PACIFIC RIM APPLICATIONS AND GRID MIDDLEWARE ASSEMBLY

COLLABORATION OVERVIEW 2009-2010

Above (left to right): Jefferson Hang, Jacqueline Chin, and Amir Shirkhani prepare to zip line in New Zealand—courtesy of Jefferson Hang

INSTITUTIONS AND CONTACTS

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*

ASIA-PACIFIC ADVANCED NETWORK (APAN): Seishi Ninomiya, *snino@affrc.go.jp*; Kento Aida, *aida@nii.ac.jp*

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*

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

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*

COMPUTATIONAL SCIENCE AND ENGINEERING ONLINE (CSE-Online): Thanh Truong, *Thanh.Truong@utah.edu*

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

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

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*

HIGH PERFORMANCE COMPUTING CENTER (HPCC) OF HANOI UNIVERSITY OF TECHNOLOGY (HUT): Thanh-Thuy Nguyen, *thuynt-fit@mail.hut.edu.vn*; Huu-Duc Nguyen, *ducnh-fit@mail.hut.edu.vn*

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*

PRAGMA is an institution-based organization governed by a Steering Committee* that invites new members, determines locations of workshops, and sets overall direction. This year, one new member joined PRAGMA: the University of Hong Kong (UHK).

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

INSTITUTE OF INFORMATION TECHNOLOGY-VIETNAM (IOIT-VN): Vu Duc Thi, *vdthi@ioit.ac.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*

MALAYSIAN INSTITUTE OF MICROELECTRONIC SYSTEMS (MIMOS): Thillai Raj T. Ramanathan, Luke Jing Yuan, *jyluke@mimos.my*; Irdawati Ab Rahman, *irdawati@mimos.my*

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

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

NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING (NCHC), NATIONAL APPLIED RESEARCH LABORATORIES: Whey-Fone Tsai*, *cooweftoo@nchc.org.tw*; Fang-Pang Lin*, *fplin@nchc.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 GRID OFFICE (NGO): Hing Yan Lee, *hingyan@ngp.org.sg*

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

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

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

THAI NATIONAL GRID CENTER (TNGC OR THAIGRID): Somsak Sriprayoonsakul, *somsak_sr@thaigrid.or.th*

TransPAC2, INDIANA UNIVERSITY: James Williams*, *william@indiana.edu*; John Hicks, *jhicks@iupui.edu*

UNIVERSITI SAINS MALAYSIA (USM): Habibah A. Wahab*, *habibahw@usm.my*; Huah Yong Chan, *hychan@cs.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 (NBCR); Peter Arzberger*, *parzberg@ucsd.edu*; Philip Papadopoulos*, *phil@sdsc.edu*; Mason Katz, *mjk@sdsc.edu*; Teri Simas, *simast@sdsc.edu*

Above: "Sunset Dock" in Taiwan—courtesy of Michael Nekrasov.

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*

* Denotes Steering Committee Member. More information may be found at www.pragma-grid.net/about/committee.
More information about each of the PRAGMA Institutional Members may be found at www.pragma-grid.net/about/institutions

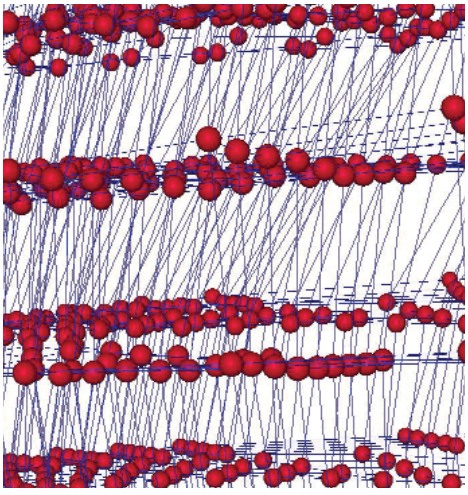
OVERVIEW

THE PACIFIC RIM APPLICATION AND GRID MIDDLEWARE ASSEMBLY (PRAGMA) was established in 2002 as an open organization which focuses on practical solutions to create, support, and sustain international science and technology collaborations via multidisciplinary and multi-institutional teams of researchers who develop or apply grid technologies to advance science. PRAGMA is a demonstrated framework for collaboration. Through its activities, PRAGMA has been a conduit of advancing technologies, ideas, and people. It has launched new groups (e.g., the Global Lake Ecological Observatory Network (GLEON), the Coral Reef Environmental Observing Network (CREON), and the Avian Flu Grid). It has also created programs for undergraduate and graduate students (PRIME, PRIUS, and MURPA), as well as launching Institutes to help train other researchers around the Pacific Rim in both grid technologies and in their applications.

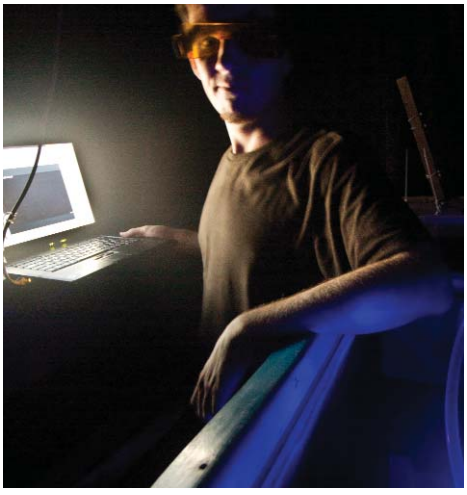
Over the seven and a half years since PRAGMA's inaugural workshop, technologies have continued to evolve. The notion of grid computing is being supplanted by virtualization over local- and wide-area networks. Commercial models for accessing computing and data resources exist now that were not available in 2002. As documented in this Collaborative Overview, PRAGMA researchers are experimenting with virtualization as a means to conduct high-throughput *in silico* screening experiments as required in the Avian Flu Grid.

Another trend that will continue to accelerate is the deployment and use of sensor networks to observe key variables in the environment. This past year has seen a renewed effort to build more bridges with GLEON as a driver of PRAGMA technologies. In particular, pieces of an infrastructure are being assembled to link remote sensing data with ground sensor networks. First steps were demonstrated at PRAGMA 17 with a federation of heterogeneous data (satellite and field sensor data). Additionally, sharing of satellite data in a secure environment in the GEO Grid system demonstrates an effort of more than a year to establish a virtual organization membership service (VOMS) that reflects competing needs in PRAGMA. This activity required the integration of four different software components to work cohesively together.

p. 12



p. 26



p. 23

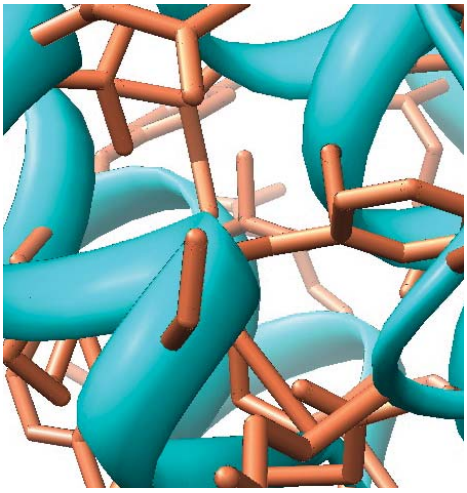
Technologies go in and out of fashion, but applications and collaborative relationships endure and grow.

Finally, this year has seen additional activities in developing the next generation of researchers able to work in international teams. The PRIME program had a record total of 33 students at 13 sites this year, and MURPA sent its first group of students to UCSD. Further, high-definition (HD) technology was used for a series of MURPA lectures. In addition, we are continuing to learn by attempting to extend the number of sites involved using this technology. For example, we attempted a three-way conference on Cardiac Modeling Day, organized by Monash University, involving Osaka University and NICT in Japan, as well as Calit2 and UCSD in the U.S. These highlights are expanded upon in the Accomplishments section. Other advances are captured in the sections on Working Groups and PRIME and MURPA.

Our twice yearly workshops continue to provide a venue to report on progress, strengthen relationships and involve new collaborators in our activities. This past year PRAGMA held workshops in Daejeon, Korea (March 2009, PRAGMA 16) and Hanoi, Viet Nam (October 2009, PRAGMA 17). In addition, PRAGMA co-hosted a PRAGMA Institute with the NCHC Southeast Asia International Joint Research and Training Program in High-Performance Computing Applications and Networking Technology (December 2008). These efforts help us build a broader and stronger community.

This year we are pleased to welcome the University of Hong Kong to PRAGMA. Two future workshops have been scheduled: PRAGMA 18 in San Diego, CA, USA 2-4 March 2010, and PRAGMA 19, in Changchun, China, hosted by Jilin University, 13 -15 September 2010.

The PRAGMA of the future will continue to develop, share, integrate, and innovate technology, as well as bring researchers together into multidisciplinary, multi-institution international collaborations. PRAGMA's success will be in the use of these technologies to advance science and the use of its collaborative framework to continue to generate new opportunities for its members and future researchers and citizens.



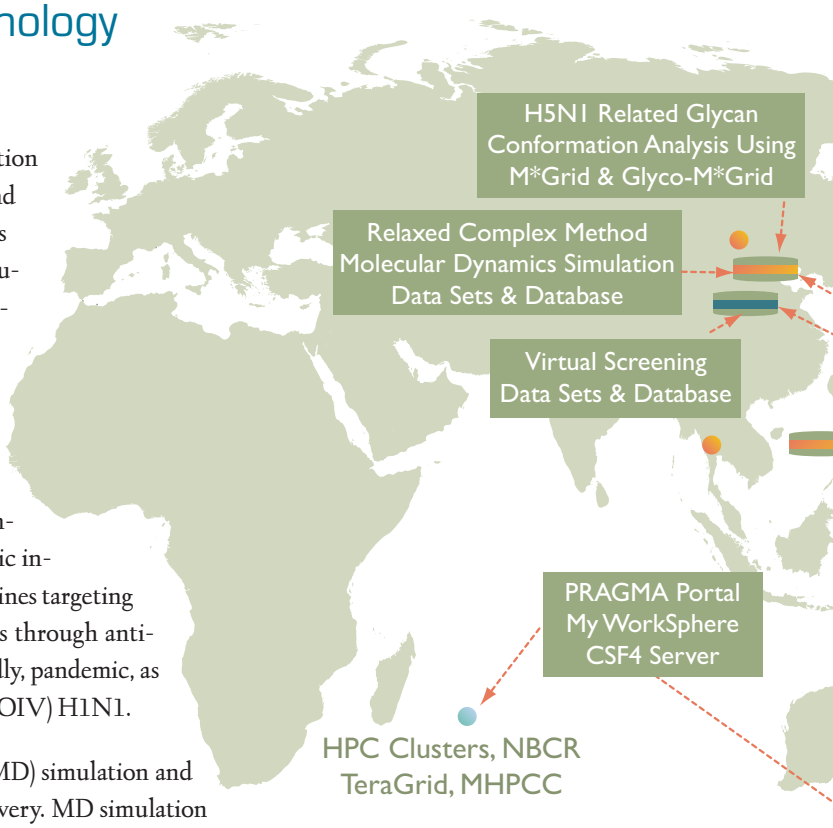
p. 21

p. 42

p. 10

ACCOMPLISHMENTS

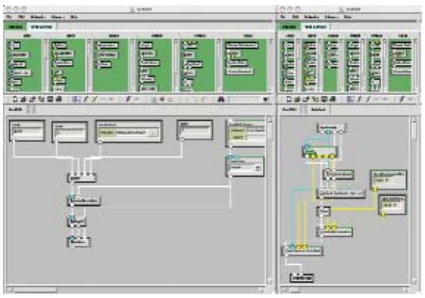
Avian Flu Grid: Science Driven Technology Integration and Advancement



AVIAN FLU GRID (AFG) is a research project and a virtual organization (VO) of the PRAGMA grid that leverages the computational tools and resources contributed by members of the PRAGMA community. Its scientific focus is to understand how the influenza virus constantly mutates, infects humans and periodically becomes pandemic. This is essential for developing new drugs and better vaccines for both prophylactic and therapeutic purposes.

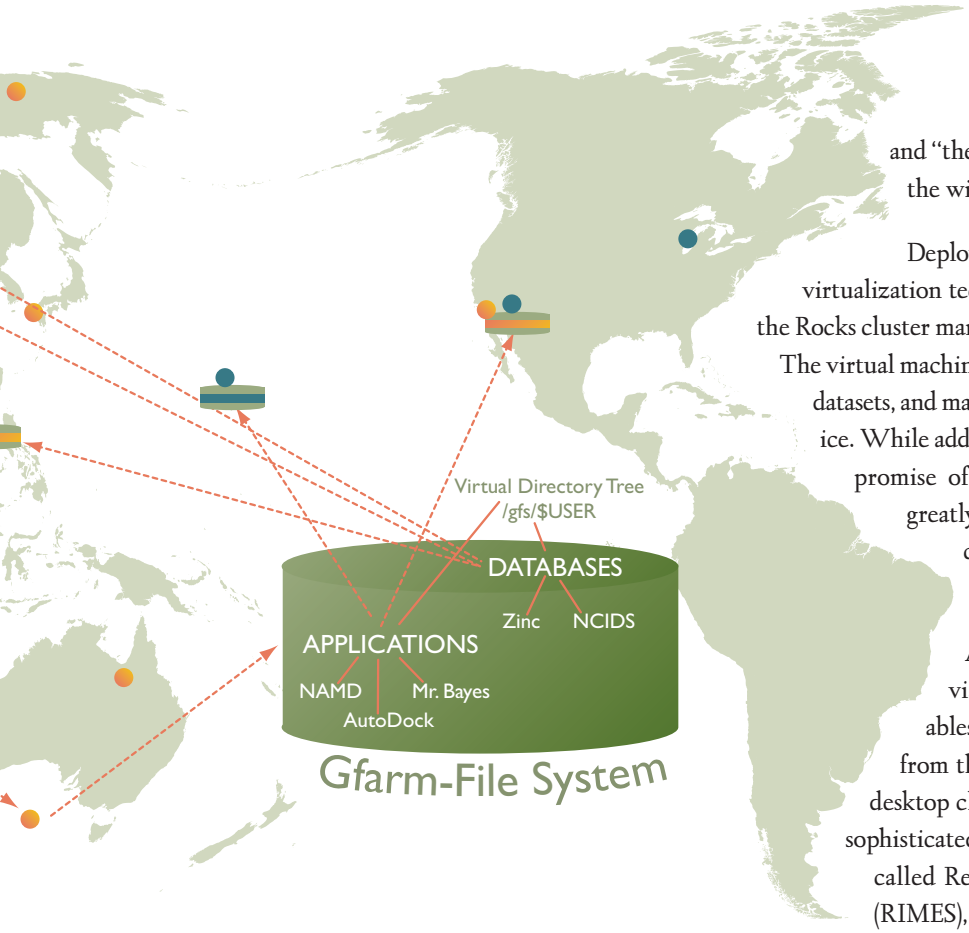
The viral envelope proteins hemagglutinin (HA) and neuraminidase (NA) are known to be responsible for the initial attachment of the virus to host-cell membrane glycoproteins and the release of newly synthesized viruses respectively. These are important points of therapeutic intervention, with known drugs such as Tamiflu targeting NA, and vaccines targeting HA. However, the virus continues to circulate in different reservoirs through antigenic drift and shift, and becomes drug resistant, and even, unexpectedly, pandemic, as in the case of the current pandemic of swine-origin influenza virus (SOIV) H1N1.

The Relaxed Complex Scheme (RCS) combines molecular dynamic (MD) simulation and virtual screening into a novel strategy for computer-aided drug discovery. MD simulation requires a low-latency network within a high-performance computing environment, whereas virtual screening requires many single CPU docking experiments. The results are shared among collaborators around the world. The application requirement and data sharing scenario are commonly found in many research settings, which makes the AFG infrastructure reusable by many other researchers and projects. Technology selection for integration and maturation in the AFG is driven by scientific requirements for best practice, by simplicity for the ease of use and maintenance, and by open source standards for longevity.



The Opal Toolkit developed by the National Biomedical Computation Resource (NBCR) and San Diego Supercomputing Center (SDSC) provides familiar application-specific web services that hide the computational infrastructure from researchers. Opal 2.1 now supports a number of job managers such as fork, Sun Grid Engine (SGE), Distributed Resource Management Application API (DRMAA), Condor, and the Community Scheduler Framework 4 (CSF4) developed at Jilin University. Opal enables AFG applications, such as NAMD, AutoDock, APBS, and PDB2PQR, to be accessed as standard web services by different clients, including workflow tools such as Vision and Kepler. Thus, many different users may use the same workflow environment and access the same application services provided by different distributed providers.

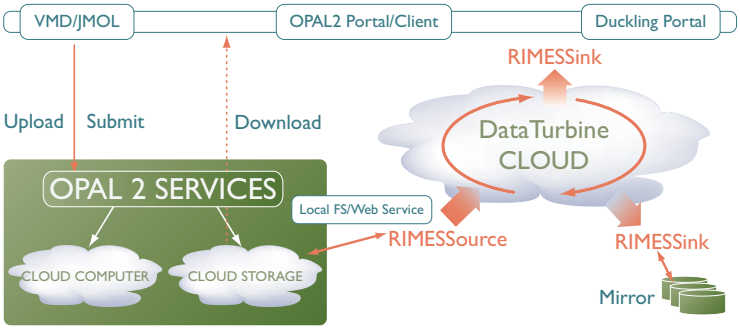
Both Condor and CSF4 allow an Opal application service to be distributed transparently in production. Data sharing has different requirements from high-performance and high-throughput computing. Gfarm, developed by AIST and CCS, has been used in the past to demonstrate high-performance data replication. Recently released Gfarm 2.3 adds additional security features that are required for production use in AFG. However, additional bandwidth challenges, such as "the first mile"



and "the last mile" problems, remain while replicating data across the wide-area network (WAN).

Deployment of the AFG infrastructure has benefited from the virtualization technology from the Xen virtual machines available from the Rocks cluster management system and the Amazon Elastic Cloud 2 (EC2). The virtual machine images contain the necessary applications and required datasets, and may be accessed the same as any other Opal application service. While additional work is required to turn this into routine use, the promise of cloud-based computing through virtualization may greatly improve the on-demand availability and scalability in case of a pandemic outbreak. A panel on the use of clouds was held at the PRAGMA 17 Workshop in Vietnam.

A new development in AFG is the on-demand real-time visualization of simulation and docking results. This enables researchers to check the progress of their experiments from their favorite visualization environment, whether it is a desktop client, a workflow tool, a tiled-display wall (TDW), or a sophisticated StarCAVE for immersive visualization. This new tool, called Real-time bioMEdical data Streaming and visualization (RIMES), utilizes the existing Opal services and adds a DataTurbine (www.dataturbine.org) cloud component that streams frame-based simulation data to different clients. This work resulted from an exchange of researcher Kejun Kevin Dong from CNIC to UCSD for three months. This close collaboration allowed the linking in of the Duckling Portal, created by CNIC, with Opal and DataTurbine.



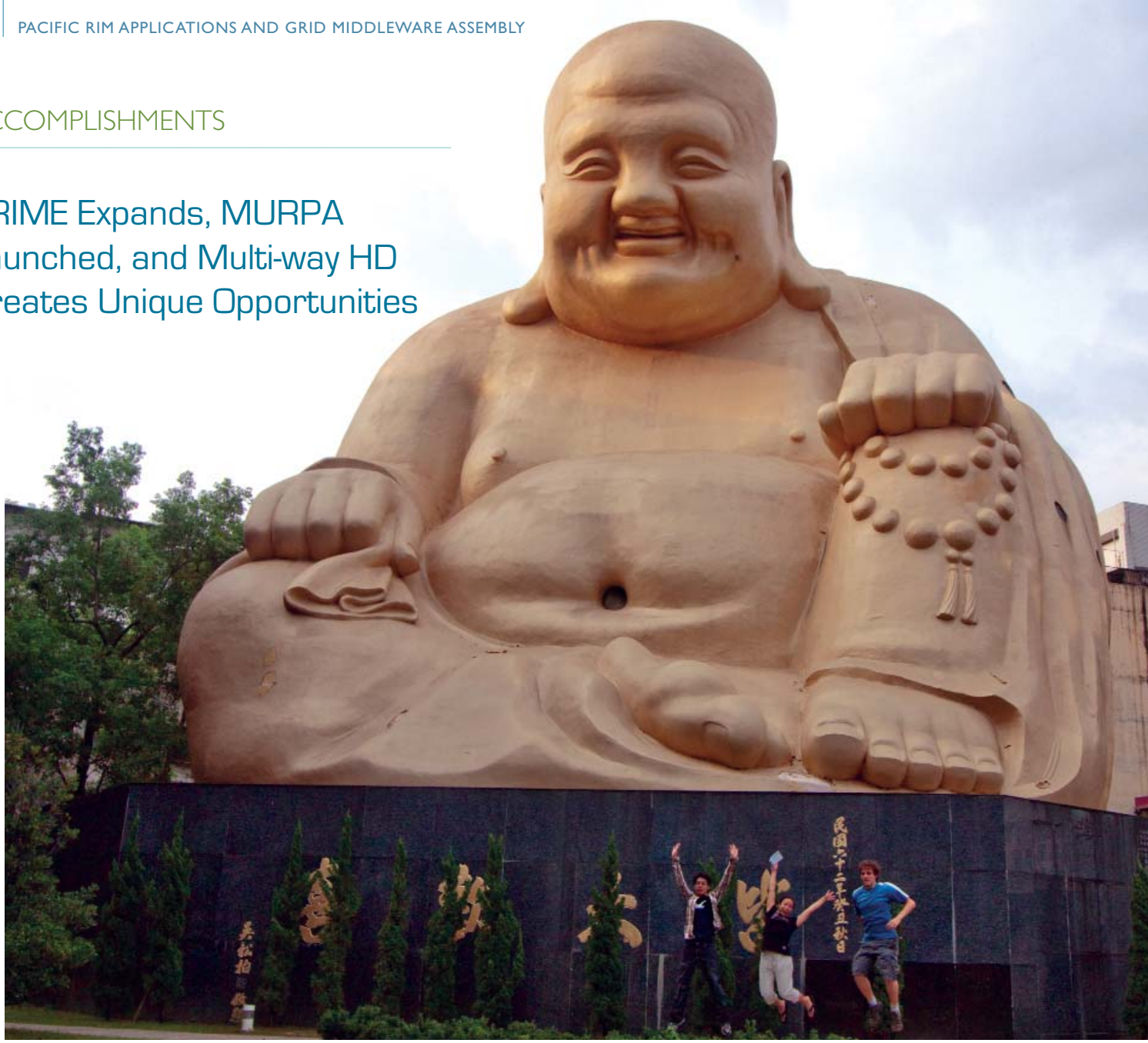
PARTICIPATING RESEARCHERS: H. Wahab, co-chair, *USM*; W. W. Li, *NBCR at UCSD*, co-chair; Jincheol Kim, *KISTI*; K. Jeong, *Konkuk U.*; H. Pao, K. Wu, and D. Xu, *UCSD*

COLLABORATORS: R. Amaro, *UCI*; M. Alam, *U. Hawaii*; J.H. Lin, *Academia Sinica, National Taiwan U.*; Doman Kim, *Chonnam Univ*

RESOURCES: K. Nan and K. Dong, CNIC, CAS; X. Wei, CCST, JLU; Y. Tanimura, AIST; O. Tatebe, *CCS, U Tsukuba*; S.T. Hwang, *Kookmin U.*; K. Cho, *KISTI*

ACCOMPLISHMENTS

PRIME Expands, MURPA Launched, and Multi-way HD Creates Unique Opportunities



THE PRAGMA FRAMEWORK has allowed participants to extend their interactions through new projects. Several of these projects focus on developing the human capital needed to use emerging grid technologies on applications in multidisciplinary and multicultural teams. Two examples are PRIME and MURPA (see section PRIME, PRIUS, MURPA in this Collaborative Overview).

The year 2009 is noteworthy for each of these projects. For PRIME, this was the largest class of students in its six-year history (33 participants); overall, PRIME has now had more than 100 students participate in its activities. This has been possible because of the eagerness of sites to mentor students. The benefit for them has been that students become the bridges between the labs involved, as well as conduits for the technologies and applications. Several new host sites participated in PRIME this year:

- Doshisha University, Kyoto, Japan (bioengineering);
- National Institute of Information and Communication Technology (NICT) (networking, cultural heritage image processing), Tokyo, Japan;
- National Museum of Marine Biology and Aquarium (NMMBA), Kenting, Taiwan, which co-hosted students with NCHC and National Taiwan University (sensors, algorithms for coral reef observing);
- National Taiwan University (NTU), Taipei, Taiwan, (virtual screening); and
- University of Hyderabad, Hyderabad, India (virtual screening, streaming sensor data).

"It is natural to present our findings to our mentors at the end of our experience, but, never before have I presented to such a global audience all [at once]. There were people from Monash University, my mentor and other audiences at University of California, San Diego and a Cardiac Modeling group from Oxford University, England. This very diverse group completely indicates the relation of one country to another and the need for more programs like PRIME to give students a feel for what they can expect in the future."
RAMYA CHITTERS, MONASH

In addition, the Department of Civil and Environmental Engineering of the University of Auckland hosted a student (earthquake engineering). The 13 institutions and the host mentors involved with PRIME broadened the opportunities and provided mentoring for the students, and in many cases, helped support aspects of the program.

More details about PRIME and some of the projects can be found in the Section PRIME, PRIUS, and MURPA.

MURPA, having been modeled on PRIME, sent its first group of four students to UCSD early in 2009. Prior to their departure the students participated in lectures from potential host researchers at UCSD, who presented using high-definition (HD) video conferencing facilities. While at UCSD, these students worked on projects that involved the Kepler scientific workflow engine, virtual screening of inhibitors for avian influenza, software simulation of the heart, and visualization of 3-D city models using large megapixel tiled-display walls. Prior to their return to Monash University, the students gave a seminar to researchers at UCSD and Monash, using the same HD video teleconferencing facilities used for the lectures by UCSD host mentors.

HD video teleconferencing will be playing an increasing prominent role in PRAGMA and in education. We note that the PRIME students in Monash gave final presentations on their work using this facility. In early October, Monash, Calit2, and Osaka University put in place a 2.5 way HD-linked experiment as part of the Monash University Cardiac Modeling Day (www.messagelab.monash.edu.au/cardiacmodellingworkshop09).

The speakers, 2009 PRIME students Anna Pham (Osaka), Ramya Chitters and Scott Revelli (both Monash), presented reports on their cardiac projects from the Calit2 HD studio (see PRIME, PRIUS, and MURPA), and Professors Peter Hunter (U. of Auckland) and Kevin Burrage (UQ and Oxford) delivered seminars on their work from Monash.

Images: (p8) "[Prime Students] Jumping for Buddha" in Taiwan—courtesy of Michael Nekrasov; (left) Teleconference view from Calit2 (at UCSD) with David Abramson and students (at Monash) and Peter Arzberger (in Virginia); (right) Lori Jue and Sabina Piras.

ACCOMPLISHMENTS

As is often the case with initial experiments, there were some glitches on the day, and Osaka could not view the Calit2 and Monash streams simultaneously (hence the 2.5 way descriptor). However, in an unplanned variation to the program, participants at Osaka were able to view the presentations from Professors Hunter and Burrage delivered from Monash via a 12 Mbps new generation H.264 stream and viewed using VLC 1.0.2.

Streams from Osaka to Monash and Calit2 were encoded through Robst (robust streaming tools) at ~25Mbps and similarly viewed using VLC (version 0.8.2—later versions struggle with the Robst Stream). Monash and Calit2 used their DVCProHD streams over which PRIME—and now MURPA seminars—are conducted. Expected issues with differing latencies in each system did not eventuate during the event. Indeed, running a DVCProHD stream and new generation H.264 stream side-by-side at the Monash venue, it was difficult to tell the streams apart and latency was a non-issue during both testing and the event.

The two main problems were matching the slightly different technologies available at each venue and procuring machines that could handle the computation required for decoding new generation H.264 streams. Machines needed to be at least dual core and preferably, handling only one computationally intensive job.

These experiments between Osaka University, Calit2 and Monash University will continue, but already have shown that it is possible to use free software to achieve very good, relatively low cost, HDV VTC results. In turn, this should allow earlier adoption by institutions that are reticent to take up HD VTC because of cost and/or doubt regarding its value.

As this technology becomes more affordable and easier to use, PRAGMA envisions having broader use of it for classes on computational science at several of the PRAGMA institutions.

Lake Observing and Modeling and the Use of Integrated and Secure Data Resources

FRESHWATER LAKES provide a number of important ecosystem services including supplies of drinking water, support of biotic diversity, transportation of commercial goods, and opportunity for recreation. There is a clear need to better understand how changes in land use, human population, and climate interact to affect lake dynamics at local, regional, continental, and global scales. Developing this understanding across such scales is a formidable challenge, in part because ecological systems are characterized by high spatial and temporal variability, non-linear dynamics, and coupled physical/biological processes. In lakes, this complexity is manifested in phenomena such as sudden and short-lived algal blooms, changes in frequency and response to disturbances such as mixing events caused by typhoons, interdependency of biota and biogeochemical processes, and the waxing and waning of fish stocks.

To understand this complexity and variability, researchers must use multiple approaches, including modeling, comparative analyses, and long-term observations. Central to all these approaches is high quality, spatially and temporally pervasive, comprehensive, well-documented and easily accessible data.

GLEON, the Global Lake Ecological Observatory Network, was launched in March 2005 as a grass-roots network of people, institutions, programs and data, all linked by cyberinfrastructure and united by a common mission. The GLEON mission is to understand and predict responses of lake ecosystems to natural processes and human activities at regional, continental, and global scales. GLEON is built on other efforts such as PRAGMA, the EcoGrid Project in Taiwan, and the North Temperate Lakes Long Term Ecological Research (LTER) program, whose researchers collaboratively

FORMOSAT 2 satellite image of the Lake Rotoiti diversion wall, which has been pan sharpened to a resolution of 2 m with a histogram equalisation contrast stretch applied. Image supplied by Taiwan's National Space Organisation (NSPO).

studied the impact of disturbances of a series of typhoons on Taiwan's Yuan Yang Lake dynamics. GLEON is funded by the Gordon and Betty Moore Foundation, the US National Science Foundation (NSF), and funding agencies associated with members; its activities have been further expanded with additional funding through NSF Research Coordination Network, CI—TEAM, and most recently a Cyber-enabled Discover and Innovation (CDI) award.

During this past year, researchers from GLEON and PRAGMA have been working together to integrate sensor data from GLEON lakes with satellite images that are being collated by PRAGMA members AIST and NCHC, as well as PRAGMA collaborator NARL. Initial efforts have been focused on Lake Rotorua in New Zealand, the second largest lake (by surface area) on the North Island. Lake Rotorua has been subject to the effects of increased land-use intensity from dairy farming, resulting in increased inflows of chemicals (nitrogen in particular), which has led to eutrophication of the lake and incidences of phytoplankton blooms from cyanobacteria that can potentially produce toxins, resulting in a number of deleterious side effects.

Monitoring the lake for basic properties is therefore of upmost importance. Currently, single point, but high-frequency, buoy data is collected and models use this data as input to predict circulation of water and thus distributions of nutrients and phytoplankton. By linking data from sensors in lakes with various remote imaging approaches (e.g., MODIS, ASTER, Formosat2), 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 is a long-term goal of the GLEON and PRAGMA interactions.

This year, PRAGMA has taken several steps to respond to this challenge. In particular, researchers from the GEO Grid project at AIST (www.geogrid.org/en/index.html) have been engaged (see GEO Working Group Section). GEO Grid is a project to integrate various data types about the earth systems, in a way that allows for access, but also provides protection of specific data as needed (e.g., licensing). Two steps have been taken by GEO Grid to fulfill needs of the GLEON researchers. First, GEO Grid has federated heterogeneous distributed databases (e.g., satellite data, field sensor data, GLEON data) using OGC standards, and pulls data from different sources in response to a single query. Second, they have appropriate access controls using GSI+VOMS for sensitive data (e.g., ASTER data and Formsat-2 data). In particular, GEO Grid federates MODIS (satellite) data and PEN (field sensor) data. Both data are provided via OGC standards (e.g., W*S and SOS). Finally, through a formal agreement between AIST and NARL, GEO Grid will be able to link to data from Formosat2 of Taiwan's National Space Organization NSPO, which will provide high-resolution photographic images.

The VOMS effort, which is now used by GEO Grid, has been a more-than-a-year-long effort by a team of researchers in PRAGMA. This effort will be important for ensuring the security of data sharing. See the Resources Working Group section for details.

PRIME, PRIUS AND MURPA

AS SOCIETY IS FACED WITH INCREASINGLY COMPLEX PROBLEMS, *such as the sustainability of a healthy planet and combating increasing threats from infectious diseases, it will rely increasingly on science to understand these processes. In turn, the shape of science and its conduct are being changed by advancing technologies that create data at rates that can, in a single year, outstrip the total aggregate digital archives, but can also bring us together in groups, across great distances, with increasing ease. Successfully addressing global challenges will necessitate using these new types of technologies (and others not yet dreamt of) and will require researchers with abilities to use them to work together across disciplines, institutions, and cultural boundaries.*

PRIME, PRIUS, and MURPA were established to prepare students to address these and other global challenges. Successfully addressing global challenges will require using these new types of technologies and developing researchers with abilities to use these technologies, as well as work together across disciplines, institutions, and cultural boundaries.

"Penang, Malaysia is a place that I could never have really understood as a tourist. My PRIME experience has not only allowed me to appreciate Malaysian culture, with its distinctive mix of Malay, Chinese, and Indian cultures, but also to turn inwardly and really be grateful for my own Taiwanese American culture as well." JESSICA HSIEH, USM

PRIME: Pacific Rim Experiences for Undergraduates

PRIME, THE PACIFIC RIM UNDERGRADUATE EXPERIENCES PROGRAM (prime.ucsd.edu), was established in 2004 with nine participants and three sites. Since that time, the program has grown to a class of 33 students at 13 host sites in the summer of 2009. More than 100 students have participated in PRIME to date.

Students conduct research with mentors both at UCSD and at the host site. The nature of these projects has generally been the development of technologies (e.g., visualization tools for datasets) or the application of grid and other technologies to science and/or engineering problems. Often, PRIME students will build on, or extend, projects from other students' work in previous years. Examples of that include work on cell and cardiac modeling, virtual screening with the avian flu virus, as well as protein phosphatase, and developments in the visualization of data (see Collaborative Overview 2008 for details).

This year, to illustrate the diversity of applications and indicate some of the scientific results that were obtained, we highlight thirteen PRIME projects, which were conducted during the northern hemisphere summer of 2009. The projects are clustered into major themes that have recurred over several years. The table includes all 33 2009 students, their mentors and host sites, as well as the topics of their research. (see pp. 14-15)

In addition to the research experience (see PRIME Research Highlights), PRIME provides students with a framework to explore and embrace the culture of their host sites. Both before students go, and after they return, they participate in seminars on cultural awareness. They talk with previous PRIME students to gain their perspectives of the experience. This year, five students agreed to write back to the broader UCSD community about their experiences outside the

Images: (left) Students at Merbok, a traditional Malay village—courtesy of Jessica Hsieh; (top) "Garbage truck passing by" in Danshey—courtesy of Michael Nekrasov; (bottom) "Water Buffalo"—courtesy of Michael Nekrasov

MENTOR(S), INSTITUTION	STUDENT, MAJOR	UCSD MENTOR (AFFILIATION)	PROJECT
DAVID ABRAMSON Monash Univ, Melbourne, Australia	ELISA ABATE bioengineering: premedical	KIM BALDRIDGE (U Zürich)	Grid Workflow Tools for Performing Molecular Computations
	RAMYA CHITTERS bioengineering: biotechnology	ANUSHKA MICHAILOVA (Bioengineering)	Using Nimrod/E to Perform Sensitivity Analysis in Cardiac Electrophysiological Models
	NICHOLAS ECHOLS CS, interdisciplinary computing and the arts (ICAM)	JÜRGEN SCHULZE (Calit2)	Tiled-Wall Plane Image Viewer for Opencover
	SCOTT REVELLI bioengineering	ROY KERCKHOFFS (Bioengineering)	Determining the Optimal Pacing Sites for Biventricular Pacing the Failing Heart with Left Bundle Branch Block
	ADI SINGER computer science	JÜRGEN SCHULZE (Calit2)	3-D City Modeling
KAI NAN Computer Network Information Center, Chinese Academy of Sciences, Beijing, China	MICHAEL SIY, biology, physiology, neuroscience	WILFRED LI (NBCR)	A Search for Small Molecule Inhibitors Against Hemagglutinin
	CHELSEA WONG human biology	WILFRED LI (NBCR)	Avian Flu: Drug Resistance and Off-Target Analysis of Neuraminidase
ARUN AGARWAL AND ANAND K. KONDAPI Univ of Hyderabad, Andhra Pradesh, India	DEE CHEN bioengineering	JASON HAGA (Bioengineering)	Virtual Screening for SSH2 Specificity via PRAGMA Grid
	MATTHEW MUI bioengineering, premedical	JASON HAGA (Bioengineering)	Identification of a Specific Inhibitor for Dual- Specificity Phosphatase Ssh-2 Using Grid Computing
ARUN AGARWAL, K.V. SUBBARAO, RAJEEV WANKAR Univ of Hyderabad, Andhra Pradesh, India	BRIAN MCMAHON computer science	TONY FOUNTAIN AND SAMEER TILAK (Calit2)	Tsunami Sensor Streaming and Event Detection
NOZOMU INOUE Doshisha Univ Kyoto, Japan	UTSAV GUPTA bioengineering, biotechnology	ROBERT SAH (Bioengineering), KOICHI MASUDA (Orthopedic Surgery)	Development of Image Processing Methods for the Quantification of Joint Fluid Volume and the Characterization of Cartilage Lesions via 3-D MRI
	ANDREW SOU microbiology	KOICHI MASUDA (Orthopedic Surgery), ROBERT SAH (Bioengineering)	Kinematics of the Lumbar Facet Joint during Extreme Phases of Flexion and Extension
	LING (KELLI) XU bioengineering:, premedical	KOICHI MASUDA (Orthopedic Surgery) ROBERT SAH (Bioengineering),	Intervertebral Disc Endplate Micro-structure
KAORI FUKUNAGA AND SHINJI SHIMOJO National Inst for Information and Comm. Tech., Tokyo, Japan	ISABELLE FANCHIU interdisciplinary computing and the arts (ICAM)	MAURIZIO SERACINI (Calit2, CISA3)	Terahertz Spectroscopy in Color and Material Analysis
KENJI MURATA, TAKU MORINOBU, AND SHINJI SHIMOJO, Nat'l Inst for Information and Comm. Tech., Tokyo, Japan	JADE KWAN cognitive science: human computer interaction	JÜRGEN SCHULZE (Calit2)	3-D Visualization of Astronomical and Network Analysis Data
JUNG-HSIN LIN National Taiwan Univ, Taipei, Taiwan	JENNIFER CHOY biochemistry and cell biology	WILFRED LI (NBCR)	Ensemble-based Virtual Screening for Non-structural Protein 1 Effector Domain Inhibitors

MENTOR(S), INSTITUTION	STUDENT, MAJOR	UCSD MENTOR (AFFILIATION)	PROJECT
SUSUMU DATE Cybermedia Center, Osaka Univ, Osaka, Japan	ALLYSON CLARK bioengineering	JASON HAGA (Bioengineering), MARSHALL LEVESQUE (PRIME 06)	Verifying Ssh-2 Inhibitor Specificity Through Virtual Screening Experiments on the Grid
	SASHA KORUGA mathematics-CS	JÜRGEN SCHULZE (Calit2)	RealTime Virtual Reality Rendering (Computer Vision)
	CHRISTOPHER LAU bioengineering	JASON HAGA (Bioengineering), MARSHALL LEVESQUE (PRIME 06)	Multiple Instance Tiled-Wall Display for Chimera
	ANNA PHAM general biology	MASAHIKO HOSHIJIMA (NCMIR), JÜRGEN SCHULZE (Calit2)	3-D Visualization of Cardiac MRI/CT and EM Tomography
	CORY STEVENSON, bio-engineering, biotechnology	RAJ SINGH (Calit2/CRBS)	Large Image Viewer Utilizing SAGE
	WEN-WAI YIM bioengineering	JASON HAGA (Bioengineering), MARSHALL LEVESQUE (PRIME 06)	Implementing DOCK on Virtual Clusters
HABIBAH WAHAB Universiti Sains Malaysia, Penang, Malaysia	JESSICA HSIEH, bioengi-neering, biotechnology	WILFRED LI (NBCR)	Optimizing the Design of SiRNAs Targeting the NP Protein
	JESSICA LIU biochemistry and cell biology	WILFRED LI (NBCR)	Hierarchical Screening & Off-Target Analysis of Neuraminidase
DAVID HAMILTON Univ of Waikato, Hamilton, New Zealand	JACQUELINE CHIN environmental engineering	JANE TERANES (Environmental Systems/ESYS)	Correction for Non-Photochemical Quenching in Chlorophyll Fluorescence
JASON INGHAM AND LIAM WOTHERSPOON Univ Auckland, New Zealand	JEFFERSON HANG structural engineering	LELLI VAN DEN EINDE (Structural Engineering/NEES)	Development of NZNEES Data Acquisition and Management Protocols
POUL NIELSEN Univ Auckland, New Zealand	AMIR SHIRKHANI structural engineering	ROY KERCKHOFFS (Bioengineering)	Development of Python Bindings in CellML API
TONY FAN AND FANG-PANG LIN National Museum of Marine Biology and Aquarium, Kent-ing, Taiwan and National Cen-ter for High-performance Computing, Taichung, Taiwan	ROBERT CHEN computer engineering	RYAN KASTNER (Computer Science and Engineering)	Application of Image Processing Techniques for the Characterization of Coral Larvae Blooms
	TSUNG HAN (HANK) LIN computer science	DOUG PALMER (Calit2)	Plankton Automatic Recognition System
	MICHAEL NEKRASOV computer engineering	TONY FOUNTAIN AND SAMEER TILAK (Calit2)	Integrating Cameras for Coral Reef Monitor-ing into Environmental Observing Systems
	WINNY WEN environmental systems (ecol-ogy, behavior, & evolution)	JENNIFER SMITH (Scripps/CMBC)	Interspecific Competition of the Coral Community
KEH-CHYUAN TSAI National Center for Research on Earthquake Engineering, Taipei, Taiwan	LORI JUE structural engineering	LELLI VAN DEN EINDE (Structural Engineering/NEES)	Research and Testing of Large-scale Steel Structures
	SABINA PIRAS structural engineering	LELLI VAN DEN EINDE (Structural Engineering/NEES)	Testing for Application of Unbonded Brace

TABLE 1: Prime research projects

PRIME, PRIUS AND MURPA

research labs. These are captured in their "Dispatches from the Field" (ucsdnews.ucsd.edu/thisweek/2009/09/28_dispatches_index.asp). In addition, when students return we hold a re-entry meeting to talk about their experiences and how PRIME has affected their lives.

Over the last two years, PRIME has begun to use a more detailed assessment instrument, the Intercultural Development Inventory (IDI), which directly measures the students' attitudes towards cultural difference. This is an instrument based upon the Developmental Model of Intercultural Sensitivity (DMSI). For our purposes, the 'stages of cultural awareness' are conceptually related to the IDI categories which rate a student's degree of sensitivity along a continua from 'ethnocentric' to 'ethnorelative' attitudes. This was applied to PRIME 2008 and PRIME 2009, both before their departure to the host site and upon their return. These results are still being analyzed.

With the increase in the number of students this year, PRIME is being forced to address challenges in scaling the program. What works in terms of logistics for 10 students begins to break down at 33. We anticipate the need to eliminate some of the unnecessary logistic options, while retaining the academic rigor. We hope that our experiences will help inform other programs.

"My experience has instilled in me a belief that through working hard many new opportunities will open. I am more determined than ever before to pursue my studies on to the graduate level, and I will do so with the greater amount of confidence and initiative that participating in PRIME has given me." WEN-WAI YIM, OSAKA

PRIUS: Pacific Rim International UniverSIty

PRIUS, PACIFIC RIM INTERNATIONAL UNIVERSITY (prius.ist.osaka-u.ac.jp/en), was established in 2005 through funding to Osaka University from Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) for "Fostering of Globally Leading Researchers in Integrated Sciences," in tight cooperation with the PRAGMA community. PRIUS aims to nurture students' abilities to take a leadership role in next-generation integrated science, as well as offer an international perspective on both academic and business societies through tight collaboration with PRIME and PRAGMA. The PRIUS program, aimed primarily at graduate students, leverages the social networks already built through R&D collaboration in PRAGMA. The educational infrastructure comprises two main components: an internship experience and a course on international integrated science, taught mainly by members of the PRAGMA community. Below is the summary of the three current PRIUS students.

Mr. Atsushi Narishige and Mr. Masashi Nakagawa are staying at Nanyang Technological University (NTU). Atsushi has worked on an observation tool to observe the Internet's global behavior through the Gulliver Project under the supervision of Professor Bu Sung Lee, Francis. He, as part of a team, has successfully used Hadoop, which implements a MapReduce Framework, to process the large amount of log data and visualizing the data. Masashi is working on a survey of the latest research on spectrum-sharing-based game theory in cognitive radio networks under the supervision of Professor Wang Ping.

Elsewhere, Mr. Tomoshige Ohno is staying at Universiti Sains Malaysia (USM) for three months. Dr. Habibah Wahab, a professor in IPharm, Malaysian Institute of Pharmaceuticals and Nutraceuticals, as well as at USM, is his supervisor. Tomoshige is working on the development of docking simulation tools together with Habibah and her students and collaborators. The lab where he works is the Pharmaceuticals Design and Simulation lab, School of Pharmaceutical Sciences. At the time of writing, he has already succeeded in developing computational tools for docking simulation using the ZINC database and is now working to develop a docking workflow.

Images: (left to right) Nicholas Echols, Scott Revelli, Adi Singer, Ramya Chitters, Elisa Abate with David Abramson (in yellow) of Monash University; Gold covered furnace for burning Ghost Money (Joss paper) at a temple on GuanShan Mountain, HengChun, Taiwan; Enjoying the Hanshin Tigers Experience with Japanese Host in Koshien Stadium, Osaka—courtesy of Wen-wai Yim; "Lion Guard"—courtesy of Michael Nekrasov

"My work at NCREE has taught me more than just seismic design - it taught me the importance of international collaboration and of communication as an engineer and member of society." SABINA PIRAS, NCREE

MURPA: Monash Undergraduate Research Projects Abroad

MURPA, THE MONASH UNDERGRADUATE RESEARCH PROJECTS ABROAD PROGRAM (messagelab.monash.edu.au/MURPA), was established in 2008 at Monash University, building on the success of PRIME and the Pacific Rim International UniverSIty (PRIUS) program at Osaka University, which extended the PRIME model to graduate students and embedded it in a curricula. MURPA modified the guest lecture concept from PRIUS and created a digital seminar series for prospective students.

The first group of MURPA students, a total of four, conducted their research in early 2009 at UCSD. Below are highlights of the projects. After three months abroad, MURPA students typically return to Monash to complete their final year of school (honors).

Hoang Anh Nguyen worked with Ilkay Altintas, deputy coordinator for Research, Cyberinfrastructure Research, Education and Development (CI-RED) at SDSC/UCSD, on a software engineering project involving parameterized simulation over a distributed network of computers managed by the Nimrod software package combined with the scientific work flow manager, Kepler. The primary goal was to implement an easy to use graphical user interface that displays results on an Optiportal and other display devices and provides a tool for users to interact with Nimrod/K via Optiportal (this functionality is not yet available).

Guanghua (Aimee) Li worked with Wilfred Li, the executive director of the National Biomedical Computation Resource (NBCR) at UCSD, on a project involving virtual screening for inhibitors against avian flu. She used cheminformatics and data mining techniques to optimize novel inhibitors against avian influenza neuraminidase N1, which is responsible for the release of newly synthesized viral particles from a host cell membrane surface. She also incorporated different analysis tools into a workflow using the Opal web services toolkit. These small molecule inhibitors complement the vaccine development effort against the influenza virus; both are important in the fight against the threat of a pandemic flu outbreak.

Nicholas Barnes worked with Andrew McCulloch, Roy Kerckhoffs and Fred Lionetti in McCulloch's Cardiac Mechanics Research Group at UCSD on a research project involving three-dimensional (3-D) hyperelastic problems in biomechanics and their solution by subdividing contiguous subregions of 3-D meshes and solving the resulting data parallel problems separately. Barnes has been largely compiling and running various solvers on different platforms and benchmarking them. The overall goal is to develop and test a partition of unity method for parallel computation of biomechanics problems in the software package, Continuity.

Peter Serwylo worked with Calit2 UCSD project scientist, Jürgen Schulze, on a visualization project, largely of his own design, involving 3-D city models, as used in many disciplines such as advertising, planning, simulations and games. Serwylo worked on developing a way to generate these models in a timely manner with minimal human input.

For more information about students and their projects, see www.calit2.net/newsroom/article.php?id=1484.

Image: Winny Wen, PRIME 2009 (left) "Night Market (Winny and our wonderful guide Lidiya)" in Taiwan—courtesy of Michael Nekrasov

HIGHLIGHTED PRIME PROJECTS

A. Cellular and Cardiac Modeling for Validation and Understanding

This area of research has been one in which current students build upon the work of previous PRIME students, thus advancing the scientific knowledge through generations of PRIME students. Also, the students learn about a powerful family of tools—NIMROD—directly from the developers at Monash.

A1. SENSITIVITY ANALYSIS IN CARDIAC ELECTROPHYSIOLOGICAL MODELS

Ramya Chitters (PRIME 2009), Anushka Michailova (UCSD), David Abramson, Tom Peachey, Colin Enticott, Blair Bethwaite, Slavisa Garic and Phillip Chan (Monash)

Cardiac excitation-contraction coupling is a sequence of well-orchestrated events. Both the excitation and the contraction processes and their interactions are required for an integrative model of cardiac electromechanical interactions. A number of mathematical models have been developed to study excitation-contraction coupling in ventricular cardiac cells. Recently, Saleh Amirriazi (PRIME 2007) integrated the Shannon et al. four compartment model and the the Michailova et al. equations describing magnesium-nucleotide regulation of ATP-sensitive potassium channels, L-type calcium channel, sodium/potassium ATPase, sarcolemmal and sarcoplasmic reticulum calcium -ATPases into the models. The updated ionic model is more complex due to the multiple domains (junctional cleft, sub-membrane space, cytosol, sarcoplasmic reticulum) and an increased number of unknown model parameters; therefore, it implies a less stable set of ordinary differential equations systems. For this reason, any information gained based on this model could yield a significant level of misunderstanding if one wants to use it to investigate excitation-contraction coupling in ventricular myocytes.

The main goal of this project was to examine the changes in total ionic flux(es) in the compartment of interest which may play the largest role in determining specific model outputs (action potential (AP) and Ca transients in the four cell subdomains) in ventricular cardiac cells isolated from the sub-endocardial tissue layer in rabbits. Our approach was to use both NIMROD/E (novel NIMROD grid computing technologies for experimental design that allows for rapid evaluation of parameter sensitivities in complex computational models), and NIMROD/G (a full parameter sweep program for preliminary testing of the parameters).

In agreement with the experiment by Shannon *et al.*, our results suggest that small changes in L-type calcium and sodium/calcium exchanger total fluxes in the junctional cleft may significantly affect the cell function (Figure 1). Surprisingly, the NIMROD/E studies suggest that small changes in the junctional chloride/calcium flux may also have a prominent effect on the model outputs. New experiments need to be perfomed to test this hypothesis. The fact that both the activation potential duration and cytosolic calcium peak are impacted by the same flux(es) distributions further strengthens our findings.

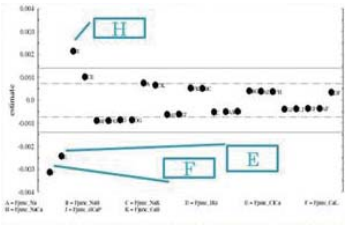


Figure 1. Length plot for action potential (AP) duration for ADP_90 test indicating F (Fjunc_CaL), E(Fjunc_ClCa) and H(Fjunc_NaCa) to have the most significant impact on the AP duration of the cell.

A2. DETERMINING OPTIMAL PACING SITES FOR BIVENTRICULAR PACING IN THE FAILING HEAT WITH LEFT BUNDLE BLOCK

Scott Revelli (PRIME 2009), Roy Kerckhoffs, Jazmin Aguado-Sierra, Fred Lionetti (UCSD), David Abramson, Blair Bethwaite, Tom Peachey, Colin Enticott (Monash)

One of the most dangerous maladies currently facing society is heart failure. The evidence of this is that each year approximately one million people in the U.S. are admitted to hospitals due to heart-related health concerns and approximately thirty-five billion U.S. dollars are spent every year on treating heart failure.

The goal of this project is to determine the placement of the leads to optimize cardiac performance for biventricular pacing of a failing heart with left bundle branch block, a cardiac conduction abnormality that results in the left ventricle contracting later than the right ventricle. This project was pursued by using Continuity 6, a finite element modeling program designed by UCSD's Cardiac Mechanics Research Group for modeling biological systems. In Continuity 6, an anatomically detailed biomechanical and an electrophysiology model of a rabbit heart were created and validated. By employing Nimrod/G, a parametric modeling system, designed and developed by the Message Lab at the University of Monash, a parameter sweep can be conducted on the fully coupled

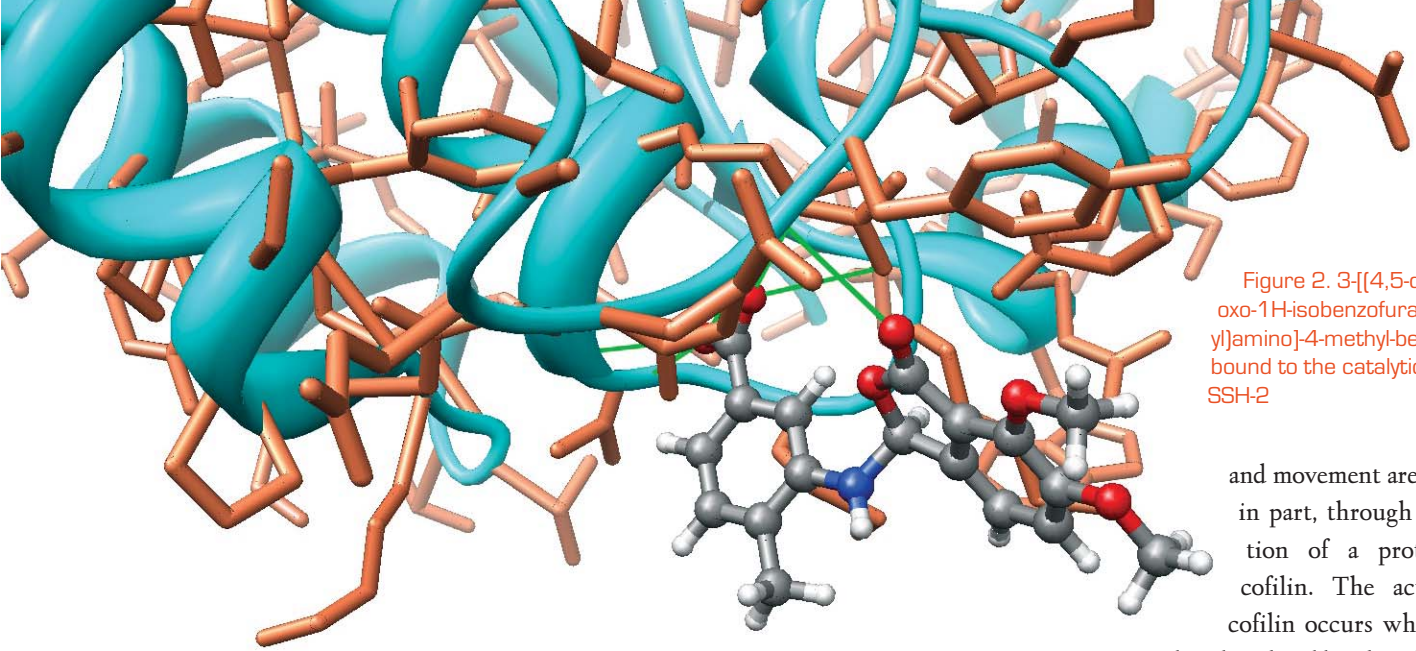


Figure 2. 3-[[4,5-dimethoxy-3-oxo-1H-isobenzofuran-1-yl]amino]-4-methyl-benzoic acid bound to the catalytic site of SSH-2

model. To produce the most accurate results as possible, an existing finite element rabbit heart model was refined, producing one of the most computationally demanding problems ever created in Continuity 6 (see image on page 28). The biomechanical and the electrophysiology models are currently in the process of being coupled, and the Nimrod/G script, which will conduct the parameter sweep, has been outlined.

These accomplishments have set the stage for future work, which will include conducting a parameter sweep on the fully coupled model using Nimrod/G. The parameter variations will simulate biventricular pacing of the heart with one lead fixed at the apex of the right ventricle, while altering the position of the other lead around the left ventricle. Future work will also involve introduction of one or more scars of varying sizes and locations into the model to produce unique information concerning the optimal pacing site for an infarct of a specified size, in a specific location.

B. Virtual Screening and Multiple Sequence Alignment in Disease Detection and Prevention

This area is of great interest to many of the PRIME host sites. This year students went to USM, CNIC, and NTU (new this year) to look at aspects of screening the avian influenza virus. In addition, students worked at Osaka University, as well as U Hyderabad (new this year), on aspects of virtual screening of SSH inhibitors. Both of these projects have had students in previous years create knowledge that has been built upon. In both cases, stronger interactions between the groups involved has emerged because of the student exchanges.

B1. *IN SILICO* IDENTIFICATION OF HIGH POTENTIAL SSH-2 SPECIFIC INHIBITORS

Matthew K. Mui (PRIME 2009), Marshall J. Levesque (U Pennsylvania and PRIME 2006), Arun Agarwal, Anand Kondapi (U Hyderbad), Jason H. Haga (UCSD)

Cellular functions are generally highly regulated activities. In particular, cell growth and movement are controlled, in part, through the activation of a protein called cofilin. The activation of cofilin occurs when it is dephosphorylated by Slingshot-2 (SSH-2). SSH-2 is a member of the dual specificity phosphatase (DSP) family and is known to contribute to the progression of cancer and Alzheimer's disease. For this reason, finding a specific inhibitor for SSH-2 may have a profound impact in clinical treatments for various diseases. Inhibitors can be found through traditional methods in which the binding affinities of each and every compound to a target are determined through direct wet bench testing. With the availability of very large chemical compound databases, however, this method is costly and inefficient. This study eliminates the need to manually test a very large chemical compound space by employing grid computer technologies in virtual molecular docking experiments to identify a specific list of high potential SSH-2 inhibitors.

Similar to the traditional method, binding affinities are determined in the *in silico* experiments using the molecular docking simulation software DOCK 6.0. Taking advantage of the built-in message passing interface (MPI) of DOCK, in combination with OpalOP, an entire experiment is split into multiple jobs that are distributed over

"Getting results after weeks of hard work is really rewarding and satisfying. I am more excited than ever about doing research in my graduate studies." MATT MUI, U HYDERABAD

HIGHLIGHTED PRIME PROJECTS

"As alien as the environment seemed at first, by the time I left I had become accustomed to it more completely than I could have realized." BRIAN MCMAHON, U. HYDERABAD

multiple clusters on the grid. Once these jobs are complete, the results are reassembled to generate a master list with all the binding scores. This list is further sorted to determine the top binding compounds. By screening SSH-2 and other DSPs, the results of each compound can be cross-analyzed to identify SSH-2 specific inhibitors.

The results from five DSP screenings (specifically: SSH-2, VHR, VH3, PTEN and KAP), suggest that 3-[(4,5-dimethoxy-3-oxo-1H-isobenzofuran-1-yl)amino]-4-methyl-benzoic acid (Figure 2) shows the highest affinity for SSH-2, but lowest affinity for the other DSPs, among the best 100 SSH-2 binding compounds. These results suggest that this compound has high specificity towards SSH-2. Although these results must be confirmed through wet bench testing, this study eliminates compounds that are not likely to bind to, or show high specificity for, SSH-2 and it provides a foundation for the analysis of molecular interactions of chemicals with SSH-2.

B2. VIRTUAL SCREENING FOR IDENTIFICATION OF SMALL MOLECULE INHIBITORS OF HA IN AVIAN FLU

Michael Siy (PRIME 2009), Wilfred Li (UCSD), Kevin Wu (UCSD and PRIME 2008), Hsing Pao (UCSD and PRIME 2007), Kai Nan, Jianjun Yu (CNIC)

The recent H1N1 pandemic has illustrated to the world the rapid spread of a virus and underscores the need for drug discovery of preventative medicines in the case of pandemics in general. This project focused on the avian flu virus—H5N1—which has a much higher estimated fatality rate than the current H1N1. The specific project aims are to search for small molecule inhibitors that have a high affinity for the hemagglutinin (HA) receptor-binding domain of the influenza virus and inhibitors against HA that stop the conformational change of the trimers needed for membrane fusion, using the relaxed complex scheme (RCS). Other studies have shown that TBHQ is a molecule that stabilizes neutral pH structures, inhibiting conformational change and preventing fusion. Thus, the TBHQ binding site was used in developing a trimeric interface inhibitor.

The screening methods used in this project's approach are based on new software, AutoDock Vina, recently made available. Vina provides higher performance and enhanced ease of use compared to the earlier version of AutoDock, AutoDock4, which was used last year by PRIME student Kevin Wu in similar screenings. After validating docking parameters using positive and negative controls (redocking), we screened specific binding sites on H5 against databases, such as the NCI Diversity Set II (NCIDSII). Significantly, both the HA receptor binding domain and TBHQ binding site had a highest binding affinity with the same molecule from the NCIDSII database. Analyzing this single molecule and the methods in which this research was performed can hopefully aid in the production of a new drug which ultimately can prevent the binding of the H5N1 virus in humans, and thus its infection.

B3. OPTIMIZING DESIGN OF NP SPECIFIC SIRNAS FOR TREATMENT OF AVIAN FLU

Jessica Hsieh (PRIME 2009), Wilfred Li (UCSD), Habibah Wahab, Sy Bing Choi, Wai Keat Yam (USM)

The avian influenza virus is capable of affecting as many as 500 million people and causing more than 500,000 deaths worldwide in a typical year. Specifically, the highly contagious subtype H5N1 virus is able to infect humans directly by infecting the human respiratory tract. Currently, the world community has also been concerned with the swine-origin H1N1 pandemic. The influenza virus is able to undergo antigenic shifts and drifts, therefore, the efficacy of present vaccines and medications is limited because not all subtypes of the influenza virus can be targeted. The focus of this project was to help improve the method of using siRNAs (small, interfering RNAs) specific to conserved regions of the viral genome to prevent the translation of viral messenger RNAs to proteins in the influenza virus replicative cycle.

In particular, this project focused on optimizing the design of small interfering RNAs that will target conserved regions of the influenza viral mRNAs coding for the NP protein. RNA interference research has the potential to find a treatment that may be able to target all subtypes. The project required the use of the Basic Alignment Local Search Tool (BLAST) to identify where successful siRNAs usually target, how specific they were, and what possible off-targets there were, as well as how many. The Influenza Virus Resource obtained influenza sequences

from different subtypes and performed multiple alignments as well. Additionally, MUSCLE, a protein multiple sequence alignment program, ran large multiple alignment jobs.

There were several outcomes from this research project. First, although a general consensus as to where an siRNA, specific to the NP protein, should target was not obtained, there is information as to how specific an siRNA can be, as well as the lowest number of off-targets an siRNA may have, based on the data obtained from the positive controls analysis. Meanwhile, although there were no regions of 19-25 nucleotides that were completely conserved for all the five subtypes studied, there were significant overlaps. Thus, it is possible to suggest a region of around 100 nucleotides—out of a segment that includes around 1500 nucleotides—in the influenza viral genome coding for the NP protein that may have the most potential with respect to optimizing the design of NP specific siRNAs.

C. Environmental Observing and Modeling in Coral Reefs and Lakes

Environmental monitoring and modeling will play an increasingly important roll in science in the future. The focus on coral reef ecology (NMMBA) is new to PRIME this year; the interaction with U Waikato (lakes) is in its second year and is part of the GLEON activity.

C1. CORAL REEF OBSERVING THROUGH DATA CAPTURE, DATA STREAMING, AND AUTOMATION

Michael Nekrasov (PRIME 2009), Robert Chen (PRIME 2009), Tony Fountain, Sameer Tilak, Ryan Kastner (UCSD), Tony Fan (NMMBA), Fang-Pang Lin (NCHC)

Coral reefs are of prime ecological and economic importance, having the highest species diversity of any marine habitat and ranking near the top of all ecosystems with respect to annual gross productivity. Sweeping environmental changes have been predicted to occur in coming decades that will affect coral reproduction and survival. Yet, we currently have only a rudimentary understanding of the processes by which coral reef ecosystems may resist change or recover following disturbances, particularly as a function of the spatial scale of a perturbation. One approach to capturing information about coral is through observing systems that are able to stream data from sensors and cameras to anywhere in the world. One of the most biodiverse coral reefs in the world is off the coast of southern Taiwan.

One goal of the project, led by PRIME student Michael Nekrasov, was to use cameras as a tool for real-time monitoring and analysis of coral. Different types of cameras and sensors were integrated into DataTurbine, a middleware for integration and monitoring of data from heterogeneous sources being developed at UCSD's San Diego Supercomputer Center (SDSC) and Calit2. In collaboration with Melissa Roth (SIO and an NSF-funded East Asia-Pacific Summer Institute recipient), a system for capturing coral fluorescence was developed. Over- and under-water cameras were set up. Excitation lights were mounted and directed at coral to induce fluorescence. Blocking filters were mounted on the cameras to block out the excitation light, thereby leaving only the light cast off by the coral (see Figure 3). The data was sent over a network to NCHC where it was stored via DataTurbine.

A significant goal for the project was to use this system for the automated counting of coral larvae as they are released (a process that is usually difficult to see because it happens during the night with each colony

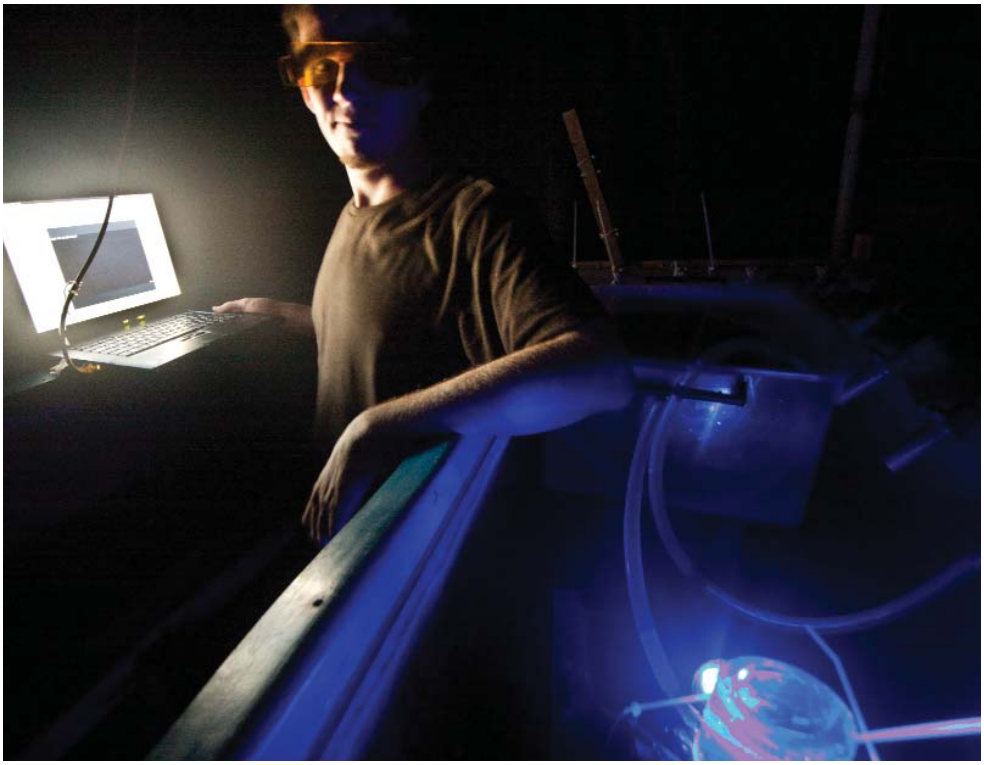


Figure 3. Michael Nekrasov at NMMBA, Kenting, Taiwan, capturing images and video to investigate coral fluorescence.

HIGHLIGHTED PRIME PROJECTS

"This summer, I felt I gained a few steps into the world of international collaborative research. Culturally, I feel I gained perspective in the traditions my parents have been teaching me and insight into the long, amazing history of China." MICHAEL SIY, CNIC

releasing larvae in the hundreds). Seriatopora hystrix, so chosen for the easy to predict monthly release cycles, were placed in custom-made flow-through tanks. Different configurations of lights and cameras were tested in an attempt to get a clear enough image for counting the millimeter-size larvae. The images were sent through DataTurbine where a parallelized image-analysis program developed by PRIME 2009 student Robert Chen was used to count the larvae. Although a live count was not accomplished, mainly due to the onset of a typhoon, the system was capable of registering larvae that were artificially introduced to the water column.

Although much more design and experimentation remains to be done, real-time systems in conjunction with fluorescence proved to have scientifically interesting applications that should be explored further.

C2. CORRECTION FOR NON-PHOTOCHEMICAL QUENCHING IN CHLOROPHYLL FLUORESCENCE

Jacqueline Chin (PRIME 2009), David Hamilton (U. Waikato), Jane Teranes (UCSD)

The Global Lakes Ecological Observatory Network (GLEON) is a group of scientists, engineers, and technologists who use sensors to collect, remotely transfer, and analyze lake data. This group aims to raise awareness of the need to protect lakes and increase knowledge of their important roles in carbon cycling, water resource supply, and biodiversity at the local level, as well as on a global scale. The New Zealand connection in GLEON at the University of Waikato, has played an important role in advancing the science and technology available to members. Their accomplishments include deploying several remote monitoring buoys in lakes, pioneering a robust buoy design, using an independent company to host its data—while still making it open access—and hosting a GLEON meeting in New Zealand in 2009.

GLEON is continually seeking improved methods for monitoring lakes. One aspect of lake monitoring that is of particular interest is microscopic algae (phytoplankton), which form the base of the aquatic food chain. Phytoplankton are important in determining the trophic level or productivity of a waterbody. One means of monitoring levels of phytoplankton is with chlorophyll fluorescence sensors; the number of monitoring buoys that measure this variable is increasing rapidly in the observatory network. In the past several years, research has proven chlorophyll fluorescence to be a good proxy for algal biomass. However, readings taken from instruments moored just beneath the water surface show strong signs of a process known as non-photochemical quenching of algae, whereby bright light leads to a physiological response in the algae that results in strongly depressed levels of fluorescence. A correction could lead to better use of the abundant data obtained via monitoring buoys, as well as the data obtained from commonly used manual profiles of chlorophyll fluorescence taken during daylight hours.

Data was readily available through regional environmental data managers: Environment Bay of Plenty (sponsors of the buoys) and iQuest, a commercial data manager which posts the data to websites. By using this open access, real-time buoy data and Microsoft Excel, plots of the chlorophyll fluorescence data were made in MATLAB. The daytime fractional decrease in fluorescence was compared with observations of instantaneous surface irradiance to examine how light affected non-photochemical quenching.

After careful plots of the fractional fluorescence with surface irradiance, and integrating seasonal variations, a regression line was generated (Figure 4). This regression is a preliminary attempt to correct for non-photochemical quenching of chlorophyll fluorescence. Further analysis of the statistical accuracy of the regression is required and will be done by examining additional data for Lake Rotorua, as well as data for different lakes and different fluorescence sensors.

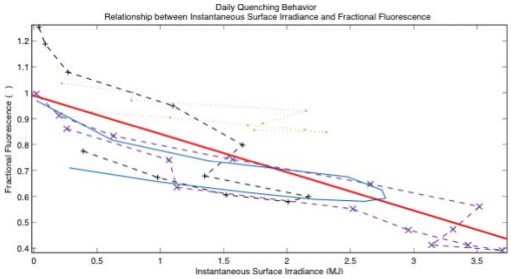


Figure 4. Relationship between instantaneous surface irradiance (I) and fractional fluorescence decrease (FF) relating to nighttime values. Green, fine dotted line is one day in Autumn (10 Apr. 08); Black '+' points is one day in Winter (27 Aug. 08); Violet 'x' is one day in Summer (31 Dec. 08); Blue solid line is one day in Spring (16 Sept. 08). Line is a regression equation for all data: FF=-0.1479I+0.99, R^2 = 0.6238, P<0.01

D. Testing Structures and Acquiring Data in Earthquake Engineering

PRIME has had interactions with NCREE over the years, and this year we sent two students to their location. In addition, we initiated an interaction with U Auckland in this area as part of a larger interaction on the Network of Earthquake Engineering Simulation (NEES) and sent two students to U Auckland as well. One earthquake engineering project from each site is discussed here.

D1. TESTING FOR APPLICATION OF UNBONDED BRACE

Sabina Piras (PRIME 2009), Lelli Van Den Einde (UCSD), Keh-Chyuan Tsai (NCREE)

In general, the large scale structural testing that occurs in any structural engineering laboratory is always investigating new seismic design strategies or retrofit strategies to strengthen existing structures during an earthquake. This work investigated the potential use of the Unbonded Brace (UBB), or Buckling-restrained Brace (BRB) (fabricated by Nippon Steel Engineering Co., LTD.), in seismic structural design. To reach this goal, a full-scale unbonded brace was tested using a Universal Testing Machine at the National Center for Research on Earthquake Engineering (NCREE) in Taiwan (see Figure 5). The cyclic performance of the UBB-12 specimen was based on the Standard Loading Protocol, which was derived from the 2005 AISC *Seismic Provisions for Structural Steel Buildings*.

The UBB-12 specimen featured a steel yielding core plate with a mortar fill in a hollow structural section (HSS) casing. Buckling-restrained braces are characterized by their ability to yield inelastically in compression and tension. It dissipates energy through stable tension-compression yield cycles. This behavior is achieved by limiting buckling of the steel core within the bracing elements.

Figure 5. PRIME students Sabina Piras (left) and Lori Jue in front of UBB-12 Test Specimen in Universal Testing Machine

Based on our test results, the specimen exceeded the four test requirements set forth by NCREE, thus verifying that these brace designs can be used in seismic regions. The first of these requirements is that the UBB-12 specimen not fracture during the Standard Loading Protocol test. It did not and the cumulative plastic ductility exceeded the required $200D_{by}$. Although the Low-Cycle Fatigue test was prematurely terminated, and only eight cycles were completed, it still met the second requirement, again exceeding the required $200D_{by}$. Hysteresis behavior was stable for both the Standard Loading Protocol and the Low-Cycle Fatigue test—with the exception of sporadic changes in axial force, which were due to the slipping of the connection bolts. The third test requirement was that the calculated K_{error} value be a maximum value of ±10%; the K_{error} calculated in this test was -6.54%. Lastly, each cycle during the Standard Loading Protocol test had a loading difference less than 10%, thereby achieving the last test requirement.

HIGHLIGHTED PRIME PROJECTS

"PRIME was an amazing experience for me. I had the opportunity to work with two premier labs, located on opposite sides of the world. The experience gave me an insider's view of what research really feels like." SCOTT REVELLI, MONASH

D2. IMPROVING DATA ACQUISITION FOR EARTHQUAKE ENGINEERING EXPERIMENTS

Jefferson Hang (PRIME 2009), Lelli Van Den Einde (UCSD), Liam Wotherspoon, Jason Ingham (U Auckland)

New Zealand's Network for Earthquake Engineering Simulation (NZNEES) is a major research program led by the University of Auckland's (UoA) Civil and Environmental Engineering Department. NZNEES has an ongoing collaboration with the U.S. NEES project which allows both groups to share data from experiments between the projects.

NZNEES is transitioning their data collection equipment from the traditional National Instruments (NI) DAQ (data acquisition) devices to the newer NI-DAQmx devices. The scope of this project was to create a set of virtual instrument (VI) programs linking NI's LabVIEW (Laboratory Virtual Instrumentation Engineering Workbench) software to work properly with NZNEES's new equipment. This included the creation of user manuals to accompany each program that was developed. In order to create these programs, knowledge of LabVIEW and its numerous functions was required, as well as a general understanding of how the data-acquisition equipment worked (Figure 6). With this knowledge at hand, a user-friendly program was created allowing researchers to view data from their earthquake engineering experiments on a graph as well as in a data table. The program also saved the data to a file of the researcher's choice and sent the data to the NZNEES NEESpop server to stream it to remote project participants. The design allowed for the creation of different variations of the program that supported varying amounts of analog input channels. Once these programs were completed, they were validated during actual

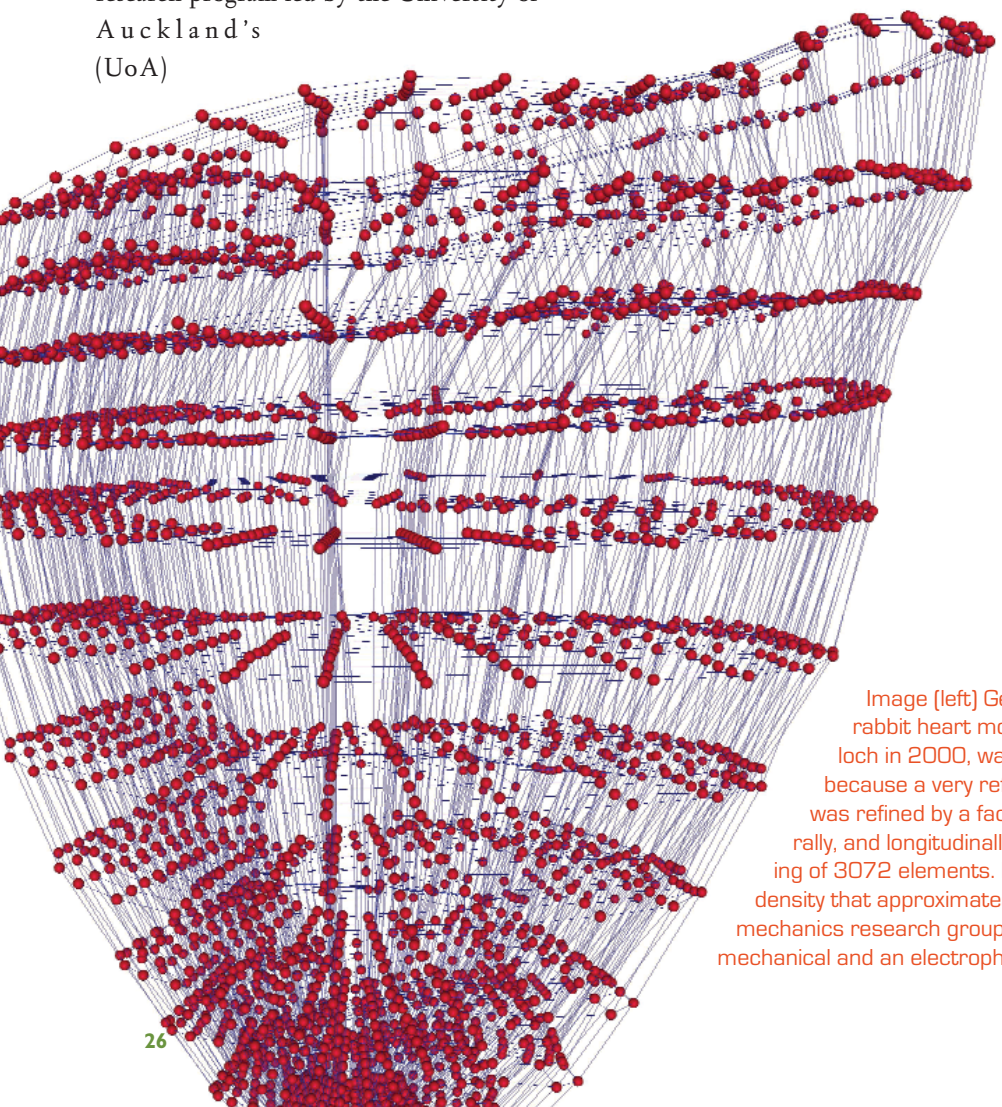


Image (left) Generated by Scott Revelli. The original 48 element rabbit heart model, which was created by Dr. Vetter and Dr. McCulloch in 2000, was refined in order to produce more accurate results because a very refined mesh is what the EP model needs. The mesh was refined by a factor of four in each direction (circumventally, transmurally, and longitudinally). This refinement process produced a mesh consisting of 3072 elements. More importantly, the refined model has a element density that approximates a model created in 2009 by the UCSD's cardiac mechanics research group, for another project that also involved coupling a biomechanical and an electrophysiology model. See PRIME Highlight A2, page 20.

earthquake engineering experiments. With these programs, researchers are able to conduct research with more powerful equipment and acquire more data. In addition, UoA is hoping that these programs will encourage more students to start using LabVIEW to collect data for their experiments since LabVIEW has the ability to acquire data from a large number of sources, is easy to use and has a simple interface, and the data can be streamed in real time so that remote researchers can participate in experiments as if they were actually present.

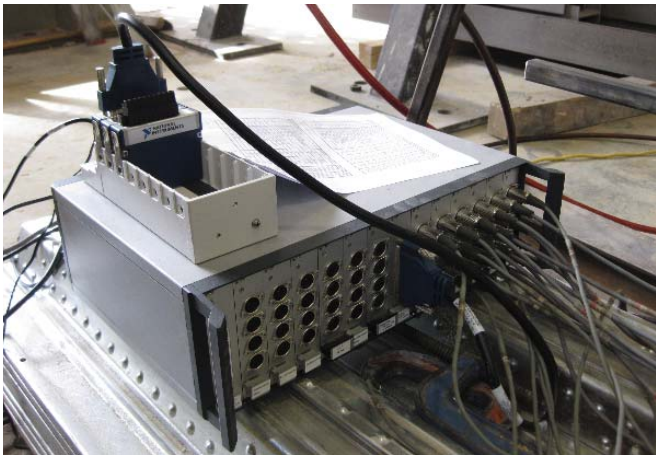


Figure 6. The new data-acquisition equipment.

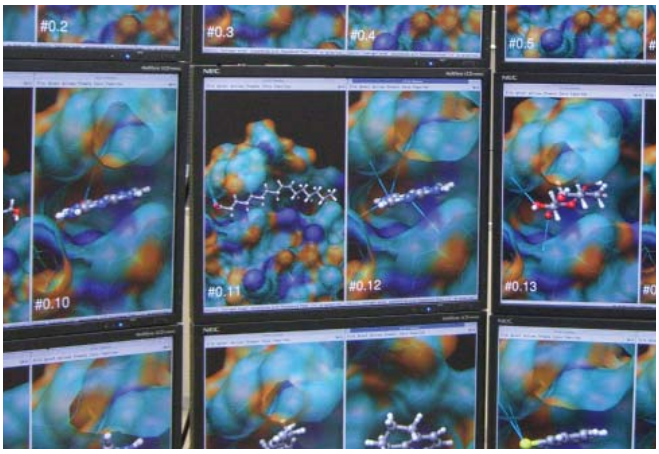


Figure 7. Ligand-receptor interactions on ViewDock TDW. The 4x3 array of displays has two instances of Chimera running on each display for a total of 24 different interactions shown. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081)

E. Visualization, Virtualization, and Workflows in Applications

The projects in this section focus primarily on testing or developing aspects of cyberinfrastructure. Several have attributes mentioned above, namely that of persistent interactions (e.g., those with Osaka and Monash). We have also started interactions with a new lab, NICT. One of the two NICT projects is highlighted below.

E1. VISUALIZATION OF VIRTUAL SCREENING RESULTS ON TILED-DISPLAY WALLS (TDW)

Christopher Lau (PRIME 2009), Marshall J. Levesque (U Pennsylvania and PRIME 2006), Susumu Date (Osaka U.), Jason H. Haga (UCSD)

Drug discovery is driven by the identification of particular ligands that have a high affinity for a target molecule. Virtual screening facilitates this discovery process through the use of molecular simulations to computationally identify these ligands. In recent years, the advancement of grid computing has allowed for the generation of a tremendous amount of data from virtual screening experiments. These results cannot be utilized in laboratory experiments until they are analyzed manually to identify anomalies, redundant chemical structures, and false positives. While grid computing is able to connect supercomputers, on an international scale, to perform virtual screening, a bottleneck in this process is the analysis of the overwhelming volume of resulting data. For this reason, a new method is needed which allows for the analysis of the results of virtual screening, thereby accelerating the drug discovery process.

A collaboration between Dr. Jason Haga at UCSD and Dr. Susumu Date at Osaka University resulted in the development of a solution: *ViewDock TDW* (Tiled-display Wall) as a solution. This tool allows a user to simultaneously visualize, group, compare, and analyze dozens of ligand-receptor interactions on a TDW. To achieve this, it was necessary to make modifications to the source code of *ViewDock* and its communication with the molecular visualization program Chimera. In addition, several other shell, python and perl scripts were developed to facilitate the user interface and commands associated with *ViewDock TDW*.

Deployment and testing of *ViewDock TDW* was successful at Osaka University in the Cybermedia Center and at UCSD in Calit2 (Figure 7). *ViewDock TDW* preserves the functionality of the Chimera software, allowing for visualizations of protein surfaces, calculations of hydrogen bonds, and many other features. The program also enables global control (e.g., rotation, zoom, view angle, etc.) over all ligand-receptor models, as well as specific control for commands to be sent to individuals or groups of ligand-receptor pairs on the TDW. This program will help greatly in the visualization and analysis of virtual screening data.

HIGHLIGHTED PRIME PROJECTS

E2. 3-D VISUALIZATION AND WORKFLOW TOOLS FOR MULTIPLE APPLICATIONS

Jade Kwan (PRIME 2009), Jürgen Schulze (UCSD), Ken Murata, Yoshinori Kobayashi (NICT), Shinji Shimojo (NICT and Osaka U), Atsuhiko Yasuda and Yoshihiro Katayama (XOOMS Co., Ltd.)

Communicating technical and scientific results to the public is increasingly important as a means to explain how public funding is used. A key tool in this dialog is visualization. The National Institute of Information and Communications Technology (NICT) in Japan conducts various types of research on topics from space environment to network analysis. In particular, NICT has the capability of displaying images and visualizations on high-definition tiled-display walls (TDW) for the general public to interact with.

Having datasets of space, planets, the solar system, or of computer network models, displayed on a TDW not only allows people to view the most minute details on a large scale, it also enables them to interact with the models in their own desired way (i.e., walking, flying, zooming, rotating). Successfully displaying these visualizations requires the knowledge and use of multiple software platforms and tools (e.g., Advanced Visual Systems (AVS), three-dimensional modeling tools like Autodesk Maya 2009 and Autodesk 3ds Max 2010, and COVISE/OpenCOVER) which provided a suitable environment to manipulate the datasets when displayed on tiled-display walls. Data visualization, also known as information aesthetics, needs proper design and function to perform successfully. This requires collabora- tions between NICT and experienced graphic companies like XOOMS Co., Ltd., and more generally, between scientists and artists.

The project integrated various tools, developed a workflow of visualization to display datasets of the Earth, Mercury, Sun, and Venus, along with a 3-D network map of a specific research group.

Image. Jade Kwan demonstrating image data on the tiled-display wall (TDW) at NICT.

E3. IMPLEMENTATION AND USE OF DOCK ON VIRTUAL CLUSTERS

Wen-wai Yim (PRIME 2009), Yasuyuki Kusamoto (Osaka U), Marshall J. Levesque (U Pennsylvania and PRIME 2006), Susumu Date (Osaka U), Jason H. Haga (UCSD)

One persistent problem in grid computing systems is the variation in cluster performance from each member organization. When performing virtual screening in the grid environment, this issue manifests itself as differences in molecular docking simulation scores across different clusters. One solution to the problem of cross-platform incompatibility is the developing area of virtual machine technology. The objectives of this project were to (1) determine the factors that contribute to inconsistent dock simulation performance across different clusters and (2) assess the use of virtual clusters for molecular docking simulations to solve possible grid heterogeneity problems.

The general approach to this project involved executing the same dock simulation on controlled physical clusters, then again on virtual clusters installed on the same physical clusters. The simulations were performed with DOCK 6.2, a molecular dock simulation program developed by the University of California San Francisco. The implementation of virtual clusters used Xen virtual machine technology. Mr. Yasuyuki Kusumoto from the Cybermedia Center at Osaka University provided virtual cluster network technology. Dock simulation experiments were also performed on PRAGMA Grid testbed clusters for comparison purposes. The various experiments included dock simulations from the DOCK 6.2 developer's test program, dock simulations on a single processor and multiple processors, dock simulations on multiple clusters using Opal OP and Jakarta Tomcat for job distribution, and

"Perhaps the biggest attribute I've found out about myself through this program is my perseverance. This program also made me realize how important international collaboration is because one country cannot possibly do everything alone." JEFFERSON HANG, U AUCKLAND

dock simulations on virtual computers through a virtual network. Results from the simulations were then compared across various clusters and correlated to their real- or virtual-platform specifications, including machine bit configuration, OS bit configuration, GCC version, Linux version, and dock installation optimization.

Preliminary results suggest that OS-bit configuration, GCC compilers, and dock installation compilation optimizations contribute significantly to the inconsistent dock simulation scores and performance. Virtual clusters do provide a solution to overcome the heterogeneous system, but sacrifice ease of set up and longer dock simulation times. One important issue to address is the communication of virtual clusters with local network security protocols on real clusters if virtualization is to be implemented successfully in the grid environment.

E4. GRID WORKFLOW TOOLS FOR PERFORMING MOLECULAR COMPUTATIONS

Elisa Abate (PRIME 2009), Kim Baldridge (U Zürich), David Abramson (Monash)

Many important and fundamental questions in organic and biological systems can be better understood by investigating them in their native environment. Such phenomena can be studied using quantum-mechanical hybrid techniques to look at the specific properties of individual molecules. One can investigate, for example, structural, energetic, electrostatic, and binding energy properties; the interactions within and between molecules rely heavily on these characteristic properties. The ability to analyze the various properties of molecules can aid our general understanding of their function in biological processes. For example, one might want to propose a more efficient drug interaction at a particular receptor site. Such investigations involve testing various molecular parameters to fully understand the variations in chemical response. Consequently, parametric computational experiments are very important to the understanding of biological processes because they provide a means to explore behavior of complex systems.

The computational tools that facilitated the project were: Kepler, GAMESS, APBS, and Nimrod/K. The molecular computations executed in GAMESS and APBS integrated quantum mechanics with continuum electrostatics strategy. The workflow and grid tools, Kepler and Nimrod/K, were then used to organize and manage the complex process. This combination allowed the extraction of proteins and ligands from databases and the preparation of GAMESS input files from coordinate files according to specified quantum mechanical methods. In addition, GAMESS ligand computations, which produced charges and radii necessary for computations with APBS software, were run and followed with APBS computations for the electrostatic field and binding energies of ligand protein interactions. These procedures were arranged into a parameter sweep and run via Nimrod/K on grid clusters. Such parametric experiments investigate changes in structure, environment, and methodology of the molecules. Also generated from this workflow were visual displays of the results. Successful implementation of the described workflow provides a functional framework that can be used easily by others for their specific protein-ligand investigations.

WORKING GROUPS 2009-2010

Resources and Data Working Group

THE GOAL OF THE RESOURCES AND DATA WORKING GROUP is to improve the interoperability of grid middleware, as well as make grids easier for scientists to use. . During the past year, the group continued its routine-use experiments with grid applications and middleware, finalized development of VOMS—now used by the GEO Grid, and also worked on new paradigms and new technologies. The group is poised to experiment with Virtual Machine and Cloud Computing.

Currently, the PRAGMA Grid has 24 sites in 16 countries/regions which provides a total of 1022 CPUs, more than 1.3 terabytes of memory, and over 24.8 terabytes of online storage. Over the last year, the Grid has gone through an improvement process. Many sites have upgraded, or are in the process of upgrading their resources.

More than 20 applications have been developed and run using PRAGMA Grid resources this year, with topics ranging from biology to image processing and from chemistry to sensor-data analysis. Most of these applications continue to utilize grid middleware, such as Ninf-G, Nimrod, Gfarm, CSF or DataTurbine. These routine-use experiments result in continuous enhancements for all of the application middleware involved. DataTurbine, for example, has provided a very essential data-streaming solution for many applications and projects, (see DataTurbine in Middleware section below). At the same time, the integration and deployment of these applications and projects also provide the means to advance DataTurbine as a software. After DataTurbine was deployed at several GLEON sites, a problem was revealed wherein data was lost due to network connectivity interruption or power outages. After in-depth research, the DataTurbine team discovered the cause and gained an understanding of the problem, then created and tested "a push mirror" scheme which pushes all of the data collected during the interrupt, once connectivity was re-established. This led to a new DataTurbine software release—V32.B2. See Table 1 for more application examples.

PRAGMA Grid has become an essential resource and support for the PRIME and MURPA programs. This year, 11 students ran applications and worked on their projects using the PRAGMA Grid—getting support and learning from the experts in the PRAGMA Grid team. In turn, the projects that the PRIME students worked on further contributed to the research and development of the PRAGMA grid. For example, one PRIME student worked on a "Virtual Machine for Docking" project this summer on the PRAGMA grid and gained meaningful insights (see PRIME, MURPA, PRIUS section, Yim). Through this work, the PRIME student learned a great deal about state-of-the-art research topics and grid computing from advisors and experts in PRAGMA Grid. Since the Resources and Data Working Group is shifting its focus toward virtual machines and plans to run an Avian Flu application as its pilot case, this PRIME project and its results by the student have immediate relevancy and importance to our current and future works in this area.

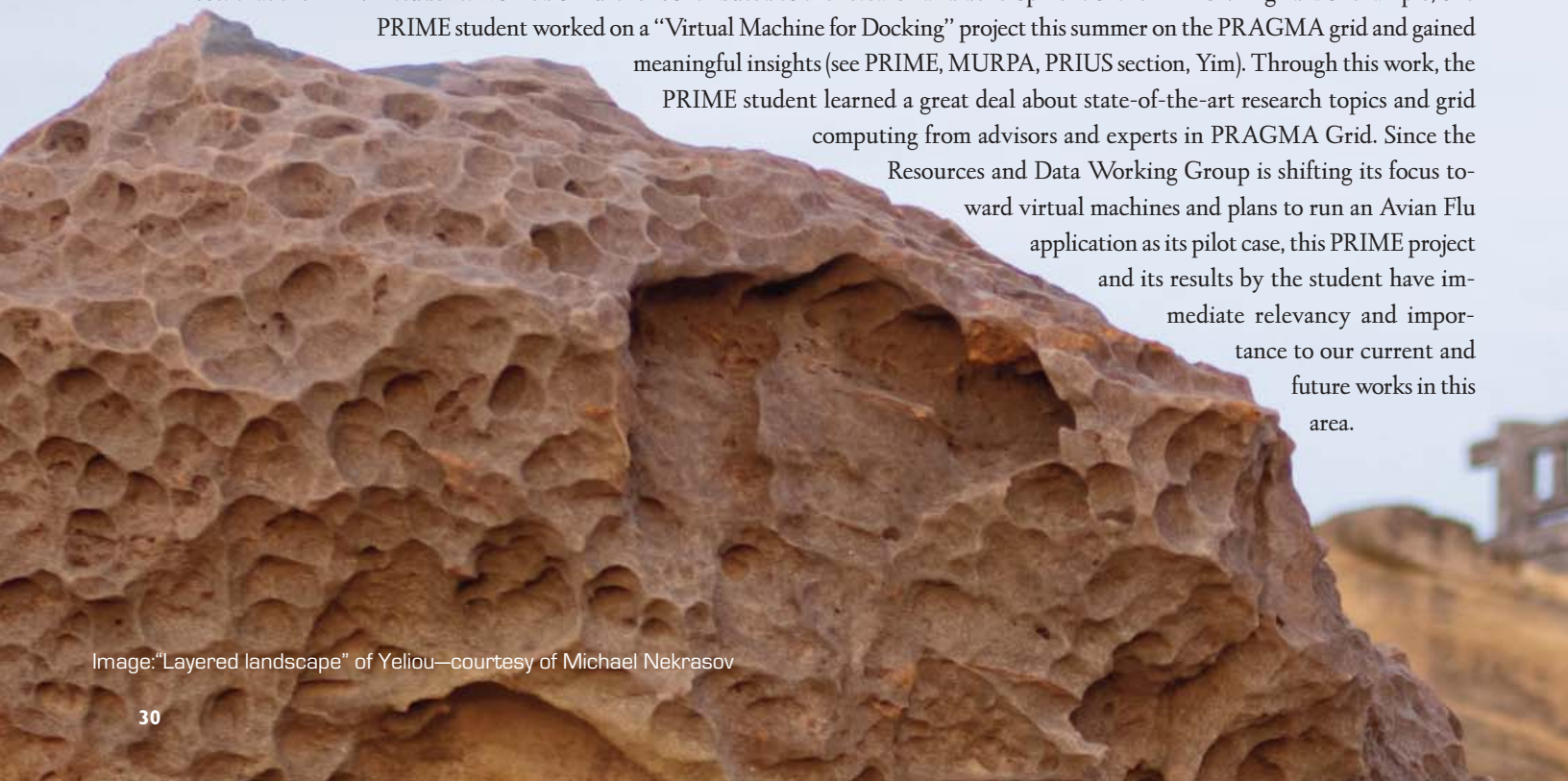


Image:"Layered landscape" of Yeliou—courtesy of Michael Nekrasov

The four working groups reflect the interests and efforts of current PRAGMA members. Each group has a lead, or coleads, who coordinate the activities of the group during and between meetings. Each group has projects that bring a focus to their activities, working towards milestones and giving demonstrations at PRAGMA workshops and other meetings. This structure allows for the inclusion of new applications, as well as new areas of interest.

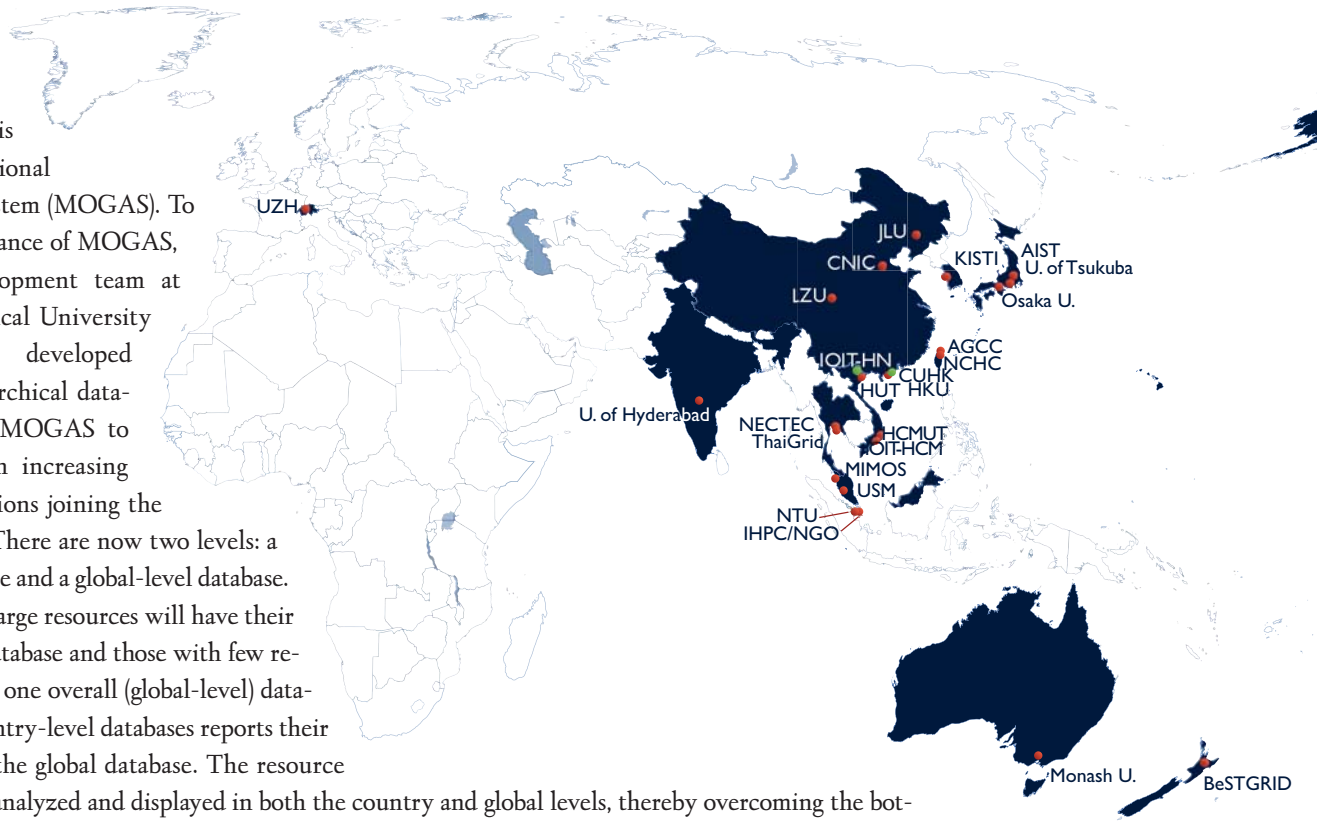
A major effort of the Resources and Data Working Group culminated this year in creating PRAGMA VOMS to scale administration of the PRAGMA Grid. Many PRAGMA Grid sites are themselves grids, or working with many other local or regional grids. We need to interoperate among all. BeSTGrid, for example, is a local grid and has joined PRAGMA Grid as one entity. Another example is a pilot project that requires PRAGMA Grid to work with GEO Grid in Japan. To ease the setup for access and authentication among these grids, while serve conflicting needs among all users—some require a shared group account, others require a persistent and dedicated individual account—the Resources group has worked for two years on development and now has a tested solution. The solution requires a set of software working cohesively: VOMS, GUMS, Globus and PRIMA. Four PRAGMA Grid sites (AIST, Japan; BeSTGrid, New Zealand; IHPC, Singapore; and SDSC, USA) formed a VOMS pilot team, then setup and tested this integrated system. The team also implemented and tested a pool account using GUMS. The pool account scheme automatically maps each member of a group to a persistent individual account without the need for system administrators to create each account one-by-one. This greatly reduces the effort needed for both system administrators and users. This work also produced step-by-step documentation for both system administrators and users and makes it easier for others to implement or use the setups. goc.pragma-grid.net/wiki/index.php/VOMRS.

Grid infrastructure middleware continues to develop and improve in the PRAGMA Grid. For example, the DataTurbine team found the PRAGMA Grid to be a realistic network environment with a large set of geographically distributed machines spanning multiple administrative boundaries and a realistic client workload, and therefore, is an ideal testbed for conducting scaling and robustness experiments under real-world conditions and at global-scale. DataTurbine has been tested in the PRAGMA Grid, is used in many application areas within the PRAGMA community and has been integrated with many key software applications. Examples include collaborations between DataTurbine, CNIC and NBCR which have resulted in the integration of docking with DataTurbine for biomedical research projects and collaborations with CREON which resulted in integration with OptiPortals, thus bringing real-time sensor data online. DataTurbine also has been deployed in a GLEON testbed and used in other PRIME projects, such as being integrated into a tsunami warning system and an underwater camera system (see PRIME, Nekrasov). The DataTurbine team also worked with commercial entities to integrate DataTurbine with open source software, such as ESPER (for event streaming and detection) and GlassFish (an application server). (goc.pragma-grid.net/wiki/index.php/DataTurbine)

The Resources group continued building a data grid in the PRAGMA Grid with Gfarm software. For four years, the Gfarm development team has collaborated with many PRAGMA Grid sites, testing and improving Gfarm software. Version 2.2 was released March 2009; version 2.3 was released September 2009. Each release includes significant improvements in both functionality and performance, (see datafarm.apgrid.org). The Gfarm v2 Grid file system is now operating in the PRAGMA Grid testbed in a production phase, providing a global file system for users among Pacific Rim countries. One pilot project for Gfarm is GEO Grid. In collaboration with the University of Tsukuba, Japan, the National Institute of Advanced Industrial Science and Technology (AIST), Japan and UCSD, USA, Gfarm has been setup for GEO Grid and now supports the sharing of 150 GBytes of the GEO science data between Japan and the U.S.

WORKING GROUPS 2009-2010

Another middleware in PRAGMA Grid which made significant improvements is the Multi-Organizational Grid Accounting System (MOGAS). To improve the performance of MOGAS, the MOGAS development team at Nanyang Technological University in Singapore has developed MOGAS with hierarchical databases. This enables MOGAS to scale better with an increasing number of organizations joining the PRAGMA testbed. There are now two levels: a country-level database and a global-level database. Countries that have large resources will have their own country-level database and those with few resources can share the one overall (global-level) database. Each of the country-level databases reports their summary to update the global database. The resource usage information is analyzed and displayed in both the country and global levels, thereby overcoming the bottleneck at the MOGAS server which was experienced previously due to the large number of updates from all the sites. (For more details, see ntu-cg.ntu.edu.sg/mogas_pragma/MOGAS_Federated_Database.doc.) The team has deployed a global-level database on the PRAGMA Grid Operations Center server in the U.S. and a country-level database on the PRAGMA Grid system in Japan. The accounting data can be viewed at goc.pragma-grid.net.



Application name	Domain	Middleware	Lead Organization
PovRay Rendering	Computer Science	Ninf-G	AIST, Japan
CSTFT	Environmental Monitoring	Ninf-G	UPRM, USA
Gridjobs	Computer Science	Globus	UPRM, USA
Heterodimerization between mu and delta receptors	Biology		LZU, China
RIMES	Computation Biology	DataTurbine	CNIC, China
DEER	Biomedicine	Globus	USM, Malaysia
GridDebug	Computer Science		MU, Australia
fMRI	Medical Science		UMelb, Australia
PIAX	Computer Science		OsakaU, Japan
Avian Flu Grid	Biomedical Science	DataTurbine, CSF,	NBCR/UCSD, USA
GridWebOS	Computer Science	RB/RIP	NCHC, Taiwan
Matlab parameter sweeps	Human Computer Interaction	Nimrod/G	Monash, Australia
11 PRIME and MURPA applications	Organic Chemistry, Computer Science, Cardiac Modeling, Computational Chemistry	Globus, Opal, Nimrod/G, Nimrod/O, Nimrod/E, Nimrod/K	UCSD, USA; OsakaU, Japan; UoHyd, India; Monash, Australia

Table 1: Key application activities in PRAGMA Grid

The Resources group has always been very active regarding knowledge sharing with the broader grid communities. Many members in the group have participated in multiple conferences, such as SC09, ACOMP2009, OGF, CCGrid, and Grid Asia. In addition, PRAGMA started organizing PRAGMA Institute training programs during the past year, where many members of the Resources and Data Working Group have given numerous presentations and tutorials. Many collaboration teams in the group have published papers and documented issues, solutions and lesson-learned on

the many major projects in which they are involved. There were an increased number (and duration) of visits and exchanges among PRAGMA Grid sites this year. These visits brought people together and enabled more efficient and productive team efforts.

Overall, the Resources and Data Working Group aims to develop, deploy, and test all layers of grid software which can simplify and ease the use of grids for scientists. In the coming year, the group will focus on developing and testing metascheduling software and portal services, building datagrids and linking sensor networks to the PRAGMA Grid, as well as running more diverse applications to utilize and test datagrid and sensor networks. Future challenges include seeding real science applications and helping more scientists begin using the grid.

ACKNOWLEDGEMENTS: The Resources and Data Working Group, as well as PRAGMA in general, welcome and wish to acknowledge the contributions of institutions participating in PRAGMA and the PRAGMA Grid efforts but have not yet become official members (see Supporting Institutions).

MORE INFORMATION: See goc.pragmagrid.net and the Publications section for research publications.

RESOURCES AND DATA: *M. Katz, co-chair, SDSC at UCSD; Y. Tanaka, co-chair, AIST; Y. Tanimura, T. Ikegami, A. Iijima, AIST; E. Yen, H. Shih, ASGCC; V. Mencl, BeSTGrid; K. Nan, K. Dong, CNIC; R. Yeung, S. Tang, C. Chu, CUHK; T. Nam, T. S. Nguyen, K. D. Tran, HCMUT; W.K. Kwan, F. Cheung, HKU; T. T. Nguyen, T. T. Tran, D. T. Le, HUT; T. Hung, H. M. Chan, B. Yeo, Y. S. Foo, IHPC; T. V. Lang, V. T. Dao, A. K. Tran, IOIT-HCM; X. H. Wei, Z. H. Ding, Y. Luo, Q. N. Guo, JLU; S. Hwang, H. Kim, J. Moon, K.W. Cho, KISTI; W.B. Chen, Y. Zhang, H. Yan, LZU; H. Kauthary, J. Y. Luke, MIMOS; D. Abramson, C. Enticott, B. Bethwaite, S. Androulakis, MU; W. Huang, M. Yu, NCHC; S. Vannarat, S. Prueksaaroon, NECTEC; F. Lee, J. W. Zhang, Y. S. Tan, NTU; S. Date, S.Kuwabara, K. Kokubo, Osaka U.; P. Papadopoulos, C. Zheng, SDSC; P. Uthayopas, S. Sriprayoonsakul, S. Phatanapherom, ThaiGrid; A. Jofre, J. C. Maureira, C. B Retamal, U. Chile; S. Pandey, U. Melb; J. L. Gordillo Ruiz, E. Murrieta Leon, P. Martinez, P. Palacios, UNAM; A. Agarwal, R. Wankar, R Atlury, B. Rao, UoHyd; D. Rodriquez, W. Rivera, J. Sanabria, UPRM; H. Wahab, F. Haron, C. H. Yong, S. K. Ang, H. L.. Koh, S. Y. Teh, K. B. Tan, USM; O. Tatebe, U Tsukuba; K. Baldridge, N. Williams, M. Packard, UZH.*

COMPUTE GRID ONLY: *AIST, ASGC, BeSTGrid, CNIC, CUHK, HCMUT, HKU, HUT, IOIT-HCM, IHPC, KISTI, JLU, LZU, MIMOS, Monash, NCHC, NECTEC, NGO, Osaka U., SDSC, ThaiGrid, U. Chile, Nat. U. Mexico (UNAM), UoHyd, U. Puerto Rico Mayagüez (UPRM), USM (Computer Science, Mathematical Science and Pharmaceutical), and U. of Zürich.*

Middleware	Participating Organizations
Middleware	Participating Organizations
DataTurbine	UCSD, USA; GLEON; CREON; CNIC, China
MOGAS	NTU, Singapore; KU, Thailand; all PRAGMA Grid sites
VOMS/GUMS/PRIMA	BeSTGrid, New Zealand; AIST, Japan; IHPC, Singapore, UCSD, USA
Gfarm	UTsukuba, Japan; AIST, Japan; UCSD, USA
SCMSWeb	KU, Thailand; all PRAGMA Grid sites
Nimrod/G	Monash, Australia; UCSD, USA

Table 2: Key middleware activities in PRAGMA Grid

RESOURCES: SDSC (host CA, VOMS, Nimrod portal, Bioportal, GOC, GLEON-INCA).

APPLICATION INFORMATION:

1. goc.pragma-grid.net/wiki/index.php/Ninf-G_demo_at_SC08_--_a_distributed_PovRay_rendering
2. goc.pragma-grid.net/wiki/images/d/da/Cstft.pdf
3. goc.pragma-grid.net/wiki/index.php/User:Sanabria
4. www.ncbi.nlm.nih.gov/pubmed/19214754
5. goc.pragma-grid.net/wiki/images/7/73/RIMES.doc
6. goc.pragma-grid.net/wiki/index.php/DEER
7. goc.pragma-grid.net/wiki/index.php/GridDebug
8. www.gridbus.org
9. goc.pragma-grid.net/wiki/index.php/Grid_middleware_on_PIAX
10. goc.pragma-grid.net/wiki/index.php/Avian_Flu_Grid
11. prime.ucsd.edu/student_collections2009.htm

WORKING GROUPS 2009-2010

Biosciences Working Group

Bioscience applications are very diverse, with different computing and data management requirements. The Biosciences Working Group aims to conduct cutting-edge scientific research using available computational resources from the PRAGMA Grid, high-performance computing clusters (such as the TeraGrid), and emerging cloud-computing resources. We collaborate across working groups, interacting with computer scientists to ensure that emerging technologies meet the requirements of biological applications. By engaging diverse scientific expertise to address biological problems of common interest through international partnerships, we expedite research and development activities through seminal research and development, as well as training and outreach activities in the global community.

One major focal activity of the Biosciences Working Group has been our work on addressing the emerging pandemic threat of the avian influenza virus, in particular, the H5N1 subtype, and the currently pandemic H1N1 subtype. We have established a virtual organization, called Avian Flu Grid (See Accomplishments), to facilitate data and resource sharing and algorithmic development in the ongoing fight against infectious diseases. This activity has provided specific usage scenarios for middleware such as CSF4, Gfarm, Glyco-M*Grid, Opal, and their transparent access through various types of user applications and virtualized problem-solving environments, such as portals and workflow tools. For example, CSE-Online is an enabling online gateway for virtual screening on both the PRAGMA Grid and TeraGrid. Vision is a desktop workflow application that can access Opal-based web services for applications such as AutoDock4 and NAMD2. Opal 2.1 now supports production use of CSF4 on the PRAGMA Grid through improved plugin architecture. Other examples of member-contributed technologies include DrugScreener-G from KISTI, Grid Application Platform from ASGCC, and RIMES for real-time simulation data streaming and visualization from CNIC.

The Biosciences Working Group has actively participated in the training and dissemination activities offered by projects such as PRIME, PRIUS, and MURPA. (See PRIME, PRIUS, MURPA section for details.) The students from these projects have contributed to various aspects of the avian flu research, from the discovery of novel hits for potential avian flu inhibitors to middleware components that are reusable by other researchers. They have fostered close ties between researchers around the Pacific Rim working on similar problems and promoted both cultural and scientific exchanges. Many students continue to participate in avian flu research and coauthor seminal research publications in high-quality journals.

In addition, this year members of the Biosciences Working Group organized a one-day workshop in advance of the PRAGMA 16 Workshop. The workshop, the First International Workshop for Infectious Disease Research in Cyberinfrastructure, brought together researchers from the diverse fields of biochemical engineering, molecular biology, computational biology, bioinformatics, and grid and distributed computing. One outcome of the workshop was an agreement to collaborate between the WISDOM project, one of the leading projects in application of grid computing research, in the EGEE (Enabling Grids for E-sciencE) project of Europe and the Biosciences Working Group on the application of grid computing to biological sciences, especially in the area of large-scale virtual screening and drug discovery.

Together with the Resources and Data Working Group, we are constantly refining the requirements for the Avian Flu Grid identifying newly emerging technologies that may prove beneficial to biosciences research, reinforcing the routine use of distributed resources, and making them accessible to the broader biomedical community.

PARTICIPATING RESEARCHERS: *H. Wahab, co-chair, USM; W. W. Li, NBCR at UCSD, co-chair; Jincheol Kim, KISTI and Pohang University of Science and Technology; K. Jeong, Konkuk U.; K. Baldridge, SDSC at UCSD and UniZH; H. Pao, K. Wu, and D. Xu, UCSD; T. N. Truong, CSE-Online, U. Utah*

COLLABORATORS: *R. Amaro, UCI; M. Alam, U. Hawaii; J.H. Lin, Academia Sinica, National Taiwan U.; Doman Kim, Chonnam University; Phineus Markwick, UCSD*

RESOURCES: *K. Nan and K. Dong, CNIC, CAS; X. Wei, CCST, JLU; Y. Tanimura,AIST; O. Tatebe, CCS, U Tsukuba; S.T. Hwang, Kookmin U.; J. Lee, K. Kee, K. Cho, KISTI; H. Lee, ASGCC*

Telescience Working Group

The Telescience Working Group aims to create and develop new information technology that allows scientists to remotely use advanced technology and devices and demonstrate future science on an advanced cyberinfrastructure. Examples of such devices include high-accuracy scientific measurement devices such as ultra-high-voltage electron microscopes, widely distributed devices such as sensors and highly sophisticated visualization facilities, such as tiled-display walls (TDWs). The working group's activities have progressively gained in importance, expanding year by year because of the ubiquity and maturity of high-speed network technology.

The R&D activities in the Telescience Working Group this year fall into four types of projects. The first is the use and control of electron or other microscopes and the related analysis of the data. The second is the development of environmental observing systems, starting with NCHC's Ecology Grid (EcoGrid) which inspired other grass roots efforts and collaborations such as the Global Lake Ecological Observatory (GLEON) and the Coral Reef Ecological Observatory Network (CREON). The third component is the use and dissemination of tiled-display wall (TDW) technology as a means for sharing information and collaborating. A fourth and final area involves the use of high-definition (HD) video to conduct distributed seminars, with a long-term vision of creating course content for PRAGMA members. We summarize key progress made in the past year for two of these areas: environmental observing systems and TDWs.

Environmental observing systems are growing in use and importance as means to take measurements on greater spatial and temporal scales. Members of the Telescience Working Group have been working with the members of the GEO Working Group to begin to link data from sensors with those from satellites via the GEO Grid technology as a way to improve modeling of lakes (see Lake Observing and Modeling Accomplishment). In addition, NCHC members have extended EcoGrid to new sites, such as Kenting's National Museum of Marine Biology and Aquarium (NMMBA), as well as to Orchid Island, off the coast of Taiwan, and installed an underwater video camera. In addition, NCHC supported the government's rescue and planning tasks in the aftermath of Morakot disaster (typhoon in August 2009). To complement the features of EcoGrid, a PRIME student worked with members of the Telescience Working Group (NCHC), as well as NMMBA in Taiwan, to extend modules for the Resources and Data Working Group's DataTurbine to include an underwater camera observing coral (see PRIME, Nekrasov). The NARC team continues to deploy Field Servers (for meteorological data collection), and their software MetBroker can now handle data from databases, predicted models, and real-time data.

Tiled-display walls (TDWs) continue to play a major role in the efforts of the Telescience Working Group. In addition to having PRIME students work on specific features (see PRIME section), the working group explored a new application: the visualization of cultural heritage objects in order to better explore approaches to scientific preservation and restorations. TDW technology is capable of visualizing various features in both 2-D and 3-D, such as the structure of an archeological site or different layers in paintings (looking at locations of previous restoration, as well as changes in original design). Working with Calit2 at UCSD, the collaborating institution, the National Institute for Information and Communication Technology (NICT) digitized a Renaissance-era Italian painting by Giotto, using multiple imaging technologies, such as ultra-violet (UV) and terahertz. All the images were compiled and shown on the TDW. The TDW can be a tool of restoration for curators and specialists; high speed networking allows researchers around the world to analyze these images and take steps towards a science-based restoration. At PRAGMA 16 in Daejon, Korea, members of the Telescience Working Group showed Giotto's painting via SAGE and TDW hardware prepared by KISTI, thus demonstrating the possibility of content exchange on the TDW.

Finally, NICT has developed the I-TDW (Interactive Tiled-display Wall) which allows a user to explore the content interactively with a mobile phone equipped with a motion sensor and a touch panel. The user can move the cursor by tilting the mobile phone, choose an item by tapping the touch screen, and zoom in/out by scratching the screen. The layered images switched to another by swinging the mobile phone (see adjacent image). A PRIME student (I Fanchiu, with R Kadobayashi and M Chikama) developed the interactive museum contents which give rich cultural information behind the visible image by integrating the I-TDW and scientific analysis of the art. In addition, she also (with K.Fukunaga) worked at NICT on aspects of the terahertz technology for art analysis.

Using the I-TDW, NICT has analyzed Giotto's (1267-1337) masterpiece, "Polittico di Badia" (ca. 1300) at the Uffizi Gallery by using multi-spectral analysis, including the terahertz imaging which gives non-invasive cross-section images of the art.

WORKING GROUPS 2009-2010

The Telescience Working Group has also helped incubate education and human development of internationally collaborative research groups. Through the PRIME and PRIUS projects, students from UCSD and Osaka University joined research teams from the Telescience Working Group in different counties for nine weeks. (See PRIME, PRIUS, and MURPA sections.)

In addition, we discovered several instances in which members of the group utilized high-definition (HD) cameras to communicate amongst themselves. In particular, the MURPA program is using HD video to hold lectures between UCSD and Monash. Furthermore, the Telescience Working Group demonstrated the capability to hold a three-way HD meeting for the Cardiac Modeling Day at Monash (see Accomplishments).

PARTICIPATING RESEARCHERS: *F-P Lin, co-chair, NCHC; S. Shimojo, co-chair, S. Date, T. Akiyama, E. Sakane, K. Nozaki, S. Kuwabara, A. Nakazawa, H. Takemura, CMC, Osaka U.; S. Kato, Hyogo U. of Health Sciences; M. Lee, KISTI; B. Durnota, Complexibotics; M. Ellisman, S. Peltier, A. Lin, T. Molina, R. Singh, NCMIR at UCSD; J. Schulze, Calit2 at UCSD; J.H. Woo, Konkuk U.; H. Chou, S-I Lin, S.Cheng, NCHC; S. Ninomiya, NARC, APAN; T. Fountain, S. Tilak, UCSD; K. Nan, K Dong, CNIC; D. Abramson, Monash*

Geosciences Working Group

The GEO Working Group is the youngest of the four current working groups in PRAGMA. The goal of the GEO Working Group is to develop a grid testbed to support geosciences and practical applications based on huge datasets of satellite imagery and GIS data, for uses such as disaster prevention and environmental monitoring. The group's development of the testbed will be coordinated with other international efforts, in particular, with the Global Earth Observation System of Systems (GEOSS), which seeks to achieve comprehensive, coordinated, and sustained observations of the Earth system, in order to improve monitoring of the state of the Earth, increase understanding of Earth processes, and enhance prediction of behavior.

The GEO Working Group is building on the strengths of its members. In particular, AIST has a very strong effort on GEO Grid, which is developing an infrastructure to integrate remote sensing and GIS data from various sources. The preliminary design of GEO Grid involves integrating all the relevant data virtually as a set of services. The software architecture and its preliminary implementations are based on grid computing, web service technologies and GIS standards defined by the Open Geospatial Consortium (OGC). The National Applied Research Laboratory (NARL) in Taiwan is a nonprofit organization, sponsored by the National Science Council, with nine research institutions. The NARL GEO (Global Earth Observatory)-Grid project is deploying a high-resolution 3-D GIS platform for Taiwan. Disaster event GIS data, as well as typhoon forecasting, rainfall, and flooding data can all be displayed over the platform. NARL includes NCHC, which is integrating NARL's GEO-related projects including operating an earth-observing satellite Formosat-2 imagery, GIS information, 3-D visualization, GLEON, and typhoon monitoring and prediction. The GEO Working Group will continue to promote collaboration between these projects, bringing more participants to this activity, and will begin technological integration of various components, with a particular focus on integration of data from various systems.

This past year the GEO Working Group members focused on several areas that overlapped with other PRAGMA working groups, in particular the Telescience and Resources working groups.

At PRAGMA 17, the GEO Working Group demonstrated the integration of data from two different modes: ASTER and FORMOSAT-2. This integration required technical database federation between AIST and the National Space Program Office, which is part of the National Applied Research Laboratory (NARL) of Taiwan. The value of being able to federate this type of data is that it provides a richer resource for researchers studying areas of the earth where the images overlap. Furthermore, it required ongoing collaborations on policy and agreements. Also demonstrated at PRAGMA 17 was the ability to integrate sensor data into this system. This was accomplished with a collaboration of the GEO Grid group and Dr. Seiichi Kato at Hyogo University of Health Science. This effort was a collaboration with the Telescience Working Group.

For another application, the GEO Grid team, again working with members of the Telescience group, demonstrated their ability to link GLEON data into the GEO Grid system (see Accomplishments). The interaction with a specific application group also helps disseminate the impact of GEO Grid technology and to extend the scientific impact of the effort. Further, members of GLEON include individuals that help shape policy on land use around lakes, thus allowing the technology to have a broader societal impact. It is expected that this interaction will increase over the next year.

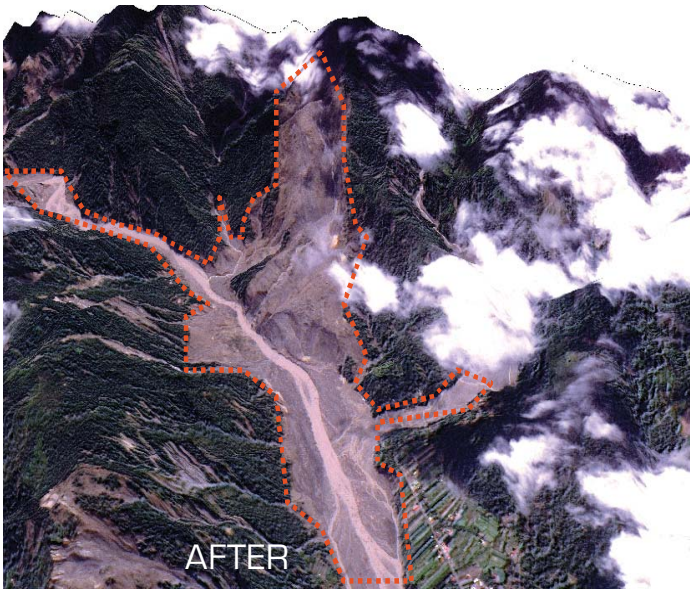


Image: 3D Stereo Images of Hsiao-lin Village-Disaster-Hit Village Before and After Typhoon Marakot (8 August 2009)

Given the often sensitive nature of environmental data, it is important to have secure data sharing. Working with the Resources and Data Working Group, the GEO Working Group has integrated the VOMS (Virtual Organization Management Service) into the GEO Grid. (See the Resources and Data Working Group section for a description of VOMS and its development this year.)

NARL has been implementing the "Taiwan Integrated Earth Observation System" (TIEOS) in order to integrate national environmental information related to nine social-benefit areas proposed in GEOSS: Disaster, Health, Energy, Climate, Water, Weather, Ecosystems, Agriculture and Biodiversity. The plan is to gather information related to disaster prevention, to assist in systematic monitoring of the country's environmental changes, and to improve forecasting accuracy in support of the government's decisions on disaster prevention. NARL hosted the first TIEOS international forum (TIEOS 2009) in June, which brought experts from the United States, Japan, Australia, and Korea to provide constructional recommendations for Taiwan.

The "3-D GIS Taiwan" project, led by NCHC, is a research project supporting TIEOS. Recent progress includes integration of satellite remote sensing and airborne image data for Taiwan, following the OGC standards. The August 2009 Typhoon Morakot brought a near-historic record of rainfall, caused landslides and floods, and killed hundreds of civilians in southern Taiwan. The 3-D GIS VR platform and Web 3-D GIS platform were used for decision support for disaster recovery, this clearly illustrated the need for near real-time image processing for disaster rescue and recovery.

In the aftermath of Typhon Morakot, NCHC extended the idea of real-time environmental observation to the idea of "be-there," which incorporates 3-D and panorama photography technologies with the 3-D GIS platform. The result generates multi-scale, multi-level scenarios that allows decision makers to have better insight into what is happening onsite.

This year, through the PRIME program, activities of the Telescience and Resources working groups were used to demonstrate streaming data between the Indian National Center for Ocean Information Services (INCOIS) and the University of Hyderabad, using DataTurbine (see Resources and Data Working Group for a description).

As part of its outreach activities, the GEO Grid group organized a successful workshop that was held in conjunction with the PRAGMA 17 workshop in Vietnam.

Future activities include extending the Satellite-Field data integration service from inland water (GLEON) to sea water (CREON) next year, with the possible additional application of coastline erosion monitoring.

PARTICIPATING RESEARCHERS: *R. Nakamura, co-chair, AIST; F. Cheng, co-chair, NARL; A. Memon, UCSD; S. Vannarat, P. Srichaikul, NECTEC. W-F Tsai, FP Lin, NCHC; A. Agarwal, U. Hyderabad; K. Nan, CNIC; K. W. Cho, KISTI; Y. Tanaka, S. Sekiguchi, AIST*

COLLABORATORS: *C-C. Liu, National Cheng Kung U.; P. Bajcsy, NCSA*

PUBLICATIONS

Resources and Data

Abramson, D, Bethwaite, B, Enticott, C, Garic, S, Peachey, T, Michailova, A, Amirriazi, S and Chitters, R, "Robust Workflows for Science and Engineering", to appear, *ACM MTAGS09, 2nd Workshop on Many-Task Computing on Grids and Supercomputers*, co- located with ACM/IEEE SC09 (International Conference for High Performance, Networking, Storage and Analysis), Portland, OR, Nov 16, 2009. [Amirriazi and Chitters were PRIME students in 2007 and 2009, repectively]

Abramson D, Bethwaite B, Dinh M, Enticott C, Firth S, Garic S, Harper I, Lackmann M, Nguyen H, Ramdas T, Russel A.B.M, Schek S, Vail M. "Virtual Microscopy and Analysis using Scientific Workflows." Accepted by *e-Science 2009*, Oxford, UK, Dec 9-11, 2009.

Ding Z, Wei X, Tatebe O, Papadopoulos PM, Arzberger PW, Li WW. "Cyberinfrastructure for Biomedical Applications: Metascheduling as Essential Component for Pervasive Computing." In: Cao J, editor. **Cyberinfrastructure Technologies and Applications**, Nova Science, 2009. ISBN: 978-1-60692-063-3.

Kokubo K, Ichikawa K, Ishi Y, Date S. "A Proposal of P2P-based Resource Discovery Mechanism for the Grid." *Proceedings of 2009 9th Annual International Symposium on Applications and the Internet*, July 2009, pp. 288-289.

Wei X, Ding Z, Yuan Y, Li WW. "VJM: a novel grid resource co-allocation model for parallel jobs." *Int J Grid Distributed Computing*, 1 (2), 2009, pp.1-12.

Biosciences

Ding Z, Wei X, Zhu Y, Yuan Y, Li WW, Tatebe O. "Implement the Grid Workflow Scheduling for Data Intensive Applications as Scheduling Plug-ins." *IEEE eScience 2008*, Indianapolis, IN, 2008.

Newhouse EI, Xu D, Markwick PRL, Amaro RE, Pao HC, Wu KJ, Alam M, McCammon AJ, Li WW. "Mechanism of Glycan Receptor Recognition and Specificity Switch for Avian, Swine and Human Adapted Influenza Virus Hemagglutinins: A Molecular Dynamics Perspective." *J Am Chem Soc*, 2009. [Pao was a PRIME student in 2007.]

Krishnan S, Clementi L, Ding Z, Arzberger PW, Li WW. "Leveraging the Power of the Grid with Opal: A Guide to Biomedical Application Developers and Users." In: Cannataro M, editor. **Handbook of Research on Computational Grid Technologies for Life Sciences, Biomedicine and Healthcare**, Milan, 2008.

Krishnan S, Clementi L, Ren J, Papadopoulos P, Li WW. "Design and evaluation of Opal 2: toolkit for scientific software as a service." International Workshop on Cloud Services, International Conference on Web Services 2009, Los Angeles, 2009.

Liu X, Kai M, Jin L, Wang R. "Computational study of the heterodimerization between mu and delta receptors." *Journal of Computer-aided Molecular Design*, 6 (Jun23) 2009, pp. 321-32.

Allyson Clark, Sasha Koruga, Christopher Lau, Wen-wai Lim, and Cory Stevenson in Kyoto

Caving in New Zealand—courtesy of Peter Arzberger

Publications have been grouped into areas paralleling the working groups (some papers span more than one group). These publications have been offered by the authors as representing research efforts that have benefited, at least in part, from participation in PRAGMA.

Xu D, Newhouse I, Pao H, Amaro RE, McCammon JA, Li WW, Arzberger PW. "Distinct Glycan Topology for Avian and Human Sialo-Pentasaccharide Receptor Analogues upon Binding Different Hemagglutinins: A Molecular Dynamics Perspective." *J Mol Biol*, 2009, pp. 465-491. [Pao was a PRIME 2007 student.]

Telescience

Fountain T, Tilak S, Shin P, Holbrook S, Schmitt RJ, Brooks B, Washburn L, Salazar D. "Digital Moorea Cyberinfrastructure for Coral Reef Monitoring." Accepted by *ISSNIP (Dec 2009, Melbourne AU)*.

Sakamoto N, Kuwano H, Kawamura T, Koyamada K, Nozaki K. "Visualization of Large-scale CFD Simulation Results Using Distributed Particle-Based Volume Rendering, International Journal of Emerging Multidisciplinary Fluid Sciences." In press.

Nozaki K, Baba K, Noro M, Nakagawa M, Date S, Shimojo S. "Distributed E-Science Application for Computational Speech Science, Remote Instrumentation Services on the eInfrastructure: Applications and Tools." In press.

Nozaki K, Date S, Baba K, Nakagawa M, Shimojo S. "A Large Scale Visualization Platform for Speech Analysis." *IPSJ Journal*, Vol.2, No. 3, 2009, pp. 189-200.

Sakane E, Kido Y, Date S, Shimojo S. "A Data Management and Visualization System towards Online Microscopic Imaging." *Proceedings of 32nd International Convention MIPRO, Conference on Grid and Visualization Systems*, May 2009, pp.335-340.

Sanabria J, Rivera W, Rodriguez D, Yunes Y, Ross G, Sandoval C. "A sensor grid framework for acoustic surveillance applications." *Proceedings of the Fifth International Conference on Network Computing (INC2009)*, Seoul, Korea, August 25-27, 2009. In press.

Sanabria J, Rivera W, Rodriguez D. "A framework for opportunistic execution in grid environments." *The 2009 World Congress in Computer Science Computer Engineering and Applied Computing*, Las Vegas, NV, July 13-16, 2009, regular research paper (RRP) ID #: GCA9043.

Geosciences

Agarwal A, Baru C, Crosby C, Keller R, Subbarao KV, Nandigam V. "iGEON-India: An International Collaborative Activity of the GEON Project." Accepted by Earth and Space Science Informatics-1 session, Geophysical Research Abstracts, Vol. 11, EGU2009-11792-1, *European Geosciences Union General Assembly*, Vienna, Austria, April 19-24, 2009.

PRIME, PRIUS, MURPA

Arzberger P, Wienhausen G, Abramson D, Date S, Lin F-P, Nan K, Shimojo S. "PRIME: an integrated and sustainable undergraduate international research program." Submitted for publication.

Jessica Hsieh with a rhinoceros hornbill at the Penang Bird Park—courtesy of Jessica Hsieh

WORKSHOPS AND INSTITUTES

Working meetings are held biannually to allow members to review accomplishments and plan for future activities. These workshops are often held in conjunction with special topic workshops or PRAGMA Institutes.

Future Meetings

PRAGMA 20: Spring 2011, Hong Kong, China. Hosted by Hong Kong Univ.

PRAGMA 19: 13-15 September 2010, Changchun, China; Hosted by Jilin Univ.

PRAGMA 18: 2-4 March 2010, San Diego, USA; Hosted by UCSD; *Held in conjunction with a workshop on Sensing and Observing Technologies for Coral Reef Ecology*

2008-2009 Workshops and Institutes

5TH PRAGMA INSTITUTE: 30 November-4 December 2009; Hosted by NCHC; *held in conjunction with the Southeast Asia International Program*

PRAGMA 17: 28-30 October 2009, Hanoi, Vietnam; Hosted by IOIT; *Held in conjunction with a GEO Grid Workshop, 28 October 2009*

PRAGMA 16: 23-25 March 2009, Daejon, Korea; Hosted by KISTI; *Held in conjunction with the First International Workshop on Infectious Disease Research in Cyberinfrastructure (IDRiC International Workshop), 23 March 2009*

4TH PRAGMA INSTITUTE: 1-5 December 2008; Hosted by NCHC; *Held in conjunction with SEAIJRTP-HPCN ("Southeast Asia International Joint Research and Training Program in High-Performance Computing Applications and Networking Technology")*

3RD PRAGMA INSTITUTE: 21-22 October 2008, Penang Malaysia; Hosted by USM

PRAGMA 15: 22-24 October 2008, Penang, Malaysia; Hosted by USM

PRAGMA 14: 10-12 March 2008, Taichung, Taiwan; Hosted by NCHC

SUPPORTING INSTITUTIONS

In addition to our members, many institutions participate, or have participated, actively in PRAGMA. We welcome and value their contributions in extending the PRAGMA Grid to new sites.

- The Chinese University of Hong Kong (UCHK; www.cuhk.edu.hk/v6/en), located in Hong Kong, China, has contributed resources to the PRAGMA Grid and attended PRAGMA 10, PRAGMA 12, PRAGMA 13, and PRAGMA 14 workshops.
- The Ho Chi Minh City Institute of Information Technology (IOIT-HCM; www.ioit.ac.vn/pages/index.asp?lang=1), is a scientific research institution of the Vietnamese Academy of Science and Technology (VAST), located in Ho Chi Minh City, Vietnam. IOIT-HCM is one of the five leading organizations of VNGrid project. It has contributed resources to the PRAGMA Grid and developed and tested grid applications in the PRAGMA Grid. IOIT-HCM researchers have built grid-based applications including bioinformatics, geoinformatics, telemedicine, and e-Science in computational fluid dynamics research.
- LanZhou University (LZU; www.lzu.edu.cn), located in Lanzhou, China, has contributed resources to PRAGMA Grid and attended PRAGMA 12 and 13 workshops.
- Nanyang Technological University (NTU; www.ntu.edu.sg), located in Singapore, has been developing and maintaining MOGAS—a grid accounting software system in PRAGMA Grid and attended PRAGMA workshops since 2006.
- National Applied Research Laboratory (NARL; www.narl.org.tw/en) was established in 2003 to consolidate nine 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 Communication 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 for the future. NICT supported students in the PRIME program in 2009 and has participated in the activities of the Telescience Working Group through support of the high-definition video conferencing testing.
- Universidad de Chile (UChile; www.uchile.cl), located in Santiago, and is the lead of Chile's national grid initiative (CLGrid). CMM is contributing to the PRAGMA Grid and the development of a multi-user data farm (based on Gfarm) in conjunction with UPRM and CPTEC. Now, CMM is receiving funding to push CLGrid to the next level of grid deployment, testbeds and international collaboration.
- Universidad Nacional Autónoma de México (UNAM; www.unam.mx), located in Mexico City, has contributed resources to the PRAGMA Grid.
- University of Puerto Rico at Mayagüez (UPRM; www.uprm.edu), located in Mayagüez, Puerto Rico, has contributed resources to PRAGMA Grid, attended the PRAGMA 13 workshop and is running a sensor data analysis application in PRAGMA Grid.
- Organic Chemistry Institute Grid Competence Center (GC3), University of Zürich (UniZH; www.oci.unizh.ch), located in Zürich, Switzerland, has contributed resources to the PRAGMA Grid and been involved in PRAGMA meetings from early on. GC3 researchers are involved in computational chemistry activities, including projects with Nimrod, GAMESS, APBS, Unicore and Kepler. They also mentor PRIME students.

OTHER PROJECTS AND ORGANIZATIONS

Other projects and organizations have collaborated with PRAGMA

The California Institute for Telecommunications and Information Technology (Calit2; www.calit2.net) has provided space for PRAGMA activities and visitors in Atkinson Hall, the home of its UCSD division. Calit2 is promoting the use of OptIPortal Technology and has helped in the tested of multiway HD video teleconferencing. In addition, Calit2 has provided support for students of PRIME.

The National Biomedical Computation Resource (NBCR; nbcr.net) is a National Institutes of Health National Center for Research Resources project and involves researchers at UCSD, The Scripps Research Institute, and Washington University. NBCR's goals are to conduct, catalyze and enable multiscale biomedical research through development, deployment and use of advanced cyberinfrastructure. NBCR is currently collaborating in the virtualization of aspects of the avian flu grid.

National Center for Imaging and Microscopy Research (NCMIR; ncmir.ucsd.edu) is a National Institutes of Health National Center for Research Resource at UCSD. NCMIR's interdisciplinary staff develops state-of-the-art 3-D imaging and analysis technologies to help biomedical researchers understand biological structure and function relationships in cells and tissues in the dimensional range between 5 nm^3 and 50 μm^3. NCMIR researchers have helped mentor PRIME students.

The Global Lake Ecological Observatory Network (GLEON; www.gleon.org) is a grassroots network of limnologists, information technology experts, and engineers who have a common goal of building a scalable, persistent network of lake ecology observatories. GLEON is built on earlier projects in the Telescience Working Group and PRAGMA catalyzing a relationship between the EcoGrid project of NCHC's and the North Temperate Lakes Long Term Ecological Research project at the University of Wisconsin (www.lternet.edu/sites/ntl). It has been supported by the U.S. NSF, the Gordon and Betty Moore Foundation, National Science Council of Taiwan, and member institutions. An additional key partner in launching GLEON is at the Center for Biodiversity and Ecology Research, University of Waikato, New Zealand (cber.bio.waikato.ac.nz), which hosts a PRIME student. Currently GLEON researchers are interacting with the GEO working groups (see Accomplishments).

The Coral Reef Ecological Observatory Network (CREON; www.coralreefeon.org) is a collaborating association of scientists and engineers from around the world striving to design and build marine sensor networks. Similar to GLEON, it grew out of Telescience Working Group activity. Its growth has been supported by the Gordon and Betty Moore Foundation, the U.S. National Science Foundation, the National Science Council Taiwan, and member institutions. There are strong collaborations between NCHC, Academia Sinica, the Moorea Coral Reef Long Term Ecological Research site (mcr.lternet.edu), the Australian Institute for Marine Science, and James Cook University, Australia.

The Open Source DataTurbine Initiative (OSDT; www.dataturbine.org) is an NSF-sponsored project to develop and deliver streaming-data middleware to the science and engineering communities. The OSDT Initiative includes national and international collaborators from universities, private companies, and government agencies. The OSDT Initiative publishes the DataTurbine software and participates in numerous sensor network and environmental observing system projects. The OSDT central office is located at UCSD.

PRAGMA SPONSORS

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.

CCST receives major funding support from the Chinese Natural Science Foundation, in particular China NSF, No.60703024 for CSF4 development.

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 is supported by the Chinese Academy of Sciences.

AIST's sponsors include the Special Coordination Funds for Promoting Science and Technology (MEXT, Japan) and the "Design and implementation of security infrastructure for U.S.-Japan science Grid," Strategic International Cooperative Program (JST, Japan).

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 funding.

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).

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, through the Knowledge Innovation National Grid (KING) project.

PRAGMA is supported by its 35 member institutions and the 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.

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

NECTEC receives its funding through the National Science and Technology Development Agency (NSTDA).

NGO receives funding from Singapore's A*STAR, the Defense Science & Technology Agency (DSTA), the Infocomm Development Authority (IDA), Nanyang Technological Univ. (NTU), and National Univ. of Singapore (NUS).

TransLight/StarLight receives major funding from the NSF (OCI-0441094).

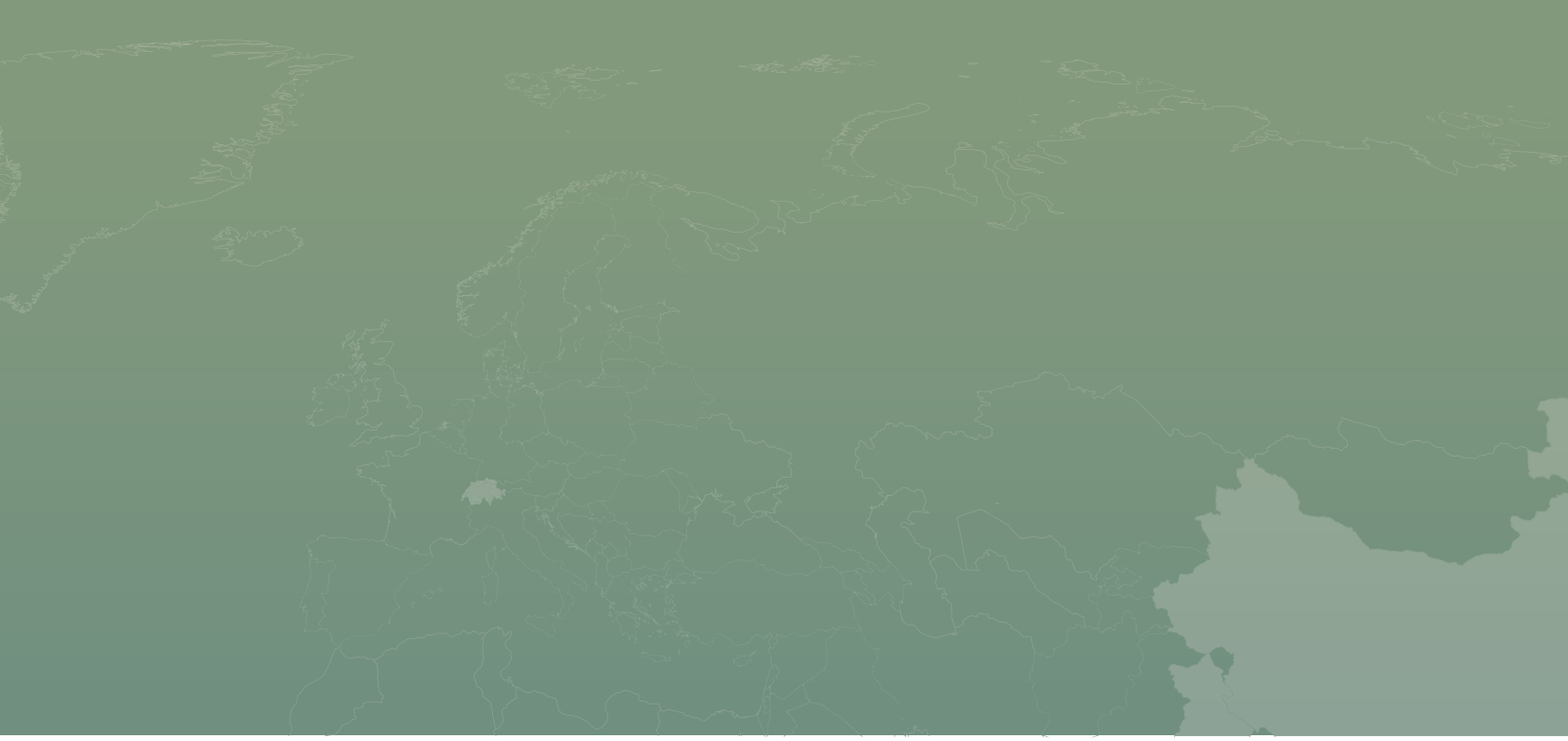
TNGC receives support under the Software Industry Promotion Agency (SIPA), Ministry of Information and Communication Technology, Thailand.

TransPAC2 receives major funding from the U.S. 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 0407508 with additional support from the Office of Cyberinfrastructure, NSF OISE 0710726 with additional support from the Division of Information and Intelligent Systems (IIS) and the Division of Chemical, Bioengineering, Environmental, and Transport Systems (CBET) and Calit2 UCSD; PI: G Wienhausen. Work to build GLEON is supported in part by an award from the Gordon and Betty Moore Foundation, which also funded some PRIME students in 2008 and 2009; Collaborative Research: Automating Scaling and Data Processing in a Network of Sensors: Towards a Global Network for Lake Metabolism Research and Education DBI-0446802; National Biomedical Computation Resource (NIH, NCRR P 41 RR08605), which funded a PRIME student in 2008.

Group Photo from the PRAGMA 15 meeting in Penang, Malaysia



www.pragma-grid.net