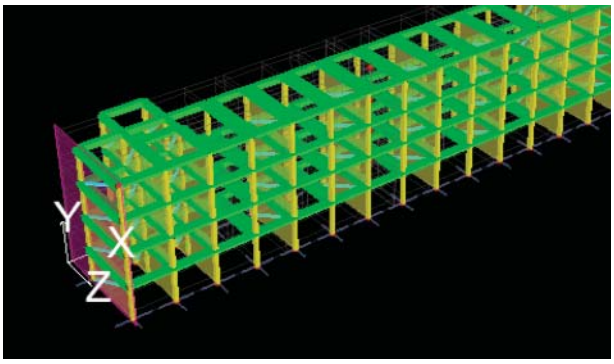


Fig. 3. Modeled Building in PUSHZ Direction

MURPA Advances 2010 - 2011

As with PRIME, the Monash Undergraduate Research Projects Abroad (MURPA) program supports summer (i.e., January-March down under) internships with leading research groups overseas. It not only provides a research experience at the undergraduate level, but does so in an international context. Students are placed for a period of eight weeks, allowing them to integrate into the research groups as team members. MURPA also involves an advanced seminar scheme, in which students attend seminars given by world-leading experts, which is particularly helpful before they embark on their research adventures abroad. The seminar structure is novel, because it uses a cutting-edge high-definition interactive video link to the UC San Diego division of Calit2 making it feasible to attract some of the world's best researchers "virtually" to Monash. These seminars also allow students to "meet" potential UCSD mentors and learn more about potential projects.

In March 2010, MURPA and Calit2 (and their respective universities) were recipients of the 2010 Innovations in Networking Award in the category of High Performance Research Applications from the Corporation for Education Network Initiatives in California (CENIC) for "Enhancing Student Exchange Experiences with High Definition Video Conference." The award was made by the organizing committee for the CENIC 2010 conference. The awards are given annually for the innovative use and expansion of high-performance networking. This HD video streaming also allows mentors at UCSD to attend the final student presentations of the PRIME students who are hosted by the MeSsAGE lab at Monash. More on our use of HD interactive streaming is below.

MURPA has entered its third year; eight students have gone through its ranks to date. In 2010-11, the range of sites available for students will expand. Adding to UC San Diego (previous and continuing) as a host site will be the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign. NCSA projects are expected to focus on the Blue Waters supercomputer and advanced applications. During 2011, we will also extend MURPA-related seminars to include the Technion (Israel Institute of Technology, Haifa) and the University of Warwick (UK), and may offer internships at these universities as well.

Some MURPA students have continued their summer internship projects into their fourth-year honors research project after they return to Australia, providing a much longer, and more internationally relevant, research project than the norm. This year we further leveraged research projects across the two undergraduate programs (MURPA and PRIME): a PRIME student (hosted at Monash) began work in June that was a continuation of a project conducted by a MURPA student in January and February (at UCSD). This allowed more than one student tocontribute to the project's outcome, while still providing continuity and progress on the project during the year.

In 2009, MURPA held a Cardiac Modeling Workshop at Monash which included speakers from Oxford University, the University of Queensland, the University of Auckland and UCSD. Some speakers were physically present at Monash, while others attended via interactive HD video streaming. A number of PRIME students, who had already returned to the US, also gave presentations at the workshop, alongside very senior presenters.

Typically, around 30-40 students and staff attend MURPA lectures, at times, the attendance has been as high as 80. Archives oftalks are available on the MeSsAGE Lab website (link below). In 2010 we added a series of lectures from NCSA, mostly focused on the Blue Waters project, in addition to the lectures from UCSD/Calit2.

Recognition of MURPA is growing, and the program now attracts students from the Faculties of Engineering and Science, in addition to Information Technology students. In 2010, MURPA had a conference paper (below) accepted based on their summer project; in addition, a MURPA graduate has started a PhD at Monash.

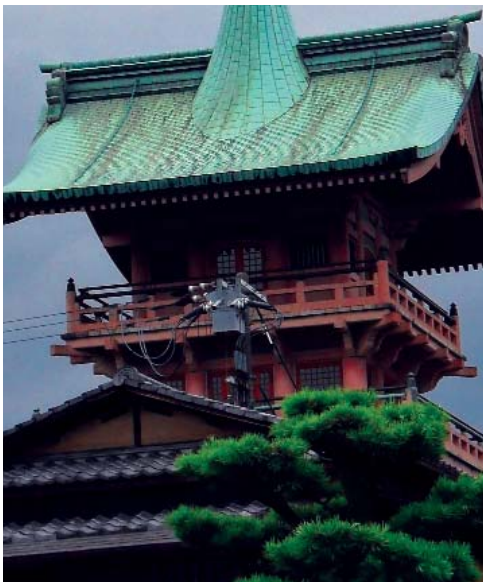
Aisa Na'im, Daniel Crawl, Maria Indrawan, Ilkay Altintas, Shulei, Monitoring data quality in Kepler, in the *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, Chicago, Illinois, Managing Data Quality in Collaborative Science (MDQCS) workshop*, pp 560-564, 2010.

We look forward to continuing and growing the success of MURPA. There is a clear need for external funding; it is critical for the continuation of MURPA. Therefore, two external funding applications were submitted this year.

ADDITIONAL INFORMATION: Monash Undergraduate Research Projects Abroad (MURPA): **messagelab.monash.edu.au/MURPA**; Monash eScience and Grid Engineering (MeSsAGE) Lab : **www.messagelab.monash.edu.au**

Three students hosted at Osaka University and NICT participated in adapting today's state-of-the-art technologies to showcase the cultural heritage of a city more than a thousand years old at the 1300th Anniversary Celebration at Nara.

Five students hosted at three different PRAGMA sites, CNIC, USM, and NTU, conducted computer-aided drug discovery for influenza virus prevention and treatment.



Images: (clockwise from upper left) The pagoda in the Kyomisu, Kyoto "Skylight"—courtesy of Nita Chen; PRIME students in yukata enjoying food from the festival street vendors during the Gion Festival—courtesy of Kelly Lo; Ajanta Caves monkey surviving off donated food—courtesy of Keith Lee; PRIME alumni, faculty and staff welcome PRIME 2010 students as they return to life and studies at UCSD—courtesy of Teri Simas

WORKING GROUPS, WORKSHOPS AND INSTITUTES

PRAGMA workshops are meetings of all members of the PRAGMA community. They are the major vehicle where information is exchanged between working groups, researchers, and institutions and they provide excellent opportunities to engage new researchers and students at the host sites. They are organized by the four working groups in PRAGMA, which are as follows:

- RESOURCES WORKING GROUP: working to make the distributed resources of PRAGMA useful to diverse applications. Co-leaders: *Yoshio Tanaka* (AIST), *Mason Katz* (SDSC/UCSD), and *Cindy Zheng* (SDSC/UCSD)
- TELESCIENCE WORKING GROUP: focusing on a variety of activities that require access to, or use of, remote equipment, such as tiled-display walls and sensors. Co-leaders: *Shinji Shimojo* (NICT and Osaka University) and *Fang-Pang Lin* (NCHC)
- GEO WORKING GROUP: creating an infrastructure to share and integrate data on global earth observations, including remote sensing data and data from land-, lake-, and ocean-based sensors. Co-leaders: *Ryosuke Nakamura* (AIST), *Franz Cheng* (NARL), and *Sornthep Vannarat* (NECTEC)
- BIOSCIENCES WORKING GROUP: focusing much of its efforts over the last several years on integrating technologies to create an infrastructure to advance the screening of potential compounds to combat infectious diseases. Co-leaders: *Wilfred Li* (NBCR/UCSD) and *Habibah Wahab* (USM)

In 2010 two PRAGMA Workshops were held.

- PRAGMA 18, hosted by the University of California, San Diego, 2-4 March 2010 in San Diego
- PRAGMA 19, hosted by Jilin University, 13-15 September 2010 in Changchun

PRAGMA Workshops are also used to host other targeted activities, such as the case at PRAGMA 18, where a Workshop on Coral Reef Monitoring Technologies was held and later, a coordinated workshop on CREON Information Technology was held prior to PRAGMA 19 (see page 15 for more information).

In addition to workshops, PRAGMA has a tradition of disseminating information and tools to a broader community through its Institutes. The two most recent PRAGMA Institutes have been held jointly with NCHC's *Southeast Asia International Joint Research and Training Program in High-Performance Computing Applications and Networking Technology*; another joint program will take place December 2010 (see below). This joint meeting has been instrumental in engaging more institutions in PRAGMA and has helped expand the reach of CREON to Thailand.

Future Meetings:

- 6th PRAGMA Institute, hosted by NCHC, 6-10 December 2010 in Taichung; joint Institute with NCHC's Southeast Asia Program (event.nchc.org.tw/2010/southeast_asia)
- PRAGMA 20, hosted by the University of Hong Kong, 2-4 March 2011 in Hong Kong
- PRAGMA 21, hosted by AIST in Fall 2011 in Sapporo

Background Image: Port of Kobe night scene. The Kobe Port Tower to the left and the Maritime monument to the right are illuminated at night along with other buildings around the bay—courtesy of Francis Duldulao

PRAGMA is an institution- and people-based organization governed by a Steering Committee that invites new members, determines locations of workshops, and sets overall direction. More information about Steering Committee members (denoted with an asterisk * in the listing below) may be found at www.pragma-grid.net/about/committee.

INSTITUTIONS AND THEIR SPONSORS

Active Members

A key component of PRAGMA is active involvement, by participation in workshops, contributing resources, hosting workshops, and/or promoting and supporting student and researcher exchanges. The following institutions have contributed to PRAGMA activities in the past year.

ACADEMIA SINICA GRID COMPUTING CENTRE (ASGCC): Simon Lin, *sclin@gate.sinica.edu.tw*; Eric Yen, *eric@sinica.edu.tw*

ADVANCED SCIENCE AND TECHNOLOGY INSTITUTE (ASTI): Denis Villorente, *denis@asti.dost.gov.ph*; Grace Dy Jongco, *gracedj@asti.dost.gov.ph*

BeSTGRID NEW ZEALAND (BeSTGRID): Nick Jones, *n.jones@auckland.ac.nz*

CENTER FOR COMPUTATIONAL SCIENCES (CCS), UNIVERSITY OF TSUKUBA: Osamu Tatebe, *tatebe@cs.tsukuba.ac.jp*; Taisuke Boku, *taisuke@cs.tsukuba.ac.jp*; Mitsuhisa Sato, *msato@cs.tsukuba.ac.jp*

CENTER FOR HIGH PERFORMANCE COMPUTING, HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY (HUT): Thanh-Thuy Nguyen, *thuynt@soict.hut.edu.vn*

COLLEGE OF COMPUTER SCIENCE AND TECHNOLOGY (CCST), JILIN UNIVERSITY (JLU): Xiaohui Wei*, *weixh@jlu.edu.cn*

COMPUTER NETWORK INFORMATION CENTER (CNIC), CHINESE ACADEMY OF SCIENCES (CAS): Baoping Yan*, *ybp@cnic.ac.cn*; Kai Nan*, *nankai@cnic.ac.cn*

CYBERMEDIA CENTER (CMC) AND RESEARCH CENTER FOR ULTRA-HIGH VOLTAGE ELECTRON MICROSCOPY, OSAKA UNIVERSITY: Shinji Shimojo*, *shimojo@cmc.osaka-u.ac.jp*; Susumu Date*, *date@ais.cmc.osaka-u.ac.jp*

HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY (HCMUT): Thanh-Son Nguyen, sonsys@hcmut.edu.vn; Tran Vu Pham, *t.v.pham@cse.hcmut.edu.vn*

INSTITUTE OF HIGH PERFORMANCE COMPUTING (IHPC): Terence Hung, *terence@ihpc.a-star.edu.sg*

INSTITUTE OF INFORMATION TECHNOLOGY-VIETNAM (IOIT-VN): Thai Quang Vinh*, *qvthai@ioit.ac.vn*

KASETSART UNIVERSITY (KU): Putchong Uthayopas*, *pu@ku.ac.th*

KONKUK UNIVERSITY (Konkuk): Karpjoo Jeong, *jeongk@konkuk.ac.kr*

KOREA INSTITUTE OF SCIENCE AND TECHNOLOGY INFORMATION (KISTI): Kum Won Cho*, *ckw@kisti.re.kr*

MONASH UNIVERSITY (Monash): David Abramson*, *david.abramson @infotech.monash.edu.au*

NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING (NCHC), NATIONAL APPLIED RESEARCH LABORATORIES (NARL): Whey-Fone Tsai*, *wftsai@nchc.narl.org.tw*; Fang-Pang Lin*, *fplin@nchc.narl.org.tw*

NATIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS (NCSA), UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN (UIUC): Radha Nandkumar*, *radha@ncsa.uiuc.edu*; Danny Powell, *danny@ncsa.uiuc.edu*

NATIONAL ELECTRONICS AND COMPUTER TECHNOLOGY CENTER (NECTEC): Piyawut Srichaikul, *piyawut.srichaikul @nectec.or.th*; Sornthep Vannarat, *sornthep.vannarat@nectec.or.th*

NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY (AIST): Satoshi Sekiguchi*, *s.sekiguchi@aist.go.jp*; Yoshio Tanaka*, *yoshio.tanaka@aist.go.jp*

UNIVERSITI SAINS MALAYSIA (USM): Habibah A. Wahab*, *habibahw@usm.my*

UNIVERSITY OF CALIFORNIA, SAN DIEGO (UCSD): including the CALIFORNIA INSTITUTE FOR TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY (Calit2), SAN DIEGO SUPERCOMPUTER CENTER (SDSC), CENTER FOR RESEARCH IN BIOLOGICAL SYSTEMS (CRBS), NATIONAL CENTER FOR MICROSCOPY AND IMAGING RESEARCH (NCMIR), NATIONAL BIOMEDICAL COMPUTATION RESOURCE; Peter Arzberger*, *parzberg@ucsd.edu*; Philip Papadopoulos*, *phil@sdsc.edu*; Mason Katz, *mjk@sdsc.edu*; Teri Simas, *simast@sdsc.edu*

UNIVERSITY OF HONG KONG (HKU): W.K. Kwan*, *kwk@cc.hku.hk*; P.T. Ho, *hpt@cc.hku.hk*

UNIVERSITY OF HYDERABAD (UoH): Arun Agarwal, *aruncs@uohyd.ernet.in*


Sunset near Asakusa station in Tokyo—courtesy of Kai En Tang

Networking Members

Networking partners provide access to expertise to improve the efficiency of the resources groups in running distributed experiments and applications.

ASIA-PACIFIC ADVANCED NETWORK: (APAN): Seishi Ninomiya, *snino@isas.a.u-tokyo.ac.jp*; Kento Aida, *aida@nii.ac.jp*

PACIFIC WAVE: Jacqueline Brown, *jbrown@ms.uw.edu*

STARLIGHT AND TRANSLIGHT/STARLIGHT INITIATIVES: Maxine Brown, *maxine@uic.edu*; Tom DeFanti, *tom@uic.edu*

TRANSPAC2, INDIANA UNIVERSITY: James Williams*, *william@indiana.edu*; John Hicks, *jhicks@iupui.edu*

Industry Member

We gratefully acknowledge the contribution & support of our industrial partner.

CRAY INC.: Andrew Wyatt, *andrew.wyatt@cray.com*

Other Members

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN SUPERIOR DE ENSENADA (CICESE): Salvador Castañeda, *salvador@cicese.mx*; José Lozano, *jlozano@cicese.mx*

GLOBAL SCIENTIFIC INFORMATION AND COMPUTING CENTER (GSIC), TOKYO INSTITUTE OF TECHNOLOGY (Titech): Satoshi Matsuoka*, *matsu@is.titech.ac.jp*; Hidemoto Nakada, *hidenakada@aist.go.jp*

MALAYSIAN INSTITUTE OF MICROELECTRONIC SYSTEMS (MIMOS): Thillai Raj T. Ramanathan, Ng Kwang-Ming, *kwang.ming@mimos.my*; Luke Jing Yuan, *jyluke@mimos.my*

NATIONAL AGRICULTURAL RESEARCH CENTER (NARC): Seishi Ninomiya, *snino@affrc.go.jp*; Masayuki Hirafuji, *hirafuji@affrc.go.jp*

NATIONAL GRID OFFICE (NGO): Hing Yan Lee, *hingyan@ngp.org.sg*

More information about each of the PRAGMA Institutional Members can be found at **www.pragma-grid.net/about/institutions**.

Additional Organizations Active in PRAGMA

Instituto Tecnológico de Costa Rica (ITCR) is a national autonomous institution of higher education, dedicated to the teaching, research and extension of technology and associated sciences aimed at the development of Costa Rica. Researchers there have been steadily building the cyberinfrastructure of Costa Rica and building ties with PRAGMA partners in particular regarding cloud computing. Concerted efforts to create a massively distributed PRAGMA Grid/cloud workflow for volcanic ash-dispersion simulation and risk management, named NG-TEPHRA, in close collaboration with the MeSsAGE Lab at Monash University are in progress (see Key Collaborative Successes for more information).

Lanzhou University (LZU; **www.lzu.edu.cn**), has contributed resources to the PRAGMA Grid, attended the PRAGMA 12, 13, 17, 18, and 19 workshops, and hosted PRAGMA visitors this year.

National Applied Research Laboratory (NARL; **www.narl.org.tw/en**) was established in 2003 to consolidate 11 national laboratories into a single nonprofit organization to construct, operate, and maintain the large-scale R&D facility and platform in support of academic research, and foster the necessary manpower in various advanced fields focused by the nation. NCHC is one of the laboratories in NARL. NARL has provided leadership in the GEO Working Group, and can bring to bear several other laboratories at NARL for PRAGMA collaborations.

National Institute for Information and Communications Technology (NiCT; **www.nict.go.jp/index.html**), is an incorporated administrative agency that conducts general research and development on information technology supporting the ubiquitous society of the future. NiCT supported students in the PRIME program in 2009 and 2010 and has participated in the activities of the Telescience Working Group through support of the high-definition video conferencing testing.

INSTITUTIONS AND THEIR SPONSORS

PRAGMA is supported by its member institutions and the U.S. National Science Foundation (NSF OCI-0627026, PI: Papadopoulos and co-PI: Katz) and involves support from NSF's Office of Shared Cyberinfrastructure, Office of International Science and Engineering, Division of Information and Intelligent Systems, and Division of Biological Infrastructure.

PRAGMA Sponsors

ASGC is an e-Science center funded by both the Academia Sinica and National Science Council of Taiwan.

ASTI'S grid activities are funded by the Department of Science and Technology (DOST), Philippines.

CMC/OSAKA UNIVERSITY is supported by JGN2 of the National Institute of Information and Communications Technology (NiCT), Japan. Support for "Fostering of Globally-leading Researchers in Integrated Sciences" (PRIUS) is provided under the MEXT framework of "University Education Internationalization Promotion Program," and is promoted by the Graduate School of Information Science and Technology at Osaka University.

CCS'S PRAGMA participation is partially supported by the MEXT Grant-in-Aid for Scientific Research on Priority Areas, "New IT Infrastructure for the Information-explosion Era" (Grant number 21013005).

CCST receives funding support from the Chinese Natural Science Foundation (60703024) and the Chinese Ministry of Education (NCET-09-0428).

CNIC receives funding from the Ministry of Science and Technology of China through the China National Grid (CNGrid). Scientific Data Grid's (SDG) major funding and the Duckling software are supported by the Chinese Academy of Sciences.

AIST'S sponsors include the Special Coordination Funds for Promoting Science and Technology (MEXT, Japan) and the Core Research of Evolutional Science and Technology Program (JST, Japan).

HUT receives support from the Hanoi University of Science and Technology and Ministry of Science and Technology, Vietnam.

IHPC receives its funding mainly from Singapore's Agency for Science and Technology Research (A*STAR).

KU's PRAGMA participation has been partly funded by an SRU Grant, Kasetsart University Research and Development Institute (KURDI) and the National Research Council of Thailand.

Background Image: Mt. Fuji covered with clouds in the backdrop of the valley down below, Hakone, Japan—courtesy of Francis Duldulao

KISTI receives major funding from MOST through the K e-Science and Ministry of Information and Communication (MIC) through Korean Grid Infrastructure Implementation and Middleware Development Project (K*Grid).

MONASH UNIVERSITY PRAGMA activities (including MURPA) are supported by a range of grants from the Australian Research Council and Monash University internal funding.

NARC receives major funding from the Ministry of Agriculture, Forestry and Fishery, Japan, and the Japan Science and Technology Agency.

NCHC receives major funding support from the National Science Council, Taiwan.

NCSA is a high-end computing center funded by NSF, the state of Illinois, the University of Illinois, industrial partners, and additional U.S. federal agencies.

NECTEC receives its funding through Thailand's National Science and Technology Development Agency (NSTDA).

NGO receives funding from Singapore's A*STAR and the Infocomm Development Authority (IDA).

TRANSLIGHT/STARLIGHT receives major funding from NSF (OCI-0441094).

TRANSPAC2 receives major funding from NSF and Japan's National Institute of Information and Communications Technology.

USM'S grid activities in Malaysia are funded mainly through E-science and Marine Genomics and Natural Product Discovery National Top Down Projects and USM Central Funding.

PRIME is funded by NSF (OISE 0710726, PI: G Wienhausen) with additional support from the Office of Cyberinfrastructure, the National Biomedical Computation Resource (NBCR: NIH, NCRR P 41 RR08605), the Gordon and Betty Moore Foundation, and Calit2 UCSD. Work to build GLEON is supported in part by an award from the Gordon and Betty Moore Foundation and the NSF Research Coordination Network award (PI: P Hanson).

Many individuals contributed to the production of this Collaborative Overview, from those who conducted the work, hosted the students, participated actively in PRAGMA, to those who helped draft the materials. In addition several individuals made special contributions to the production of the brochure: Jennifer B. Matthews (SIO/UCSD) for the brochure's design and compilation; Maureen Curran (Calit2/UCSD) and Teri Simas (PRAGMA/UCSD) for the editing of the document. Their extraordinary efforts helped make this publication possible.

www.pragma-grid.net