PRAGMA

2016-2017

PACIFIC RIM APPLICATIONS AND GRID MIDDLEWARE ASSEMBLY

COLLABORATIVE OVERVIEW

CONTENTS

PRAGMA

COLLABORATION OVERVIEW

Global society faces many challenges, such as meeting pressing water quality and usage demands across the planet, enhancing the livability of smart and connected communities, understanding and preserving the biodiversity and environmental heritage of the planet for future generations, and responding to natural and man made disasters that threaten our communities. These challenges transcend institutions, disciplines and nations and demand global, collaborative approaches to address. Small- to medium-sized international, multi-disciplinary groups—part of the long tail of science—play an increasingly important role in addressing these challenges. The Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) community was established in 2002 to help these multidisciplinary groups make rapid progress in conducting research and education by providing and co-developing international, experimental cyberinfrastructure to address some of the many challenges facing global communities today.

PRAGMA has created a collaborative framework to bring international, multidisciplinary groups together, engaging both domain and technical expertise, assisting in sharing data and software, providing access to resources and creating opportunities to share progress and refocus efforts based on progress. In addition, PRAGMA has focused its efforts on the development of software to improve access to computing resources and data that can help advance the scientific efforts of these groups.

In this PRAGMA Collaborative Overview 2016–2017, we highlight the advances made by our:

- Scientific expeditions, i.e., teams of domain scientists and cyberinfrastructure experts brought together by PRAGMA, to address key scientific questions through collaboration and development of technologies;
- Technology working groups to develop software and both experimental and stable testbeds for science;
- Student organization to grow students' expertise and build their professional networks; and
- New collaborative activities to engage new researchers and define new directions for PRAGMA.

This year's advances, described more fully in the following pages, include the following:

- Extending the lake expedition's development of the GRAPLEr modeling system to include additional compute resources, improving its usability, using it to gain insights into the seasonal dynamics of algae and developing and implementing training modules to teach new generations of lake ecologists of the power of ecological modeling in lake science;
- Increasing the scalability and portability of Lifemapper, a biodiversity modeling tool, to run on a larger resource and use larger data sets, allowing it to migrate to locations where data are "pinned";
- Enhancing the PRAGMA Experimental Testbed (PRAGMA-ENT) through the deployment of an OpenFlow switch in the US, thereby improving PRAGMA-ENT's performance, then extending the testbed



to new sites such as Thammasat University, providing end-user connectivity through the use of IP over P2P (IPOP) software and deploying monitoring tools, including PerfSONAR, on the testbed;

- Developing a lightweight scheduling software for the PRAGMA Cloud and demonstrating the ability to easily extend with PRAGMA software tools the PRAGMA Cloud to include additional resources;
- Modifying a virtual machine workflow to enable rice genome researchers to enhance rice breeding;
- Piloting a Persistent ID (PID) approach to scientific data management of research data products from the PRAGMA Cloud, by leveraging Research Data Alliance (RDA) recommendations, thereby establishing PID-oriented data services that mesh with the trusted network vision of PRAGMA-ENT;
- Expanding and experimenting with an air quality monitoring testbed using the AirBox system in partnership with the Collaborations to Enable Transnational Cyberinfrastructure Applications (CENTRA);
- Prototyping a collaborative visualization platform via PRAGMA-ENT for use in disaster response situations, in partnership with CENTRA;
- Advancing computational science learning through providing various in-house simulation software on the cyber-learning open platform called EDISON (EDucation-research-industry Integration through Simulation On the Net), developed by the Korea Institute for Science and Technology Information (KISTI);
- Increasing opportunities for students in PRAGMA through exchange programs and a student Hackathon;
- Continuing to build community by engaging new researchers through activities and partnerships, both existing (e.g., Southeast Asia International Research and Training Program, (SEAIP), and new (e.g., CENTRA) and disseminating PRAGMA results more broadly, e.g., in a special edition of **Concurrency and Computation: Practical Experience.**

As we look forward to the coming year, we will be addressing specific challenges that will move us to one of our overarching goals, namely to build and use a user-invoked, software-defined network trust envelope of compute and data resources in which to conduct science using data that are often not allowed to migrate from its source. This will entail the following:

- Deploying and using software, such as Lifemapper and GRAPLEr, to utilize data (often pinned to a location) based on a scientific research question;
- Using PRAGMA-ENT and the technologies from PRAGMA Cloud and extending these environments to include sensors;
- Engaging more application users of the PRAGMA-ENT and Cloud, based on the lessons we have learned;
- Developing a PRAGMA data repository for advancing scientific research and for developing and testing tools of data science, e.g., provenance; and
- Strengthening student activities and engagement in PRAGMA.

AT A GLANCE

PRAGMA was established in March 2002, with the mission of enabling the long tail of science through scientific expeditions and infrastructure experimentation for Pacific Rim institutions and researchers

MEMBERS: 29 Institutional Members

GOVERNANCE: Steering Committee, and PRAGMA's Operating Principles and Procedures

WORKSHOPS OF PARTICIPANTS: Twice a year to share progress and plan future activities; open to all, hosted by PRAGMA members, often in conjunction with other activities

2016 WORKSHOPS

- **PRAGMA 30:** January 27–29, 2016, Advanced Science and Technology Institute (ASTI), Manila, Philippines. Held in conjunction with the 41st Asia-Pacific Advanced Network (APAN) meeting
- **PRAGMA 31:** September 7–9, 2016, Thammasat University, Bangkok, Thailand. Held in conjunction with three additional events: the ASEAN International Virtual Organization Meeting; the Baseline Environmental Data Capture Project in the Lower Mekong Basin Meeting; and the 2016 Asia Pacific Open Data Summit

2017 WORKSHOPS

- **PRAGMA 32:** April 12–14, 2017, University of Florida, Gainesville, Florida, United States. Held in conjunction with the 2nd Annual CENTRA meeting
- **PRAGMA 33 & beyond:** Discussions are underway to identify locations

SCIENTIFIC EXPEDITIONS

- **LIMNOLOGY:** Predicting lake eutrophication and training the next generation of lake scientists
- **BIODIVERSITY:** Understanding biological adaptation in extreme environments
- **ENT:** Developing an experimental network testbed for experimenting with software-defined networks and monitoring impacts of choices

WORKING GROUPS

- **RESOURCES:** Making the distributed resources of PRAGMA useful for diverse applications
- **TELESCIENCE:** Making and improving access to or use of remote equipment (e.g., tiled-display walls or sensors)
- **BIOSCIENCES:** Creating stable infrastructure to perform computational genomics analyses with a focus on rice breeding and integrating technologies to create an infrastructure to advance the screening of potential compounds to combat infectious diseases
- **CYBERLEARNING:** Developing simulation-based learning technology

STUDENTS

- **PRAGMA STUDENTS:** Stimulating international cross-disciplinary collaboration among students and junior researchers

PARTNER ACTIVITIES

- **GLEON:** Global Lakes Ecological Observatory Network, whose mission is to understand, predict and communicate the role and response of lakes in a changing global environment (gleon.org)
- **CENTRA:** Collaborations to Enable Transnational Cyberinfrastructure Applications, whose long-term goal is advancing scientific understanding of distributed software-defined cyberinfrastructure through partnerships and evolving a framework for collaborations amongst research centers, institutes and laboratories across the world. CENTRA engages junior researchers in software-defined infrastructure, with initial drivers of environmental modeling, disaster management and smart cities (*www.globalcentra.org*)
- **SEAIP:** Southeast Asia International Research and Training Program, hosted annually by the National Center for High-performance Computing, to promote collaborations in cyberinfrastructure among researchers in Southeast Asia and between those researchers and others around the world (*event.nchc.org.tw/seaip*)

SPONSORS

- Multiple, often associated with members and funded through many different national science foundations

WEBSITE

- *www.pragma-grid.net*

EXPEDITIONS

Global Approach to Lake Ecology

Harnessing Computational Tools and High-Frequency Data as a Global Approach to Lake Ecology

The global demand for energy, food and drinking water is driving unprecedented pressures on the world's lakes and reservoirs. The reduction in water availability and the degradation in water quality that the global community is experiencing today will shape its capacity for providing the natural resources needed in the decades ahead. There is no doubt that increasing demands for goods and services in one region of the world can alter the water quality in other regions that meet those demands. The global hydroscape supports the global economy.

Smart management of our freshwater resources requires near real-time, high-frequency information about our lakes and reservoirs, knowledge of how our resources respond to increasing global pressures and technologies that enable us to scale this knowledge to the continental and even global domains to predict future water quality and quantity. In short, our understanding of freshwaters needs to be at a scale commensurate with the global pressures placed upon our water resources. However, there remains a sizable gap between the emergent global water crisis and our scientific capacity for understanding the controls of that crisis and devising smart management strategies.

Lake ecologists and computer scientists in the PRAGMA Lake Expedition are teaming up to confront an initial, and until now intractable, part of the problem—how to accurately simulate the biophysical dynamics of lakes around the world under changing land use and climate conditions. Our interdisciplinary team is working together to overcome challenges that have long impeded progress towards this goal. These challenges include running the data-hungry and compute-intensive models that ecologists have developed to simulate the dynamic biophysical environments of lakes and reservoirs. Compounding the problem is the need to run thousands of simulations for each lake to ensure that the models are properly calibrated and that the uncertainties around predictions are represented. This is both a data problem and a compute problem.

Figure 1. GRAPLEr Training at GLEON 17—courtesy of Grace Hong.

To solve the compute problem, computer scientists in PRAGMA have created a service that provides users with easy access to a vast, and largely untapped, distributed set of computers. The challenges that this inter-disciplinary team has addressed go beyond harnessing the computational power of high-throughput distributed computing resources to run large numbers of model simulations. In addition, it is also essential to present a powerful, yet intuitive, user interfaces that integrates with existing desk-

HIGHLIGHTS

top environments used by lake ecology researchers and students. A major outcome of the collaboration between PRAGMA and GLEON (the Global Lake Ecology Observation Network, *gleon.org*, an organization of more than 500 scientists world-wide) researchers is the GRAPLEr system (GLEON Research And PRAGMA Lake Expedition, see Subratie et al, code available at github), which provides an integrated, end-to-end environment for users to setup, run, retrieve results and visualize/post-process results from thousands of simulations, using simple API (Application Programming Interface) calls directly from R/R-Studio, a software application widely used by PRAGMA's target audience of researchers and students. At the backend of the distributed system that powers this unique cyberinfrastructure, GRAPLEr logically connects together cloud resources from different institutions, including resources at the University of Florida, the San Diego Supercomputer Center's Comet cluster and PRAGMA-Cloud virtual machines, using the IPOP (IP-over-P2P) overlay software developed by PRAGMA. Then, GRAPLEr uses the open-source HTCondor system to schedule and manage large numbers of model runs. GRAPLEr also integrates open-source Web service software (Flask, Celery) to expose the interface that users can access through R.

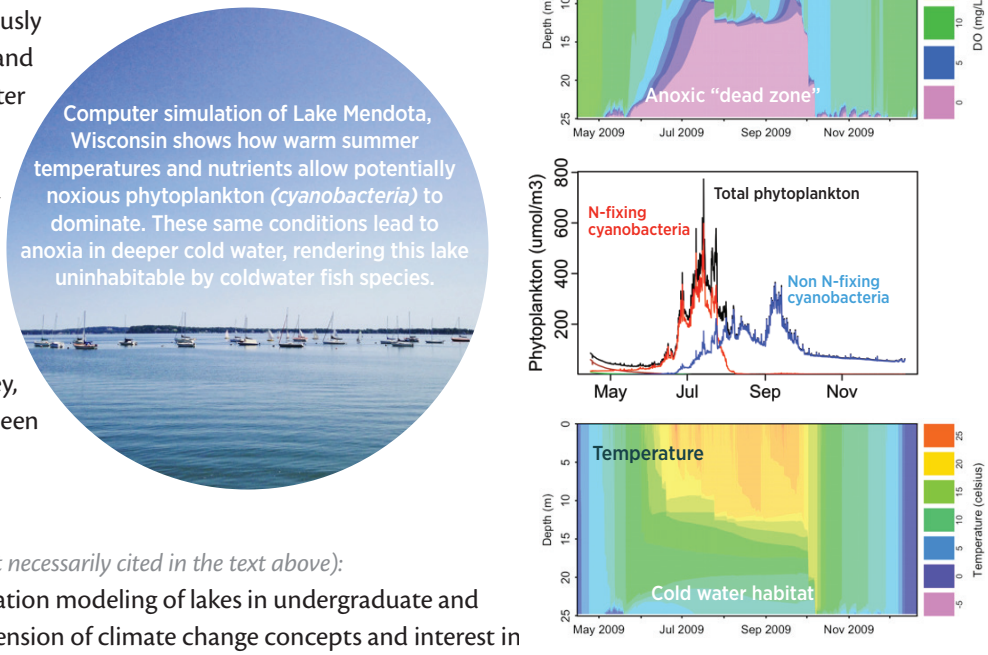
Additionally, for scientists to reap the rewards of new technological capacity, they need training. Ecosystem modeling is a critically important tool for environmental scientists, yet is rarely taught in undergraduate and graduate classrooms. To address this gap, Prof. Cayelan Carey developed a teaching module (*cemast.illinoisstate.edu/data-for-students/modules/lake-modeling.shtml*) that exposes students to a suite of modeling skills and tools, including computer programming, numerical simulation modeling and distributed computing, that the students apply to study how lakes around the globe are experiencing the effects of climate change. This module has been developed as part of a collaboration between PRAGMA, GLEON, and Project EDDIE (Environmental Data-Driven Inquiry and Exploration [*projecteddie.org*]) to develop open-source teaching materials to train undergraduate students climate change concepts and quantitative literacy. In the teaching module, students develop hypotheses about the effects of different climate scenarios on lakes, and then test their hypotheses using hundreds of model simulations in the GRAPLEr environment. The module was taught in two workshops, and it was found that participation in the module significantly increased both undergraduate and graduate students' understanding about climate change effects on lakes. Moreover, participation in the module also significantly increased students' perceived experience and interest in different software, technologies and modeling tools. By embedding modeling in an environmental science context, non-computer science students were able to successfully use and master technologies that they had previously never been exposed to (see Carey and Gougis).

PRAGMA has enabled the essential ingredients of scientific advancement—ideas, data, models, compute infrastructure and training—to converge in an international collaborative environment. It is beginning to pay off. Ecologists have learned how to recreate the seasonal dynamics of algae, including harmful species, and have identified critical shortcomings in predicting harmful blooms. The rapid increase in cyanobacteria that can lead to toxic water quality in less than a day cannot be reproduced by conventional models, inspiring lake ecologists to look to other disciplines, such as meteorology, to recreate these surprises. Our work also indicates that increasing temperatures due to climate change and increasing nutrients due to land use changes may interact synergistically to promote blooms (see Figure 2).

Together, the PRAGMA-GLEON Lake Expedition is fundamentally advancing our understanding of how global change is altering our freshwater resources while simultaneously developing new compute infrastructure and training tools to transform how freshwater ecology is studied.

PARTICIPANTS: *Univ Wisconsin-Madison, Center for Limnology:* Paul C. Hanson, Hilary Dugan, Craig Snortheim; *Univ Florida:* Renato J. Figueiredo, Ken Subratie, Saumitra Aditya, Satish Mahesula, Jaikrishna Tanjore; *Virginia Tech:* Cayelan C. Carey, Amy Hetherington, Ryan McClure, Kathleen Hamre, Alexandra Gerling

Figure 2. The Power and Limit of the Model: Output from computer simulations that provide a plausible explanation of observed fish die-offs.



PUBLICATIONS

Through and related to PRAGMA activities (not necessarily cited in the text above):

Carey, C.C. and R.D. Gougis. *In Press.* Simulation modeling of lakes in undergraduate and graduate classrooms increases comprehension of climate change concepts and interest in computational tools. Journal of Science Education and Technology.

Gerling, A.B. 2015. Hypolimnetic oxygenation mitigates the effects of nutrient loading on water quality in a eutrophic reservoir. Masters Thesis. Virginia Tech.

GRAPLE GitHub repository: *https://github.com/GRAPLE/GRAPLEr*

GRAPLE Web Site: *http://graple.org/GRAPLEr*

Hanson, P.C., C.C. Carey, M. Hipsey. *In prep.* Tail of two distributions: the probability of harmful algal blooms in model simulations.

Ruan, G., P.C. Hanson, H.A. Dugan, and B. Plale. *In review.* Mining lake time series using symbolic representation. Ecological Informatics.

Snortheim, C.A. 2015. Meteorological drivers of oxygen depletion in Lake Mendota. Masters Thesis. University of Wisconsin-Madison.

Snortheim, C.A., P.C. Hanson, K.D. McMahon, J.S. Read, C.C. Carey, H.A. Dugan. *In press.* Meteorological drivers of hypolimnetic anoxia in a eutrophic, north temperate lake. Ecological Modeling.

Subratie, K., S. Aditya, R. Figueiredo, C.C. Carey, P.C. Hanson. 2015. GRAPLEr: A distributed collaborative environment for lake ecosystem modeling that integrates overlay networks, high-throughput computing, and Web services. PRAGMA Workshop on International Clouds for Data Science (PRAGMA-ICDS'15). arXiv:1509.08955 [*cs.DC*]

Images: (left to right) Paul Hanson and Cayelen Carey; Jing Yuan Luke and Jen-Gaw Lee; Kohei Ichikawa and Susumu Date—all courtesy of Teri Simas

Lifemapper

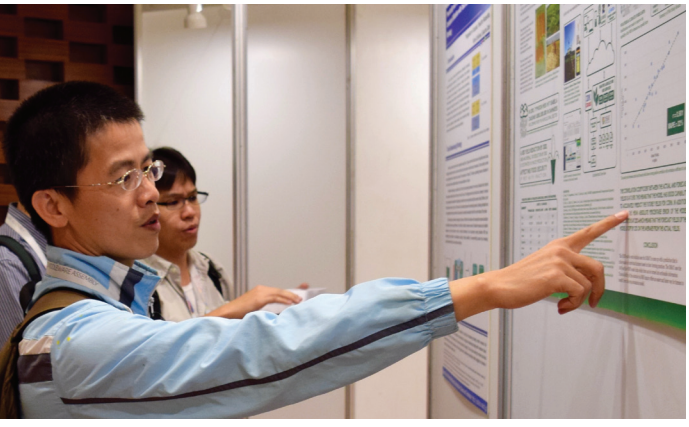
Scaling Up the Virtual Biodiversity Expedition through Improved Efficiency and Portability

The Virtual Biodiversity Expedition (VBE) has taken a complex, clustered, domain-specific, data analysis and visualization software system, Lifemapper, and made it portable by enabling its operation on a variety of system configurations. This portability is critical to many use cases where data are pinned to specific locations because of institutional, local and/or national policies. This year, the VBE team has continued its work on optimizing and moving the virtualization to the newest the San Diego Supercomputer Center's (SDSC) high-performance computing resource, Comet.

Lifemapper was developed by the University of Kansas (KU) Biodiversity Institute and includes a suite of data and tools for biodiversity researchers and an archive of species distribution maps calculated from public specimen data. Lifemapper analysis and modeling tools predict the potential geographic distribution of individual species based on where they have been found and analyze the relationship between multi-species distributions, geographic features and evolution over large-scale landscapes.

Lifemapper starts with geographic data for individual species and takes it through a complex workflow with environmental, phylogenetic and geographic data to assemble the data structures for these multi-species, multi-dimensional analyses. The complexity of the workflow made it a priority for code optimization. The KU team leveraged tools from the Cooperative Computing Laboratory (CCL, *ccl.cse.nd.edu*) at Notre Dame to dramatically simplify the job management process. One tool, *Makeflow,* manages dependencies between connected processes allowing Lifemapper, or a user, to expand or modify workflows without needing to handle new dependencies. Another tool, *Work Queue,* reduces the total overhead incurred by running only one job per process by allowing more than one "worker," i.e., process, each of which can pull jobs from a workflow (one Makeflow instance) until there are no more left. *The CC Tools Catalog Server* manages multiple Makeflow instances allowing workers to pull jobs from multiple Makeflows rather than being tied to only one. This allows a set number of workers to stay busy until everything is complete, rather than shutting down a worker when its Makeflow is complete, then starting another (with its corresponding overhead) for a new Makeflow.

The increased efficiency of job management with CCL tools feeds directly into the next major step of the VBE this year – supporting researcher-driven Lifemapper instances. Existing installations of Lifemapper have used species occurrence datasets from public aggregators, such as the Global Biodiversity Information Facility (GBIF, *www.gbif.org*), the United States Geological Survey BISON (USGS, *bison.usgs.ornl.gov*) and the US National Science Foundation project iDigBio (*www.idigbio.org*), as inputs. Predictions and computations on these instances are updated on a rolling basis as new data becomes available. Datasets are large, but there is no requirement for the entire archive to be computed in a short period of time. As Lifemapper matures, it is supporting installations for a researcher to study the distribution of a very large number of species, both individually and as a group. In these cases, a researcher needs all computations to be finished in a relatively short time, at which point the researcher can do comparative analyses. The first researcher who has requested



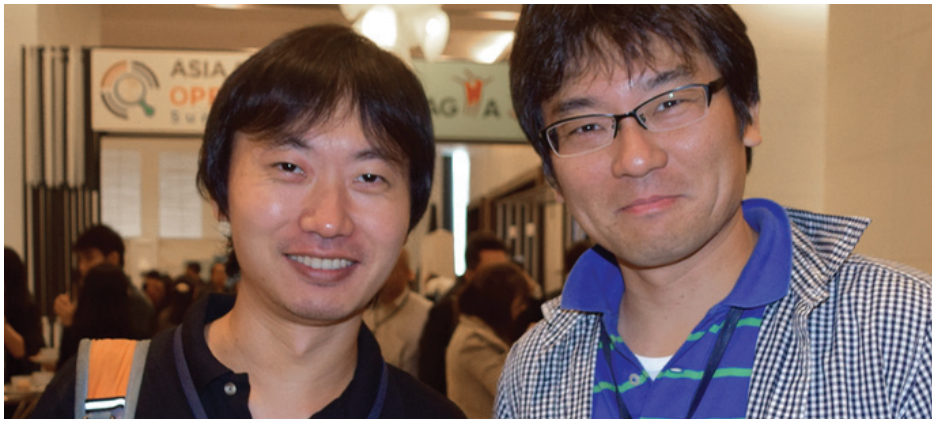


Image: In front of Fort Santiago, Manila, after PRAGMA 30

a large analysis from Lifemapper has data for approximately 12,000 species, producing over 1 million jobs, each of which takes an average of 30 minutes on the existing Lifemapper hardware.

A new opportunity arose for handling the large scale and more time-dependent computations when SDSC brought online a new Extreme Science and Engineering Discovery Environment (XSEDE) resource, Comet. Deploying Lifemapper software on a virtual cluster on SDSC's Comet cluster leverages both Lifemapper's existing Rocks-based software stack and the virtual cluster support unique to Comet. Comet is intended to provide a support for high-performance virtualization at the multi-node cluster level. This means we can install our software stack on a virtual cluster capable of the high-performance usually provided by a supercomputer. Each virtual cluster node has 120Gb of memory and 24 cores running at 2.6GHz. In addition, the Comet hardware has a full bisection InfiniBand FDR interconnect and access to SDSC's Data Oasis parallel file storage system. This translates into storage access from all virtual cluster nodes and will allow nodes to create terabytes of data and move them out upon completion.

For the virtual cluster installation, we use a specific CloudMesh client interface to create, configure and access the virtual cluster. The interface also provides a browser-based console to install, administer and connect to the cluster nodes. The start of the virtual cluster installation can be seen in the following figure:

The timeline for the installation of the virtual cluster is as follows:

1. Frontend install from the first splash screen (shown in Figure 3) to the first boot took 33 min 35 sec;
2. First frontend boot, including waiting for the automated post-boot processes until the frontend is ready to build the compute nodes, took 14 min 55 sec.
3. Building eight compute nodes took 4 min 15 sec;

As a result, a basic Rocks 6.2 virtual cluster was ready in less than 53 min. The next steps will be setting up an IP over Infiniband (IPoIB) network interface and Lifemapper software via rolls. After these steps are completed, the cluster will be ready for integration of large-scale Lifemapper runs.

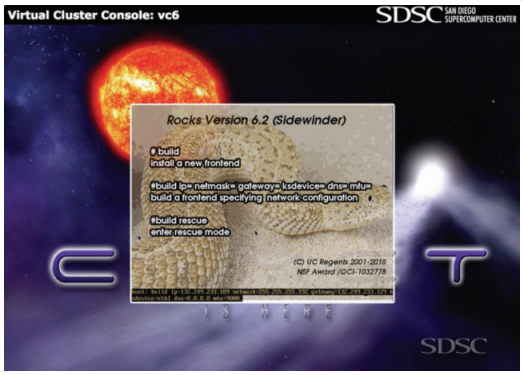


Figure 3. Using Virtual Cluster Console for frontend installation.

PARTICIPANTS: *UC San Diego:* Nadya Williams, Philip Papadopoulos; *KU Biodiversity Institute & Natural History Museum:* Aimee Stewart

PRAGMA-ENT

Expanding and Improving the PRAGMA Experimental Networking Testbed

The PRAGMA Experimental Networking Testbed (PRAGMA-ENT) expedition has the goal of constructing a breakable international Software-defined Network (SDN)/OpenFlow testbed for use by PRAGMA researchers and collaborators. PRAGMA-ENT is breakable in the sense that it will offer complete freedom for researchers to access network resources to develop, experiment with, and evaluate new ideas without concerns of interfering with a production network. PRAGMA-ENT also will provide networking support to the PRAGMA multi-cloud and user-defined trust envelopes. This will expose SDN to the broader PRAGMA community and facilitate the long-tail of eScience by creating new collaborations and infrastructure among institutes in the Pacific Rim area.

The PRAGMA-ENT team has been connecting resources in the US (University of Florida, UC San Diego and Indiana University), Japan (Nara Institute of Science and Technology, National Institute of Advanced Industrial Science and Technology and Osaka University) and Taiwan (National Applied Research Laboratories). The **R**esearch **I**nfrastructure for large-**S**cale network **E**xperiments (RISE) service offered by Japan Gigabit Network (JGN)-X is used to connect our OpenFlow switches deployed at the participating sites. Since all OpenFlow switches are interconnected through RISE, it is possible to develop our own controllers to manage the entire network testbed. This year, we successfully deployed an OpenFlow switch of RISE at the Pacific Wave exchange point in Los Angeles and started using the switch to interconnect the links of US, Japan and Taiwan. This switch has helped optimize the routing among these countries and thus has significantly improved the performance of PRAGMA-ENT.

We are currently working with researchers at Thammasat University (TU) to expand the backbone of ENT to Thailand. We have already established generic routing encapsulation (GRE) links (virtual point-to-point links over the Internet) from PRAGMA-ENT to TU. Our next goal is to establish an international VLAN (virtual local area network) between TU and ENT. We have also started a collaborative project with researchers and students of Kasetsart University (KU) in Thailand to monitor and visualize the OpenFlow network of PRAGMA-ENT. The monitoring and visualization tool of ENT would be very helpful for the researchers to understand the state of PRAGMA-ENT.

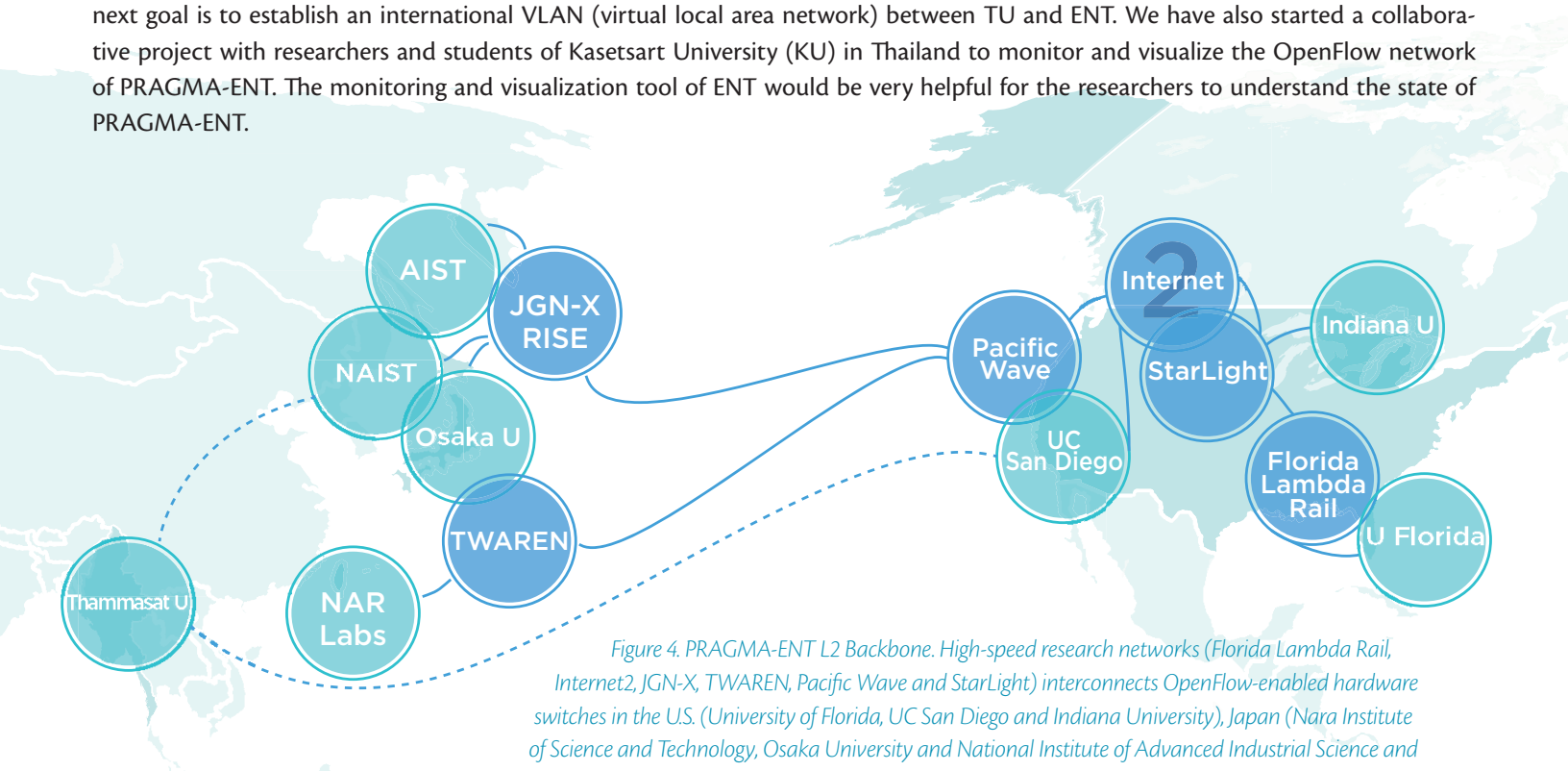


Figure 4. PRAGMA-ENT L2 Backbone. High-speed research networks (Florida Lambda Rail, Internet2, JGN-X, TWAREN, Pacific Wave and StarLight) interconnects OpenFlow-enabled hardware switches in the U.S. (University of Florida, UC San Diego and Indiana University), Japan (Nara Institute of Science and Technology, Osaka University and National Institute of Advanced Industrial Science and Technology), Taiwan (National Applied Research Laboratories) and Thailand (Thammasat University). The solid and dashed lines indicate physical connections and virtual connections respectively.

We have been also working with the IPOP project team at the University of Florida to extend the PRAGMA-ENT backbone using the IP over P2P (IPOP) overlay network. The challenge faced by many end-users is the difficulty establishing a direct connection to the international VLAN backbone of ENT just for their temporary experiments. IPOP is an overlay network technology that enables user-controlled virtual networks to be defined and managed by end-users themselves. We plan to deploy IPOP access points at the participating sites of ENT and thereby enable end-users to establish connections to ENT through the nearest access point of IPOP. Once a connection has been established, the end-users can use high-speed network of PRAGMA-ENT through the IPOP overlay.

PARTICIPANTS: *Nara Institute of Science and Technology:* Kohei Ichikawa, Pongsakorn U-chupala, Chawanat Nakasan, Che Huang, Yasuhiro Watashiba; *Univ Florida:* Matthew Collins, Kyuho Jeong, Renato Figueiredo, Jose Fortes; *National Institute for Information and Communication Technology (NICT):* Hiroaki Yamanaka, Jin Tanaka; *UC San Diego:* Nadya Williams, Shava Smallen, Philip Papadopoulos; *Osaka Univ:* Yoshiyuki Kido, Susumu Date, Shinji Shimojo; *National Institute of Advanced Industrial Science and Technology (AIST):* Jason Haga, Ryousei Takano, Atsuko Takefusa, Yoshio Tanaka; *National Applied Research Laboratory (NARL):* Li-Chi Ku, Fang-Pang Lin, Jen-Wei Hu, Te-Lung Liu; *Indiana Univ, Bloomington:* Quan (Gabriel) Zhou, Beth Plale, Jennifer Schopf; *Internet2:* John Hicks, Rick McMullen; *Computer Network Information Center (CNIC):* Kevin Dong, Yongmao Ren; *Jilin Univ:* Xiaohui Wei; *Kasetsart Univ:* Putchong Uthayopas; *Thammasat Univ:* Wanida Putthividhya, Prapaporn Rattanatamrong

Cloud Testbed

Lightweight Scheduling for the PRAGMA Cloud Testbed

The Resources Working Group has been developing a lightweight scheduler for the PRAGMA Cloud Testbed as a way to automate access to different PRAGMA Cloud resources and allow users to run scientific experiments on their own virtual clusters. Since our target user base is small and does not need to scale to hundreds or thousands of users, the PRAGMA Cloud Scheduler was designed to have a simple architecture that prioritized ease of use, and low installation and maintenance overhead, as shown in Figure 5.

For users, the PRAGMA Cloud Scheduler has an easy-to-use Web front-end based on a calendar reservation system called Booked from Twinkle Toes software. Booked was chosen after evaluating a number of scheduler options, such as batch schedulers and other custom cloud testbed tools, and was deemed the best suited to our needs. It was easy to setup, had a good look and feel, and was customizable. It had some nice additional features such as usage reporting, support for different time zones for each user, a REST API interface, fine-grained roles and permissions management, and user and group quotas. A screenshot of the PRAGMA Cloud Scheduler web interface is shown in Figure 6. The user sees a calendar table layout where the header rows are days of the week and the columns are hours of the day. Under each day, available resources are shown and can be filtered using the box on the left side of the screen. To create a new virtual cluster, the user clicks on the time of day to open a new pop-up box where the desired the timeframe, virtual cluster image, and configuration details can be requested, and then submits them to the Cloud Scheduler. Once the reservation is verified and confirmed, the user will receive an email confirmation. When the reservation is ready to be activated, the cloud scheduler will use SSH to remotely

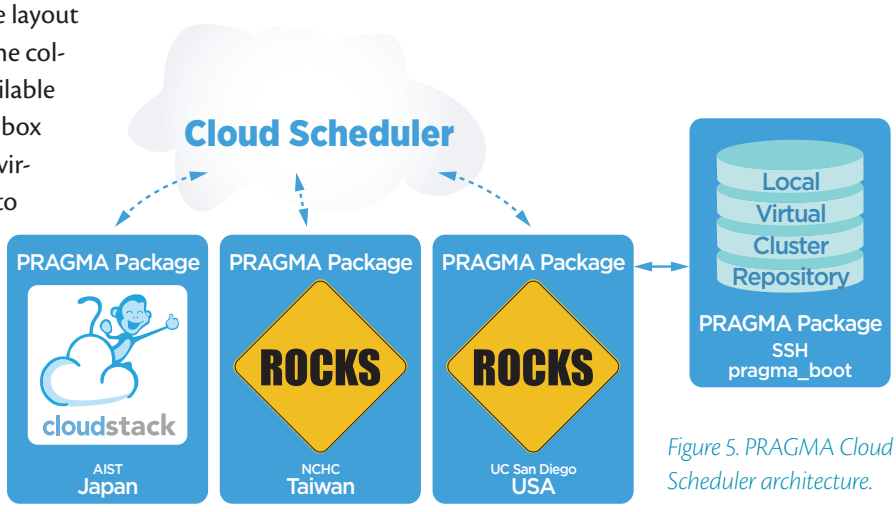


Figure 5. PRAGMA Cloud Scheduler architecture.

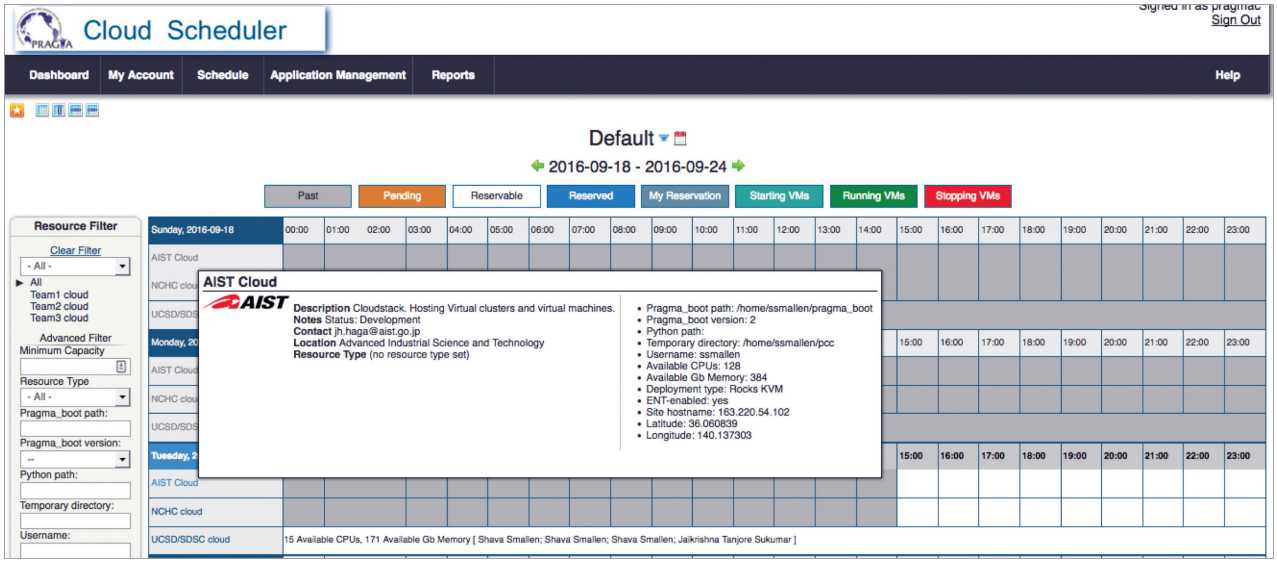


Figure 6. Screenshot of the PRAGMA Cloud Scheduler web interface. Resource availability is shown in a calendar layout with more details shown about the AIST in Japan.

launch pragma boot to deploy the virtual cluster, automatically configure the private and public network configuration on the virtual cluster nodes, and email the user when the re-configured cluster is ready.

Each PRAGMA Cloud Testbed site is able to use their own cloud VM provisioning software (e.g., Rocks) and just needs to install pragma_boot , a set of Python scripts used to launch a virtual cluster, and storage for their local virtual cluster image repository. We first made and demonstrated a prototype implementation during PRAGMA 27 at Indiana University in late 2014. Since then we have continued to develop it and have made it available to users in early 2016 with one available resource at UC San Diego in the US and another resource at the National Center for High-performance Computing (NCHC) in Taiwan. Both resources currently utilize Rocks as the VM provisioning software. In late 2016, we added a Cloudstack resource from the National Institute of Advanced Industrial Science and Technology (AIST) in Japan, which required writing a new Cloudstack driver for pragma_boot. Currently, the available virtual cluster images include basic Rocks clusters with and without Sun Grid Engine (SGE), a BioLinux cluster, and a basic CentOS 7 virtual cluster. More virtual images are under development including one for GRAPLEr (see highlight *Global Approach to Lake Ecology*).

Future work includes integration of PRAGMA-ENT, a central Google drive virtual cluster image repository, automatic virtual cluster image conversion using the Clonezilla tool, and a pragma_boot driver for Openstack.

PARTICIPANTS: *UC San Diego:* Shava Smallen, Nadya Williams, Philip Papadopoulos; *National Institute for Advanced Industrial Science and Technology (AIST):* Jason Haga, Ryousei Takano, Hiroki Ohashi, Yoshio Tanaka; *National Center for High-performance Computing (NCHC):* Weicheng Huang, Steven Shiau, Ceasar Sun

Computational Genomics Infrastructure

Expanding Computational Genomics Infrastructure to Enhance Rice Breeding

The Biosciences Working Group continues to look for new opportunities to provide cyberinfrastructure for proteomic and genomic research. Although virtual screening research to identify potential novel therapeutics continues to be an active area in PRAGMA, our efforts to deploy a platform for genomics analysis has identified rice genomics as a new research area with global importance. Discussions held at PRAGMA 30 in Manila created opportunities to collaborate with the International Rice Research Institute (IRRI). This institute aims to help rice farmers improve the yield and quality of their rice in an environmentally sustainable way. IRRI research results will contribute to its overall mission to reduce poverty and hunger, improve the health of rice farmers and consumers and ensure environmental sustainability through collaborative research, partnerships and the strengthening of national agricultural research and extension systems. Much of the IRRI activities are centered on data acquisition, data standards, data reproducibility and data management for a variety of workflows. Some examples include unique identifiers for germplasm studies, breeding management systems, standardizing phenotypic observations and sharing high-throughput sequencing and genotyping data.

Our deployment of a virtual machine VM with genomics analysis tools, Biolinux8, combined with a user-interface, Galaxy, for the creation of reproducible analysis workflows was modified to include the Tassel-GBS pipeline for rice genomics studies (see Figure 7). This pipeline uses a single nucleotide polymorphism (SNP) genotyping dataset from the Three Thousand Rice Genome (3KRG) project, which, when combined with phenotyping data, enables researchers to do genome-wide association studies (GWAS) to examine the relationship between genetic variants (i.e., SNPs) and traits (i.e., desirable physical characteristics of rice). This analysis framework will not only allow rice researchers to validate experiments and aid in decision making using a common platform, but will also facilitate IRRI collection of results from different researchers using this computational infrastructure.

As a culmination of this effort over the last year, the functionality of this workflow was demonstrated at PRAGMA 31 in Bangkok, Thailand. This genomics workflow was successfully deployed on a portion of the stable PRAGMA Cloud infrastructure, creating a scalable, easy-to-use platform for intensive genomics analyses among IRRI bioscientists. Moreover, this is an important step toward creating workflows that utilize the PRAGMA Data Identity Service and PRAGMA Data Repository (see next highlight) for computational experimental reproducibility and data sharing for long-tail scientists.

We believe this is a critical step toward the goal of developing technologies that help to create a unified workflow for the processing and analysis of large biologically related datasets. This distributed infrastructure and scientific tools can be extended for use by other biomedical scientists in a variety of disciplines and continue to push the requirements of the stable PRAGMA Cloud infrastructure.

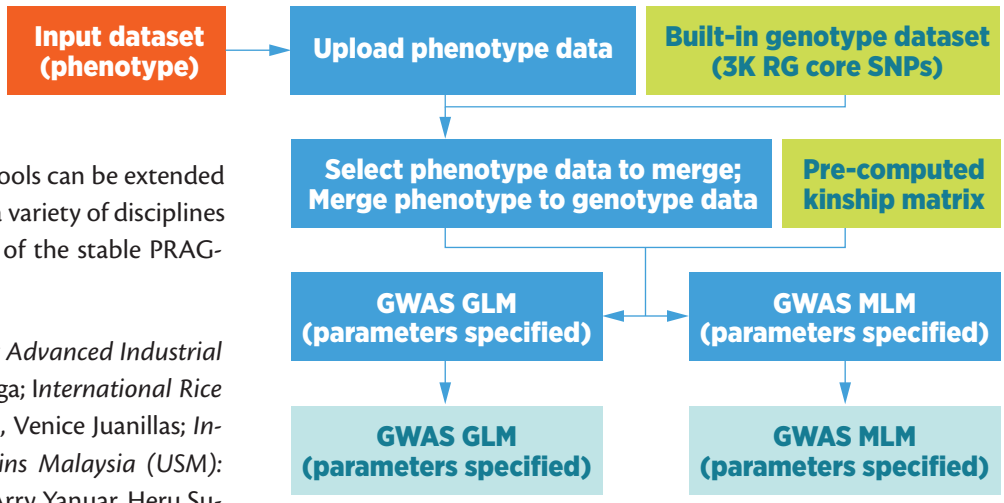


Figure 7. Overview of TASSEL-GBS workflow for rice genomics studies.

PARTICIPANTS: *National Institute for Advanced Industrial Science and Technology (AIST):* Jason Haga; *International Rice Research Institute (IRRI):* Ramil Mauleon, Venice Juanillas; *Indiana Univ:* Gabriel Zhou; *Universiti Sains Malaysia (USM):* Habibah Wahab; *Universitas Indonesia:* Arry Yanuar, Heru Suhartanto

Towards an International Data Sharing Testbed

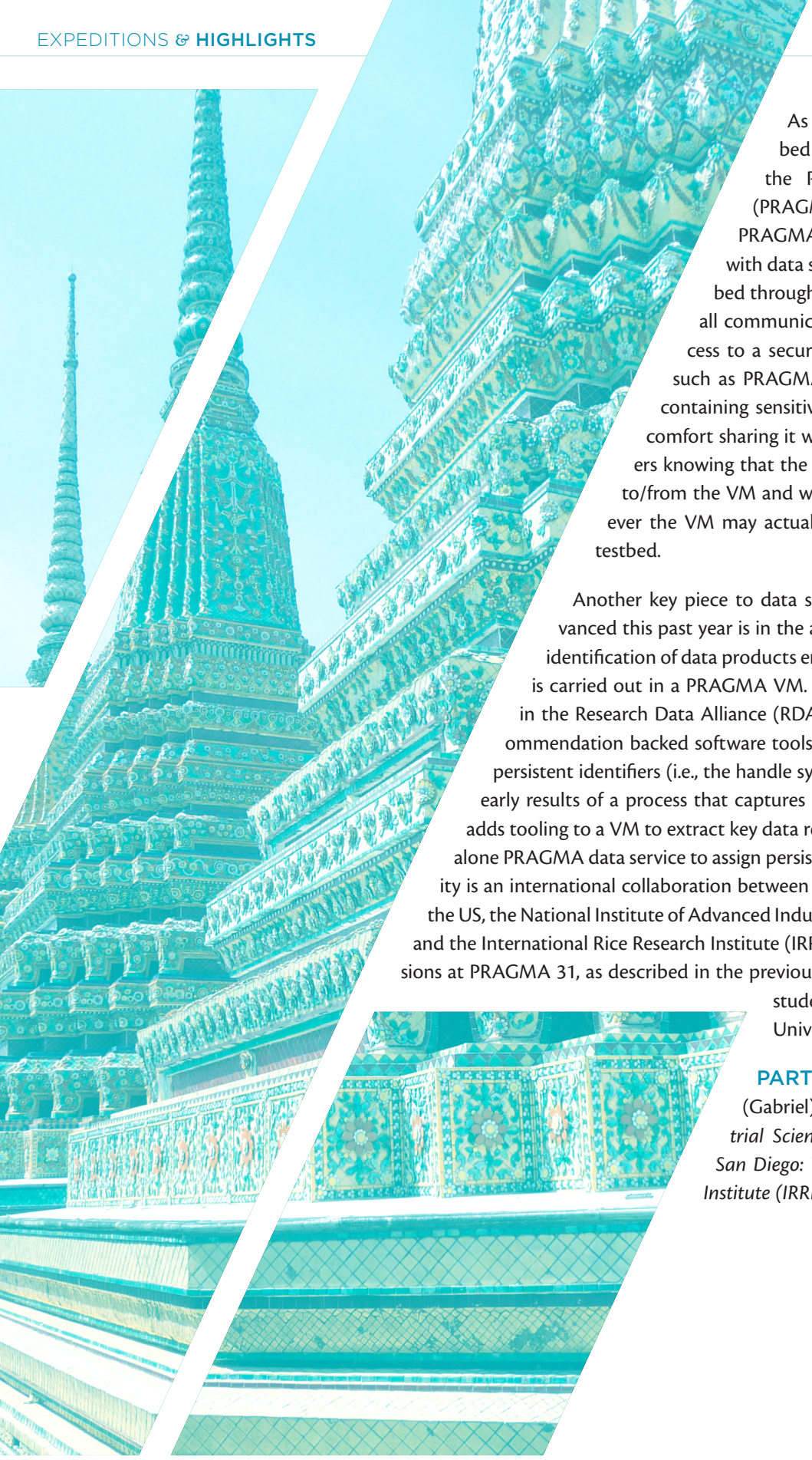
Developing and Implementing the Technological Pieces for Trusted Data Analysis

Digital data is a driving force in scientific research and discovery due, in large part, to recent advances in technology. Drones can be deployed to collect topographical and landcover data in difficult to access terrain, such as the richly biodiverse Mount Kinabalu in Malaysia. Local cell phones, purchased cheaply, can be put into the hands of farmers in drought-stricken regions who then respond to surveys administered through text messages and cloud services. Sequenced rice genomes of the indica rice genome facilitate a wide array of research and annotation. While the sequencing may originate in a country that has rice as a major food product, genomics research can be carried out anywhere.

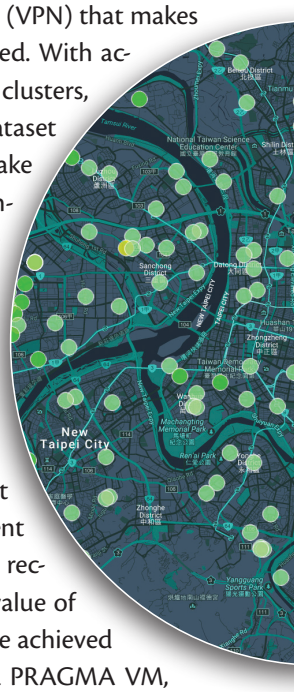
Researchers who collect data for their own research purposes are frequently willing to share the data for other research purposes, but they encounter numerous barriers to doing so. They are faced with daunting questions: *Will my data be properly cared for if I share it? Will I get proper credit for sharing it?* How can I get my data into a usable form (annotation, curation) without too much effort on my part? Over the last year, PRAGMA has made important headway in several pieces of the problem through novel technological solutions that lower the barrier to research data sharing in service of international research discovery. These pieces include advanced networking, data provenance, cloud computing, and, most recently, robust persistent data identification.

The PRAGMA model of computing employs Rocks virtual machines (VMs) to facilitate research that can be started up and run anywhere on the distributed PRAGMA testbed. These VMs are mobile data analysis engines that allow researchers to carry out their analysis anywhere on the PRAGMA testbed. As an analysis engine, a PRAGMA VM is part of the research process. Therefore, we investigated data provenance capture of PRAGMA VMs so that the lineage of the data that is created within the VM is captured and preserved.

Images: (left to right) Shinji Shimojo; Peter Arzberger and Kevin Dong; Renato Figueiredo and Wanida Putthividhya: Handing the PRAGMA Workshop Torch from PRAGMA 31 co-chair to PRAGMA 32 co-chair
—all courtesy of Teri Simas



As a software defined network (SDN) testbed for cyberinfrastructure in the Pacific Rim, the PRAGMA Experimental Network Testbed (PRAGMA-ENT) contributes to secure data analysis. PRAGMA VMs that need to talk with one another or with data sources can do so on the PRAGMA-ENT testbed through a virtual private network (VPN) that makes all communication private and protected. With access to a secure distributed network of clusters, such as PRAGMA-ENT, an author of a dataset containing sensitive biodiversity data can take comfort sharing it with his/her fellow researchers knowing that the data is protected enroute to/from the VM and when inside the VM wherever the VM may actually live on the PRAGMA testbed.



Another key piece to data sharing that was also advanced this past year is in the area of uniform persistent identification of data products emerging from research that is carried out in a PRAGMA VM. Leveraging our engagement in the Research Data Alliance (RDA), we stood up two RDA recommendation backed software tools designed to extend the value of persistent identifiers (i.e., the handle system *www.handle.net*). We achieved early results of a process that captures data provenance within a PRAGMA VM, adds tooling to a VM to extract key data results from the VM and creates a new standalone PRAGMA data service to assign persistent identifiers to key data results. This activity is an international collaboration between Indiana University and the UC San Diego in the US, the National Institute of Advanced Industrial Science and Technology (AIST) in Japan and the International Rice Research Institute (IRRI) in the Philippines and grew out of discussions at PRAGMA 31, as described in the previous highlight The project has funded graduate students and early career researchers at Indiana University and at the IRRI.

PARTICIPANTS: *Indiana Univ:* Beth Plale, Quan (Gabriel) Zhou; *National Institute of Advanced Industrial Science and Technology (AIST):* Jason Haga; *UC San Diego:* Nadya Williams; *International Rice Research Institute (IRRI):* Ramil Mauleon, Venice Juanillas

Air Quality Testbed

Expanding a Grassroots Urban Air Quality Monitoring Testbed

Air pollution is a crucial concern in modern life. It has been proven that fine particulate matter measuring less than 2.5 micrometers, i.e., PM2.5, in diameter is particularly harmful to health and can be lethal in the extreme. How to effectively monitor PM 2.5 in the air, especially in urban areas, remains a challenge. One way to tackle the problem is through crowdsourcing. However, most data collected from the crowd is not useful due to its poor and unstable quality. A low-cost automatic and connected sensor device that can be distributed to layman with a shared platform is another possible answer that is actively being investigated. Academia Sinica of Taiwan, Taipei and Taichung city governments; Osaka University of Japan; University of Technology and Design of Singapore; National Chung-Hsin University and National Center for High-performance Computing are all jointly working to develop a solution utilizing Airbox sensors, location-aware sensor system (LASS) infrastructure and LoRa technology for low-power, long-range data communications (see Figure 8). Their research will incorporate the study and prediction of air pollutant behaviors in the future using climate simulation modeling.

This work was initiated after the kickoff CENTRA meeting (see *CENTRA* highlight) and was pursued in the PRAGMA Telescience Working Group. At the PRAGMA 31 workshop, sensors were distributed for deployment in Thailand and the United States (Indiana University, University of Florida and the UC San Diego).

PARTICIPANTS: *National Center for High-performance Computing (NCHC):* Fang-Pang Lin; *Osaka Univ:* Shinji Shimojo, Susumu Date, Yoshiyuki Kido; *Univ Florida:* Jose Fortes; *Academia Sinica:* Ling-Jyh Chen

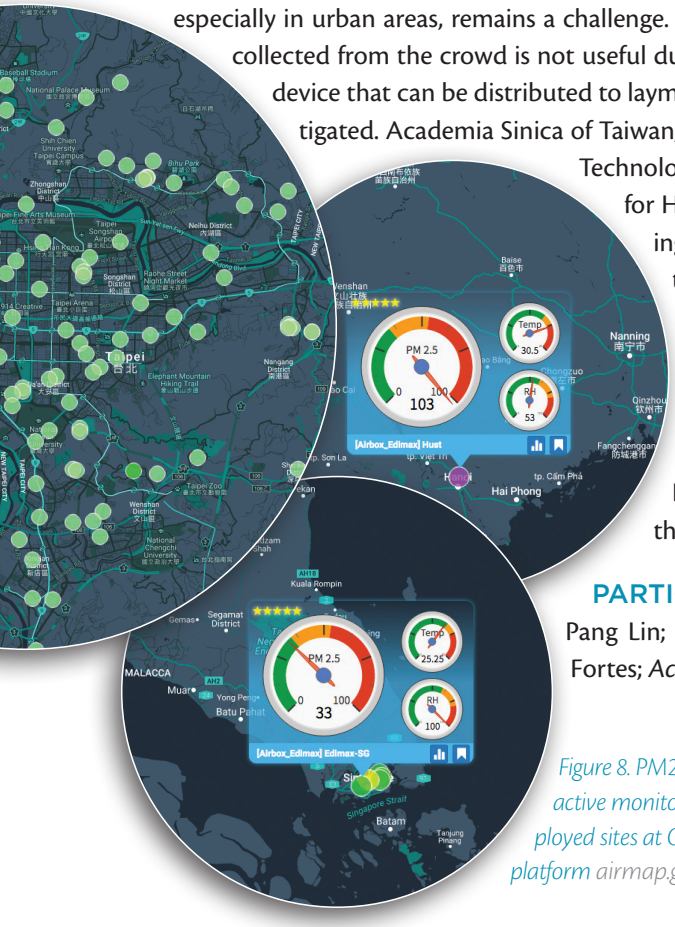


Figure 8. PM2.5 monitoring using Airbox and LASS. Left hand side is the active monitoring site at Taipei City and right hand side is the newly deployed sites at Osaka and Singapore. The data is visualized on g0v.tw shared platform airmap.g0v.asper.tw.

Visualization Environment

Prototyping a Collaborative Visualization Environment for Disaster Management

These days, visualization plays a greater role in the better understanding of data in scientific research. The proliferation of scientific data, often obtained through technological advancements in sensing devices, and its analysis by high-performance computing has accelerated the tendency to display and analyze information visually. Furthermore, aggregation of scientists' expertise and knowledge from a diverse range of scientific domains has become increasingly essential to today's scientific research. From such perspective, we have been working on the research and development of an easy-to-use portal system that allows scientists to form a collaboration environment that leverages tiled display systems but does not force them to have a higher degree of knowledge, technique or know-how.

Figure 9. A collaborative visualization Web-based environment based on SAGE-2.

The Telescience Working Group also pursued an activity in collaborative visualization for disaster management. Discussions held at PRAGMA 30 in Manila and later reinforced at the CENTRA kickoff meeting in March 2016 (see *CENTRA* highlight) resulted in steps toward a resilient software-defined infrastructure to support disaster management applications. Specifically, this involves connecting multiple tiled displays between organizations using PRAGMA-ENT and software-defined network concepts to support a disaster management visualization tool. This tool is designed to provide a continuous, uninterrupted level of information to decision makers during a natural disaster, thus enabling them to take informed actions. Currently, this effort has linked several Japanese institutions, and there are plans to expand the connections to other international sites in PRAGMA, in order to test the resiliency in a more global scenario.

PARTICIPANTS: *National Institute of Advanced Industrial Science and Technology (AIST):* Jason Haga; *Osaka Univ:* Susumu Date, Yoshiyuki Kido; *Univ Tsukuba:* Hirotake Abe; *NARA Institute of Science and Technology (NAIST):* Kohei Ichikawa, Yasuhiro Watashiba; *National Center for High-performance Computing:* Whey-Fone Tsai, Jo-Yu Chang, Chen-Kai Sun, Weicheng Huang; *Univ Hawaii:* Jason Leigh; *National Chung Hsing Univ:* Ming-Der Yang

EDISON

Advancing Computational Science and Engineering Learning through Public Release of the Cyber-Learning Platform EDISON

EDISON (EDucation-research-industry Integration through Simulation On the Net) is an innovative cyber-learning platform developed by the Korea Institute of Science and Technology Information (KISTI). Through the EDISON platform, a variety of high-performance computing (HPC) simulation programs (called Science Apps) are distributed and accessed online by researchers and students in Korea and in the PRAGMA community. EDISON, funded by the Ministry of Science, Technology and Future Planning of Korea, is a multi-year joint research project between KISTI and several domain-specific research clusters, each consisting of academic institutions; KISTI is responsible for developing and managing the platform while the specialized groups supply the Science Apps for the platform. Currently, the groups represent five different computational science and engineering (CSE) domains: Computational Fluid Dynamics (CFD), Nano Physics (Nano), Computational Chemistry (Chem), Computational Structural Dynamics (CSD) and Computer-aided Optimal Design (Design). EDISON will continue to expand by hosting another new field, Biology, later this year and supporting one more yet-to-be-decided discipline next year.

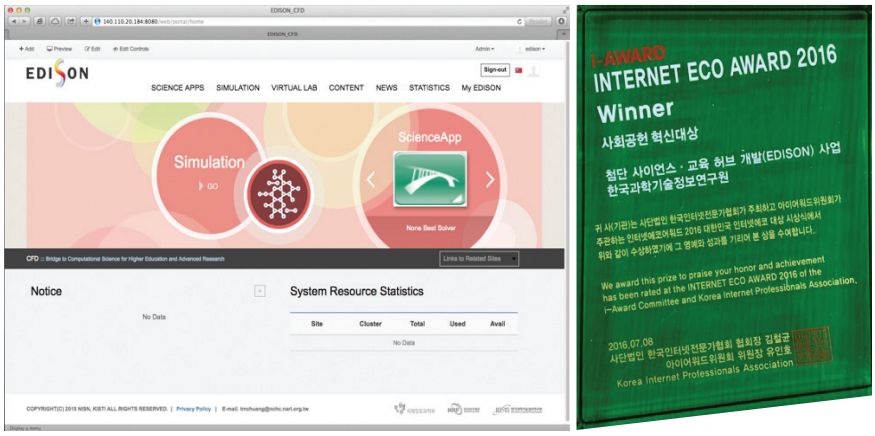


Figure 10. (left) The EDISON portal site deployed at NCHC, (right) iEco Award Korea.

The EDISON platform is widely used and is technically robust enough to serve the chosen Computational Science and Engineering community in Korea. Over the past five years, the system has provided a total of 349 Science Apps and 546 educational contents, servicing around 40K domestic and international users (cumulative). The growing popularity and technical excellence of the platform has enhanced the international recognition of the EDISON system, resulting in several breakthroughs: 1) 2013 International Data Corporation High-Performance Computing (IDC HPC) Innovation and ROI Excellence Award at SC 2013 in Denver, Colorado (USA) (first-ever for Korea); 2) copy site deployment at one of the PRAGMA participants, National Center for High-performance Computing (NCHC) in Taiwan, January 2016; 3) Best Poster Award at the Institute of Electrical and Electronics Engineers and Association for Computing Machinery conference on Cluster Computing and the Grid (ACM/IEEE CCGrid), May 2016, held in Cartagena, Colombia; and 4) Internet Eco Award Korea on innovation for a society contribution, July 2016. In particular, KISTI and NCHC will launch a joint lab for more collaboration at the annual Korea SuperComputing conference (KSC) to be held in October 2016.

The EDISON system is free and open source. Recently, KISTI made an official release of the EDISON framework on GitHub: *github.com/sp-edison/edison-2016*. A total of 21 portlet project directories are available on this release site. The portlet directories include all the technical components needed to run and operate the EDISON platform. In addition, the site provides users with a variety of user guides for platform installation, general use, Science App development and portal management. These guides will help users become quickly familiar with the system. The source code of the platform is also available at the EDISON project website (*www.edison.re.kr/project-download*). EDISON is licensed under the GNU Lesser General Public License. Please contact *edison@kisti.re.kr* for possible collaboration opportunities.

PARTICIPANTS: *Korea Institute for Science and Technology Information (KISTI):* Ruth Lee (Lead), Young-Kyoon Suh (Coordinator)

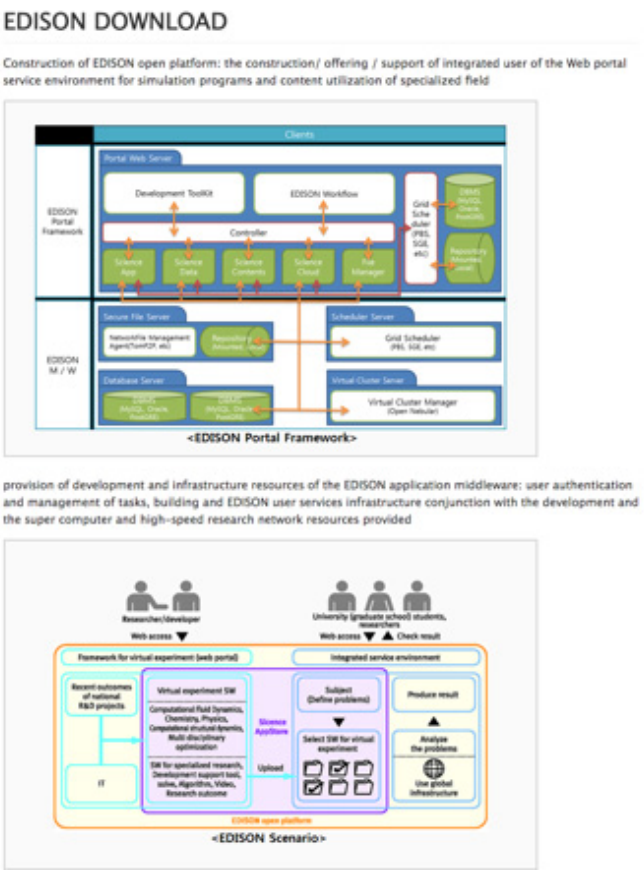


Figure 11. The official release site of EDISON on GitHub.

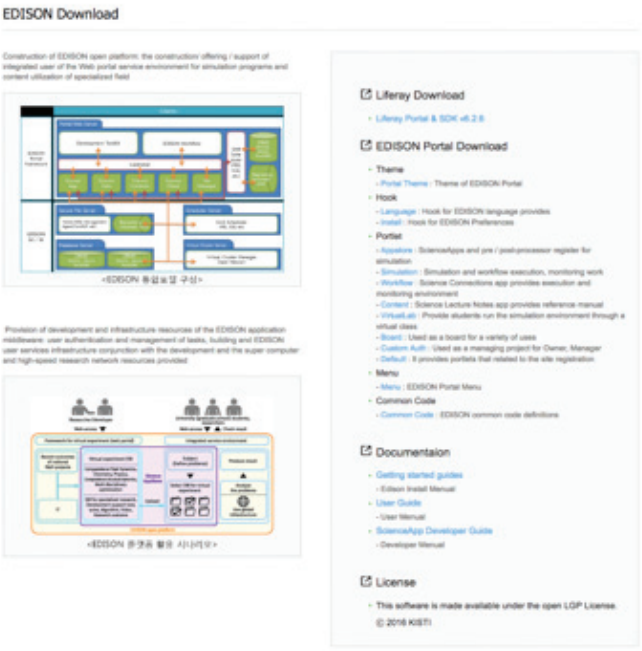


Figure 12. The project download page at the EDISON site.

PRAGMA is committed to engaging new communities of researchers. To do this, we focus on two broad efforts: One is to engage students in PRAGMA activities; the other is to work with communities of researchers to broaden the impact of the work we do and to leverage the students' various interests. Both of these activities enrich PRAGMA by infusing ideas, questions and energy into our workshops and activities between workshops.

STUDENTS

In this section we will talk about two specific activities to engage students: PRAGMA Students and MURPA/QURPA

PRAGMA Students

Building a Future Generation of International Researchers

PRAGMA Students, an organization formed in 2012, aims to help students gain opportunities for professional experiences within PRAGMA's trusted social and technical networks. As a student organization inside PRAGMA, the group is led by a student committee and advised by senior PRAGMA researchers.

To date, activities of PRAGMA Students have included organizing PRAGMA-affiliated student workshop and poster sessions as part of the biannual PRAGMA Workshops; hosting online seminars with lecturers drawn from the broad PRAGMA community; and developing a unique model to provide multiple opportunities for students to participate in PRAGMA's collaborative scientific research. PRAGMA provides a trusted network of people and opportunities in leadership that help students gain valuable professional experience.

Members of PRAGMA Students and the PRAGMA community enjoy many advantages and opportunities. These opportunities include, but are not limited to, the following:

1. Participating in workshops and conferences in international settings, which stimulates inspiration, information sharing and collaboration;
2. Working with mentors and advisors throughout PRAGMA on challenging, state-of-the-art research projects, which strengthens research skills and ability for scientific exploration;
3. Experiencing short-term residential research opportunities at other PRAGMA sites, which provide both scientific and cultural experiences.

Figure 13. (left) Instructors Nadya Williams and Shava Smallen with one of the Student Hackathon teams; in background instructor Aimee Stewart with another team; (right) Screenshot of map visualization tool for the PRAGMA Cloud Testbed developed during student hackathon.

Image: PRAGMA Student Leadership: left to right: Pongsakorn U-chupala, Quan (Gabriel) Zhou, Meilan Jiang, Chawanat Nakasan—courtesy of Teri Simas

For the January 2016 PRAGMA 30 workshop in Manila, Philippines, PRAGMA Students organized the poster session for the entire meeting. Members of PRAGMA Students committee reviewed and selected posters from all workshop participants. Around 20 posters were carefully reviewed, and detailed review feedback was provided to the submitters.

To increase interaction of PRAGMA 30 participants with the poster presenters, PRAGMA Students devised a voting system that allowed each workshop participate to identify the "best" posters by affixing stickers to posters. This voting process stimulated the attendees to be more involved in the poster presenter's work, provide feedback and suggest potential research collaborations. Furthermore, the top four posters (based on the number of stickers) were awarded prizes, and their authors were given another opportunity to present their work in the main PRAGMA workshop. The four poster winners hailed from institutes in the US, Japan and Philippines.

In addition, PRAGMA Students organized a "lightening talk" pre-workshop session that gave students the opportunity get feedback on their work and gain experience in presenting research ideas and results. Ten of these students were then invited to give short presentations regarding their respective research posters, based on their innovative and impactful projects.

At PRAGMA 31 in Bangkok, PRAGMA Students continued their contributions to the PRAGMA workshop by organizing student sessions, poster sessions and lightning talks. Particularly noteworthy at PRAGMA 31 was the Student Hackathon. Four teams of students from Kasetsart and Thammasat Universities tackled three challenges provided by the Hackaton organizers (Putchong Uthyopas, Wanida Putthividhya and instructors Nadya Williams, Shava Smallen and Aimee Stewart). The students demonstrated the exceptional capability to address the challenges posed by the tasks, learned the materials and produced outcomes that benefit PRAGMA. The Hackathon topics included: 1) Extension of the existing PRAGMA Cloud Testbed site; 2) Applications that add more visibility and analytical ability to PRAGMA Cloud Testbed services (*fiji.rocksclusters.org/cloud-scheduler/testbed-map*); and (3) Applications that search, query, discover or generally mine biodiversity data through Lifemapper APIs. Additional information associated with these topics is available at: *www.pragma-grid.net/hackathon.*

Outside of the workshop venue, some students in PRAGMA have short-term residential research opportunities. Quan (Gabriel) Zhou, a Ph.D. student from Indiana University Bloomington, was a visiting scholar at The National Institute of Advanced Industrial Science and Technology (AIST) in Tsukuba, Japan from June to July 2016. Zhou worked on the Research Data Alliance MacArthur Adoption project with collaborations between AIST, Japan and the International Rice Research Institute (IRRI), Philippines (see highlight *Computational Genomics Infrastructure*). Additionally, five undergraduate and graduate students from Mahidol University, Thailand also worked at AIST as summer interns on several cutting-edge topics, including augmented reality (AR) and networking data plane optimization. These short-term pro-

grams serve as examples of collaborations between PRAGMA institutions that strengthen research by involving mentors and advisors from multiple disciplines and multiple institutions. In addition to academic and cultural benefits, the students worked closely with host professors and senior researchers and have built many lasting connections. The students from the residential research program will also have an opportunity to present their work and seek comments from senior researchers during a PRAGMA workshop. Wassapon Watanakeesuntorn, an undergraduate student from Kasetsart University (Thailand), also carried out his internship in July to August 2016 at Nara Institute of Science and Technology (NAIST) in Japan. He worked on visualization of OpenFlow control messages and flow entries, which is expected to facilitate examination and debugging in various OpenFlow networks, including PRAGMA-ENT itself. He presented his work during the demo and poster presentation session at the PRAGMA 31 workshop..

CURRENT LEADERS OF PRAGMA STUDENTS: *Indiana Univ:* Quan (Gabriel) Zhou (chair); *Konkuk Univ:* Meilan Jiang (co-chair); *Nara Institute of Science and Technology (NAIST):* Pongsakorn U-chupala (co-chair), Chawanat Nakasan (co-chair)

PRAGMA STUDENT ADVISORS: *Indiana Univ:* Beth Plale; *Konkuk Univ:* Karpjoo Jeong; *Kasetsart Univ:* Putchong Uthayopas

MURPA

International Research Internships at the University of Queensland and Monash University

Image: PRAGMA Program Manager Teri Simas (center) with MURPA Students Lachlan Smith, Tyson Jones, William Donohoe, and Lucina Togno

In today's educational arena, universities must provide students with opportunities to work and study abroad to prepare them for global citizenship and professional competence in a multi-cultural workplace. Numerous reports have challenged universities to develop educational programs that provide an integrated academic basis for developing students' cultural/global competencies.

Since 2008, 42 Monash University students have travelled to international partner sites under the Monash Undergraduate Research Projects Abroad program (MURPA). They have travelled to the University of California, San Diego, the National Center for Supercomputing Research in Illinois, The Technion in Israel, the Institute for Infocomm Research (I2R) in Singapore and the University of Warwick. Students are placed for a period of 8 weeks, allowing them to integrate into the research groups as team members. Students have a local mentor in Australia as well as one in the remote site, and often bridge international research projects. In 2015/16 four Monash students and one University of Queensland student in the Queensland University Research Program Abroad (QURPA) program travelled to UC San Diego and two to Warwick University in the UK. Two examples of their projects are highlighted.

Tyson Ray Jones from Monash University joined Professor Frank Wuerthwein at the Department of Physics, UC San Diego in 2016. The project undertaken by Tyson concerned developing new us-

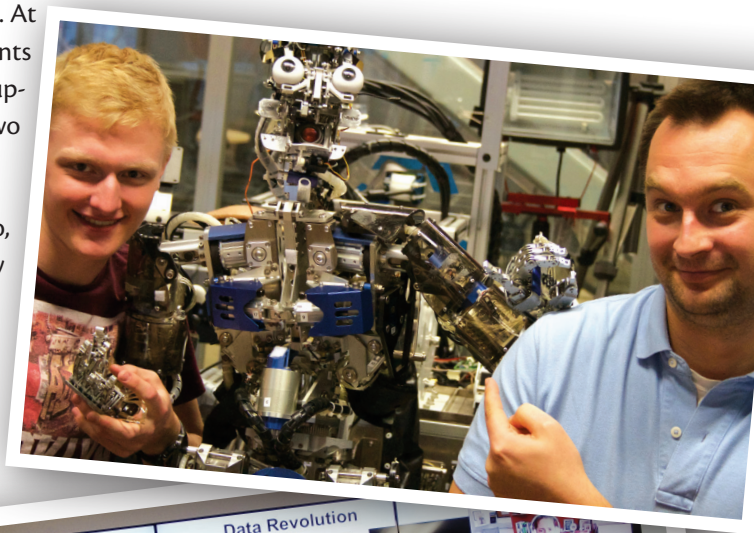
Figure 14. Roger Riordan (Sponsor from the Cybec Foundation), with MURPA 2015/2016 students: Lucina Togno, Tyson Jones, William Donohoe, Joshua Collins, Tuan Tran, and Lachlan Smith, and MURPA organizers David Abramson and Maria Indrawan-Santiago.

age and performance monitors to support the expensive roll-out of jobs to the global queues supporting the computational infrastructure of LHC (Large Hadron Collider), which is a geographically distributed, high-throughput supercomputing network, primarily used for large simulation and data analysis pertaining to particle physics coming out of the LHC project. Specifically, the tools Tyson built during his MURPA project involves Telegraf, python daemons, InfluxDB and Grafana to collect and display usage metrics of the UCLHC Condor and XRootD systems.

Joshua Riddell, a UQ Mechatronic Engineering student, spent four weeks at the UC San Diego, from early January to early February 2016. He worked on the child-friendly robot "Rubi" (see Figure 15) as part of a QURPA placement in a team with UQ PhD candidates Kristyn Hensby (developmental psychology) and Nikodem Rybak (software engineering) sponsored by the Temporal Dynamics of Learning Center. The team programmed RUBI to build simple toys, demonstarating the steps required for young children to copy. He programmed an easy to use interface for scheduling and running speech utterances and changing facial expressions in synchronization with RUBI's physical movements. He also strengthened RUBI's arm joints to suit the new mechanical challenge this experiment presented.

Josh wrote about his trip. "The highlight of the trip to UC San Diego was definitely meeting world-class researchers in their labs, as well as hearing their talks at the last Temporal Dynamics of Learning Center All Hands Meeting. As I'm a mechatronics engineering student, Mike Tolley's soft robotics lab was particularly intriguing…I would definitely recommend the placement to other UQ students."

MURPA and QURPA involve an advanced seminar scheme, in which students can attend virtual seminars given by world leading experts before they leave. These seminars also allow students to "meet" potential UC San Diego mentors and gather information regarding potential projects. In 2016 the seminar series focused on an all female cast of presenters. At the time of writing this report, six students have been selected at Monash for the upcoming MURPA 2016/2017 round, and two UQ students are yet to be chosen.



Seminars were sourced from UC San Diego, Intel, University of Delaware, the University of Notre Dame, the University of Southern California, the University of Chicago, the University of Colorado Boulder, Brown University, the University of Tennessee, the University of Utah and SDSC. As done in the past, seminars were broadcast simultaneously to Monash (in Melbourne) and UQ (in Brisbane), with audiences able to ask questions from either venue. The seminar infrastructure supports a wide range of video conference technologies (both open source and commercial), and is displayed on a 20 MPixel OptiPortal (see Figure 16).

Figure 15 (above). QURPA-funded undergraduate engineering student Josh Riddell (left) and UQ PhD student Nikodem Rybak (right) with a new friend at UC San Diego—courtesy of Kristyn Hensby.

Figure 16 (right). The view of a MURPA seminar at the University of Queensland. The presenter is Dr. Rommie Amaro from UC San Diego.

BUILDING COMMUNITY

A fundamental component of PRAGMA's activities is to engage new ideas and people, in part through PRAGMA Workshops and PRAGMA Students. In addition, we take intentional steps to co-organize other meetings and new activities that focus on topics of keen interest to PRAGMA and to the individuals and communities we seek to engage. This year we highlight four activities:

- The launch of a new organization to enable transnational cyberinfrastructure applications;
- Publication of PRAGMA accomplishments for broader dissemination;
- The establishment of new partnerships in Southeast Asia through training and sharing advances on big data and applications to disaster management; and
- The initiation of discussions to establish a central data repository for the Lower Mekong Region.

Image: At Rizal Monument, Manila, after PRAGMA 30

Bringing Together International Research Teams to Advance Transnational Software-Defined Cyberinfrastructure Solutions to Global Problems

Figure 17. Participants at the 1st CENTRA Workshop held March 2016 in Taipei, Taiwan.

CENTRA stands for Collaborations to Enable Transnational Cyberinfrastructure Applications and refers to a partnership and evolving framework for collaborations amongst research centers, institutes and laboratories across the world. The CENTRA concept and partnerships are the result of PRAGMA activities in building collaboration networks and identifying opportunities to advance cyberinfrastructure research in the Pacific Rim. The founding members of CENTRA are the Advanced Computing and Information Systems (ACIS) Laboratory of the University of Florida with support from the National Science Foundation (NSF) of the USA, the Center of Excellence for Cyber-Enablement of Applications (CECEA) of the National Center for High-performance Computing (NCHC) of Taiwan with support from the Ministry of Science and Technology (MOST) of Taiwan, and the ASEAN International Virtual Organization (IVO) funded by the National Institute of Information and Communications Technologies (NICT), Japan.

CENTRA's mission is to

- Harness software-defined technologies to address transnational societal needs in domains that include, but are not limited to, disaster management, environmental modeling and smart and connected communities;
- Advance the science needed to design and use these technologies to build effective and efficient transnational IT systems;
- Train the next generation of junior researchers and innovators to work in transnational settings through
  - Convening experts in workshops to identify solutions to key problems;
  - Providing junior researchers an immersive experience in collaborative, multidisciplinary teams that address the transnational problems (and provide them opportunities for entrepreneurial experiences);
  - Facilitating the use and development of prototypes and testbeds to demonstrate solutions.

CENTRA engages international entities in collaborative research activities in several ways. International team formation around specific project ideas are enabled by CENTRA's funding of travel to meetings organized by CENTRA and other entities. In addition to annual meetings, CENTRA takes advantage of other conferences, such as PRAGMA 31 and SEAIP in 2016 to advance specific projects, disseminate information about CENTRA and generally engage more researchers in CENTRA's philosophy and activities. Also funded are face-to-face interactions by CENTRA project team members via short visits to other members' sites. Institutional members provide access to facilities and testbeds to host researchers and enable experimental work. Individual researchers who belong to member institutions, or are invited by these institutions, can participate in CENTRA projects.

The 2016 CENTRA Kickoff Meeting took place in Taiwan from March 30 through April 1, 2016 (Figure 17). The Kickoff was opened by Dr. Chung-Liang Chien, Deputy Minister of the Ministry of Science and Technology, Taiwan (MOST) and Dr. Tzi-Dar Chiueh, Vice President of National Applied Research Laboratories (NARLabs). More than 80 participants from 28 institutions from the United States, Japan, Taiwan, Singapore, Philippines, Indonesia,

and Malaysia attended the three-day Kickoff and formed collaboration groups in the following topic areas: Smart Cities and Connected Communities, Environmental Monitoring and Disaster Management, Biodiversity and Social Communication IT, and Federated Testbeds.

The CENTRA inaugural activities also included a visit to the Central Taiwan Science Park (CTSP) where CENTRA researchers were hosted by Director Wayne Wang and his team at the CTSP Bureau. CENTRA researchers introduced collaborative cyber-infrastructure research projects initiated by CENTRA members at and beyond the Kickoff meeting. Director Wang expressed interest in collaborating with CENTRA researchers in the near future and provided them with access to CTSP activities and facilities.

Subsequent to the inaugural activity, CENTRA hosted a webinar on the air monitoring system AirBox titled "AirBox: a participatory ecosystem of PM2.5 monitoring" (see highlight *Air Quality Testbed*), which was given by Professor Ling-Jyh Chen of Academia Sinica (see Figure 18). The webinar was attended by more than 30 people.

At PRAGMA 31, there was a two day session organized by CENTRA members the ICT Virtual Organization of ASEAN Institutes and NICT (ASEAN IVO) and MIMOS. The session focused on software-defined systems on disaster mitigation and smart cities. It was well attended, with participants from Japan, Malaysia, Myanmar, Philippines, Singapore, Thailand and Vietnam.

In addition to playing a key role in the conceptualization and development of the CENTRA project and partnerships, PRAGMA is an Invited Organization of CENTRA and plays an active role via a PRAGMA representative in CENTRA's Steering Committee. In addition, PRAGMA scientists are involved in several CENTRA projects and other joint activities (see highlights *Air Quality Testbed* and *Visualization Environment*).

The next CENTRA Meeting will be hosted by the ACIS Lab, University of Florida in Gainesville, Florida on April 10-12, 2017. In conjunction with this CENTRA Meeting, the PRAGMA 32 Workshop will be held at the same venue on April 13-14, 2017.

More information about CENTRA can be found at *www.globalcentra.org*.

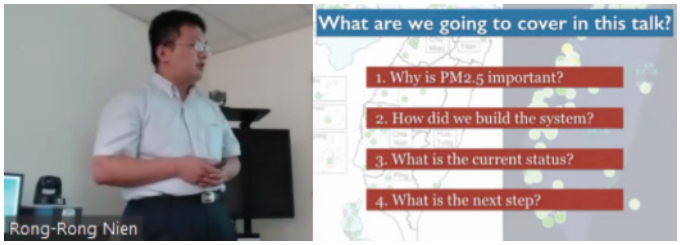


Figure 18. Virtual Presentation by Professor Ling-Jyh Chen

CENTRA CO-LEADS: *Advanced Computing and Information Systems (ACIS) laboratory, Univ Florida:* Jose Fortes; *Center of Excellence for Cyber-Enablement of Applications (CECEA), National Center for High-performance Computing (NCHC):* Fang-Pang Lin; *ASEAN International Virtual Organization (IVO) of the National Institute of Information and Communications Technologies:* Shinji Shimojo

Images: (top) Ta Prohm Temple, Siem Reap, Cambodia—courtesy of Peter Arzberger; (left) Beth Plale and Yoshio Tanaka; (right) A small group of researchers at PRAGMA 31— courtesy of Teri Simas

Disseminating PRAGMA's Technologies

In conjunction with the PRAGMA 29 Workshop in October 2015, PRAGMA organized the first International Clouds for Data Science (PRAGMA ICDS'15), a research-oriented workshop drawing from the broad PRAGMA community. PRAGMA ICDS'15 leveraged PRAGMA's long history of experimentation to showcase the latest research on the design, implementation, evaluation and use of cloud technology, networking and data management, which enable new forms of research that span international boundaries.

Seven papers from the conference were selected for publication in a special issue of the journal *Concurrency and Computation: Practice Experience (CCPE)*. The papers fall into three broad themes: research that takes advantage of the research networks and connected local clusters that connect the PRAGMA institutions; research inspired by the science expeditions defined as part of PRAGMA; and location-specific solutions that came about because of the long-standing and nurturing collaborative community that PRAGMA has built and fostered over a long period of time.

These papers are being edited for publication and are going to be available at the CCPE page: *www.cc-pe.net/journalinfo*. The list of titles and authors are as follows:

- "Lightweight Scheduling for the PRAGMA Cloud Testbed" by S. Smallen, N. Williams, and P. Papadopoulos.
- "A Simple Multipath OpenFlow Controller Using Topology-based Algorithms for Multipath TCP" by C. Nakasan, K. Ichikawa, H. Lida, and P. Uthyopas.
- "PRAGMA-ENT: Exposing SDN Concepts to Domain Scientists in the Pacific Rim" by K. Ichikawa, P. U-chupala, C. Huang, C. Nakasan, T-L Liu, J-Y Chang, L-C Ku, W-F Tsai, J. Haga, H. Yamanaka, E. Kawai, Y. Kid, S. Date, S. Shimojo, P. Papadopoulos, M. Tsugawa, M. Collins, K. Jeong, R. Figueiredo, and J. Fortes.
- "GRAPLEr: A Distributed Collaborative Environment for Lake Ecosystem Modeling that Integrates Overlay Networks, High-throughput Computing and Web Services" by K.C. Subratie, S. Aditya, S. Mahesula, R. Figueiredo, C.C. Carey, and P.C. Hanson.
- "Virtualizing Software Infrastructure for Biodiversity Expedition" by N. Williams, A. Stewart, and P. Papadopoulos.
- "In-Place Query-Driven Big Data Platform: Applications to Post Processing of Environmental Monitoring" by J-G Lee, W-F Tsai, L-C Lee, C-Y Lin, H-C Lin, and B-J Tsuang.
- "Interactive Museum Exhibits with Microcontrollers: A Use-Case Scenario" by L. Wong, S. Shimojo, Y. Teranishi, T. Yoshihisa, and J.H. Haga.
- "A Reflection on the Origins, Evolution and Future of PRAGMA" by P. Arzberger.

EDITORS FOR THE SPECIAL ISSUE: Beth Plale, *Indiana Univ,* Renato Figueiredo, *Univ Florida,* and Miao Chen, *Indiana Univ*

Engaging New Southeast Asia Members

Southeast Asia International Research and Training Program (SEAIP) and PRAGMA Institute: Taichung and Keelung December 7–11, 2015

The SEAIP's ongoing series of workshops (see *http://seaip.narlabs.org.tw/*) has opened doors for collaborations between researchers in Southeast Asia and the rest of the world and has formed the basis for growing PRAGMA collaborations in Southeast Asia. The theme of the 11th SEAIP workshop, held December 7–11, 2015, was "Cloud Computing and the Internet of Things." Organized by the National Center for High-performance Computing (NCHC), with funding from the Ministry of Science and Technology in Taiwan as well as from the National Applied Research Laboratory, the atmosphere of the workshop was informal and conducive to learning and developing collaborations. In addition to sessions at NCHC's Taichung branch, there was a session held at the National Museum of Marine Science & Technology in Keelung. The 2015 workshop also was used to plan other activities, including laying the groundwork for US-EA CENTRA.

The next SEAIP meeting will be held in Tainan and Kenting, Taiwan, December 5–9, 2016.

Chair of SEAIP: National Center for High-performance Computing (NCHC): Fang-Pang Lin.

Image: UC San Diego. Visit by USM researchers Habibah A. Wahab, Ezatul Ezleen, and Nornisah Mohamed—courtesy of Teri Simas

An Opportunity for PRAGMA and CENTRA

Establishing a Central Research Database for the Lower Mekong Region

At PRAGMA 31 researchers from the Lower Mekong Region, PRAGMA and CENTRA met to discuss the need for data sharing, as well as possible approaches to creating a capable and accessible database to support data sharing. The data under discussion covered areas of hydrology, biodiversity (indicator species), land use coverage and population distribution.

Some of the needs for data sharing that were discussed included the following concepts:

- The Mekong conditions, dynamics, and systems are rapidly changing due development and climate change;
  - Research is increasingly focusing at the nexus of domains and systems, and thus data need to be shared across domains and systems;
  - Researchers want to share data to ultimately return value to the Mekong;
  - Sharing data helps with awareness of what else is happening and could lead to better coordination of efforts.

While past efforts attempted to establish a central data repository, a key challenge in bringing data "owners" to the table is the lack of control over the data after it is shared.

With respect to approaches for a capable and accessible database, several key attributes were discussed:

- The architecture would be distributed, with the goal of bringing the database to data instead of sending data to the database. This approach would allow the data to be kept at the home institution, enabling complete control over the data by the owners;
- Workflows would need to support custom protection control over datasets;
- The architecture supports a bottom-up approach (e.g., crowd-source data) by providing data-sharing values to the contributors while addressing "ownership" concerns; and
- Tools and processes are need to be in place to support data curation.

Image: Participants in the Workshop on Lower Mekong Region —courtesy of Tho Nguyen

The Lower Mekong region is a critical geographic area for Southeast Asia that offers an opportunity for future PRAGMA and CENTRA collaborations to help make progress toward the ultimate goal of creating a data resource for the region.

ORGANIZERS OF THE WORKSHOP: *Univ Virginia:* Tho Nguyen; *Univ Washington:* Gordon Holtgrieve

WORKSHOPS

PRAGMA STUDENT POSTER AWARDS

PRAGMA 30

The Best Poster Award

Rhia Trogo-Pantola, Jed Barry Ebardaloza, Delfin Jay Sabido IX, Edgardo Tongson, Gerry Bagtasa, Orlando Balderama. IBM. Enabling Philippine Farmers to Adapt to Climate Variability Using Seasonal Climate and Weather Forecast with a Crop Simulation Model in an SMS-based Farmer Decision Support System

The Best Poster Award, 2nd Place

Lemuel Clark P. Velasco, Joshua Dave E. Gomora, Leonard N. Aguanta, Joanna Marie D. Acibar, John Paul S. Postrano, Lyka Cristina T. Diaz, Leo Anthony T. Catane. Mindanao State University - Iligan Institute of Technology. Data Visualization of Schistosomiasis Community Health Data Using Google Maps and Google Charts

The Best Poster Award, 2nd Place

Pongsakorn U-chupala, Yasuhiro Watashiba, Kohei Ichikawa, Susumu Date. Nara Institute of Science and Technology and Osaka University. Practicality and Feasibility of Improving Linux Container Utilization with Task Rebalancing Strategy

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 perspectives, applications, technologies, students, and resources to these events. These workshops are an essential opportunity to demonstrate progress on projects and to create action plans for accomplishing tasks prior to the subsequent workshop.

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.

Working Groups

The workshops are organized according to the activities of 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 (National Institute of Advanced Industrial Science and Technology [AIST]) and Philip Papadopoulos (UC San Diego).
- **TELESCIENCE WORKING GROUP:** Focusing on a variety of activities that require access to, or use of, remote equipment, such as tiled-display walls and sensors. Co-leaders: Shinji Shimojo (National Institute for Information and Communication Technology [NICT] and Osaka Univ) and Fang-Pang Lin (National Center for High-performance Computing [NCHC]). This group also includes activities on disaster management.
- **BIOSCIENCES WORKING GROUP:** Focusing much of its efforts over the last several years on infrastructure development that allows performing virtual screening experiments and computational genomics analyses with an emphasis on combating infectious diseases and, more recently, on rice breeding. Leader: Jason Haga (AIST).
- **CYBER-LEARNING:** Focusing on use of technologies to improve understanding in several areas of computational science through the use and improvement of EDISON. Co-leaders: Ruth Lee (Korea Institute of Science and Technology Information [KISTI]) and Hsi-ching Lin (NCHC).

WORKING GROUPS

Workshops

In 2016, two PRAGMA Workshops were held:

- **PRAGMA 30:** January 27–29, 2016, hosted by Advanced Science and Technology Institute (ASTI), Manila, Philippines. PRAGMA 30 was held in conjunction with the 41st Asia-Pacific Advanced Network (APAN) meeting.
- **PRAGMA 31:** September 7–9, 2016, hosted by Thammasat University, Bangkok, Thailand with additional support from the e-Government Agency. PRAGMA 31 was held in conjunction with three additional events:
  - the ASEAN International Virtual Organization (IVO) Meeting (see *CENTRA* in *Building Community* section);
  - the Baseline Environmental Data Capture Project in the Lower Mekong Basin Meeting (see *Building Community*); and
  - the 2016 Asia Pacific Open Data Summit (APODS2016), the second such meeting, which brought together representatives from the ASEAN countries to discuss the value of open data and the specific activities each is undertaking.

More information about the PRAGMA workshops can be found at *pragma30.pragma-grid.net* and *pragma31.pragma-grid.net*. The website is supported by the collaborative software Duckling, developed by PRAGMA partner the Computer Network Information Center (CNIC), Chinese Academy of Sciences.

In addition, training programs such as the Southeast Asia International Joint Research and Training Program (SEAIP, *seaip.narlabs.org.tw*) provided PRAGMA with new members (see *Building Community*).

Looking to the future, we will continue to employ these venues 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. Listed below are our planned upcoming workshops:

- **PRAGMA 32:** April 12–14, 2017, University of Florida, Gainesville, FL, United States. To be held in conjunction with the 2nd Annual CENTRA meeting.

We acknowledge the many contributions of members in hosting PRAGMA workshops and thank the organizers and host institutions for their efforts to ensure PRAGMA's continued success.

PRAGMA STUDENT POSTER AWARDS

PRAGMA 31

Best Poster Award

Chawanat Nakasan, Kohei Ichikawa, and Hajimu Iida. Nara Institute of Science and Technology. Implementing and Testing Ceph Distributed File System with Multipath TCP

Best Poster Award, 2nd Place

Takuya Yamada, Keichi Takahashi, Masaya Muraki, Susumu Date, and Shinji Shimojo. Osaka University. A Proposal of Access Control Mechanism Towards User-dedicated PRAGMA ENT for IoT Era

Best Poster Award, 3rd Place

Phuripat Akarasiriwong. Mahidol University. EddvisAR: Augmented Reality Within Web Browser Technology

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 for PRAGMA initiatives. More information about steering committee members (denoted with an asterisk * in the member list) can be found at *www.pragma-grid.net/people.php*.

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.

ADVANCED SCIENCE AND TECHNOLOGY INSTITUTE (ASTI): Jelina Tanya H. Tetangco*, *jeng@asti.dost.gov.ph*

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

DATA TO INSIGHT CENTER, INDIANA UNIVERSITY (IU): Beth Plale*, *plale@indiana.edu*

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

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

NARA INSTITUTE OF SCIENCE AND TECHNOLOGY (NAIST): Kazutoshi Fujikawa*, *fujikawa@itc.naist.jp,* Kohei Ichikawa*, *ichikawa@is.naist.jp*

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;* Weicheng Huang, *whuang@nchc.narl.org.tw*

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

NATIONAL INSTITUTE OF SUPERCOMPUTING AND NETWORKING (NISN), KOREA INSTITUTE OF SCIENCE AND TECHNOLOGY INFORMATION (KISTI): Kumwon Cho*, *ckw@kisti.re.kr;* JongSuk Ruth Lee, *jsruthlee@kisti.re.kr*

THAMMASAT UNIVERSITY: Prapaporn Rattanatamrong, *rattanat@gmail.com;* Wanida Putthividhya, *wanidap@cs.tu.ac.th*

UNIVERSITAS INDONESIA: Heru Suhartanto*, *heru@cs.ui.ac.id;* Wisnu Jatmiko, *wisnuj@cs.ui.ac.id*

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

UNIVERSITY OF FLORIDA (UF), in particular the Advanced Computing and Information Systems Laboratory and the Florida Museum of Natural History: José Fortes*, *fortes@acis.ufl.edu;* Renato Figueiredo*, *renato@acis.ufl.edu*

UNIVERSITY OF HONG KONG (HKU): W.K. Kwan*, *hcxckwk@hku.hk;* P.T. Ho, *hcxchpt@hku.hk;* Lilian Y.L. Chan, *lilianyl@hku.hk*

UNIVERSITY OF WISCONSIN-MADISON (UW-Madison), in particular the Center for Limnology: Paul Hanson, *pchanson@wisc.edu*

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): Markus Buchhorn, *markus@apan.net*

PACIFIC WAVE: John Silvester, *jsilvest@usc.edu*

STARLIGHT AND PACIFIC RESEARCH PLATFORM INITIATIVES: Maxine Brown, *maxine@uic.edu*

TransPAC, INDIANA UNIVERSITY: Jennifer Schopf, *jmschopf@indiana.edu;* Andrew Lee, *leea@indiana.edu*

Other Members

CENTER FOR COMPUTATIONAL SCIENCES (CCS), UNIVERSITY OF TSUKUBA: Osamu Tatebe, *tatebe@cs.tsukuba.ac.jp;* Taisuke Boku, *taisuke@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): Ze Luo, *luoze@cnic.cn;* Yu Chen, *cy@cnic.cn*

INSTITUTE FOR COMPUTATIONAL SCIENCE AND TECHNOLOGY (ICST): Thanh N. Truong, *Thanh.Truong@ICST.org.vn,* and *TNTruong.ICST@gmail.com;* Lam K. Huynh, *Lamhuyng.us@gmail.com;* Tuan Huynh, Cyber Lab, *vantuanspkt@gmail.com*

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

MIMOS: Hong-Hoe Ong, *hh.ong@mimos.my;* Jing-Yuan Luke, *jyluke@mimos.my*

MONASH UNIVERSITY (Monash): David Abramson*, *david.abramson@monash.edu,* Paul Bonnington, *paul.bonnington@monash.edu;* Maria Indrawan, *maria.indrawan@infotech.monash.edu.au*

NATIONAL ELECTRONICS AND COMPUTER TECHNOLOGY CENTER (NECTEC): Chalee.Vorakulpipat, *chalee.vorakulpipat@nectec.or.th*

UNIVERSITI SAINS MALAYSIA (USM): Habibah A. Wahab*, *habibahw@usm.my;* Chan Huah Yong, *hychan@cs.usm.my;* Mohd Azam Osman; *azam@cs.usm.my*

UNIVERSITY OF HYDERABAD (UoH): Rajeev Wankar, *wankarcs@uohyd.ernet.in, rajeev.wankar@gmail.com*

For more information about each of the PRAGMA Institutional Members, visit www.pragma-grid.net/members-partners.php

Images: (left) UCSD PRAGMA 31 participants—courtesy of Teri Simas; (right) Wat Po Temple, Bangkok, after PRAGMA 31—courtesy of Peter Arzberger

Additional Organizations Active in PRAGMA

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.

NATIONAL APPLIED RESEARCH LABORATORY (NARL; www.narl.org.tw/en) was established in 2003 to consolidate nine national laboratories into a single nonprofit organization to construct, operate and maintain the large-scale R&D facility and platform in support of academic research and foster the necessary manpower in various advanced fields focused on by the nation. NCHC is one of the laboratories in NARL. NARL 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 research and development on information technology supporting the ubiquitous society of the future. NICT supported PRIME students 2009–2015 and has participated in the activities of the Telescience Working Group through support of the high-definition video conferencing testing.

UNIVERSITY OF QUEENSLAND (www.uq.edu.au) has recently become involved in PRAGMA through David Abramson's move there. David remains actively involved in PRAGMA (see *MURPA* section).

VIRGINIA TECH Cayelan Carey and colleagues in Project EDDIE (Environmental Data-Driven Inquiry and Exploration; projecteddie.org, an NSF-funded project) have developed sensor-based and time series data analysis activities that can be integrated into classrooms to improve quantitative skills, reasoning, and increase student engagement. Prof. Carey has expanded upon this effort by developing additional open-source teaching materials that use the overlay network developed as part of PRAGMA to run hundreds of lake simulations of climate change scenarios. Students develop hypotheses about how climate change is affecting lakes, use the overlay network to run climate change experiments *in silico*, then examine the output to evaluate their hypotheses. To date, more than a hundred undergraduate students have participated in the modeling module across multiple universities. All teaching materials are publicly available online. Assessment results demonstrate participation in this curriculum significantly increases students' comprehension about how climate change affects lakes, as well as students' experience level working with modeling and computing tools.

Partners

GLEON, the Global Lakes Ecological Observatory Network, is a grassroots network of limnologists, ecologists, information technology experts and engineers who use GLEON's network of people, sensors and data to understand issues such as eutrophication or climate change at regional to global scales. GLEON, which was established based on an early PRAGMA expedition to place sensors on a lake in Taiwan in 2004, has grown to a network of more than 500 members. It has developed new knowledge and insights, created new data products and developed a very successful Graduate Student Association (GSA). There are several ties between GLEON and PRAGMA, including shared personnel, shared learning from the GLEON GSA to develop the PRAGMA Student Group, the shared scientific expedition on Lake Eutrophication, and the joint hosting of a workshops on big data (Taiwan, December 2012), distributed computing (Quebec, 2014), and numerical simulation (South Korea, 2015). For more about GLEON, see gleon.org.

The **Network Startup Resource Center** (nsrc.org), based at the University of Oregon, was established in 1992 to provide technical assistance to organizations setting up computer networks seeking to connect to the NSFNET in support of international scientific research and education. Through coordinated technical training programs and direct engineering assistance in more than 100 countries over the past couple of decades, the NSRC builds institutional capacity to support the research community, and leverages U.S. research infrastructure to bring value to international researchers and educators. NSRC has worked with the PRAGMA community for several years in the Southeast Asia region to fund researchers from Myanmar, Philippines and Thailand to attend PRAGMA meetings and engage with the regional science community. The NSRC is partially funded by the International Research Network Connections (IRNC) program of the National Science Foundation and Google, with additional contributions from dozens of public and private organizations.

SPONSORS

PRAGMA is supported by its member institutions and the U.S. National Science Foundation (NSF OCI 1234983, PI: Papadopoulos), involving researchers at UC San Diego, Univ Florida, Indiana Univ, Univ Kansas, Virginia Tech and Univ Wisconsin-Madison. 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.

AIST's sponsors include the Strategic International Research Cooperative Program (JST, Japan) and Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures.

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

CMC/Osaka University is supported by JGN-X of the National Institute of Information and Communications Technology (NiCT), Japan.

Osaka University and National Institute of Information and Communications Technology (NICT) research was supported in part by institutional funding for their collaborative research project title "Research on High Functional Network Platform Technology for Largescale Distributed Computing."

CCS's (at the University of Tsukuba) PRAGMA participation is partially supported by the JST CREST award titled "Development of System Software Technologies for post-Peta Scale High Performance Computing."

CCST at Jilin University receives funding support from the Chinese Natural Science Foundation (51627805) and the Chinese Ministry of Education (2016YFB0201503).

CNIC's sponsors include the Chinese Academy of Sciences, the Ministry of Science and Technology of China and the Natural Science Foundation of China.

DATA TO INSIGHT CENTER, INDIANA UNIVERSITY (IU): IU's involvement in PRAGMA is enabled by the Indiana University Pervasive Technology Institute, and the School of Informatics and Computing. PRAGMA activity is funded in part through the NSF PRAGMA grant OCI 1234983, and a grant to RDA/US from the MacArthur Foundation.

Kasetsart University's PRAGMA participation has been partly funded by Faculty of Engineering, Kasetsart University.

KISTI receives major funding from the Korean government (MSIP) through EDISON (**ED**ucation-research-industry **I**ntegration through **S**imulation **O**n the **N**et) Project and KISTI through its Supercomputing R&D Program.

Image: Participants PRAGMA 31, Bangkok Thailand September 8, 2016

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 for 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 Ministry of Science and Technology, Taiwan.

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

StarLight receives major funding from NSF (ACI-1450871) for "IRNC:RXP: StarLight SDX A Software Defined Networking Exchange for Global Science Research and Education.". StarLight consortium members University of Illinois at Chicago and Northwestern University work with member University of California, San Diego, on the NSF (ACI-1541349) "CC*DNI DIBBs: The Pacific Research Platform."

TransPAC receives major funding from NSF and collaborates closely with PacificWave, Japan's National Institute of Information and Communications Technology, APAN, TEIN, and other Asian networking groups.

Universitas Indonesia's PRAGMA participation has been partly funded by the Faculty of Computer Science, and research grants from The Ministry of Research Technology and Higher Education, Republic Indonesia.

USM's grid activities in Malaysia are funded mainly through E-science and USM Central Funding.

University of Florida's involvement is supported by the NSF PRAGMA grant OCI 1234983.

University of Wisconsin's involvement is supported by the NSF PRAGMA grant OCI 1234983.

PRAGMA would like to acknowledge and thank those who helped prepare this report: Jennifer Matthews for graphic design; Tiffany Fox, Molly Wofford and Doug Ramsey for editing; Teri Simas for copy editing, and the section authors and photographic contributors!


PRAGMA
www.pragma-grid.net