

PRAGMA 35: Overview and Updates

Shava Smallen
Interim Co-chair, PRAGMA Steering Committee
University of California, San Diego

Shinji Shimojo
Interim Co-chair, PRAGMA Steering Committee
Osaka University

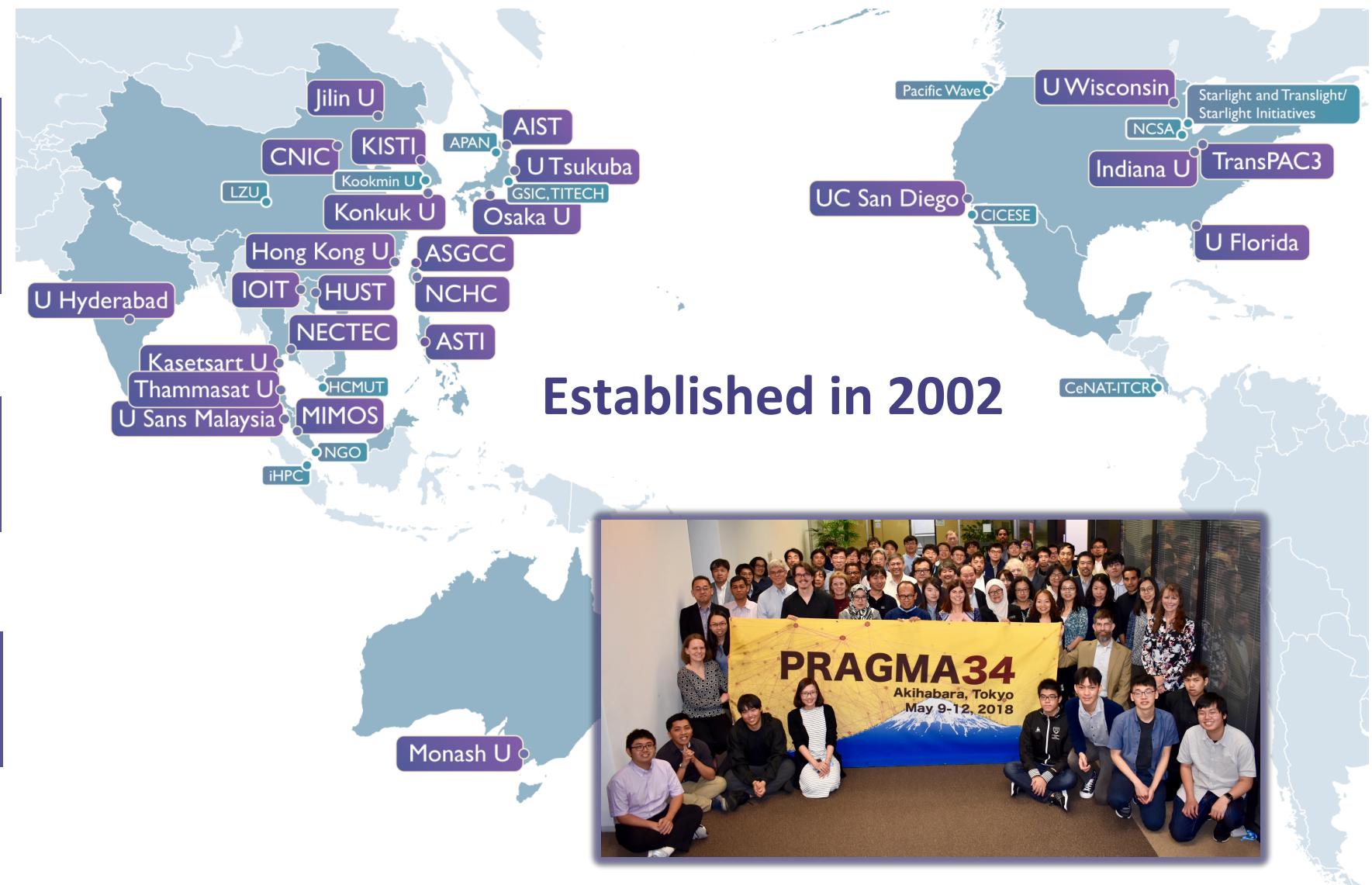


PRAGMA is an Open Community of Practice

Focused on researchers
and institutions on the
Pacific Rim

Engages “Long Tail”
science communities

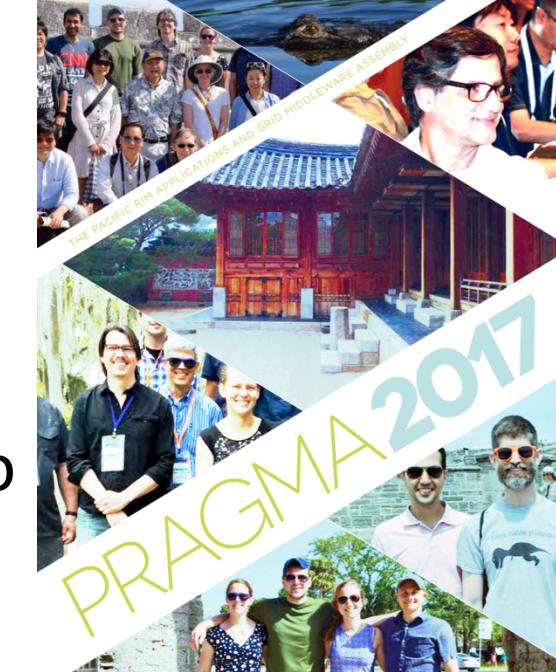
Twice yearly meetings
for measured progress



PRAGMA: Co-designing useful Cyberinfrastructure (CI), transforming long-tail science communities

- Science is **inherently international** and requires collaboration.
- Collaboration is **enabled by sharing and exchanging data, algorithms and codes**
- Cyberinfrastructure is only one dimension; **people and trust are also essential**
- There are fundamental challenges in matching existing CI to hundreds of communities. **Deep interactions with long-tail communities has the potential to transform both science domains and CI development.**

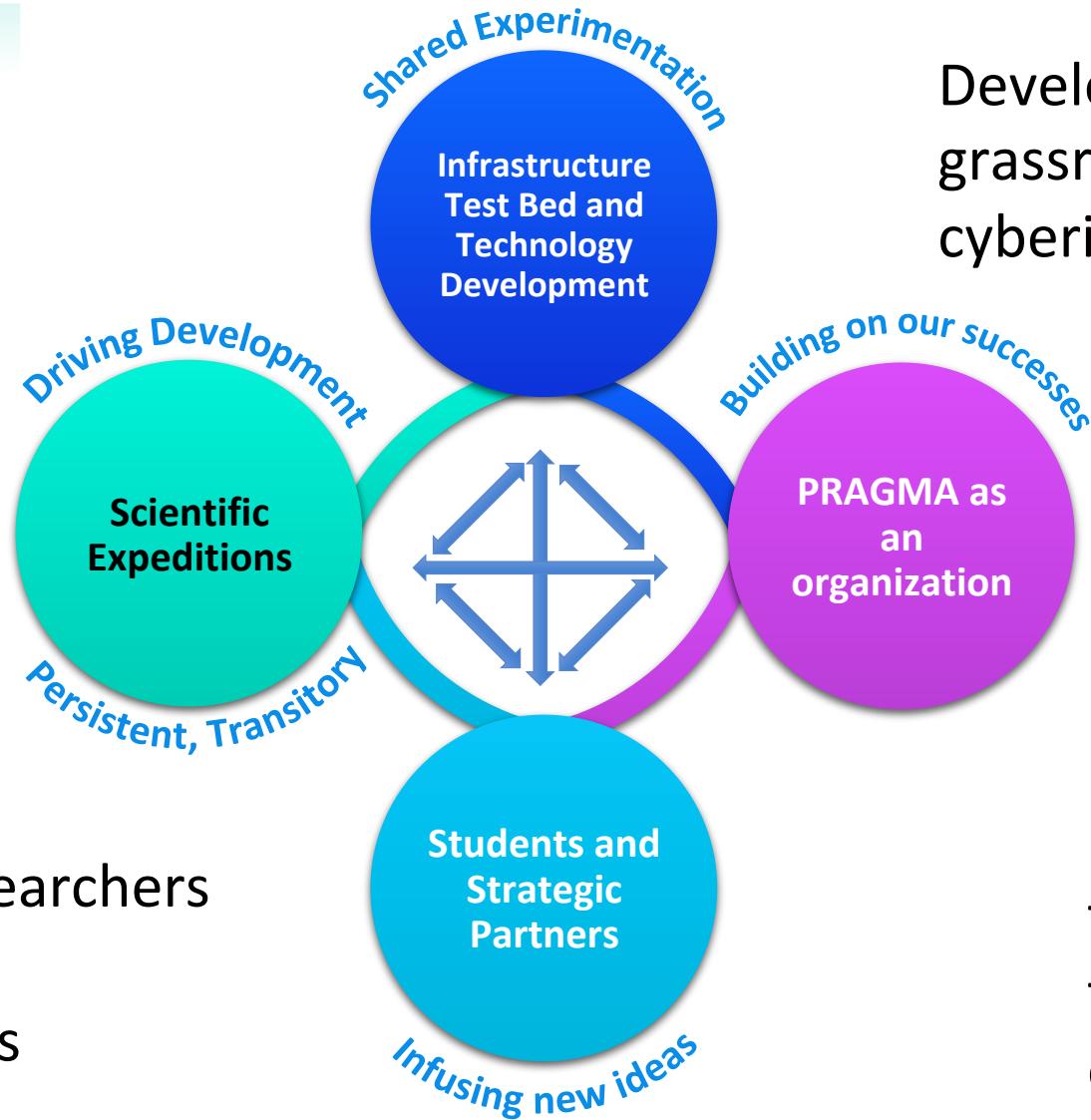
PRAGMA seeks to address these challenges to improve *CI and scientific impact via practical implementation*. We focus on international collaborations that uniquely team technology specialists and domain scientists



PRAGMA – Building Trusted Community of Practice through four strategies of collaborating

Fostering international scientific expeditions by forging teams of domain scientists and CI researchers

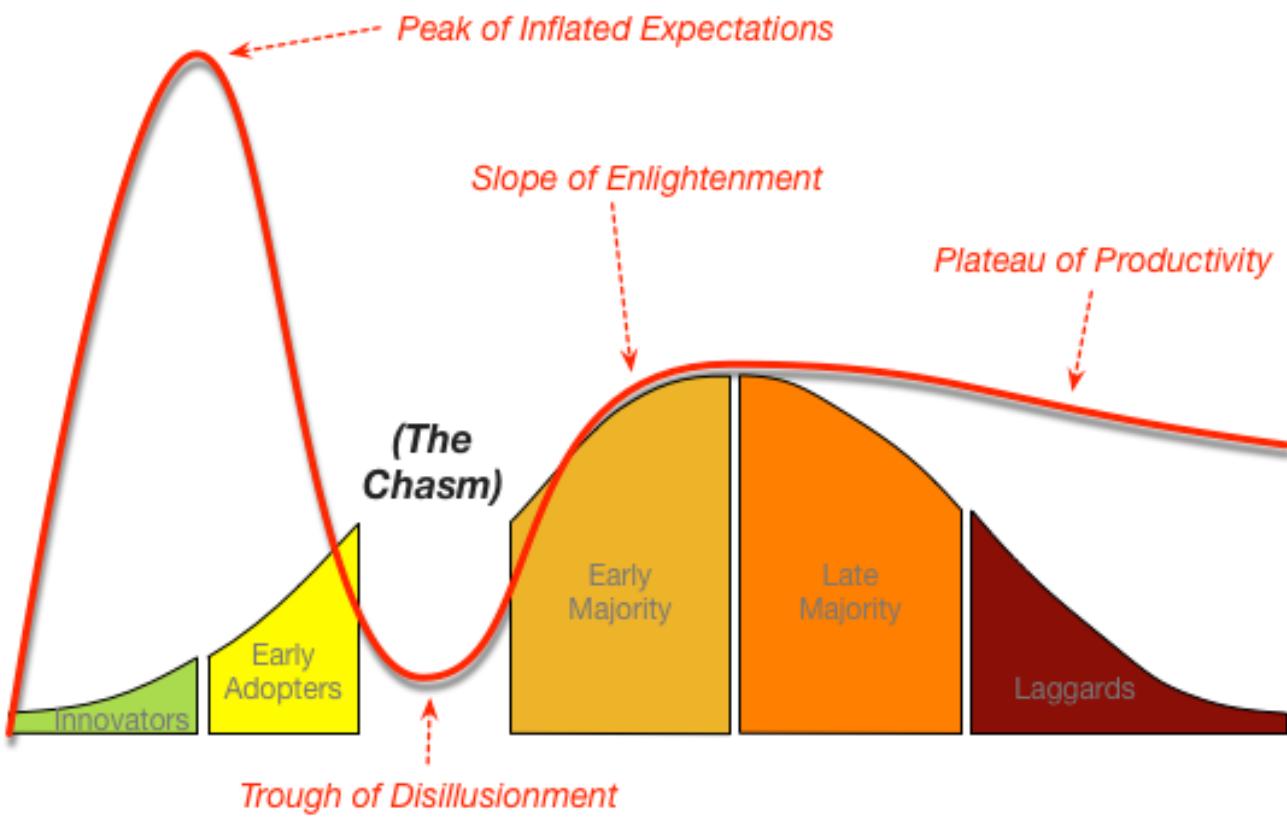
Infusing new ideas by developing young researchers and by engaging with strategic partners



Developing and improving a grassroots, international cyberinfrastructure

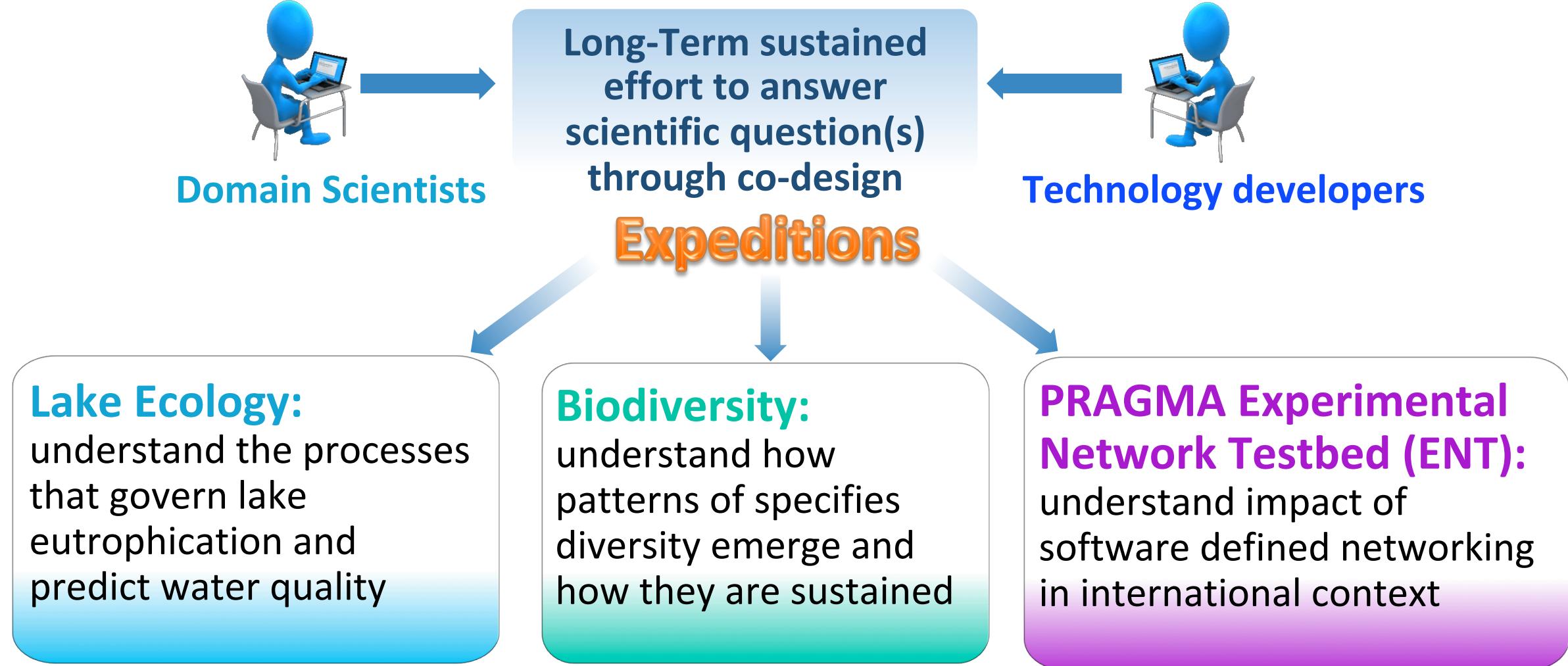
Building and enhancing the essential people-to-people trust and organization

Cyber technologists invent – but then how does one make long-term impact on domain science?



- Nearly all cyber (information) technologies follow the Gartner “hype” cycle
 - Grids, Clusters, Clouds, Big Data, Machine Learning, Software-Defined Networking, Workflows, IoT, AI, Cloud Orchestration...
- For domain scientists (non-CI)
 - What are the *potentially transformative* CI innovations
 - If I'm an early adopter or early majority, who can help me cross the chasm?

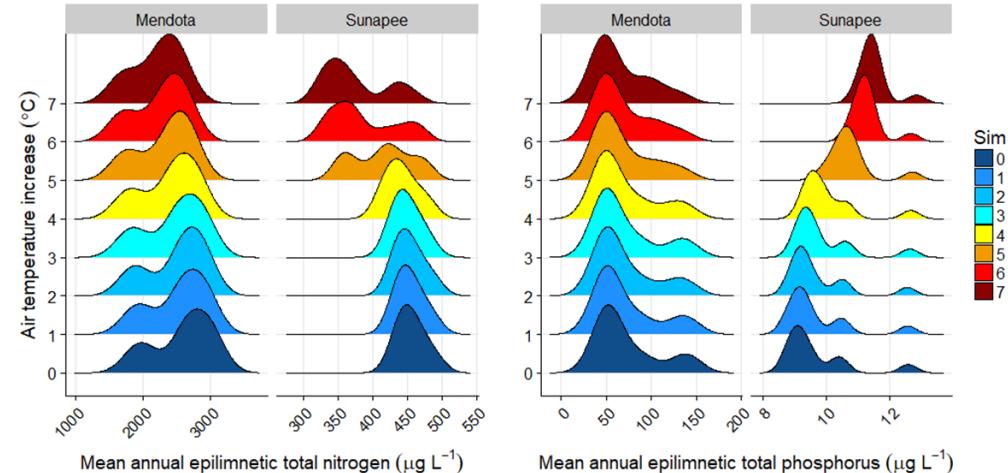
PRAGMA's Expeditions: A Model of Collaboration



Lake Expedition: Predicting Water Quality in Lakes

Paul Hanson (U. Wisconsin), Cayelan Carey (Virginia Tech), Renato Figueiredo (U. Florida)

- **Eutrophication:** excessive richness of nutrients in a lake or other body of water, frequently due to run-off from the land, which causes a dense growth of plant life.
- **Lake eutrophication** is global issue, results in **degraded water quality**
- Goal: Integrate sensor data as inputs to **computational lake models** → science goal is to ***predict water quality***



- *Renato Figureido will be giving keynote presentation shortly with more detail: **ON LAKES AND CLOUDS: A RETROSPECTIVE ON THE PRAGMA/GLEON LAKE EXPEDITION***

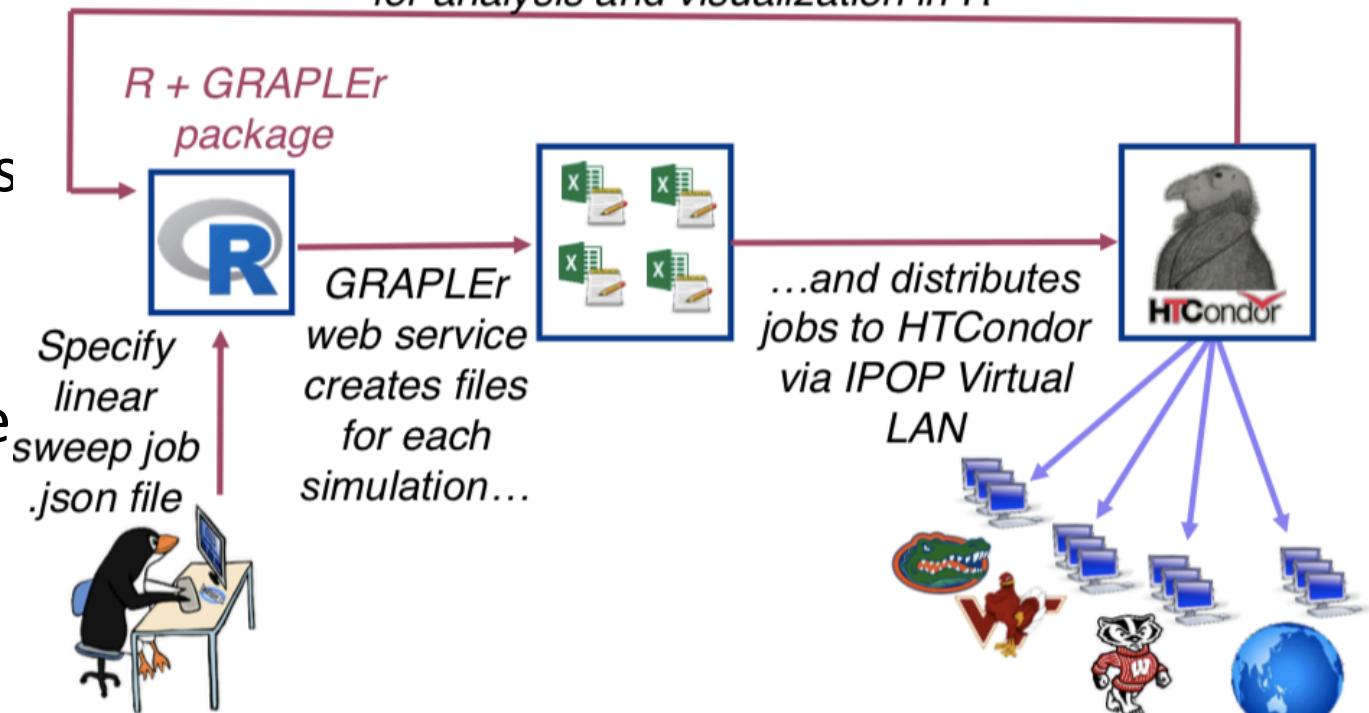
Lake Expedition: Developing Predictive Models using High-Throughput Computing and Overlay Networking

- **Sensor gateways** link sensors to internet (connecting to storage and models)
- **IPOP overlay network** create virtual private network for sensor data and computation
- **GRAPLER distributed computing** uses an R interface to allow ecologist access to distributed computing
- **Educational Modules** educate future generations in computational skills (currently at 12 universities)

Prof Carey teaching students to use GRAPLER at VaTech



Outputs from model runs are aggregated and returned to user for analysis and visualization in R



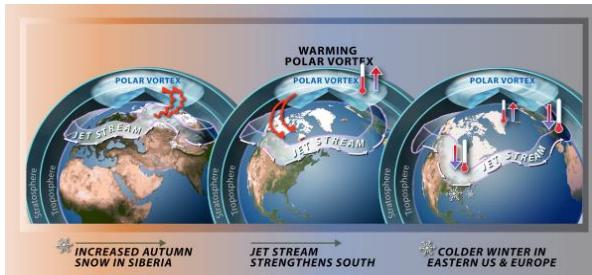
Virtual Biodiversity Expedition: Research and Predict Biodiversity of Terrestrial Species



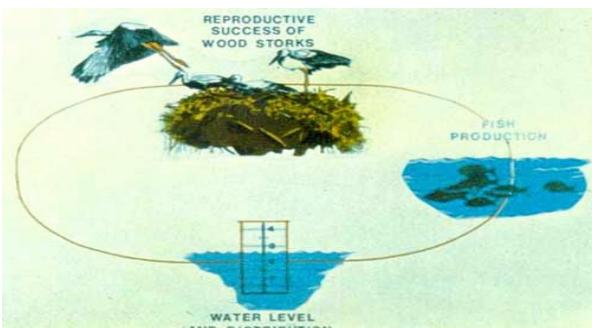
Biodiversity Inventory



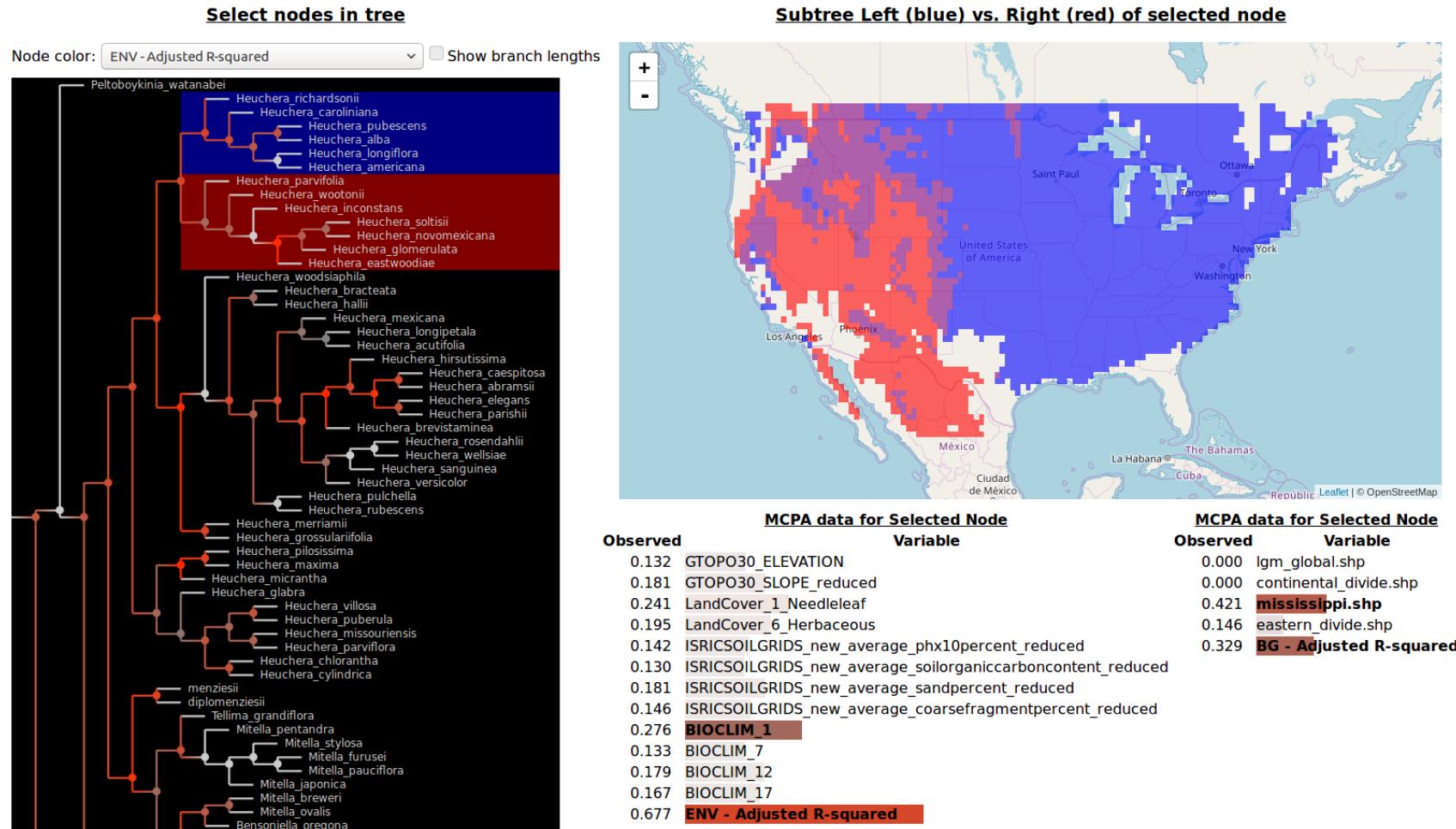
Global Climate Change



Macro-Ecological Modeling

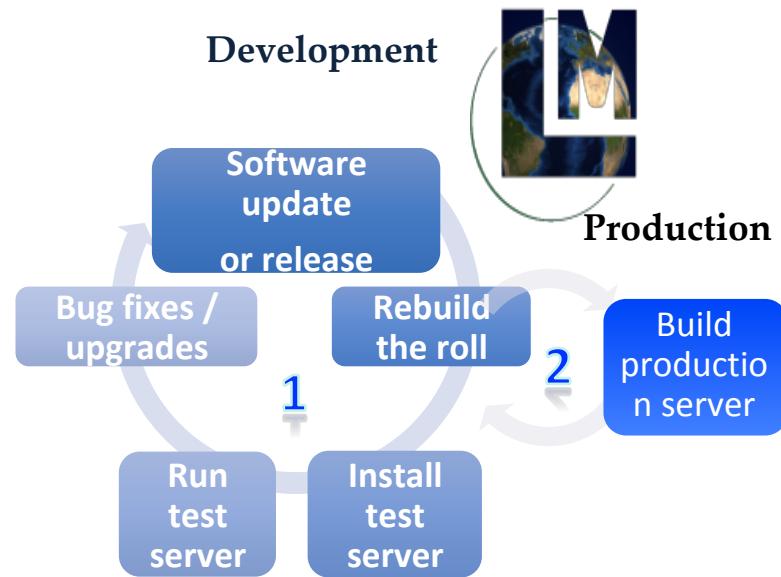


Jim Beach, Aimee Stewart (U. of Kansas), Nadya Williams (UCSD)



Linked view of phylogenetic tree, map of highlighted species distributions, and variable correlations displayed in Lifemapper.

Virtual Biodiversity Expedition: Expanding and Enhancing Lifemapper Deployments



- **Software engineering and Rocks** increased availability and flexibility of Lifemapper
- **Lifemapper deployed in a variety of environments** such as researcher laptops, project servers, HPC resources (US XSEDE Comet)
- **Working to facilitate data ingestion** to enable regional Lifemappers that use locally-available high resolution data
- **Training**
- **Updates:** GUI improvements, streamline workflows, and tune the software environment for optimal resource allocation
- Aimee will be giving **demo** during Resources and Data WG breakout

PRAGMA Experimental Network Testbed (ENT) Expedition

Motivation:

an International Software-defined Networking (SDN) Testbed for use by PRAGMA researchers and collaborators

Accomplishments:

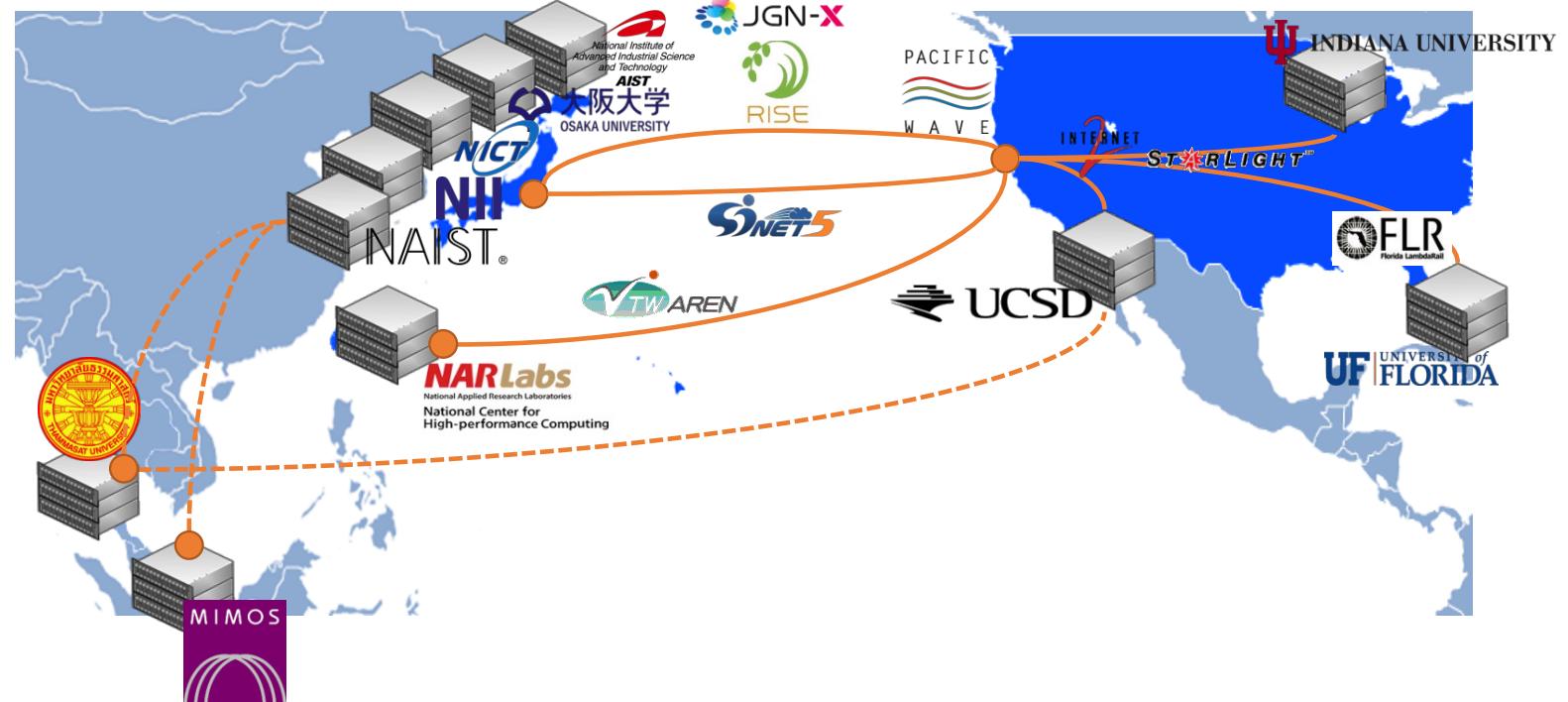
4 countries and 11 institutions have been involved in building the large scale SDN testbed

Using NAIST's AutoVFlow as a openflow meta controller

Numerous Publications

- Efficient Packet Header Rewriting (PARES) – (U. Florida, NAIST)
- Multipath TCP – NAIST with student visitation to UCSD.
- SMOC – Simple Multipath Openflow Controller
- Overseer: Latency and Bandwidth aware routing.

Kohei Ichikawa (NAIST)



IPOP Integration

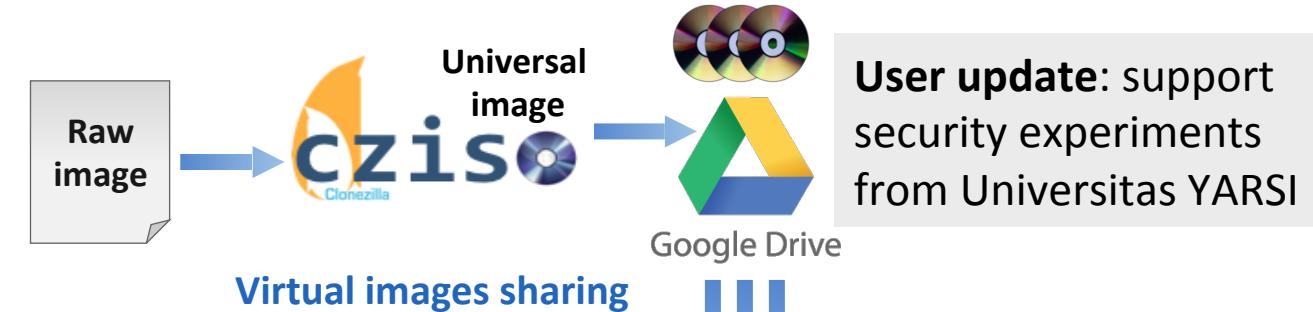
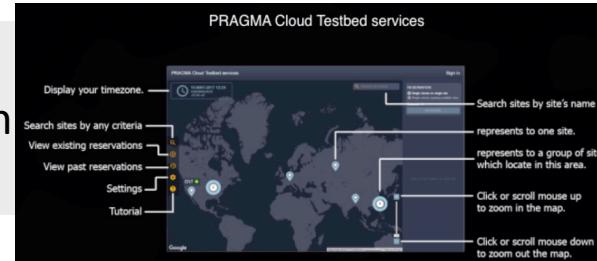
- OpenFlow hardware switches no longer required to participate in ENT
- IPOP overlay network allows tunneling of traffic from/to SDN-programmed switch ports across the Internet
- Software-based SDN switches (Open vSwitch)

Demos (later today)

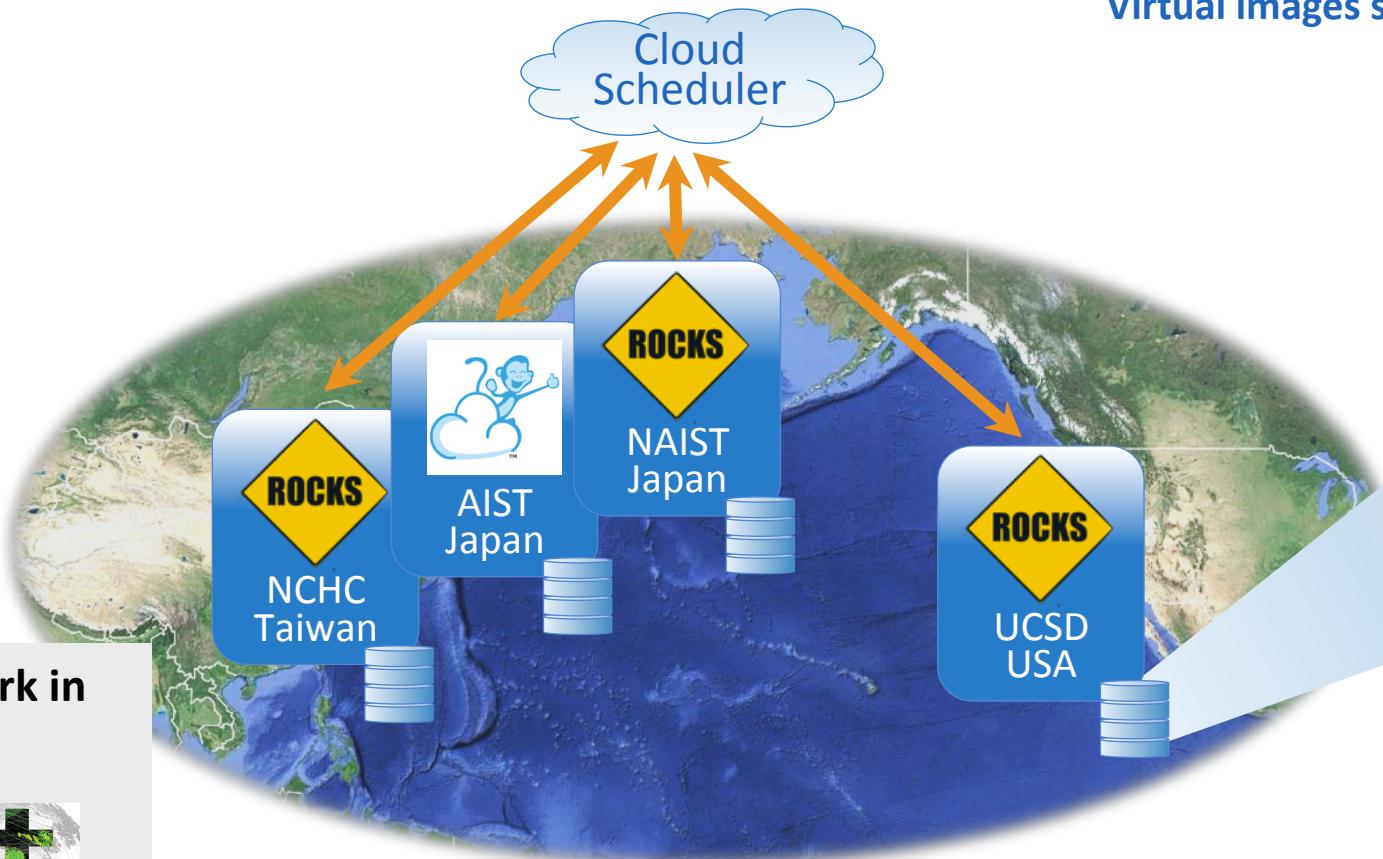
- **Kohei Ichikawa**, Integrating PRAGMA-ENT and InterCloud Platform using Dynamic L2VLAN Service
- **Kensworth Subratie**, Extending SDN Networks from Cloud-toEdge using Virtual Private Networks with Peer-toPeer Overlay Links

Building on the International Development and Collaboration to Create PRAGMA's Multi-Cloud Testbed for Advancing Science

GUI Update: enable multi-cluster reservation
(Thammasat University)



User update: support security experiments from Universitas YARSI



Monitoring: work in progress

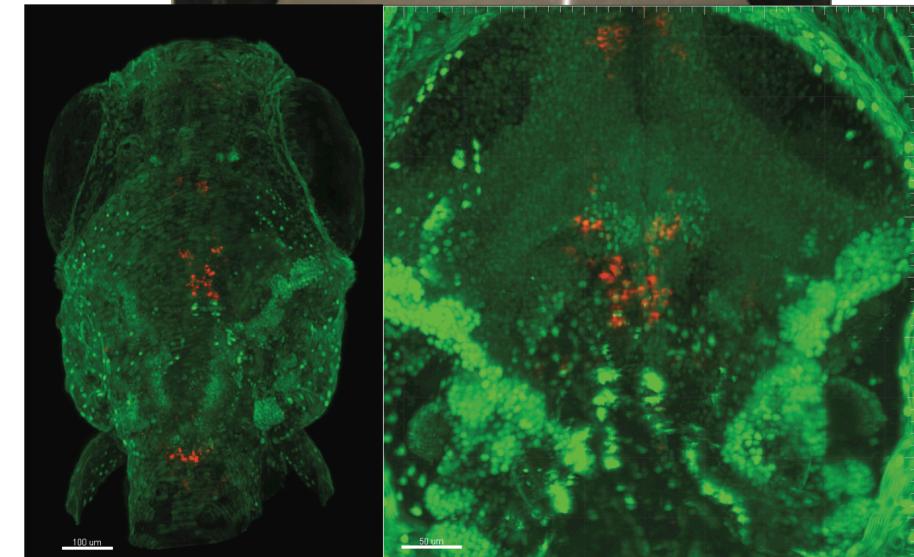
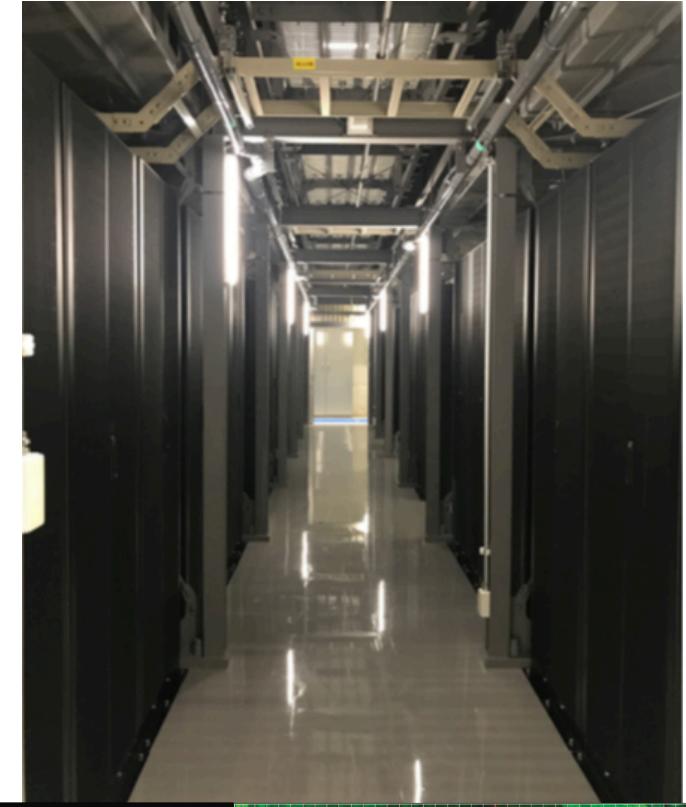


Work in progress



AI Bridging Cloud Infrastructure Resource

- New 19.9 petaflop resource dedicated to AI at National Institute of Advanced Industrial Science and Technology (AIST)
- Early access to some PRAGMA members to help test the infrastructure
- More detail provided in next keynote by Ryousei Takano: **ABCI: AN OPEN INNOVATION PLATFORM FOR ADVANCING AI RESEARCH AND DEPLOYMENT**
- Similar resource being deployed in Taiwan



Student Updates

- Special PRAGMA Student Presentation session yesterday
 - Forum for students to present research and receive feedback on content and presentation from mentors
 - Thank you to **Wassapon Watanakesuntorn (Boom)** and **Jason Haga** for organization
- Several PRAGMA 35 demos and posters
 - Best poster awards
 - Please talk to students and provide feedback



How to join PRAGMA?

1. Join a working group during PRAGMA 35 breakout sessions
2. Join the PRAGMA Student Committee
3. Consider internship or remote projects

PRAGMA Working Groups

- Expeditions provide key long-term science drivers, but Working Groups provide an organizational structure. Each group organizes its activities, goals, structure.
- Working Groups
 - **Resources and Data:** Investigates current technology trends and evaluates their potential beneficial impact on applications from PRAGMA's applications. Current projects include the PRAGMA Cloud Testbed, the Experimental Networking Testbed, Open Data Platform, Containers/Kubernetes, GPUs/Machine learning, and Monitoring. **Chairs:** Nadya Williams (UCSD), Hsiu-Mei Chou (NCHC)
 - **Telescience:** Making and improving access to or use of remote equipment (e.g., tiled-display walls or sensors). Current application areas of the group include environmental monitoring and traffic flow. **Chairs:** Shinji Shimojo (Osaka University), Fang-Pang Lin (NCHC)
 - **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. **Chair:** Jason Haga (AIST)
- Meet twice during this meeting to review progress and decide on **action items** for next meeting. Please share your ideas!

Joint session with MYREN and NSRC

- **Malaysian Research & Education Network (MYREN)** enables high speed dedicated connectivity to the education sector
- **Network Startup Resource Center (NSRC)** currently presenting (Oct 2 - 4) a Campus Network Design workshop for campus network & systems engineers from MYREN's member universities
- Some attendees will join PRAGMA 35 on Friday to foster more direct engagement between our organizations
- Keynote on Friday by director of MYREN, Kamal Hisham Kamaruddin: **ENABLING GLOBAL HPC COLLABORATION THROUGH MYREN AND NSRC**



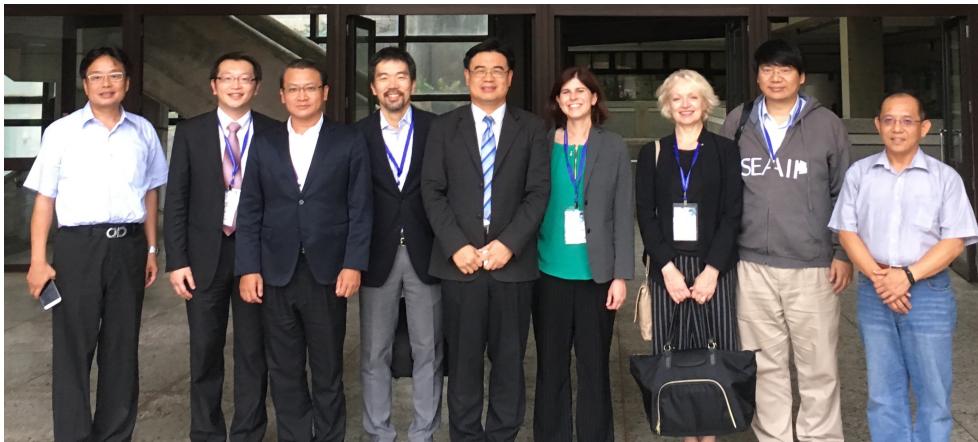
Lower Mekong Research Update

- Previous workshop held in 2014 and attended by Phil and Peter
- Recent workshop in Ho Chi Minh City led by Tho Nyugen from University of Virginia in early Sept 2018
 - Nadya and Jim attended from PRAGMA
 - PRAGMA Action items
 - Help with inventory of datasets, particularly socioeconomic data
 - Leverage students to identifying, cleaning, and indexing important data sets in the Mekong.
 - Collaboration with local professor Ly Le / Biosciences
- Tho plans to attend PRAGMA 36 meeting



Most Beautiful Bays in the World Club (MBBW) Update

- Fang-Pang Lin leading discussions for collaboration opportunities
- MBBW Congress leaders presented their organization at PRAGMA 34 in May 2018
- Fang-Pang, Shava, Nadya, and Jason just attended **2018 International Oceanic Environment and Climate Change Conference** at XIV World Bays Congress held on Penghu Bay in Taiwan
 - Presentations on environmental health and sustainability, climate change, and renewable energy
 - Met so many people that we ran out of business cards!
 - Great discussions with **Penghu County Government representative (Daniel, Hung Tung-Lin)**, **National Penghu University of Science and Technology**, and **Alston Ho (iPromise)**
 - Student involvement (undergraduate and master's students)
 - Interested in projects related to “**blue economy**”. E.g., environment and tourism projects
 - Penghu can be a starting point for collaboration with MBBW
- Next MBBW Congress meeting in Toyama Bay, Japan



PRAGMA Steering Committee Items this meeting

1. Membership
2. Proposed PRAGMA 37 dates
3. Shared responsibility model, 4 Functional Area Responsibility (FAR) chair nominations and vote
 - Membership
 - Workshops
 - Communications
 - Mentoring
4. Collaborative Overview



Updates on US side

- 1-year NSF supplement proposal awarded, expires in Sept 2019, Smallen (PI), Williams (Co-PI)
- Follow on proposal for support past 2019
 - Proposal effort led by Dr. Renato Figueiredo, University of Florida
 - Shava and Renato traveled to NSF in August 2018 and had informal conversation with program officer – thanks to Peter and Bill for excellent coaching!
 - Renato, Shava, Nadya, Paul, Cayelan and Jim planning visit to NSF on Nov 8th to give formal presentation and meet with several program officers

Big thank you to Phil and Peter ...



Director, Research Cyberinfrastructure Center



... and to PRAGMA community

Terima Kasi!

Thanks to our local chair Nurul Malim and hosts Universiti Sains Malaysia and MDEC

Sponsors

- Vitrox
- Silverlake
- Hilti
- Fusionex
- Novorient

Patrons

- Prof. Dr. Ahamed Tajudin Khader
- Prof. Dr. Rosni Abdullah
- Prof. Dr. Mohd Nazalan Mohd Najimudin
- Prof. Dr. Habibah Abdul Wahab
- Dr. Peter Arzberger

Terima Kasi!

Big Data Summit 2 Program Committee

- Assoc. Prof. Dr. Bahari Belaton
- Assoc. Prof. Dr. Chan Huah Yong
- Dr. Mohammed F.R Anbar
- Dr. Mohd Heikal Husin
- Dr. Nur Syibrah Mohd Naim
- Dr. Mohd Nadhir Ab Wahab
- Dr. Azleena Mohd Kassim
- Dr. Mohd Halim Mohd Noor
- Dr. Anusha Achuthan
- Dr. Chew Xin Ying
- Mr. Iznan Husainy Hasbullah
- Mr. Mohd Azam Osman
- Mr. Mohd Azam Osman
- Mr. Mohamad Azhar Mustapha
- Mr. Mohamad Hadzri Yaakop
- Mr. Shik Abdulla and team
- Mdm. Sheela Muniandy
- Mdm. Noor Aida Lob Abu Bakar
- Mdm. Badriyah Che May
- Mdm. Halizah Abdul Razak
- Mdm. Nurul Nadiah Zambri
- Mdm. Siti Zainura Abdul Kadir
- Mdm. Rohana Omar
- Mdm. Nur Sadrina Abdul Rahim

PRAGMA 35 Program Committee

- Prof. Renato Fuigeredo
- Prof. Shinji Shimojo
- Prof. David Abramson
- Prof. Heru Suhartanto
- Assoc. Prof. Kohei Ichikawa
- Assoc. Prof. Putchong Uthayopas
- Dr. Jason Haga
- Dr. Ryousei Takano
- Dr. Yoshiyuki Kido
- Dr. Fang Pang Lin
- Dr. Prapaporn Rattanatamrong
- Nadya Williams
- Aimee Stewart
- Hsiu-Mei Chou
- Weicheng Huang
- Wassapon Watanakesuntorn

***Special thanks to
Grace Hong for photos
and logistics help!***

Terima Kasi!

We look forward to more collaboration so please talk to us about your ideas.



Future Meetings

- PRAGMA 36 Jeju, South Korea, April 26-27, 2019
- PRAGMA 37 San Diego, USA, September 2019

Software: <https://github.com/pragmagrid>

Web: www.pragma-grid.net

Info: pragma-grid-team@googlegroups.com

Our Most Heartfelt Thanks

- Please help me Thank Outgoing Steering Committee Members
 - Dr. Wing Keung Kwan, Hong Kong University
 - Dr. Patchong Uthayopas, Kasetsart University

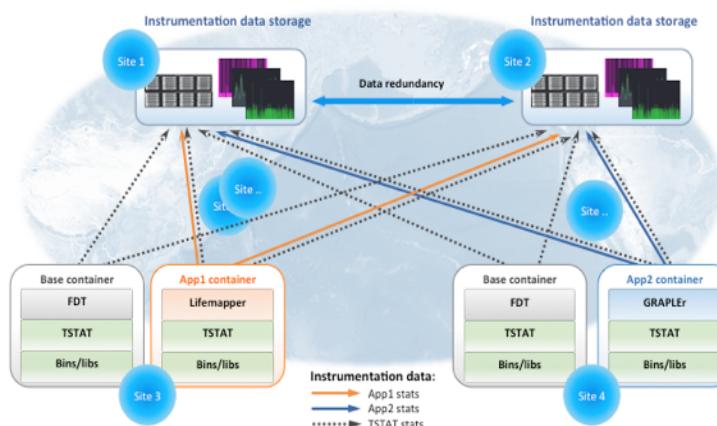
Toward a PRAGMA data infrastructure

Integration of data servers from international partners to support a notion of FAIR data:

- findable
- accessible
- interoperable
- reusable

Current efforts

- Draw on RDA notion of an international fabric composed of data objects
 - Rice Galaxy
 - AirBox project



Plans for the future

- Infrastructure monitoring
- PIDs assignment
- Data storage - AIST's, NCHC's interest in AI computing platform, expertise with CEPH storage
- Work with Artificial Intelligence Bridging Cloud Infrastructure (ABCI) at AIST on practical data transfers

Using deep learning to detect museum specimens that had been contaminated with mercury salts

Contaminated



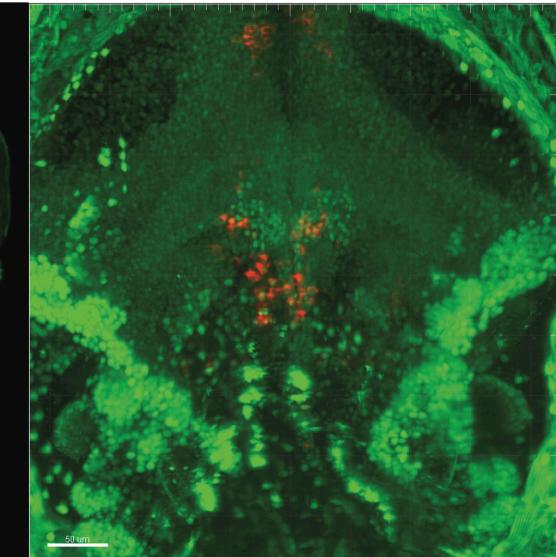
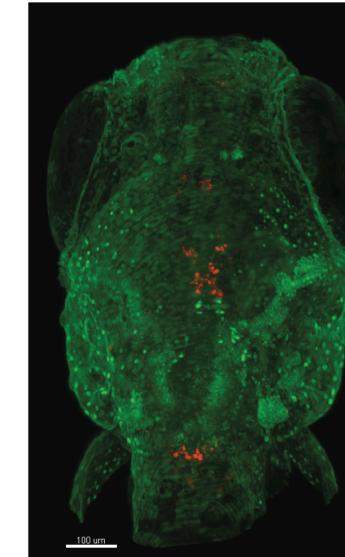
Unsure



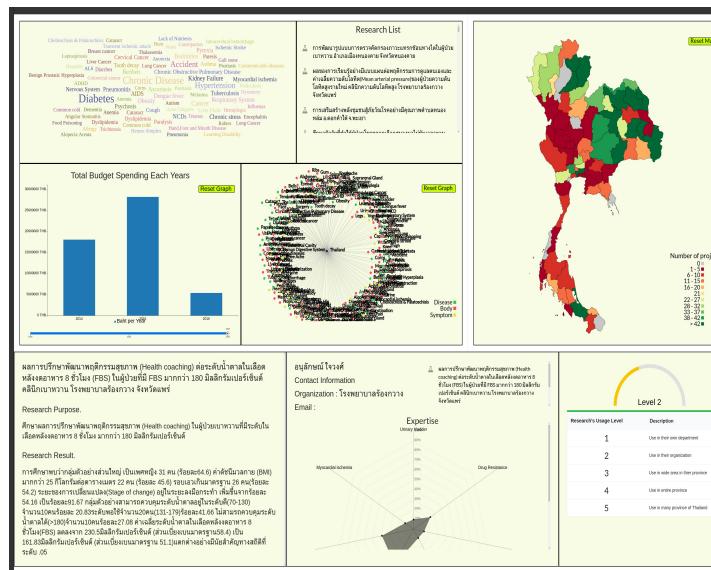
Clean



Using neural deep learning to learn architectures of the zebra fish brain



Visualizing data using SAGE2 display walls



Enabling FAIR (findable , accessible, interoperable, reusable) access to Airbox pollution data using persistant identifiers

