COLLABORATION OVERVIEW

2012-2013

Robust, international research scientists &

THE PACIFIC RIM APPLICATIONS AND GRID MIDDLEWARE ASSEMBLY

network of engineers

PRAGMA INTRODUCTION

The Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) is a robust, international network of research scientists and engineers from more than 30 institutions who address application driven questions through cyberinfrastructure development.

PRAGMA pursues activities in four broad interdependent areas:

- **Fosters international "scientific expeditions" by forging teams of domain scientists and cyberinfrastructure researchers** who develop and test necessary technologies to solve specific scientific questions and create usable, international-scale, cyber environments;
- **Develops and improves a grassroots, international cyberinfrastructure** for testing, computer science insight and advancing scientific applications by sharing resources, expertise, and software;
- **Infuses new ideas** by working with and supporting junior researchers to gain experience in cross-border science and by engaging strategic partners;
- **Builds and enhances the essential people-to-people trust and organization** developed through regular face-to-face meetings, a core component of PRAGMA's success.

None of these activities can stand alone. Ideas become stale without new people. Infrastructure that has impact must be built to meet the specific needs of applications, and scientific expeditions and data-sharing require organizational support to succeed.

PRAGMA's Past Successes

PRAGMA began in 2002 as a workshop series to explore the technical, organizational, and trust elements needed to enable these small- to medium-sized international networks of research scientists. It has since expanded its efforts: responding to various disasters or emergencies including SARS in 2003 and the Japanese Tsunami of 2011, as well as creating a persistent cyberinfrastructure test bed for international collaboration and enhancing multiple software packages.

PRAGMA has also incubated new scientific networks such as GLEON (Global Lakes Ecological Observation Network), and international projects such as the GEO Grid, and played a critical role in the formation of programs that help train the next generation of cyber-scientists through research exchanges, e.g., PRIME (Pacific RIM Experiences for Undergraduates, University of California San Diego), PRIUS (Pacific Rim International UniverSities, Osaka University) and MURPA (Monash University Research Project Abroad).

The Future of PRAGMA

Over the next several years, PRAGMA will focus on several expeditions that are crucial to its mission. One of these **sustained scientific expeditions** will address the societal issues inherent to **predicting lake eutrophication** (i.e., defined as excessive plant, algal and bacterial growth due to nutrient enrichment)—one of the greatest water quality challenges facing freshwater ecosystems throughout the world.

30.

5.

A second project seeks to **understand biological adaption in extreme environments** while researching the specific biotic, abiotic, and evolutionary factors that affect patterns of diversity, distribution and endemism in ultramafic (i.e., high magnesium- and iron-oxide concentration) regions in Southeast Asia. PRAGMA researchers will also **develop approaches for computer-aided drug discovery in infectious diseases,** focusing on neglected tropical diseases that affect about one billion people on this planet.

Each of these expeditions will advance the specific research goals of these groups and develop useful infrastructure. For example, the lake and biodiversity expeditions require developing tools that track usage of data and protect sensitive data by using private networks. The drug discovery expedition integrates tools to create useful workflows for virtual drug screening. In all cases, we will work with existing scientific networks (e.g., Global Lakes Ecological Observatory Networks) or grow nascent networks of researchers to ensure our developments have a broader impact.

In addition to these three scientific expeditions, PRAGMA will explore many other shorter-term expeditions as well as possible expeditions in disaster recovery of information systems and services (building on the work of AIST, NCHC, and UCSD) and via GEO Grid's network in: Japan, Taiwan, Vietnam, Thailand. Other planned expeditions include cultural heritage sharing (Osaka, NICT, and UCSD) and environmental observatory deployments coupled with other scientific expeditions in biodiversity, lake science, and disaster mitigation.

PRAGMA also intends to build on its successes in sharing virtual machines to **create a multi-provider cloud** (or multi cloud) with some characteristics of federated clouds. Our approach uses virtual, software-defined overlay networks to simplify application infrastructure by solving various connectivity issues among cooperating virtualized resources. These software-defined overlay networks also **form trust envelopes** so that PRAGMA can practically address fundamental issues in selective data sharing, provenance and raw data acquisition. Furthermore, the PRAGMA multi-cloud can be used as a test bed and resource to evaluate cybeinfrastructure components, gain insights into long-term data use, and advance science.

An additional goal for PRAGMA is to engage new participants and generate new ideas by developing student activities and forming strategic partnerships. Building on the activities of PRIME, PRIUS and MURPA, and motivated by the GLEON Student Association (www.gleon.org/Students.php) success, we have **created the PRAGMA Student organization**, to give graduate students in PRAGMA research and leadership opportunities. Also, we have begun to implement plans to develop strategic partnerships in general, with a focus on three key countries or geographic regions: India, China, and Southeast Asia. Through this activity, PRAGMA members will engage new partners in targeted workshops and interactions of mutual benefit. Efforts will also include continuing our workshops that are open to members and others interested in participating, contributing to, and learning from PRAGMA activities. These workshops will rotate among member institutions to engage researchers and students new to PRAGMA and to educate members about specific applications and technologies developed by the member institution and others in the region. These workshops are critical for demonstrating progress between the meetings and for planning activities between workshops.

In this year's PRAGMA Collaborative Overview (2012-2013) we highlight activities that exemplify these directions:

- Building a work environment to share data and to develop predictions in the biodiversity expedition;
- Developing redeployable sensor networks and placing them in lakes, coral reefs and forests
- Continuing to develop approaches and the PRAGMA infrastructure to share virtual machine images;
- Experimenting with overlay networks on the PRAGMA cloud;
- Taking steps to develop a PRAGMA data cloud;
- Delivering technology developed by PRIME students to the broader public;
- Developing a student network in PRAGMA;
- Initiating collaborations with strategic partners.

Our ability to advance this work reflects funding received over the last year, as well as funding received prior to this year. This funding is acknowledged in the section on *Members, Partners and Sponsors*, p. 32. In particular the US National Science Foundation designated its award for PRAGMA activities as a *Science Across Virtual Institute* (SAVI, www.nsf.gov/savi). PRAGMA engenders the SAVI spirit that engages the best minds in the world, accelerates scientific advances by working together across international boundaries, and connects researchers with shared interests and goals to solve important societal challenges. In particular the NSF SAVI funding, together with other PRAGMA partner funding, will launch or extend three expeditions on biodiversity, lake ecosystems, and infectious diseases; develop the multi-cloud environment with provenance and user defined trust envelopes through software-defined overlay networks; create an international student association; focus on common questions that affect India, China and Southeast Asia; and broaden engagement of US researchers to include researchers at University of Florida, Indiana University, University of Wisconsin and the University of California, San Diego.

Over the next several years, our success will be measured by the extent we: help **advance the scientific goals** of our extant and future expeditions; **provide tools and infrastructure** to these groups and the broader community; **harness technology trends** in mobile sensing, smart software, advanced data capture and storage, ubiquitous computing and expanded networks; **create a new generation of scientists and engineers** able to participate in and lead networked groups; **expand participation** on PRAGMA and more generally in networked science; and ultimately **enable small-to-medium size international groups to make rapid progress in conducting research and education by providing and developing international, experimental cyberinfrastructure.**

Note: In the following sections we use abbreviations for members' and partners institutions. The abbreviations can be found in the section on Members, Partners and Sponsors.

Images (above): The Eternal Spring Shrine, Taroko Gorge, Hualian—courtesy of Sara Taghizadeh; (below left to right): Colorful Tourist Rides, Melaka, Malaysia—courtesy of Jenny Liang; PRAGMA 22 participants in Melbourne, Australia—courtesy Philip Chan

HIGHLIGHTS

Biodiversity in (Southeast) Asia: A Scientific Expedition to Understand Adaptation in Extreme Serpentine (Ultramafic) Environments

THE PACIFIC RIM ENCOMPASSES GLOBALLY SIGNIFICANT AREAS OF BIODIVERSITY. Southeast Asia in particular—where Wallace's Line marks a transition zone between Asia and Australia—offers unique opportunities to study island effects and geological and tectonic complexity as drivers of adaptation and evolution. Southeast Asia is also a region of high human and economic growth; biodiversity loss and climate change impacts increase the urgency of using efficient and effective infrastructure methods in research.

A broad scientific challenge in biodiversity research is to understand how plants, animals, and microbes adapt to extreme environments, changing climate, and toxic conditions. Within this context, we investigate the specific biotic, abiotic, and evolutionary factors that affect patterns of diversity, distribution and endemism (native to a specific geographic region) in ultramafic regions. Ultramafic (high magnesium- and iron-oxide concentration) outcrops are not uncommon around the Pacific Rim, where tectonic activity is high, and fragments from the upper mantle of the earth or oceanic crust have been uplifted and exposed. These are deficient in essential macronutrients, but rich in metals and rare earth elements.

We will initially focus on data from and about Mount Kinabalu (4,095m), in Sabah, Malaysia. Northern Borneo, including Kinabalu, is ranked among the six highest global diversity centers. Kinabalu's high floral diversity and endemism is likely attributable to multiple factors, including elevation and corresponding climatic extremes, isolation effects of precipitous topography, and diverse geology, including some of the most geochemically extreme (pers. comm. van der Ent) ultramafic substrates known.

Our approach will involve species distribution modeling and phylogenetic, and phylogeographic analysis requiring species occurrence data from biological collections. Also included will be geospatial data, and molecular (sequence) data, available from large-scale resources (e.g., DataOne, GBIF) and from individual projects.

As a first step, we carried out an initial experiment using the PRAGMA cloud to conduct a simple analysis of species distribution modeling. The data used for this experiment are specimen collection records compiled from worldwide herbarium resources and maintained in a database at the University of Florida. These data illustrate plant diversity of ultramafic soils of Mount Kinabalu. Metadata files describing nine species distribution data sets are uploaded to a ESRI GeoPortal server running at Universiti Teknologi Malaysia (UTM).

We developed a workflow (depicted in Figure 1) that has a script running at the Indiana University (IU) PRAGMA node and queries the ESRI GeoPortal server running at UTM to retrieve metadata describing raster data sets of Mount Kinabalu specimen collection. By analyzing the resulting metadata records, the script downloads the data sets stored in a Dropbox location and posts the data sets as occurrence sets to the Lifemapper instance running at University of Kansas. Using the returned occurrence set identifiers, the script submits post-experiment requests to Lifemapper and retrieves metadata results returned for each Lifemapper experiment. The script then catalogs metadata describing each experiment, along with the URL for Lifemapper prediction outputs, in a ESRI GeoPortal server instance running at IU.

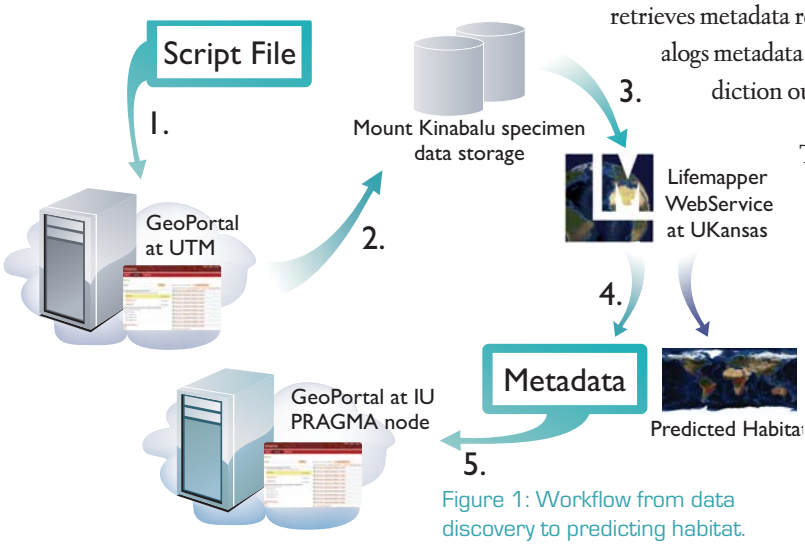


Figure 1: Workflow from data discovery to predicting habitat.

This is the first step in the longer-term biodiversity expedition. This first step came from discussions at PRAGMA 22. The success of this expedition requires the ability to share data securely, create new knowledge and scientific understanding, and develop tools and approaches useful in other research efforts in biodiversity.

PARTICIPANTS: Indiana U: *Umashanthi Pavalanathan, Yuan Luo and Beth Plale;* U Kansas Biodiversity Institute: *Aimee Stewart, C. J. Grady;* U Florida: *Reed Beaman, Andreina Weichselbaumer;* Universiti Teknologi Malaysia: *Shahir Shamsir, Suzanna Azmy, Chew Teong Han*

Image: Gradual Entrainment Lake Inverter (GELI) on Crystal Lake—courtesy of Carol Warden and Joe Sanfilippo.

Redeployable Cyberinfrastructure for Environmental Observing Systems: Setting the Stage for Future Expeditions

MANY IMPORTANT PROBLEMS IN ENVIRONMENTAL SCIENCE, NATURAL RESOURCE MANAGEMENT, AND DISASTER RESPONSE REQUIRE SIMILAR TYPES OF CYBERINFRASTRUCTURE FOR DEPLOYING AND MANAGING SENSOR NETWORKS. Researchers must be able to quickly and efficiently deploy sensor networks in remote and sometimes hostile environments to provide real-time data that will inform scientific analyses and policy decisions. These systems must support a variety of sensors, and need to provide sustainable power and communications services in areas where line power and wired networks are rare.

Although applications may have different science objectives (thus requiring different sensors and monitoring regimes), the underlying architectures and cyberinfrastructure is common. Our approach to address these application needs was to develop the DataTurbine Android Sensor Pod, a relatively inexpensive, easy-to-deploy, and robust end-to-end system (i.e., sensor to user) for environmental monitoring.

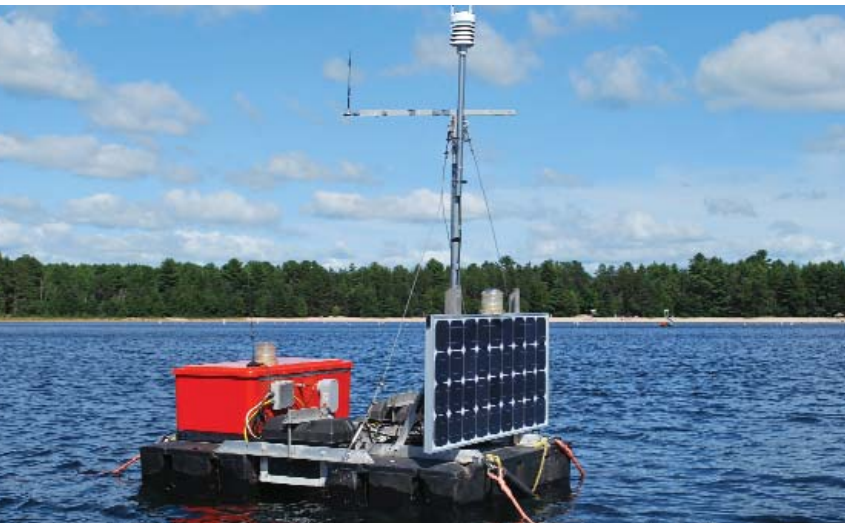
Our approach is two-fold: 1. Utilize the latest mature and stable technologies for mobile/embedded computing, communications/networking (including WiFi, satellite modems, and cell phones), streaming data middleware (Open Source DataTurbine), and cloud computing (the Amazon EC2 and the Open Science Data Cloud); 2. Integrate these technologies with custom software developed by our team into deployable hardware and software (Tilak et al 2011). By design, the system minimizes the amount of ground/field support needed for a successful deployment by streaming the data in near real-time directly from sensors to a cloud computing platform, which is scalable to adapt to changing loads (either from the source/sensor side or the demand/users side). The data can then be accessed and shared simultaneously by multiple users. These features support scientific expeditions by enabling sensor networks to be deployed to remote and sometimes hostile environments quickly and with a minimum of on-site infrastructure.

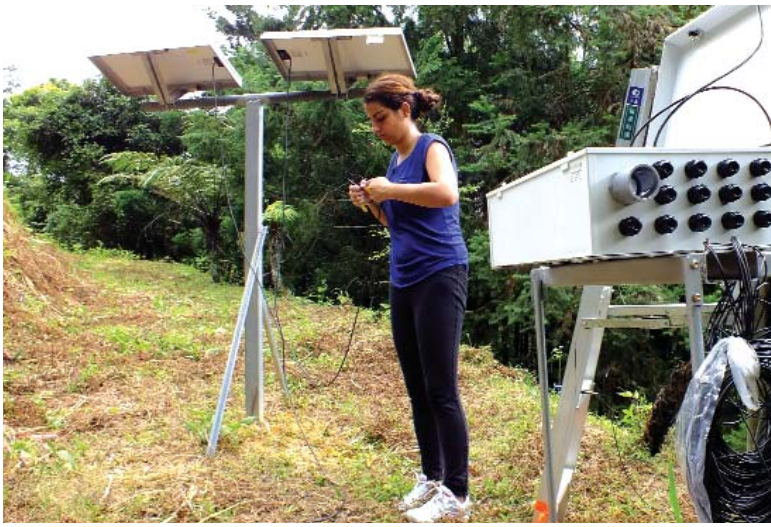
Our development approach involves working with scientists to design, prototype, and test, deploy and test, and improve the design for future deployments. We provide three examples where this system has been deployed and improved to support scientific research. The specific deployments are at sites that support a larger network of scientists and thus amplify the impact of the development (e.g., GLEON in the case of lakes). This system will assist future PRAGMA expeditions.

Lake Deployment at Crystal Lake in Wisconsin

Rainbow smelt, a cool-water fish species that was once introduced and became established in non-native lakes, led to the demise of native fish species such as yellow perch, cisco (or lake herring), or walleye due to competition for food resources or direct predation. In the case of walleye, the introduction of rainbow smelt led to degradation of recreational fishing, an ecosystem service. A new approach to eliminating rainbow smelt—whole lake thermal manipulation—takes advantage of the differences in thermal tolerances between smelt on the one hand, and yellow perch and walleye on the other. By experimentally mixing a lake via multiple large diaphragms stationed over the deepest spot in the lake during mid-summer, researchers expect that the deeper waters of a lake will warm sufficiently to eliminate the cool water habitat required only by the smelt.

Crystal Lake is a primary study lake of the Long-Term Ecological Research (LTER) network's North Temperate Lake (NTL) site, which is also a key GLEON site.





Crystal Lake Deployment Sensors: WEATHER STATION: Air Temp, Barometric Pressure, Relative Humidity, Rainfall (Amount, Duration, Intensity), Wind Speed and Direction); HYDROLAB MULTIPROBE: Water Temp, Conductivity, pH, Fluorescence; THERMISTOR CHAIN: 27 Temperature Sensors at Various Depths; TFRI DEPLOYMENT: Weather Station, Video Camera, Microphone

In August 2012, this expedition deployed an additional buoy over a secondary deep spot in the study lake, Crystal Lake, away from the diaphragms and the existing observational buoy. This was done to determine the horizontal extent of mixing in the lake as well as to collect associated measures of lake attributes, including dissolved oxygen, chlorophyll, and temperature at multiple depths. To date, we have observed that the diaphragms successfully warmed the bottom waters throughout the entire lake, stressing the rainbow smelt. During the winter, the lessons learned from this deployment will be integrated into refined design of the hardware and software, and the experimentation will continue in summer 2013.

Forest Deployment at Lienhauchih Research Center, Taiwan

The impact of climate change on forest dynamics represents an important research topic for scientists at the Taiwan Forest Research Institute (TFRI). Understanding the processes that affect both plant and animal communities is critical for advancing scientific understanding and crafting rational management policies for these natural resources. Ecological science research at the Lienhuachih (LHC) Research Center of TFRI relies on numerous heterogeneous, integrated suites of sensors for observing physical and biological processes. During the summer of 2012, two UCSD PRIME students (Ariana Tsai, Sara Taghizadeh) worked in Taiwan with TFRI researchers to customize and deploy the sensor system for terrestrial ecology applications. That system included a weather station (air temperature, barometric pressure, relative humidity, rainfall, rainfall duration, rain intensity, wind speed, wind direction), video camera, and microphone. The project involved working with the Android operating system for the cell phone that was used in the communication system. In addition, the students created a tutorial for assembling a hardware component of the deployed system.

Coral Reef Deployment in Moorea, French Polynesia

While coral reefs have undergone unprecedented changes in community structure in the past 50 years, they now may be exposed to their gravest threat since the Triassic era. This threat is increasing atmospheric CO2, which equilibrates with seawater and causes ocean acidification (OA). Key to understanding this process is measurements by OA-related sensors including pH, PCO2, temperature, and pressure. The deployment of OA-related sensors at Moorea and Palmyra Atoll and the collection of measurements from these sensors will provide valuable insights into OA

The Taiwan Forestry Research Institute (TFRI) is more than one hundred years old, manages several research centers such as the one at Liehuachih, which specializes in medicinal plants, and has been instrumental in the coordination of the Taiwan Ecological Research Network (TERN).

Images (left to right): redeployable buoy on Crystal Lake—courtesy of Grace Hong; TFRI student—courtesy of Yu-Huang Wang; Deploying the buoy at MCR—courtesy of Tony Fountain

The MCR deployment involved the following sensors: SeaFET (temperature, pH); SeapHOx (temperature, pH, controller temperature, pressure, O2 saturation); CTD (conductivity, temperature, salinity); Pro-Oceanus CO2-Pro (Raw absorbance, CO2, gas pressure, humidity)

and lay the foundation for larger, more rigorous OA studies in the future. The ability to measure these phenomena in-situ and in real-time represents a significant advance in our ability to monitor key environmental processes.

In January 2012, members from UCSD and University of California Santa Barbara (UCSB)/Moorea Coral Reef (MCR) travelled to Moorea to test the DataTurbine Android Sensor Pod system on the coral reef off of Gump Station. This deployment provided important experiences for the research team in deploying and managing the autonomous buoy system for ocean acidification research. The lessons learned from this deployment led to changes in the system design that improved system robustness and reliability. These improvements were manifest in subsequent deployments, described above. The team plans to return to MCR in 2013 to perform longer and more in-depth monitoring experiments.

Related Deployments at Coral Reefs

With an earlier version of this technology, a PRAGMA Expedition helped field a conductivity, temperature and depth (CTD) device, using the DataTurbine server to stream data, on Racha Island (Thailand) with the assistance of NECTEC, UCSD, and Mullica Jaroensutasinee, Krisanadej Jaroensutasinee (Walailak University, Thailand). Scott Bainbridge (AIMS, Australia). Knowledge gained from the MCR deployment will be readily shared in CREON.

PUBLICATION: Tilak, S., P. Shin, P. Arzberger, S. Holbrook, R. J. Schmitt, A. Brooks, and K. Seydel, The Open Source DataTurbine (OSDT) Android Sensor Pod: Embedded cyberinfrastructure for smart buoy controllers and experiments in ocean acidification and limnology, Santa Barbara, EIM 2011, 2011. Google Scholar

Jaroensutasinee, J., M. Jaroensutasinee, S. Bainbridge, T. Fountain, S. J. Holbrook and M. Nekrasov. 2012. CREON - Integrating disparate sources of remote coral reef sensor data. *Proceedings of the 12th International Coral Reef Symposium*, Cairns, Australia, 9-13 July 2012, 5B Sensor Networks and their Applications.

PARTICIPANTS: Lake Deployment, U Wisconsin: *Corinna Gries, Tim Kratz, Jordon Read, Tim Meinke, Ken Morrison;* Forest Deployment, TFRI: *Chau-Chin Lin, Yu-Huang Wang, Sheng-Shan Lu,* UCSD, PRIME 2012: *Sara Taghizadeh, Ariana Tsai;* Coral Reef Deployment: *Sally Holbrook, Russell Schmitt, Andy Brooks, Robert Carpenter;* All Deployments: *Peter Shin, Sameer Tilak, Gesuri Ramirez, Tony Fountain;* Cycronix: *Matt Miller*

Moorea Coral Reef (MRC) is part of the US LTER network, and a founding member of the Coral Reef Environmental Observatory Network (CREON, creon.sdsc.edu:8080/confluence/display/creon/Welcome).

VM Image Sharing Experiment: From Grid to Cloud – Phase 3

BEGINNING IN 2010 PRAGMA SHIFTED ITS FOCUS FROM GRID TO CLOUD COMPUTING. The PRAGMA cloud infrastructure is heterogeneous, consisting of different architectures based on local needs and decisions. We continue to pursue the goal of making new technology easy-to-use for researchers, enabling them to advance science and global-scale collaborative research. Pursuing this goal requires us to seek out and develop methods to ensure interoperability of many different architectures and clouds. Our approach allows users to author their own application virtual machines (VMs) with their preferred VM hosting platforms (e.g. Xen, KVM), and then run their VM at any PRAGMA cloud site. We are developing the infrastructure to automatically re-configure these custom VMs to be compatible with the remote site.

Our progress in porting (a form of sharing) VM images has taken place in three phases. In the first phase, (completed in early 2011), we successfully demonstrated a manual port of three application VM images among three different VM hosting environments (i.e., a pairing of a VM hosting platform with VM hosting managing software, e.g. Open Nebula). Phase 2 began at PRAGMA 20 (March 2011, Hong Kong), where, based on what we learned from the phase 1 experiments, we designed a PRAGMA cloud infrastructure to use Gfarm for VM image depository and sharing and to automate VM deployment process on various virtualization platforms that make up the PRAGMA multi-cloud. At PRAGMA 21 (October 2011, Japan), we demonstrated the automated deployment of a GEO Grid VM image from Gfarm to three pilot sites—AIST, NCHC and UCSD.

After PRAGMA 21, we started Phase 3 of the VM sharing experiment with four objectives: 1) to expand PRAGMA cloud resources; 2) to enhance Gfarm functionality and performance; 3) to author and run more application VMs and 4) to develop an easy-to-use user interface for creating VM images.

Inst.	Region	Nodes	CPUs	Mem.	Disk Space	Platform	Manager
AIST	Japan	32	64	768GB	600TB	KVM	OpenNebula
ASTI	Phillipines	10	80	160GB	5TB	Xen	OpenNebula
CNIC	China	1	2	128GB	2TB	Xen	Rocks 5.4.3
HKU	Hong Kong	3	3	16GB	1TB	Xen	Rocks 5.4.3
IU	USA	8	16	128GB	12TB	Xen	Rocks 5.4
JLU	China	1	4	34GB	1TB	Xen	Rocks 5.4.3
KU	Thailand	1	4	16GB	0.25TB	KVM	Rocks 6.0
KU	Thailand	3	8	6GB	0.6TB	KVM	OpenStack
LZU	China	1	8	16GB	1TB	KVM	Rocks 6.0
MIMOS	Malaysia	6	12	192GB	1TB	KVM	Eucalyptus
NCHC	Taiwan	8	64	128GB	2TB	KVM	OpenNebula
OSAKA U	Japan	3	24	34GB	3TB	Xen	Rocks 5.4
SDSC	USA	18	25	380GB	10TB	Xen	Rocks 5.4
SDSC	USA	32	64	512GB	16TB	Xen	Rocks 5.4.3
SDSC	USA	4	16	32GB	6TB	KVM	Rocks 6.0
UoHyd	India	4	4	16GB	4TB	Xen	Eucalyptus
13	9	135	398	2566GB	663.31TB		

Figure 3. PRAGMA cloud resources, with virtualization hosting platforms (platforms) and VM hosting management software (manager). The combination of the platform and manager make the "environments."

Expand PRAGMA Cloud Resources

The successful demonstration of the VM image sharing experiments at PRAGMA 20 and PRAGMA 21 excited researchers at many PRAGMA sites and motivated them to join this effort. In this last year, the PRAGMA cloud grew from three to 13 sites and formed a rich variety of virtualization hosting platforms and management software (i.e., VM environments), creating a diverse set of resources for our interoperation experiments. These 13 sites are displayed in Figure 2, with the resources and environments summarized in Figure 3.

Deploy and Enhance Gfarm File System

Working closely with the University of Tsukuba (U Tsukuba), the Gfarm development team at UCSD built a Gfarm metaserver, tested the Gfarm file server and Gfarm client setup, and developed a Gfarm roll for Rocks systems that packages the Gfarm software to install easily and be operated in the Rocks environment (goc.pragma-grid.net/wiki/index.php/Gfarm). This facil-

itated rapid deployment of Gfarm among PRAGMA sites. By Supercomputing 2011 (SC11, November 2011), a total of seven sites (UCSD, AIST, NCHC, IU, LZU, Osaka U, and CNIC) had installed and set up Gfarm software; of those, four sites (UCSD, LZU, Osaka U, and IU) used the Gfarm roll.

With Gfarm, we are able to securely store and share VM images among PRAGMA sites located around the Pacific Rim. Since the VM image files are typically GB sized, file transport performance is important. The Gfarm replication scheme allows Gfarm sites to store files locally but still keep them synchronized among Gfarm sites, thereby speeding up the usage performance when accessing a file from a Gfarm site.

Figure 2: PRAGMA cloud sites map.

The experience of supporting the PRAGMA cloud has fueled the continued development of Gfarm. In the last two years, Gfarm software has been improved and enhanced through frequent new software releases, beginning with version 2.4.2 and continuing to the most recent version of 2.5.7. In addition, several bugs have been fixed and new functionalities have been added to support data sharing for both cloud and grid. Furthermore, Gfarm is being used in the HPCI storage, a Japan nation-wide file system, initially accessed from nine national universities and the "K" computer site.

Authoring Real Science Application VMs

By PRAGMA 21, we had authored four real science application VMs, which were successfully run on various PRAGMA cloud sites and Amazon's EC2 sites. In 2012, each of UCSD and LZU authored bioscience application VMs.

UCSD's bioscience application VM was authored on a Rocks/Xen system. This bioscience application provides an environment designed to support molecular dynamics simulations and virtual screening experiments for in silico drug discovery, with a special focus on supporting the use of the Relaxed Complex Scheme. It includes Web-based access (using Opal Web services) to popular molecular docking and analysis applications such as AutoDock, AutoGrid, AutoDock Vina, Gromacs, MGLTools and couples them in a flexible and scalable fashion.

LZU authored two bioscience application VMs on a Rocks/KVM system. One is a spatially explicit species competition simulation to investigate, in a homogeneous habitat, the effects of dispersal and recruitment limitation on delaying competitive exclusion. This application examines the

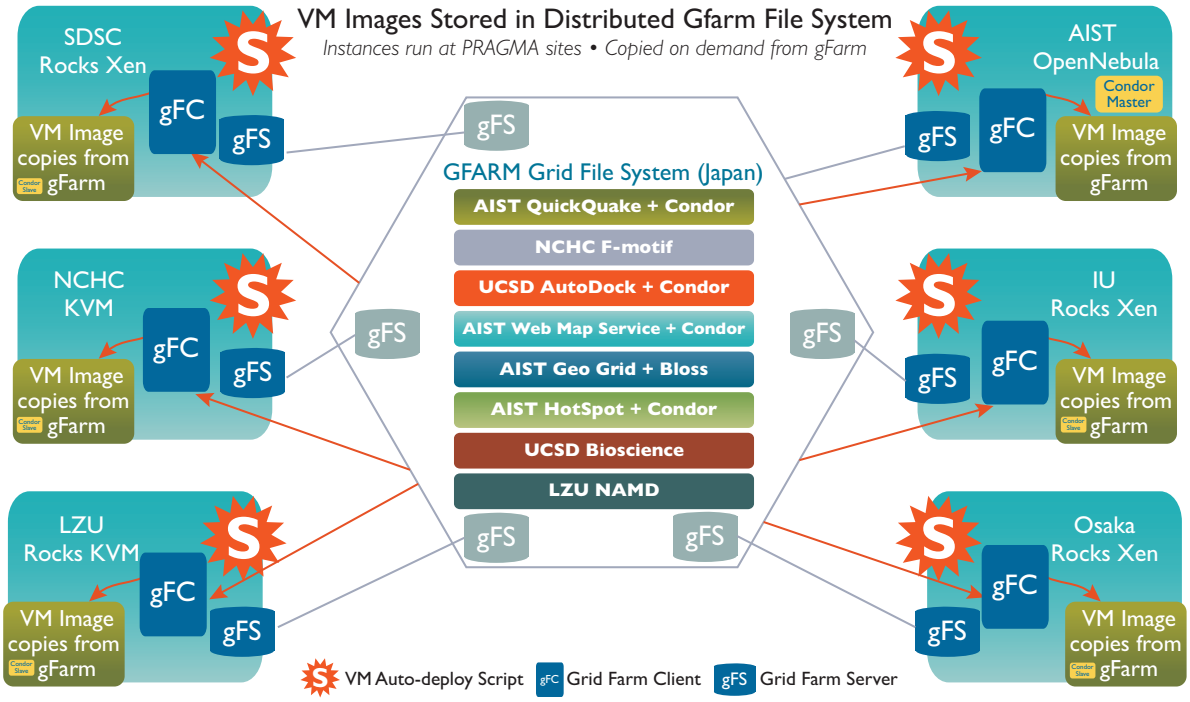


Figure 4: Automated VM deployment in PRAGMA cloud.

species diversity (measured by Simpson's diversity index), species abundance distribution and species area relationships under the conditions of slight to moderate competitive asymmetries among species, and with dispersal and recruitment limitations. The second VM application simulates the movement and function of ubiquitin, a small regulatory protein, using the molecular dynamics package NAMD. Ubiquitin is a protein that plays a key role on eukaryotic intracellular protein degradation, whose importance has been widely recognized in the scientific community.

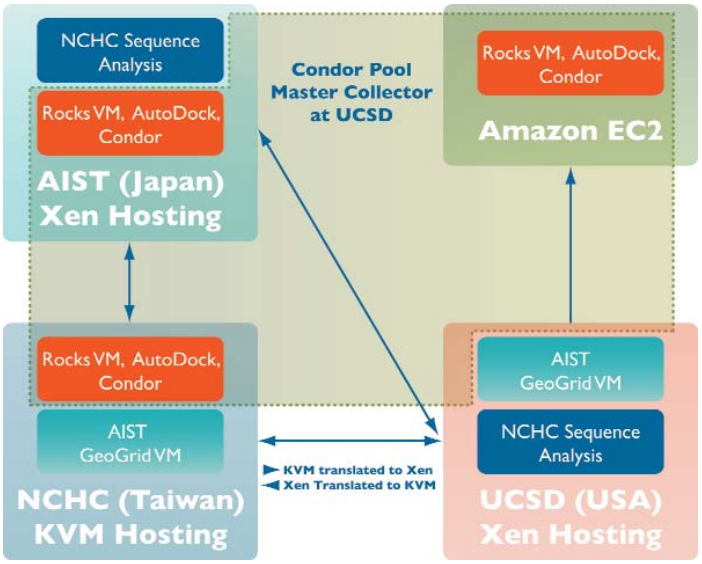


Figure 5: Running Application VMs in PRAGMA cloud and Amazon EC2.

Automating the VM Porting Process

Starting after the PRAGMA 21 Workshop, we deposited six application VM images in Gfarm, and deployed four of them on-demand among many PRAGMA cloud sites. In the four weeks between PRAGMA 21 and SC11, seven sites were able to setup their VM hosting systems with Gfarm and deploy a GEO Grid application VM. This was made possible through well-documented VM hosting server setup procedures, easy setup/install Rocks systems with Gfarm rolls, and automated deployment tools such as vm-deploy script and other mechanisms including cluster toolkit and cloud middleware (e.g., Rocks, OpenNebula, etc.). At SC11 we successfully demonstrated the GEO Grid applications.

In 2012, we started to deploy the Bioscience application VM created by UCSD. This application VM is more complex in its requirements. Using the same setup and with some enhancements to our automated deployment scripts, we successfully demonstrated the running of the bioscience applications in the PRAGMA cloud at the PRAGMA 22 (Melbourne) Workshop. To date, we have successfully deployed the UCSD bioscience application VM to seven sites (UCSD, AIST, CNIC, JLU, IU, OsakaU, NCHC). (See Figure 4 for an example.) In addition, we deployed the UCSD bioscience VMs as a virtual cluster on Amazon's EC2 resources and demonstrated the bioscience applications running on the virtual cluster at EC2 and on VMs running on PRAGMA cloud sites (see Figure 5).

Before PRAGMA 23, we also deployed LZU's bioscience application VM to six PRAGMA cloud sites (LZU, UCSD, Osaka U, AIST, JLU, and CNIC). At PRAGMA 23, we successfully demonstrated the running of LZU's bioscience applications in the PRAGMA cloud.

Develop VM Authoring Tools and User Interface

Recently, NCHC received an award from the National Science Council Taiwan to continue to develop Ezilla, an important component in our approach to port virtual machine images between PRAGMA sites, thus to expand and ensure persistence of cloud services. The project focuses on the software development of cloud Interoperability (using an approach similar to phase 2 of the sharing experiment). Working with domestic and international partners, NCHC developed VM authoring tools and a web-based, easy-to-use user interface. NCHC worked on automating the process of packaging applications into VM images, so that domain scientists can easily create application VM images without having to learn about VM technology. QEMU Copy On Write (qcow) is used to implement the differential storage backup for the application VM image and to reduce the storage space needed for the VM storage. In addition, this approach increases significantly the boot up speed of the virtual machine.

UCSD and AIST have helped test the Ezilla software. At PRAGMA 22, NCHC demonstrated the software. Future tests will involve more PRAGMA members.

The new version of Ezilla has further enhanced the user interface features (such as Drag&Drop) with improved administrative tools aimed at creating a production cloud service. The new version was released at PRAGMA 23, and an Ezilla based production service has been announced for public testing.

To address the barrier to adapting high-end computing applications in the cloud, NCHC has proposed a Scientific Cloud Marketplace for VM images, which would be present in Ezilla as a software interface between scientific applications and the cloud infrastructure. Finally, to expand

the type of applications using the cloud infrastructure, Ezilla recently implemented the Volatility Information Platform used to provide on-line and instant estimates of dynamic risk asset volatilities. As a first step to expand interest in the business community, both S&P 500 index (SPX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) are provided as examples (see Figure 6).

The Future: On To Virtual Cluster Sharing

Prior to Fall 2012, our efforts have been focused on successfully porting single-node VMs. Leading up to PRAGMA 23, we shifted our focus to virtual cluster porting. Sharing virtual clusters among different cloud systems is a complex problem, requiring retaining both the relationship between head and worker nodes and many cluster functions. Following the successful path of our prior collaborative experiments, we, again, started Phase 1 of this effort with the three pilot sites (UCSD, AIST, NCHC). Through many discussions we concluded that we need to formulate the information required for virtual cluster porting and to design a common user interface (a command line). Then each site can develop an automated implementation of virtual cluster porting. The first phase of this task is to manually port virtual clusters. This step allows us to learn the differences among various virtualization environments and to discover issues and requirements for virtual cluster porting. AIST and NCHC sites created two simple virtual clusters. The UCSD team manually ported them to a Rocks/KVM system, documented and shared the required info and changes with AIST and NCHC. Meanwhile, UCSD also successfully ported one of its own virtual cluster with bioscience applications to EC2, providing much needed initial experience to advance our effort on virtual cluster sharing. At PRAGMA 23, the pilot teams demonstrated these advances, shared their findings and engaged in further detailed discussions with Resources Working Group and PRAGMA community. As a result, the Resource Working Group has set its course to develop a standard virtual cluster porting template and to implement automated virtual cluster porting solutions for each virtualization environment.

With these technologies now developed and tested, we will work to encourage users to author virtual machines (VMs) and virtual clusters (VCs) and use the growing PRAGMA cloud infrastructure. Furthermore, we will continue to develop solutions for the general issue of cloud interoperability and will experiment with approaches to expanding the handling of larger data sets. The approach described above for VM porting has implications for recovery of critical services during major disruptive events.

PARTICIPANTS: AIST: *Yoshio Tanaka, Naoaka Yamamoto, Ryosei Takano, Akihiko Ota, Akihiro Iijima;* NCHC: *Weicheng Huang, Serena Pan;* SDSC/UCSD: *Phil Papadopoulos, Cindy Zheng, Nadya Williams;* LZU: *Wenbo Chen, Zhang Yang;* IU: *Beth Plale, Yuan Luo, Felix Terkhorn;* Osaka U: *Susumu Date, Kohei Ichikawa, Taiki Tada;* JLU: *Xiaohui Wei, Simon Lee, Qinqin Zhou, Xiaoru Shen;* ASTI: *Mary Grace C. Dy Jongco, Emeterio D. Casera, Jr.;* MIMOS: *Jing Yuan Luke;* CNIC: *Kai Nan, Kevin Dong;* UoHyd: *Arun Agarwal, Babu Rao Singathi, Rajeev Wankar;* HKU: *Kwan Wing Keung, Lilian Chan;* KU: *Putchong Uthayopas, Chawanat Nakasan;* U Tsukuba: *Osamu Tatebe*

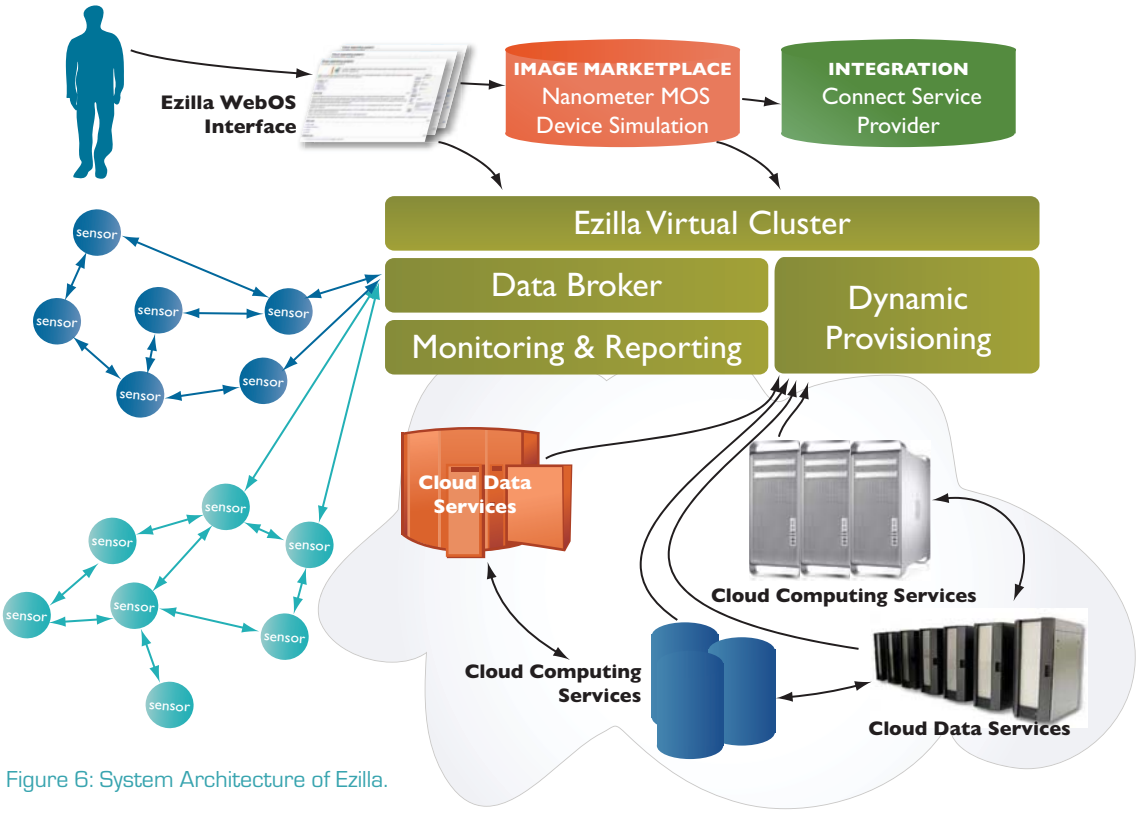


Figure 6: System Architecture of Ezilla.

Software-defined Overlay Networks

VM SHARING SCHEMES described in the previous section always involve changing the VM's network settings to enable the VM to adapt to a different network location and environment. An alternative way of porting a VM without changing the VM network settings is to set up a virtual private network (VPN) among cloud sites, so VM instances deployed within the VPN work as if they were deployed on the same private network. The VM is constructed to contact a master node on its private network and get an IP address via the Dynamic Host Configuration Protocol.

There are several software products aimed at achieving this goal. As always, PRAGMA experiments with several different technologies, based on the members' interest. These tests, often involving several PRAGMA members, are then demonstrated at PRAGMA workshops.

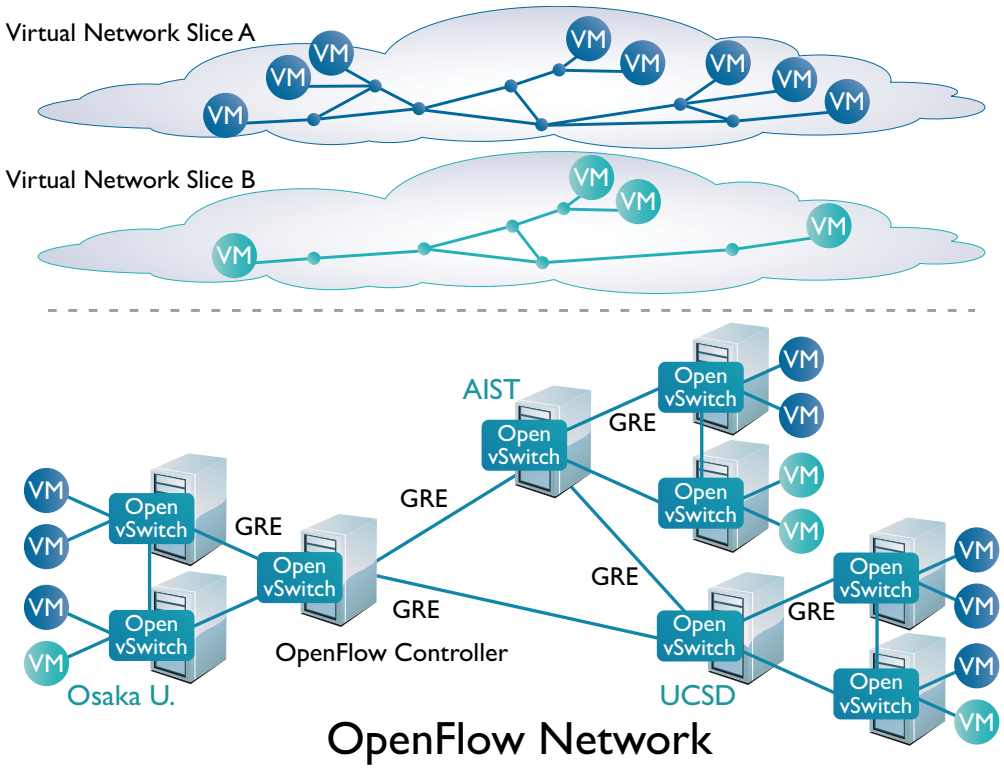


Figure 7: OpenFlow Network Environment.

The first network overlay software we tested is Open vSwitch, a software switch in OpenFlow, which is an emerging virtual network technology for creating software programmable networks. Lead by the Osaka University team, we formed a pilot team with the UCSD and AIST teams (see Figure 7). After the Osaka team installed and set up its site as the Open vSwitch control site and documented and shared their installation experiences with AIST and UCSD, the AIST and UCSD teams quickly installed and

Images (left to right): Small town near Lake Biwa of Shiga Prefecture—courtesy of Joshua Wei; Torii gate of Itsukushima Shrine on Miyajima—courtesy of Olivia Yang; PRAGMA Student workshop participants, at PRAGMA 23

set up their own local Open vSwitch. AIST also constructed a GEO Grid VM to run within this VPN. In three weeks, we tested the VPN among the three pilot sites, deployed the GEO Grid VM, and were able to run the GEO Grid application successfully. This experience convinced us that this scheme is very easy and effective. At PRAGMA 22 (April 2012, Melbourne), we successfully demonstrated this OpenFlow-based virtual network environment for PRAGMA cloud virtual clusters. The work involved deploying multi-site virtual clusters over a virtual network using Open vSwitch.

Following PRAGMA 22, we have expanded this OpenFlow VPN to more PRAGMA cloud sites, and are encouraging other application groups to use the OpenFlow VPN in the PRAGMA cloud. For example, the Biodiversity expedition is planning to use this for their applications, in part because it helps ensure limited access to sensitive data sets.

In addition to expanding the OpenFlow VPN, we are also preparing to experiment with another network overlay technology—ViNe. ViNe is developed by a team at the University of Florida. It enables symmetric connectivity among Grid resources and allows existing applications to run unmodified. Novel features of the ViNe architecture include easy virtual networking administration; support for physical private networks and support for multiple independent virtual networks in the same infrastructure. At PRAGMA 23, the University of Florida team demonstrated the software to inform participants and solicit interested partners in PRAGMA to participate in a new experiment in PRAGMA cloud. A pilot team has formed and will implement ViNe among a few sites, using same GEO Grid application (as used in the OpenFlow VPN) to test the implementation.

PARTICIPANTS: Osaka University: *Kohei Ichikawa, Taiki Tada, Sumumu Date, Shinji Shimojo;* AIST: *Yoshio Tanaka, Akihiko Ota, Tomohiro Kudoh;* UCSD: *Philip Papadopoulos, Cindy Zheng;* U Florida: *Jose Fortes, Mauricio Tsugawa*

Toward Building a PRAGMA Data Cloud

EARTH SCIENCE APPLICATIONS SUCH AS ENVIRONMENTAL MONITORING AND DISASTER MITIGATION ARE A SHARED INTEREST AMONG PRAGMA MEMBERS. Since the Earth's ecosystem is a spatially and temporally complex global system, it is not sufficient to observe events and phenomena only at the local scale. As an information technology community in Asia Pacific Rim, PRAGMA can support Earth science communities—such as flux monitoring, lake ecosystems (e.g., GLEON), and biodiversity communities—by providing an IT infrastructure that enables federated access to and operation on distributed and heterogeneous Earth observation data, such as satellite and in-situ observation data.

The PRAGMA data cloud addresses access, operation, and discovery of federated data resources. It is specialized initially for biodiversity and Earth science data. The basic design of the PRAGMA Data cloud leverages to the extent possible the Open Geospatial Consortium (OGC) standards; for example, Catalogue Service for Web (CSW) for catalogue search and Web Mapping Service (WMS) for providing map information. In PRAGMA, GEO Grid (AIST), ESRI GeoPortal (used in the Biodiversity Expedition, see first Highlight, p. 6, Indiana University (IU), University of Kansas, University of Florida, and Universiti Teknologi Malaysia (UTM)), and Environment Informatorium (E-Rium, developed by NECTEC, 203.185.96.70/ERiumPortal/index.html) are being tested for interoperability to provide an easier to use infrastructure to gain a more insightful and accurate understanding about our Earth. Since all these platforms support the OGC standards, we expect bridging of these platforms to form a foundational data infrastructure for PRAGMA.

As the first step in this bridging, we conducted an interoperability test of CSW between GEO Grid and ESRI GeoPortal. GEO Grid has its own implementation of CSW, and the metadata catalogue of various Earth observation data such as ASTER and MODIS satellite data is provided by the CSW server running at AIST. ESRI GeoPortal running at IU and UTM, is an open-source metadata server that enables discovery and use of geospatial resources. GeoPortal implements CSW, and IU experimented with metadata catalogue on PRAGMA by federating GeoPortal instances and using the IU Voretex2 collection. In the interoperation test, we confirmed that AIST's CSW client could search a metadata catalogue via GeoPortal's CSW server, and GeoPortal's CSW client could search metadata catalogue via GEO Grid's CSW server. Currently, E-Rium, a sustainable web portal for observation data, does not support CSW; however it is already confirmed that E-Rium could retrieve in-situ observation data provided by GEO Grid via Sensor Observation service (SOS).

Successfully building the PRAGMA data cloud will enable scientific expeditions in biodiversity, lake ecosystems, and disaster mitigation.

PARTICIPANTS: AIST: *Yoshio Tanaka, Isao Kojima, Hirokazu Yamamoto;* Indiana U: *Beth Plale, Umashanthi Pavalanathan, Yuan Luo;* NECTEC: *Sornthep Vannarat, Naiyana Sahavechaphan, Apivadee Piyatumrong*

Creating a Crowd-sourced Museum Exhibition

AS TECHNOLOGY'S IMPACT ON SOCIETY INCREASES, MUSEUMS ARE MOVING AWAY FROM PROVIDING A PASSIVE VISITOR EXPERIENCE TO CREATING AN EXPERIENCE THAT ACTIVELY ENGAGES MUSEUM VISITORS. New technologies are providing platforms to create these new participatory experiences. In 2011, a unique collaboration began between the Pacific Rim Undergraduate Experiences (PRIME) program at UCSD, the Museum of Photographic Arts (MOPA) in San Diego, and the National Institute of Information and Communications Technology (NICT) in Tokyo and Kyoto.

The first interface created, My Gallery Interactive, premiered at the 2011 Knowledge Capital Trial in Umeda, Osaka. It allowed more than 12,000 visitors to view images from MOPA's permanent collection and invited visitors to become museum curators by creating their own favorite gallery of photographic works.

The project generated high social impact and visibility. My Gallery Interactive received the American Association for Museums 2012 MUSE Gold Award, in the Honeysett and Din Student category, the highest award in its category. The project was also featured on The Center for the Future of Museums (CFM)

Using Soapbox can be child's play.

blog, a nationally recognized forum that assists museums in shaping a better tomorrow by exploring cultural, political, and economic challenges.

Based on audience feedback from the Knowledge Capital Trial, PRIME 2011 students Wesley Hsu and Lance Castillo created Soapbox, which is a modified, simpler version of My Gallery Interactive. Using a commercially available Samsung touch table, and the same Open Exhibits software platform as My Gallery interactive, the Soapbox interface gives visitors the ability to rate more than 120 photographs covering a broad spectrum of MOPA's permanent collection. The touch table was on display during the summer of 2012 at MOPA and communicated with a Web interface that allowed online visitors to rate the same images. The combined data from the touch table and website will result in MOPA's first crowd-sourced photo exhibition of the 40 highest rated works, premiering in October 2012.

This unique collaboration leveraged the expertise and resources of all three partners: programming students from PRIME, hardware and technical expertise from NICT, and artistic content from MOPA. The success of this project continues with PRIME, NICT, and the Balboa Park Online Collaborative (BPOC), an organization assisting non-profits with their technology initiatives (see *Training the Next Generation of Research Leaders*, p. 20). Moreover, this interdisciplinary collaboration continues to present a viable and exciting model for collaboration between museums and academic institutions interested in merging art and technology.

PUBLICATIONS: Haga, VK. et al. "My Gallery Interactive: Engaging Museum Audiences with Technology," *Museums and the Web* 2012, San Diego, CA.

Haga VK, JH Haga, S Shimojo, "Leveraging International Collaborations to Create Interactive and Educational Experiences," The 40th Annual Museum Computer Network Conference, Seattle, WA, 2012.

PARTICIPANTS: PRIME 2011 Students: *Wesley Hsu, Lance Castillo;* MOPA: *Amber Lucero-Criswell, Joaquin Ortiz, Vivian Kung Haga;* NICT: *Masaki Chikama, Yoshinori Kobayashi, Tomoaki Takata, Shinji Shimojo;* UCSD: *Jason H. Haga*

Image: Lanterns at Kitano Tenmangu Shrine, Kyoto—courtesy of Peter Arzberger

Infusing New Ideas into PRAGMA

NEW IDEAS SUSTAIN THE VIBRANCY OF A FIELD AND OF AN ORGANIZATION. PRAGMA'S EFFORTS TO CREATE A NEW GENERATION OF RESEARCHERS AND TO DEVELOP STRATEGIC PARTNERSHIPS BRING IN NEW IDEAS AND ARE AN INTEGRAL PART OF ITS SUCCESS.

Developing the Next Generation of International Researchers

PRAGMA's student members imbue the program with both new ideas and the energy and excitement to advance projects. PRAGMA's extensive network of global partners provides a unique environment for students to develop the collaboration skills necessary for future leadership in networked science. Through the PRIME and MURPA programs for undergraduate students, PRAGMA has had extensive experience utilizing this network of partners to help create technically knowledgeable, culturally inclusive, and interdisciplinary scientists capable of leading global research. (See *Training the Next Generation of Research Leaders*, p. 20)

More recently, at the PRAGMA 22 workshop, a group of graduate students and their mentors met to discuss the value, role, activities, and goals of a student group inside of PRAGMA. Institutes represented included Indiana University, Konkuk University, Osaka University, University of Tsukuba, and Monash University, and Kasetsart University. To provide an example of the value of an active student program in a research network, the president elect of the GLEON Student Association (GSA; www.gleon.org/Students.php), Kohji Muraoka (Waikato University) attended.

GLEON, the Global Lakes Ecological Observatory Network, is a grassroots network of limnologists, ecologists, information technology experts, and engineers who uses the network of people, sensors, and data to understand issues such as eutrophication or climate change at regional to global scales. GLEON, established based on an early PRAGMA expedition to place sensors on a lake in Taiwan, evolved a very successful Graduate Student Association. In particular, Kohji emphasized the role of the GSA to provide mentoring, information, training, research and networking opportunities for students. An interim steering committee has been established with representatives from Konkuk University (Meilan Jiang), Osaka University (Taiki Tada), and Indiana University (Umashanthi Pavalanathan and Yuan Luo).

Since that meeting, PRAGMA students have been meeting bi-weekly, have collaborated in a joint GLEON GSA, PRAGMA Student, and ASLO (the Association for the Sciences of Limnology and Oceanography) student workshop at the July ASLO annual meeting (see GSA Newsletter Issue 2, September 2012, www.gleonrcn.org/media/GLEON_GSA_Newsletter_Issue_2_Sept_2012_final.pdf); and organized a poster session and student workshop for PRAGMA 23. The students will plan a session for PRAGMA 24, continue working on specific research projects, build stronger ties with GLEON GSA, explore ways to develop mentoring for students in PRAGMA, and work with the PRAGMA Steering Committee to develop student exchanges and a distributed seminar series for the summer of 2013.

Strategic Partnership Development

By hosting open meetings and rotating host sites, PRAGMA workshops serve as a venue for engaging new researchers. Targeted auxiliary workshops and training programs associated with larger PRAGMA workshops provide ways to focus on special topics and engage researchers from specific geographic regions. We illustrate how one sequence of focused workshops was driven in part by graduate students and resulted in the Biodiversity Expedition, highlighted on page 6.

Researcher Reed Beaman (U Florida) was interested in developing broader connections with PRAGMA and, in particular, researchers in Asia who are interested in biodiversity. With PRAGMA colleagues Fang-Pang Lin (NCHC), Habibah Wahab (USM), and Peter Arzberger (UCSD), he co-organized a companion workshop for researchers from Southeast Asia who were invited to SEAIP (Southeast Asia International Joint Research and Training Program) 2011 and extended invitations to other interested researchers. The Biodiversity and Cyberinfrastructure Workshop provided a forum for exchanging information on several topics, as well as a way for a subgroup to develop plans for future research in three areas of interest (all of which would benefit from improved cyberinfrastructure). Those areas include natural products chemistry, linking data from

Images: (left) TFRI students—courtesy of Yu-Huang Wang; (above) Fireworks at family gravestones in Nagasaki to call ancestral spirits back home for Obon—courtesy of Pooja Makhijani; (below) Initial discussions of the PRAGMA Student organization: Meilan Jiang (Konkuk) and Taiki Tada (Osaka)—courtesy Kohji Muraoka

heterogeneous resources, and building geospatial data capacity that could aid biodiversity and climate research (see report and list of participants at http://event.nchc.org.tw/2011/southeast_asia/upload/content_file/4fe90876eaa8a.pdf).

Following this workshop, Beaman focused on adaption in ultramafic outcroppings in the context of research he was beginning to explore with Antony Van der Ent of the University of Queensland. In advance of PRAGMA 22, he and others organized the "Biodiversity in (Southeast) Asia Workshop: Defining Cyberinfrastructure to Support the Study of Ultramafic Ecosystems." At this workshop and the subsequent PRAGMA 22 and 23 workshops a group continued to sharpen plans for the virtual "Biodiversity Expedition." This process brought in several new researchers to PRAGMA, provided PRAGMA with a specific set of goals, and engaged other technologies in PRAGMA.

Other workshops associated with PRAGMA are discussed in the *Workshop and Working Groups* section (p. 30). Throughout their many efforts to develop educational outreach and strategic partnerships, PRAGMA collaborators have discovered that persistent interaction is necessary for launching successful projects.

PRIME

TRAINING THE NEXT GENERATION OF RESEARCH LEADERS

MURPA

> Conducting research on my own project at a university in Australia had a huge impact on me. It confirmed for me how much I wanted to go into research...I hope that this gift will help more students have the same influential experiences. **Haley Hunter-Zinck, 2008 PRIME alum**

THE PACIFIC RIM EXPERIENCES FOR UNDERGRADUATES (PRIME) PROGRAM was created in 2004 to provide a project-based, hands-on research internship program, combined with a cultural awareness experience, for science and engineering undergraduates at UCSD. PRIME grew out of the PRAGMA collaborative framework and people network. PRIME's projects are based on PRAGMA collaborations, as well as additional collaborations between UCSD and PRAGMA researchers. This year marked the ninth year of the program, sending 18 students to the eight sites listed here, on the right. To view the PRIME 2012 students' progress and final reports on their research and cultural interactions, please visit: prime.ucsd.edu/student_collections2012.htm). To date, more than 170 students have participated in PRIME.

This year PRIME received an unexpected and generous gift from one its former alumni, Haley Hunter-Zinck (PRIME 2008, Monash University). She is currently a graduate student at Cornell University, working on a Ph.D. in computational biology and medicine. Her unrestricted gift to PRIME is a tribute to her mother, who was passionate about promoting the success of women in science. Hunter-Zinck hopes that her donation to PRIME will inspire others to follow her lead in supporting the international program. This gift is gratefully accepted for future students, and allows the program to continue to evolve to improve the experience.

In the rest of this section, students report on results from several of this year's PRIME projects. Many projects reflect the students' interest in having a positive impact on society in the work that they do. You can see this in the biomedical applications highlighted here, as well as the examples where state-of-the-art technologies are used to provide environmental observing (see *Redeployable* CI in *Highlights*, p. 7) or to bring cultural heritage to more members of society (see *Crowd Sourced* in *Highlights*, p. 16).

To strengthen its cultural component, PRIME has instituted a curriculum that puts students on a path of cultural awareness. We use the Intercultural Development Inventory (IDI) to measure the students' attitudes towards cultural differences. More about this index, as well as about PRIME, MURPA, and PRIUS programs, can be found in the following publications:

Haga J, Galvin J, Arzberger P, Wienhausen G. PRIME: A Model of Experiential Learning within a Global Research Community. In Education in Action: Experiential Learning in Higher Education, editors Forbes D, Lin J, and Losh E. UCSD Sixth College. pp 297-306. June 2012.

In Education in Action there are five articles by PRIME students: Iris Shieh pp 24-28; Wesley Hsu pp 37-42; Daniel Li pp 281-287; Brian Tsui pp 307-314; Jade Kwan pp 315-319.

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

Institutions Involved in PRIME in 2012

MONASH UNIVERSITY, MELBOURNE

DOSHISHA UNIVERSITY, KYOTO

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

OSAKA UNIVERSITY, OSAKA

UNIVERSITI SAINS MALAYSIA (USM), PENANG

UNIVERSITY OF AUCKLAND, AUCKLAND

TAIWAN FOREST RESEARCH INSTITUTE (TFRI), TAIPEI

Images (from the top): Atayal aboriginal tribe dancing in Wulai—courtesy of Ariana Tsai; Floating Torii gate of the Itsukushima Shrine in Miyajima—courtesy of Pooja Makhijani; Fisherman village jetty, Penang, Malaysia—courtesy of Jenny Liang

Prototype Mobile Gallery Guide Using Novel Location-Based Technology

MOBILE APPLICATIONS USING GLOBAL POSITIONING SYSTEMS (GPS) ARE CONSIDERED STATE-OF-THE-ART FOR LOCATION-BASED SERVICES (LBS). The technology, however, is not effective for indoor or underground environments. Wireless local area networks (WLAN) are one solution, but implementation can be problematic, especially in scenarios where the location of objects may change frequently.

This project featured a collaboration between the Balboa Park Online Collaborative (BPOC), the San Diego Museum of Art (SDMA), and the National Institute of Information and Communications Technology (NICT) to create an indoor Android-based mobile gallery guide, which allows visitors to construct their own personalized gallery tour. The visitors place their hand-held mobile device on the touch table, which synchronizes the devices. Then visitors can select exhibition artworks by dragging the artwork towards the mobile device. Once the visitors have finished their selections, they can create their tour by transferring the artwork and associated metadata to the mobile device. A listing of the artworks and locations are displayed on the mobile device.

Once the visitors are within range of the artwork they selected they receive identification and distance updates from a palm-sized device (PlaceSticker, Information Services International-Dentsu LTD, Tokyo, Japan) that uses periodic radio signals with limited information content such as a unique identifier and positional data. When the visitor is close to the artwork, the appropriate metadata is displayed on the device. Once the visitors leave the artwork location, the gallery tour is updated and the visitor can proceed to the next artwork.

The prototype developed for this project demonstrates the use of a novel, location-based technology that is usable indoors, and more importantly, provides easy deployment and flexibility in positioning for the end-user. PlaceSticker is ideal for deployment in museum art exhibitions, to help manage the delivery of metadata to the museum audiences in an engaging manner. With each PlaceSticker device uniquely assigned to an artwork, this allows for increased flexibility of positioning the artwork as well as easy implementation by museum staff.

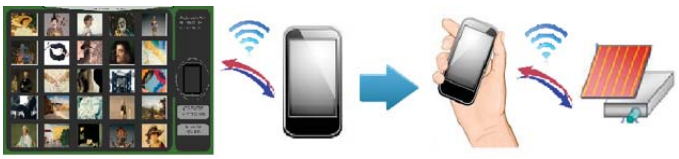


Figure 1: The flow of Mobile Gallery.cv 7.

PARTICIPANTS: PRIME 2012: *Michael Yao, Scott Mo;* BPOC: *Vivian Kung Haga, Christina DePaolo;* SDMA: *Alexander Jarman;* NICT: *Masaki Chikama, Tomoaki Takata, Yoshinori Kobayashi, Shinji Shimojo;* UCSD: *Jason H. Haga*

Images (left to right): Bag End, as seen in Lord of the Rings taken near the town of Matamata, New Zealand—courtesy of Shirley Zhang; PRIME students Katie Chuh and Esther Gudiel test seismic response to anchors in unreinforced masonry buildings, New Zealand—courtesy of Katie Chuh

> I think all students should study abroad. It pushes students out of their comfort zone, which is important for future scientists in our increasingly global society. **Haley Hunter-Zinck, 2008 PRIME alumni**

Torquay beach along Great Ocean Road Melbourne—courtesy of Harriet Hu

Earthquake Engineering

FOR CENTURIES, STRUCTURES IN EUROPE HAVE BEEN BUILT USING THE TECHNIQUE OF SECURING TIMBER DIAPHRAGMS TO THE PERIMETER WALLS THROUGH THE IMPLEMENTATION OF ADHESIVE ANCHORS. This structural component has proven to be beneficial in retrofitting buildings subject to the seismic frequencies of natural disasters by providing additional connections within the frame and therefore reducing the distance of which unsupported walls must span.

Decades ago, this application was adopted in New Zealand to help strengthen buildings along the Australian fault line with the use of external anchor plates. The technique has since evolved to a widespread use throughout the construction industry. After the devastating earthquake in Christchurch in 2011, these anchors came under examination as one of the main failures of the buildings subject to the natural disaster. It was discovered that prior to the implementation of the anchors in New Zealand's unreinforced masonry buildings, few tests of the anchors had been conducted and little was actually known about the strength of these adhesive techniques. Motivation to test these retrofits came from both a significant lack of prior research as well as the severity of building failure in the Christchurch earthquake last year.

The focus of our research, performed in partnership with the University of Auckland, was on the seismic response of anchors in unreinforced masonry buildings. Multiple tests were conducted to observe the seismic responses of different anchor parameters. Tests consisted of measuring the shear strength of masonry walls, learning the process of the anchors' installation, measuring the displacements of anchors being pulled out of the masonry walls, and observing the tensile strengths of anchors that were either installed using epoxy or grout adhesives. Photographic interpretations were also collected and organized into a database that could be used for later analysis of several Christchurch buildings that had been affected by the Canterbury earthquake swarm in 2011.

In order to display the data that was collected, we created graphs to better illustrate our claims for the main causes of multiple levels of damage. Additional analysis of the effect of anchors in unreinforced masonry buildings is necessary and will add to our findings, analyses, and conclusions.

PARTICIPANTS: PRIME 2012: *Katie Chuh, Esther Gudiel*; U Auckland: *Dmytro Dizhar, Josiah Campbell, Brad Christie, Jason Ingham*; UCSD: *Lelli Van Den Einde*

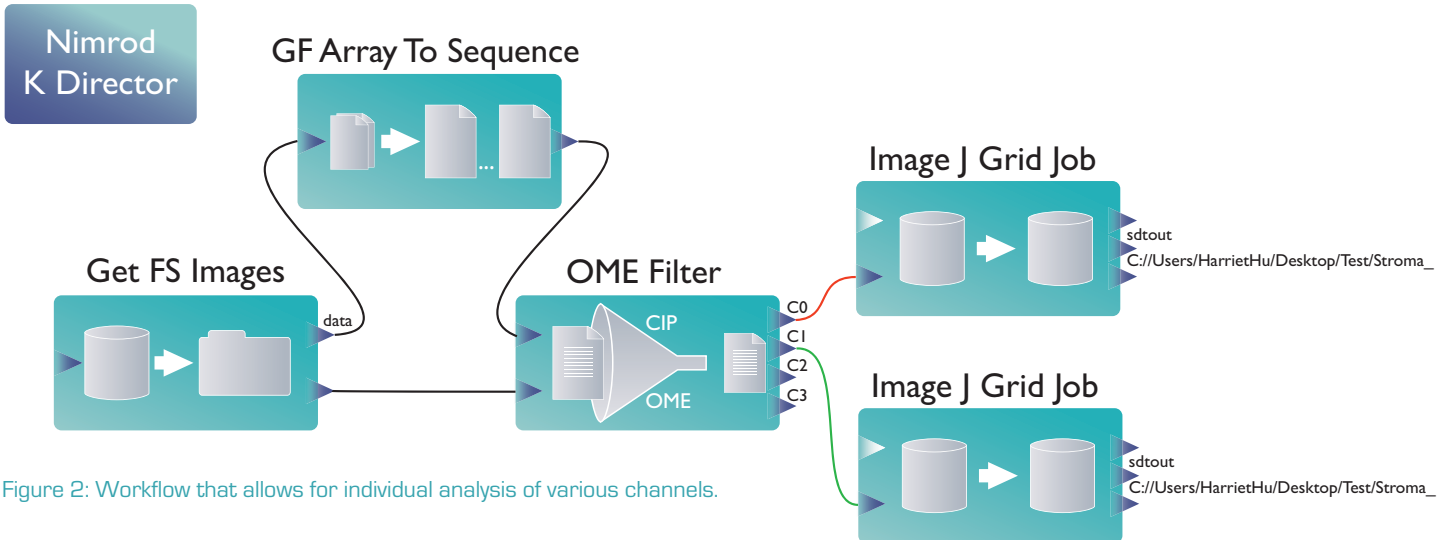


Figure 2: Workflow that allows for individual analysis of various channels.

Automating the Analysis of Microscopy Images through Nimrod/K Project in Kepler Workflow System

THE ADVENT OF MODERN BIOLOGICAL MICROSCOPY HAS LED TO THE INTEGRATION OF VISUALIZATION SOFTWARE WITH COMPUTATIONAL TOOLS USED TO ANALYZE MICROSCOPY IMAGES. Through collaboration with PRIME mentors David Abramson and Ilkay Altintas, as well as Mary Vail from the Department of Biochemistry and Molecular Biology at Monash University, PRIME student Harriet Hu designed a project to increase efficiency of everyday microscopy analysis.

Dr. Vail's work focuses on the effects of an anti-EphA3 antibody, chIIIA4, and its attenuation of tumour growth in cells (Figure 3). Dr. Vail, a biological scientist, noted that conducting routine post-imaging analysis can create a bottleneck effect—much of which can be automated through the use of the Kepler Scientific Workflow System (kepler-project.org). Kepler is a combined workflow graphical user interface and a background execution engine that assists non-computer scientists in analyzing scientific data. Using Kepler v2.3, as well as the Nimrod/K tools developed at Monash, (messagelab.monash.edu.au/NimrodK), Hu was able to construct several workflows that streamlined image processing of experimental data. Programs such as the LIF2OME, used for image extraction, and ImageJ for image processing were integrated into Kepler workflows, providing greater usability and functionality in image processing. Each workflow can be modified to accept necessary parameters in order to create composite images, analyze particles, and compute individual analysis on separate channels of an image.

In addition, these workflows can be further developed based on the analysis that needs to be conducted per experiment, and provide the foundations for more workflows that can be used for image analysis. This project allowed for future case studies of the functional uses of Kepler Scientific Workflow System, as well as a continuation for the long-lasting international collaboration between Monash University and the San Diego Supercomputer Center.

PARTICIPANTS: PRIME 2012: *Harriet Hu;* Monash/MURPA: *David Abramson;* Monash: *Mary Vail;* SDSC/UCSD: *Ilkay Altintas*

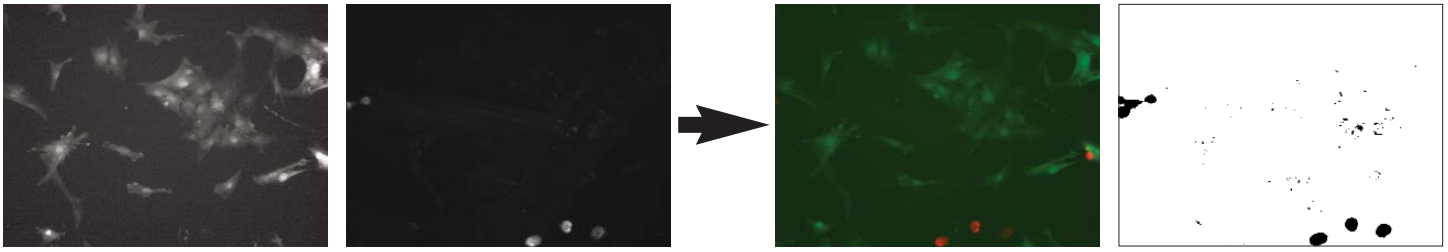


Figure 3: On the left, original microscopy images with separate channels is used as input into the workflow to create composite binary images (on the right panel) to be analyzed.

Virtual Screening for High Specificity Inhibitors of SSH-2

GRID COMPUTING HAS BECOME A POPULAR METHOD FOR DRUG DISCOVERY. Rather than test each chemical compound in wet lab experiments, virtual screening using grid computational resources can simulate the interactions between the compounds (often in a database library) and proteins. In addition, the use of grid computing shares hardware resources and facilitates the performing of multiple computational simulations to determine optimal bindings of compounds to proteins (i.e. docking) at the same time, expediting the process. This project focused on the discovery of inhibitors that would bind specifically to SSH-2, a member of the dual-specific phosphatase (DSP) family. The DSP family of proteins have roles in various cellular processes, and have been linked to cancer and Alzheimer's disease. The challenge in identifying possible inhibitors of SSH-2 is that its overall structure is similar to the structures of other family members; thus there is a need to identify compounds that only interact with SSH-2, not the other members.

Using two previously established PRAGMA grid-enabled programs—MODELLER, which approximates the three-dimensional structure of a protein, and DOCK6, which calculates and scores the binding strength between a protein and chemical compound—this project enhances previous research. To optimize the protein models and analyze how well the resultant protein structure models were folded, the online program MolProbity was used.

The protein structure was prepared for docking and the active site—where the protein-compound interaction would occur—was specifically targeted for docking. In order to provide a higher level of accuracy, two scoring systems were used to assess the binding strength: an energy grid scoring system and the more time-intensive AMBER scoring system. The resulting scores were used to obtain a final consensus score and rank of the best to worst binding compounds.

During the project, protein structure models were generated for SSH-1, SSH-3, and DSP-21. Consensus scoring and compound ranking were completed for SSH-1, SSH-3, and DSP-21. The findings suggest that modeled protein structures can be used in virtual screening experiments. Also, the results show promising chemical compounds that may specifically bind to SSH-2, with minimal interactions with the remaining DSP family members. These compounds will need to be tested in actual cellular experiments to determine if they can specifically inhibit SSH-2 in vitro.

PARTICIPANTS: PRIME 2012: *Vicky Hwang, Joshua Wei, Olivia Yang;* PRIME 2011: *Brian Tsui;* Osaka University: *Susumu Date, Shinji Shimojo;* UCSD: *Jason H. Haga*

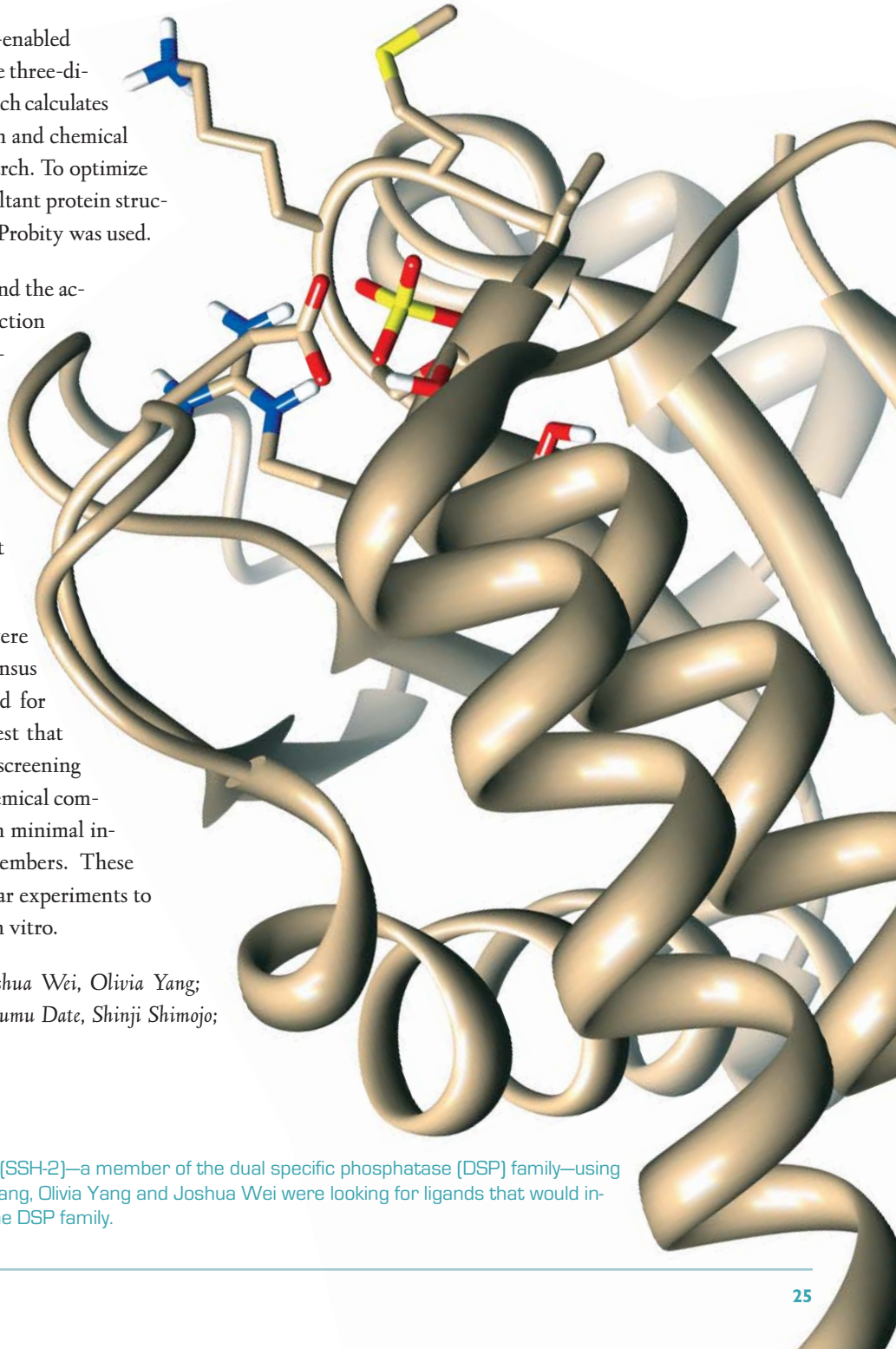


Figure 4: A snapshot taken of slingshot homolog 2 (SSH-2)—a member of the dual specific phosphatase (DSP) family—using the visualization program UCSF Chimera. Vicky Hwang, Olivia Yang and Joshua Wei were looking for ligands that would inhibit SSH-2 without binding to other members of the DSP family.

Molecular Dynamics Simulations of Membrane-bound Cytochrome P450 1A2

THIS RESEARCH PROJECT FOCUSED ON COMPUTATIONAL SIMULATION, I.E., MOLECULAR DYNAMICS, OF THE ENZYME CYTOCHROME P450 1A2 IN TWO ENVIRONMENTS: MEMBRANE-BOUND AND WATER. One of the main functions of cytochrome P450 1A2 is drug oxidation, i.e. drug digestion. Its large binding (active) site can accommodate various sizes of molecules (substrates) to oxidize different kinds of drugs.

The purpose of this project is to determine if, and how, the presence of the membrane has any impact on the stability of the enzyme, including any structural changes in one particular component of the enzyme (i.e., helix structures B, C, F, G) and within binding (active) site.

The two simulations—one, membrane-bound system and another, water system (without the membrane)—were set up using the CHARMM force field, which models the force relationship among the atoms in the systems using two topology files: CHARMM 27 for the protein and CHARMM 36 for the lipid (in this case a structure component of the membrane). The lipid is set up with a ratio of 2:1 of 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) to mimic the environment of the endoplasmic reticulum (ER), a cellular subunit, in human liver, where the 1A2s are mostly found. The massively parallelizable NAMD2 program was used for the molecular dynamics simulations, and 250 nanoseconds (ns) of sampling were gathered for each system.

A number of analyses were performed through the extension tools in VMD (Visual Molecular Dynamics) and other programs, including: root mean square deviation (RMSD), root mean square fluctuation (RMSF), GROMOS clustering, active site analysis, hydration analysis, and a heme angle analysis. The result of the baseline RMSD indicates that the membrane-bound protein with a value of 2.29 angstroms is more stable compared to the water system with a value of 2.41 angstroms. In addition, the RMSF graph helps identify specific residues with significant RMSF differences, which corresponds to regions in the protein that undergo major structural rearrangements in the presence of the membrane. The selected residues or the corresponding helices are then highlighted as areas of interest for further structural analysis, including the interaction with the substrate, the active site, and the membrane. Structural clustering helps identify the different protein structures sampled in each environment, which can be used subsequently as the input for the active site analysis and hydration analysis. In a hydration analysis we identify and track pathways of water molecules that residue in the active site for more than 2 ns to determine potential water channels in both systems.

A major result of the project shows that the size of the active site in the membrane-bound system is significantly smaller with the value of 840 cubic angstroms, whereas the volume in only water, the active site is roughly 2.5 times larger, 2168 cubic angstroms. In addition, in the membrane bound system, a major portion of a helix structure (14) unfolds back to its primary structure. Together, these results will help us understand how the P450 enzymes are able to recognize their many molecular and the atomic-level interactions governing the recognition process.

PARTICIPANTS: PRIME 2012: *Tiffany (Yu-Ting) Sun;* USM: *Sy Bing Choi, Habibah Wahab;* UCSD: *Rommie Amaro, Luke Czapla.*

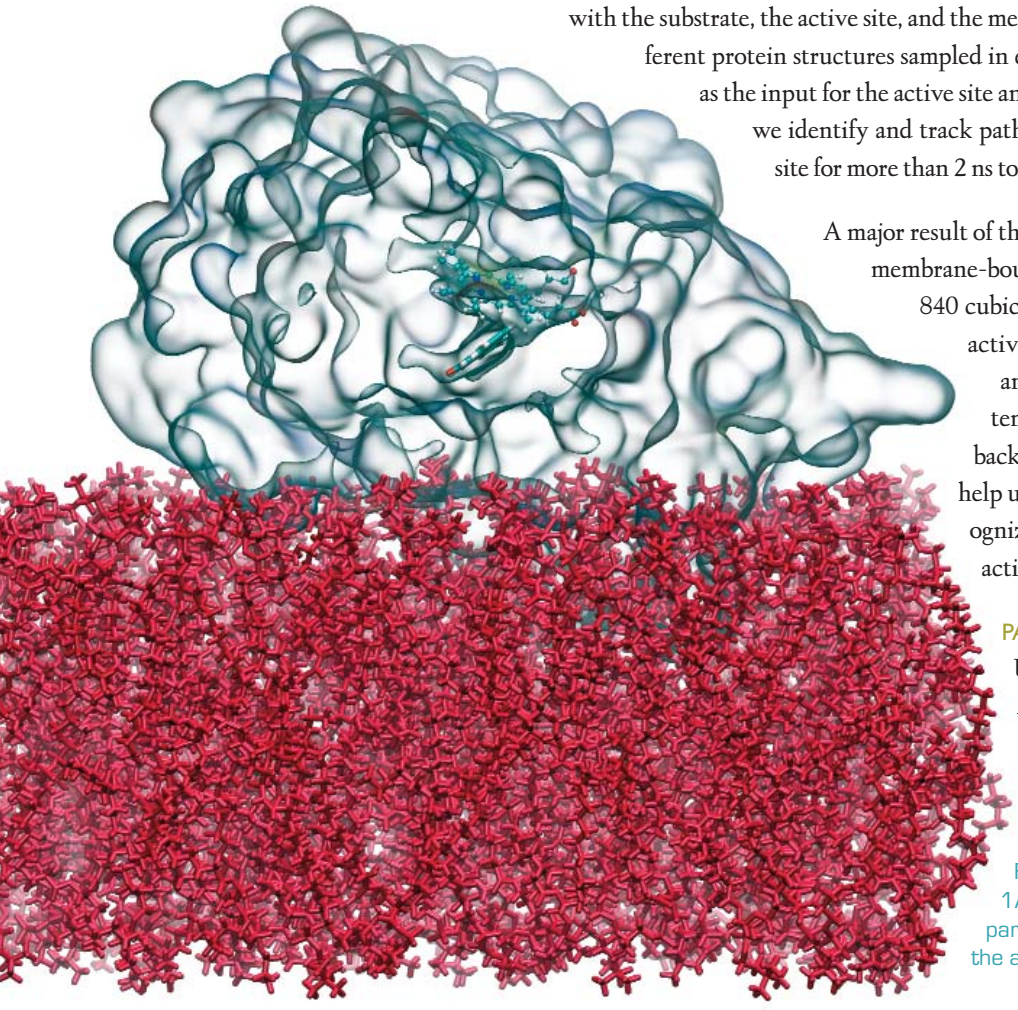


Figure 5: Membrane-bound cytochrome P450 1A2 system. The cytochrome is shown in transparent surface, with the heme and bound ligand in the active site. The membrane is shown in red.

Advancement of Chagas Disease Treatment through the Identification of Potential Natural and Synthetic Product Targets in the *Trypanosoma Cruzi* Proteome

CHAGAS DISEASE IS A TROPICAL MALADY CHARACTERIZED BY CARDIOVASCULAR AND GASTROINTESTINAL COMPLICATIONS IN THE CHRONIC STAGES. Prevalent in regions of Mexico, Central and South Americas, Chagas originates from infection by *Trypanosoma Cruzi*, a protozoan parasite transmitted via contact with triatomine insect feces and contaminated bodily fluids. While the acute phase is largely asymptomatic, chronic Chagas disease includes heart and digestive abnormalities followed by death. Existing treatments with Nifurtimox and Benznidazole are effective in acute phases, but are limited in the chronic stages and often have detrimental side effects.

The purpose of the project is to apply an in-silico approach to identify natural and synthetic compounds that may act as inhibitors of essential proteins in the *Trypanosoma Cruzi* proteome. Our focus is mainly on molecules derived from natural products which often display specific biological activity and have highly diverse stereochemistry. We docked these compounds to proteins and evaluated the possible interactions between individual amino acids comprising the protein active site and natural product derivative in hopes of contributing to the development of a more effective, efficient and safe treatment for *Trypanosoma Cruzi* infection, while halting disease progression (Figure 6).

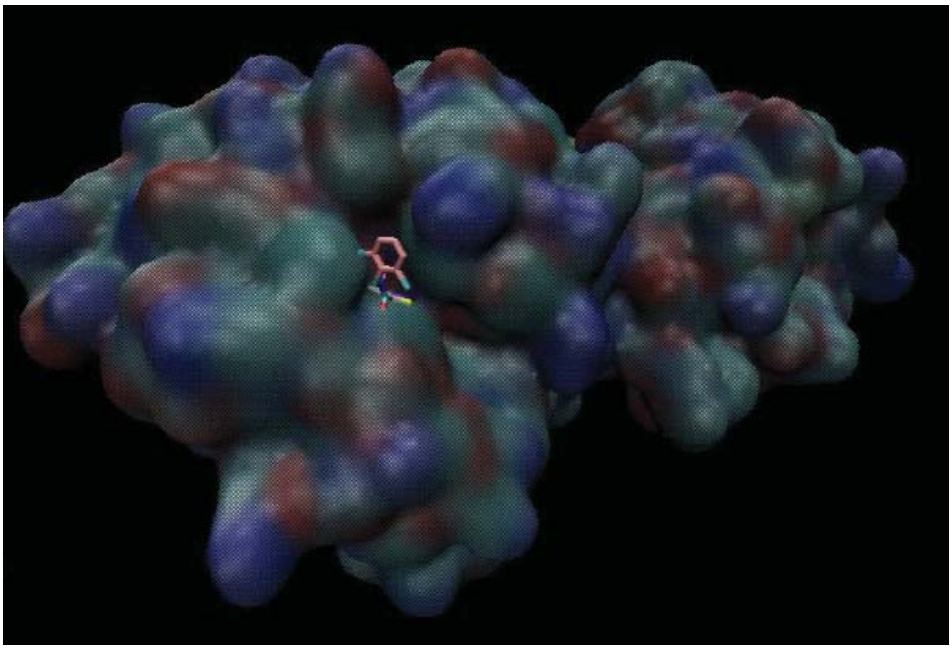


Figure 6: NSC332452 compound docked to 1MS8 (TcTS).

In this computational approach, glycosylphosphatidylinositol (GPI)-anchored surface enzymes, Trans-sialidase (TcTS) and gp63 are of interest. The literature indicates these proteins play vital roles in cell infection, including evasion from the host immune system, cell-to-cell communication, cell penetration, and the attachment of parasites to host receptors. Protein crystal structures 1MS8 (TcTS) and 1LML (gp63) were selected for analysis from the RCSB Protein Data Bank (PDB) based on molecule type, resolution, location of missing residues and complexed ligands. Using Autodock 3.0, a program used to predict the binding of flexible ligands (small molecules) to designated 3D receptors, natural compounds from the Universiti Sains Malaysia's Nature-Based Drug Discovery System (NADI) and synthetic compounds from the National Cancer Institute (NCI) databases were docked between 100 to 1,000 times to these crystal structures using a computational virtual screening approach (Figure 7).

SMAP, a software tool designed for binding pocket similarity search, determined the specificity of these compounds to TcTS and gp63 by screening the *Trypanosoma Cruzi* proteome for structurally similar proteins. However, available PDB crystal structures only represented a small portion of the entire set of proteins comprising the *Trypanosoma Cruzi* proteome. The addition of homology structures to the screened library would provide a more accurate prediction of proteome targets.

With the results of the virtual screening, in-vitro assays can test the efficacy of our predicted top NADI and NCI compounds in inhibiting *Trypanosoma Cruzi* development. The use of NADI serves as a step towards in-silico drug discovery with natural resources. Results would greatly contribute to further drug development for the treatment of Chagas disease and *Trypanosoma Cruzi* research.

PARTICIPANTS: PRIME 2012: *Jenny Liang*; USM: *Sy Bing Choi, Habibah Wahab*; UCSD: *Philip Bourne, Li Xie, Chirag Krishna*

MURPA Advances – This Year

AS WITH PRIME, THE MONASH UNDERGRADUATE RESEARCH PROJECTS ABROAD (MURPA) PROGRAM SUPPORTS SUMMER (I.E., JANUARY-MARCH "DOWN UNDER") INTERNSHIPS WITH LEADING RESEARCH GROUPS OVERSEAS. IT NOT ONLY PROVIDES A HANDS-ON RESEARCH EXPERIENCE TO UNDERGRADUATES, BUT DOES SO IN AN INTERNATIONAL CONTEXT. Students are placed for a period of eight weeks, allowing them to integrate into the research groups as team members. MURPA also involves an advanced seminar scheme, in which students attend seminars given by world-leading experts (this is particularly helpful before they embark on their research adventures abroad). The seminar structure is novel because it uses high-definition interactive video links, making it feasible to attract some of the world's best researchers "virtually" to Monash. These seminars also allow students to "meet" potential MURPA mentors and learn more about potential projects.

MURPA has entered its fifth year and over 20 students have engaged to date. Students have performed internships at UCSD; the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign; and Technion (Israel Institute of Technology, Haifa, Israel). In 2013 we will add the University of Warwick (UK) to the list of possible sites.

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

MURPA seminars are a key component of the scheme, and serve multiple purposes. In addition to allowing students to meet potential mentors, the lectures are integrated into two undergraduate courses, one in Distributed Computing, and another in Computational Science. These provide

Images (left to right): Sydney Opera House—courtesy of Harriet Hu; 18th Biennale of Sydney (city wide art exhibition), this particular piece was created by Tiffany Singhl—courtesy of Harriet Hu

a valuable research focus for courses, broadening the relevance and reach of the material taught. Typically, more than 60 students and staff attend MURPA lectures. Archives of talks are available on the MeSsAGE Lab website (messagelab.monash.edu.au/MURPA).

In 2011, five students went to UCSD, one student to NCSA and one to Technion. In particular:

- Liang Ding worked on a parallel debugger for Eclipse and visited the Blue Waters supercomputing project at NCSA.
- David Warner went to Technion, in Haifa, Israel, where he worked on building a system for assigning conference papers to reviewers.

At UCSD,

- Thomas Moore developed an automated approach for the systems definition of ad hoc wireless sensor networks;
- John Bell worked on increasing utility of and awareness in a novel publication framework called SciVee;
- Satvik Kumar explored implementing a novel data clustering algorithm (called GSOM) on a new data intensive computer at the San Diego Supercomputing Center (called Gordon);
- Minh Huynh explored approaches to authoring virtual machines for the Amazon Elastic Compute Cloud (EC2); and
- Victoria Weldon worked on electrical propagation prediction and validation for patient-specific cardiac models.

To learn more about past MURPA projects see messagelab.monash.edu.au/MURPA/PastProjects

WORKSHOPS AND WORKING GROUPS

PRAGMA workshops are meetings of all members of the PRAGMA community. They are the major vehicle for information exchange between working groups, researchers, and institutions; they also provide excellent opportunities to engage new researchers and students at the host sites. New participants bring new perspective, applications, technologies, students and resources.

Workshops are hosted by different organizations to provide a platform for PRAGMA members to meet and discuss research interests, and ideally develop new collaborations, with members of the hosting institutions.

The workshops are organized by the four working groups in PRAGMA, which are as follows:

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

In 2012, two PRAGMA Workshops were held:

- PRAGMA 22: 17-19 April 2012, Melbourne, Australia, hosted by Monash University
- PRAGMA 23: 9-11 October 2012, Seoul, South Korea, hosted by Konkuk University

PRAGMA workshops are also used to host auxiliary, targeted activities, such as at PRAGMA 22 (April 2012), where a Biodiversity workshop was held the day before the PRAGMA workshop (See Highlight "Infusing New Ideas"). This allows more dialogue between participants of the two activities, and planning for future activities (as described in the "Biodiversity" highlight).

In addition, training programs such as SEAIP (Southeast Asia International Joint Research and Training Program; event.nchc.org.tw/2011/southeast_asia/index.php?CONTENT_ID=2) have provided PRAGMA with new members.

This last year we organized several workshops associated with PRAGMA. As noted, it often takes a persistent interaction to launch successful projects. The Biodiversity Expedition, for example, was generated after a series of workshops on the topic. The following is a list of workshops sponsored by PRAGMA partners to build new collaborations:

Image: Rice Field in Muar, Malaysia--courtesy of Jenny Liang

Biodiversity

SEAIP (Southeast Asia International Joint Research and Training Program); Nov. 29-Dec. 3, 2011, Taichung, Taiwan: event.nchc.org.tw/2011/southeast_asia/index.php?CONTENT_ID=2). Biodiversity Workshop on 3-4 December (Xitou and Taichung, Taiwan). Report at event.nchc.org.tw/2011/southeast_asia/upload/content_file/4fe90876eaa8a.pdf. Both SEAIP and the Biodiversity Workshop were hosted by NCHC, with funding from the National Science Council.

Biodiversity in (Southeast) Asia Workshop: Defining Cyberinfrastructure to Support the Study of Ultramafic Ecosystems. April 17, 2012. Melbourne, Australia: pragma22.pragma-grid.net/dct/page/70005 for context and goc.pragma-grid.net/pragma-doc/pragma22/biodiversity-workshop for presentations. Held in conjunction with PRAGMA 22, hosted by Monash U.

Biodiversity, Ecosystems Services and Cyberinfrastructure: Potential, Challenges, and Opportunities. Oct. 9, 2012, Seoul Korea. Held in conjunction with PRAGMA23, hosted by Konkuk University.

New Partnerships

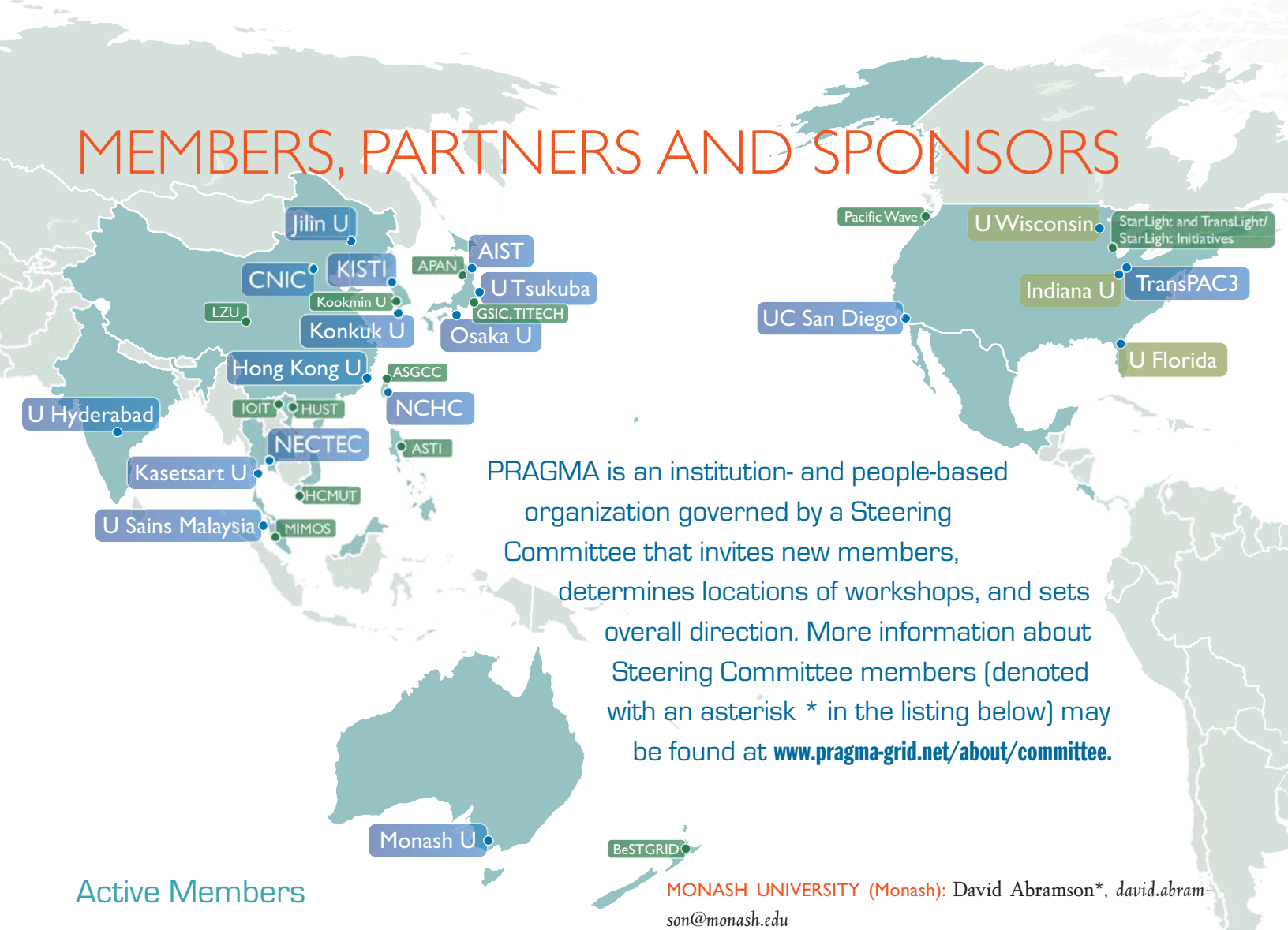
China-US Software Workshop. March 5-8, 2012. San Diego, CA, USA. Hosted by UCSD. This was the second of two scheduled workshops to build collaborations between Chinese and U.S. researchers in areas of high-performance computing, emerging architectures, and trustworthy software: www.nsf-nsfc-sw.org. This workshop was structured on the PRAGMA Workshops and lessons learned in PRAGMA. This workshop is supported by NSF and NSFC funding.

Mini-PRAGMA: Building Collaborations in Clouds, HPC, and Application Areas. 17 July 2012. Hosted by Hong Kong University: community.grid.hku.hk/hpc2/mini-pragma. Aimed at building understanding and collaborations among PRAGMA members, Hong Kong researchers, and researchers in southern China, including National Supercomputer Center of Shenzhen (NSCCSZ), Institute of Software Chinese Academy of Sciences (ISCAS), the Beijing Genome Institute (BGI), and the Shenzhen Institute of Advanced Technology (SIAT) of the Chinese Academy of Sciences. This workshop was supported by HKU Computing Centre.

Looking to the future, we will continue to employ these strategies to engage new researchers. In addition, we will work with our members to identify strategic partners and engage them through focused scientific or technical workshops. We will focus in particular on India, China, and Southeast Asia as regions where there is a growing investment in infrastructure and where there are natural partnerships of mutual interest. Listed below are our planned workshops:

- PRAGMA 24: March 22, 2013, Bangkok, Thailand, hosted by Kasetsart University. This will include workshops on GEO Grid technologies, biodiversity, and environmental observing networks. as well as one on virtualization.
- PRAGMA 25: Fall 2013. Location and Date TBD
- BIG DATA INFRASTRUCTURE: Expedition on the Network Science Landscape. Dec. 4-6, 2012. Taichung, Taiwan. Hosted by NCHC. The focus of this workshop is to bring together researchers from three distinct but overlapping areas to look for commonalities in the data infrastructure. The areas include lake ecosystems, biodiversity, and disaster management. This is supported by NSF and NSC awards and funding from NARL.

MEMBERS, PARTNERS AND SPONSORS



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

Active Members

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

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*

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

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

CYBERMEDIA CENTER (CMC), OSAKA UNIVERSITY: Shinji Shimojo*, *shimojo@cmc.osaka-u.ac.jp*; Susumu Date*, *date@ais.cmc.osaka-u.ac.jp*

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

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

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

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

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

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

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

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

UNIVERSITY OF CALIFORNIA, SAN DIEGO (UCSD): including the California Institute for Telecommunications and Information Technology (Calit2), San Diego Supercomputer Center (SDSC), Center for Research in Biological Systems (CRBS), National Biomedical Computation Resource (NBCR): Peter Arzberger*, *parzberg@ucsd.edu*; Philip Papadopoulos*, *phil@sdsc.edu*; Tom DeFanti, *tdefanti@ucsd.edu*; Teri Simas, *simast@sdsc.edu*

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

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

Networking Members

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

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

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

STARLIGHT AND TRANSLIGHT/STARLIGHT INITIATIVES: Maxine Brown, *maxine@uic.edu*

TRANSPAC3, INDIANA UNIVERSITY: James Williams*, *william@indiana.edu*; George McLaughlin, *george@mclaughlin.net*

Other Members

ACADEMIA SINICA GRID COMPUTING CENTRE (ASGC): 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*; Peter Antonio B. Banzon, *peterb@asti.dost.gov.ph*

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

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

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*

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

MIMOS: Thillai Raj T. Ramanathan, Ng Kwang-Ming, *kwang.ming@mimos.my*; Luke Jing Yuan, *jyluke@mimos.my*

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

Additional Organizations Active in PRAGMA

BALBOA PARK ONLINE COLLABORATIVE (www.bpoc.org), is an organization that provides technical support to more than 20 art, science, and cultural organizations in San Diego. They have contributed resources to the PRIME program to foster a culture of innovation and collaboration among nonprofit organizations through technology (see *Training the Next Generation*, p. 20 as well as *Creating a Crowd-sourced Museum Collection*, p. 16).

BIODIVERSITY INSTITUTE, UNIVERSITY OF KANSAS (biodiversity.ku.edu), and its researchers and students conduct research on seven continents in areas such as biodiversity informatics, systematics and ecology and evolutionary biology. They have contributed to the biodiversity expedition through participation in workshops and in the use and extension of the Lifemapper software.

INDIANA UNIVERSITY (IU), in particular the Data to Insight Center (d2i.indiana.edu) is actively involved in PRAGMA through the NSF award OCI 1234983, and responsible for the test and develop of new analysis and provenance tools to track how data are utilized by the expeditions, in particular those in lake ecosystems and biodiversity.

LAN ZHOU UNIVERSITY (LZU; www.lzu.edu.cn) has contributed resources to the PRAGMA grid, attended the PRAGMA 12, 13, 17, 18, 19, and 22 workshops. They are currently providing resources and are participating in the PRAGMA cloud, and have authored a vm, demonstrated at PRAGMA 23.

NATIONAL APPLIED RESEARCH LABORATORIES (NARL; www.narl.org.tw/en) was established in 2003 to consolidate 11 national laboratories into a single nonprofit organization to construct, operate, and maintain the large-scale R&D facility and platform in support of academic research and foster the necessary manpower in various advanced fields focused by the nation. NCHC is one of the laboratories in NARL. NARL has provided leadership in the GEO Working Group and Disaster Management, 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), is an incorporated administrative agency that conducts general R&D on information technology supporting the ubiquitous society of the future. NICT supported students in the PRIME program in 2009, 2010, 2011 and 2012 and has participated in the activities of the Telescience Working Group through support of the high-definition video conferencing testing.

UNIVERSITI TEKNOLOGI MALAYSIA (www.utm.my) researchers have been involved in the "Biodiversity Expedition" and in the workshops on biodiversity.

UNIVERSITY OF FLORIDA (UF), in particular the Advanced Computing and Information Systems Laboratory (www.acis.ufl.edu) and the Florida Museum of Natural History (www.flmnh.ufl.edu) are actively involved in PRAGMA through the NSF award OCI 1234983, and responsible for both exploring IPv4 and IPv6 overlay networks in the PRAGMA cloud as well as leading the scientific expedition on understand biological adaption in extreme environments.

UNIVERSITY OF WISCONSIN (UW), in particular the Center for Limnology (limnology.wisc.edu), is actively involved in PRAGMA through the NSF award OCI 1234983, and responsible for leading the scientific expedition on predicting lake eutrophication. Furthermore, UW is a key participant in the GLEON organization.

Partners

GLEON, the Global Lakes Ecological Observatory Network (gleon.org), is a grassroots network of limnologists, ecologists, information technology experts, and engineers who uses the network of people, sensors, and data to understand issues such as eutrophication or climate change at regional to global scales. GLEON, established based on an early PRAGMA expedition to place sensors on a lake in Taiwan in 2004, GLEON has grown to a network of more than 300 members, developed new knowledge and insights, created new data products and developed a very successful Graduate Student Association. There are several ties between GLEON and PRAGMA, including shared personnel, learning from the GLEON GSA to develop PRAGMA Student group, the shared Scientific Expedition on Lake Eutrophication, and the joint hosting of a workshop on Big Data in Taiwan in December 2012.

PRAGMA Sponsors

PRAGMA is supported by its member institutions and the U.S. National Science Foundation (NSF OCI 1234983, PI: Arzberger, co-PI Papadopoulos), involving researchers at UCSD, U Florida, Indiana U, and U Wisconsin. This involves support from NSF Office of Cyberinfrastructure, Office of International Science and Engineering, Division of Information and Intelligent Systems in the Directorate for Computer and Information Science and Engineering, and the Emerging Frontiers Office of the Directorate for Biological Science. This funding provides partial support for involvement of Indiana University, University of Florida and University of Wisconsin researchers. In addition, previous support came from NSF OCI-0627026, PI: Papadopoulos, and involved 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.

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

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

CMC at Osaka University PRAGMA activities are supported by several research grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the National Institute of Information and Communications Technology (NICT), Japan. Support for "Fostering of Globally-leading Researchers in Integrated Sciences" (PRIUS) was provided under the MEXT framework of "University Education Internationalization Promotion Program."

CCS's (at U Tsukuba) PRAGMA participation is partially supported by the Core Research of Evolutional Science and Technology Program (JST, Japan).

CCST at Jilin University receives funding support from the Chinese Natural Science Foundation (60703024 and 61170004) and the Chinese Ministry of Education (NCET-09-0428).

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

AIST's sponsors include the Special Coordination Funds for Promoting Science and Technology (MEXT, Japan).

Images (left to right): Brighton Beach huts—courtesy of Harriet Hu; At the Fushimi Inari Shrine near Kyoto Japan—courtesy of Peter Arzberger

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

Kasetsart University'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.

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

Monash University PRAGMA activities (including MURPA) are supported by a range of grants from the Australian Research Council and Monash University internal funding. In particular the MURPA program wishes to acknowledge the financial support of the Faculty of Information Technology and the Monash e-Research Centre. We are extremely grateful to a number of donations from the Cybec Foundation, the Gringlas family and Gapgemini.

NCHC is one of the institutes monitored by the National Applied Research Laboratories that receives major funding support from the National Science Council, Taiwan.

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

TransLight/StarLight receives major funding from NSF (OCI-0962997). StarLight (via Northwestern University) receives funding from NSF (OIA-0963095).

TransPAC3 receives major funding from NSF and Japan's National Institute of Information and Communications Technology.

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

PRIME is funded by NSF (OISE 0710726, PI: G Wienhausen) with additional support from the Office of Cyberinfrastructure, the National Biomedical Computation Resource (NBCR: NIH, NCRR P 41 RR08605), the Gordon and Betty Moore Foundation, and Calit2 UCSD. Host sites USM, Doshisha U, and NICT provide partial to full support students for students local expenses. In addition, PRIME received support from PRIME alumni Haley Hunter-Zinck, given in tribute to her mother.

Work to build GLEON is supported in part by an award from the Gordon and Betty Moore Foundation (#2735) and the NSF Research Coordination Network award (PI: P Hanson, U. Wisconsin, Madison).

PRAGMA would like to acknowledge those who helped prepare this report: Jennifer Matthews for graphic design; Tiffany Fox for editing; Teri Simas for copy editing, and the section authors.

www.pragma-grid.net