Resources Working Group Update

PRAGMA 30

Manila, Philippines

ributed Resources

How do we make a worldwide collection of resources usable for sharing of data/computation?

Drive infrastructure by needs of specific use cases

- Biodiversity expedition
- Lake Ecology Expedition
- Experimental Networking Testbed
- PRAGMA RDA (Research Data Alliance) Sprint

Key update on Cloud Scheduler (Demo)

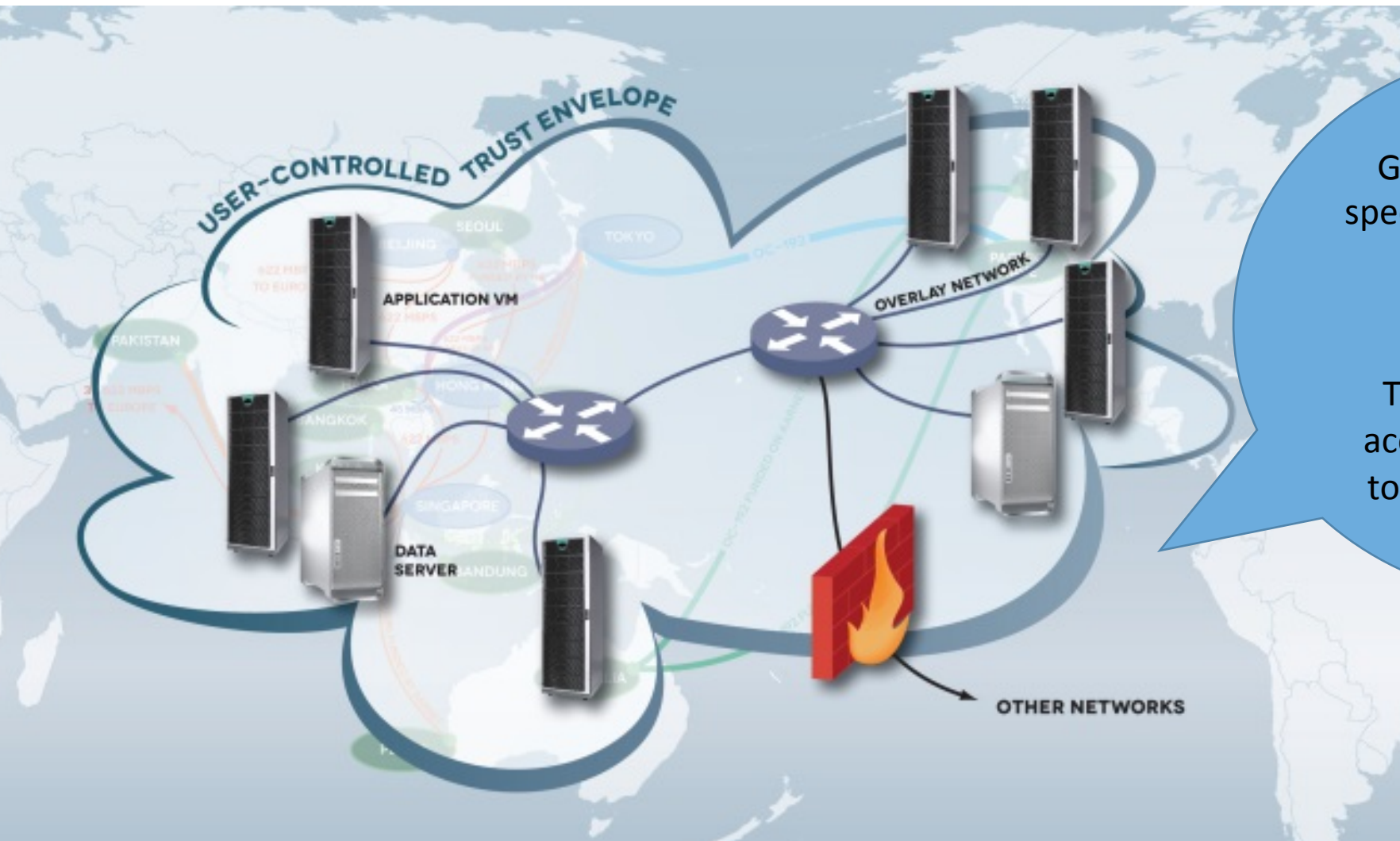Updated disk image handling dramatically improves boot times

rogress goal: Rebuilding the PRAGMA ldwide Testing Infrastructure
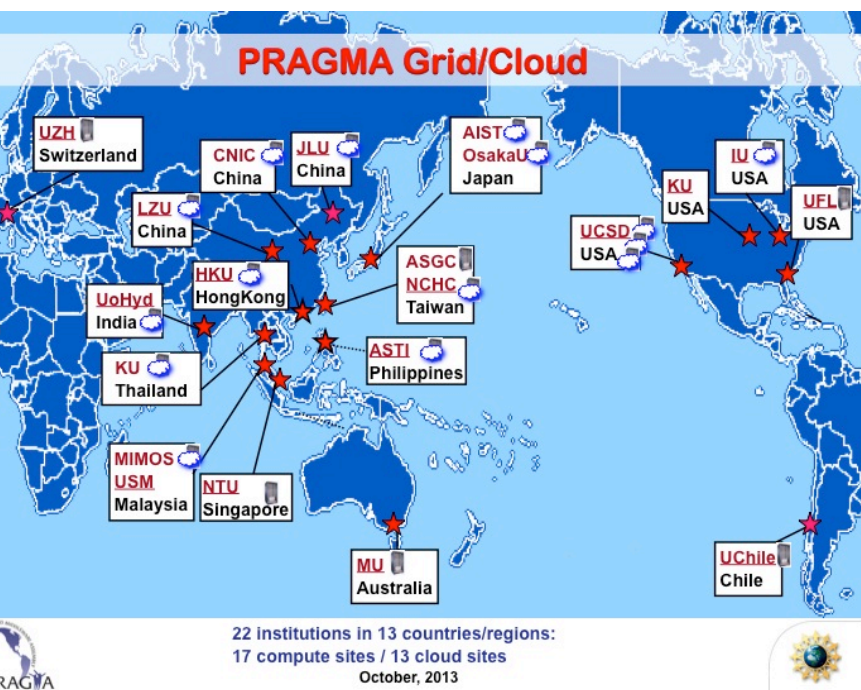
Next Generation

Trying to rebuild infrastructure AND make scientific progress
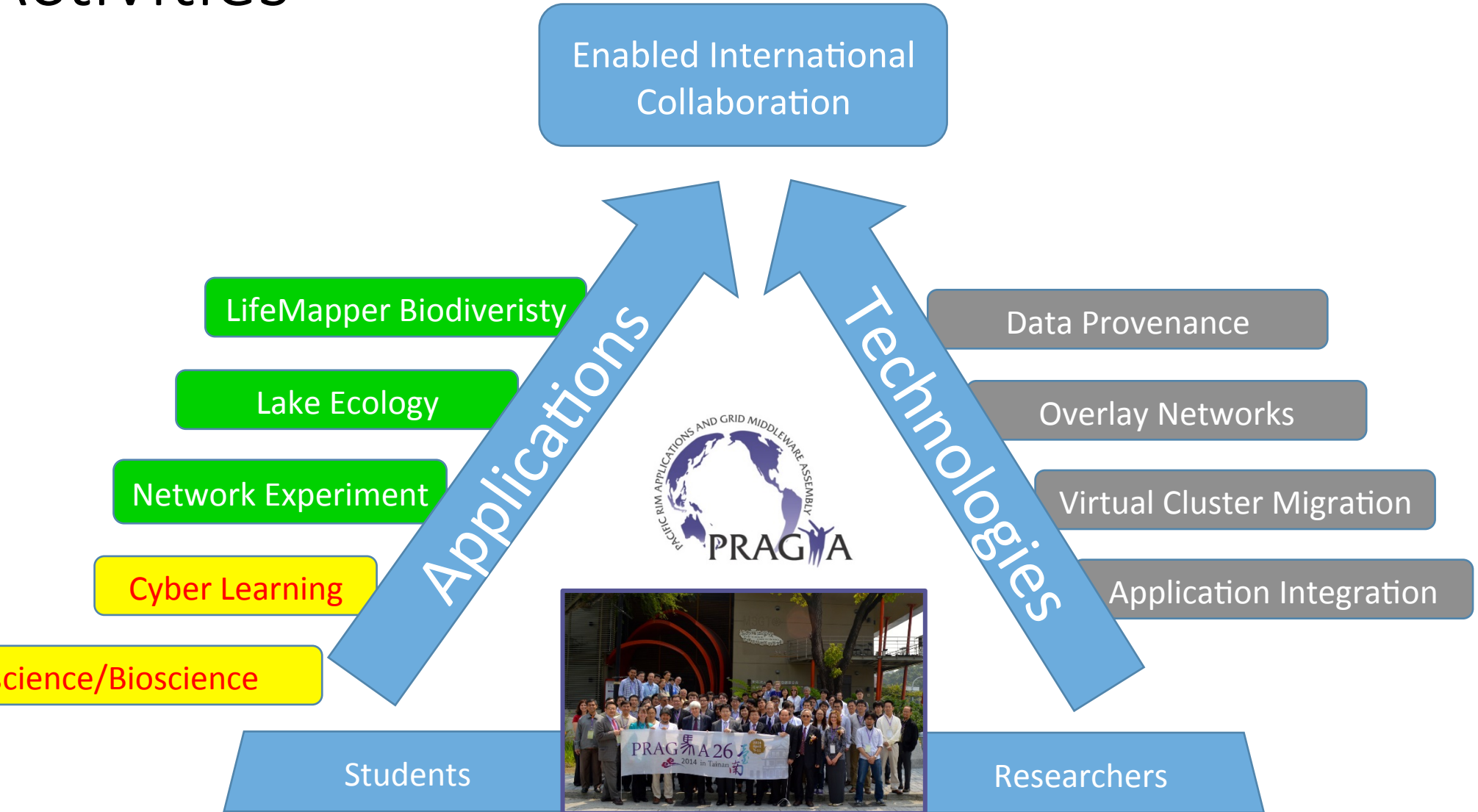
e need controlled sharing for resources

Where Resources fits in for PRAGMA Activities

Resources Working Group Update

Lifemapper PRAGMA 30 Update

KU infrastructure

- Moving towards enabling virtual clusters for beta and production systems

SDSC infrastructure

- Updated fiji PRAGMA-dedicated cluster to Rocks 6.2
- PRAGMA Boot/Image Storage capable

Lifemapper code

- Start enabling Lifemapper configuration and all setup/test invocations as a command line infrastructure. Simplify commands so that install, configuration, and testing are unified and extensible. Example:
  - **lm init db**
  - **lm star/stop pipeline**
  - **lm list users**
  - **lm update ip**
- Bug fixes
- Work on adding solr (Apache search) (for indexing)
- Continue work on formalizing requirements and code for fully described data allowing easy use of different input datasets

Lifemapper as part of the RDA Sprint

RDA

- Discussions about RDA framework to determine the trustworthiness of Lifemapper projection set through the mutation of occurrence set and Lifemapper VMs
- Identify metadata for LM objects (projections sets, occurrences sets, VMs)
- Set up 2 VMs for the demo:
  - **rocks-204** with SEA dataset and LM v.1.0.3
  - **pc-170** – use pragma_boot to instantiate rocks-204 as a new VM in a different network
  Both VMs have metadata and PIDs that can be accessed via
  http://rocks-204.sdsc.edu/lm.json
  http://pc-170.calit2.optiouter.net/lm.json
  Occurrence sets and projections are a available on both VMs at http://<Ip>/services

IPOP/Lake expedition Updates – PRAGMA-30

Renato Figueiredo, Ken Subratie, Kyuho Jeong, Paul Hanson, Cayelan Carey, Kohei Ichikawa

U. Florida, U. Wisconsin, Virginia Tech, NAIST/UCSD

Technology development

IPOP (Easy to Use VPN) Switch mode
- Handle L2 broadcasts, ARP
- SDN-based overlay bypass for fast communication within network

Revamped IPOP controller framework
- Structured P2P GroupVPN
  - Chord-based P2P self-organizing topology and routing
  - On-demand IPOP links based on traffic inspection
  - Bootstrap nodes from XMPP

PRAGMA deployments/demos
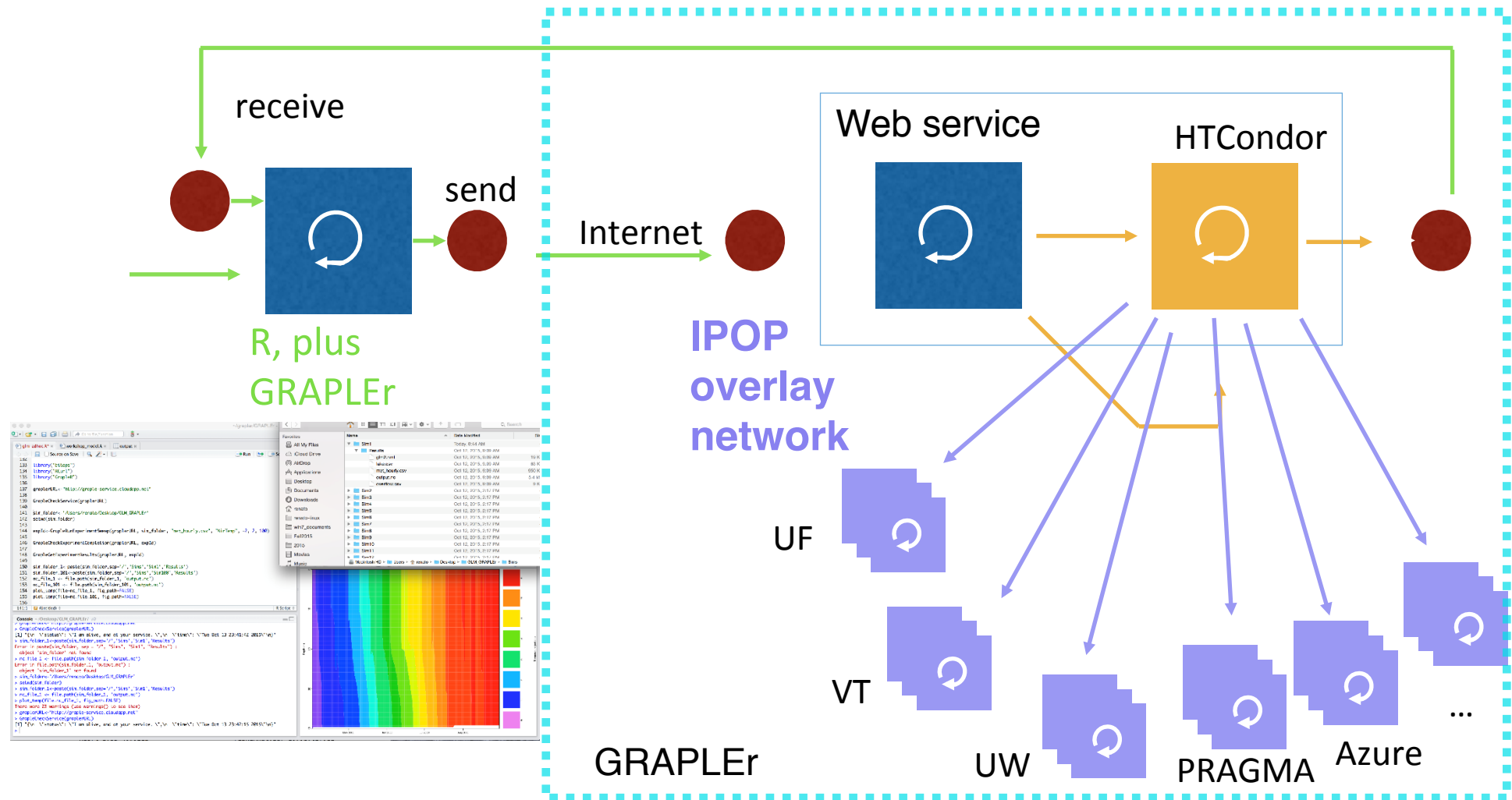
More details: demo/presentations, Friday

Extending PRAGMA-ENT with IPOP
- Collaboration with NAIST/UCSD (Ichikawa)
- Provide access to testbed through IPOP L2 mode

GRAPLEr
- Collaboration with UWisc (Hanson), VT (Carey)
- Web service + IPOP + HTCondor pool for lake modeling (GLM)
  - Development and use of R interface and Web service
  - Azure+UF; startup allocation in Comet/SDSC – will evaluate Virtual Cluster capability

GLM through GRAPLEr

receive

send

Internet

R, plus
GRAPLEr

**IPOP
overlay
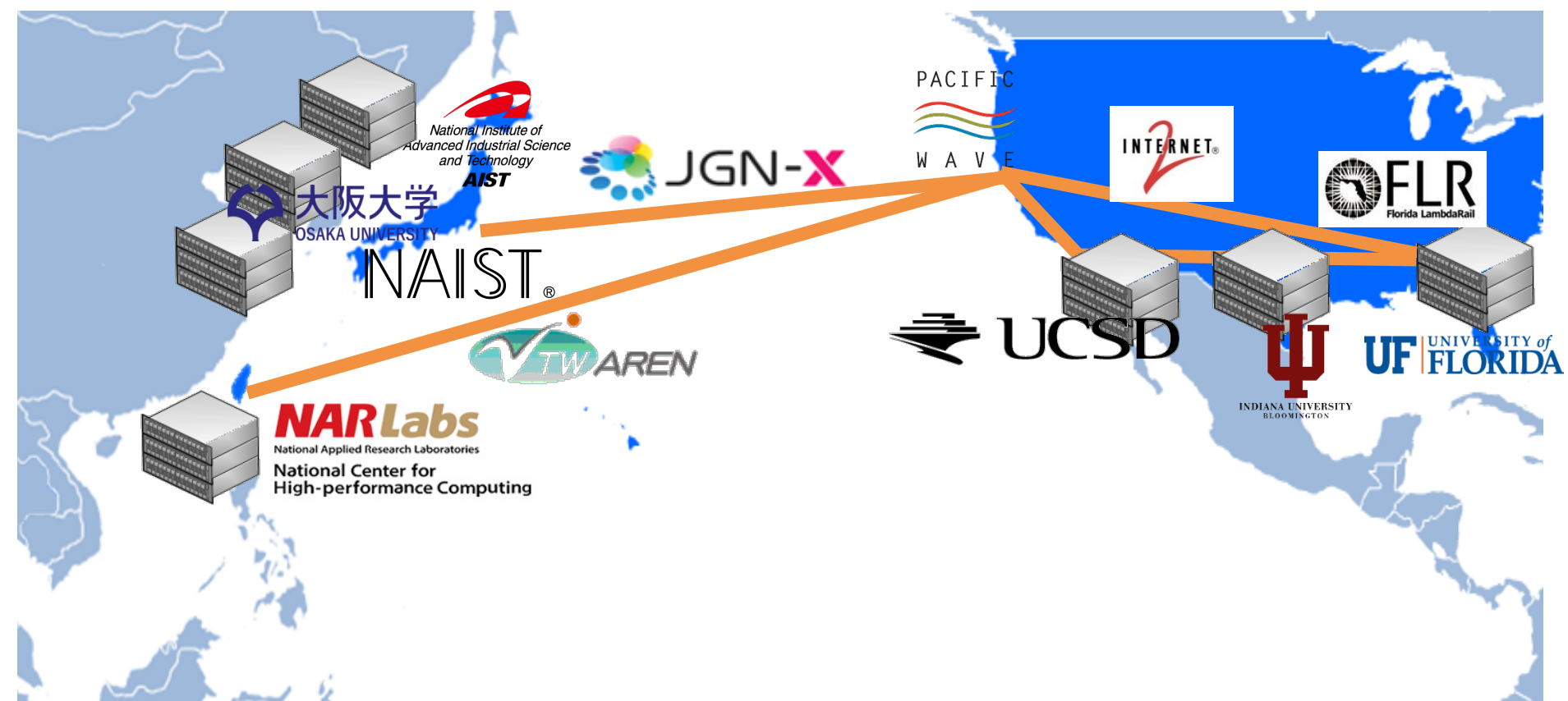network**

Web service

HTCondor

GRAPLEr

UF

VT

UW

PRAGMA

Azure

...

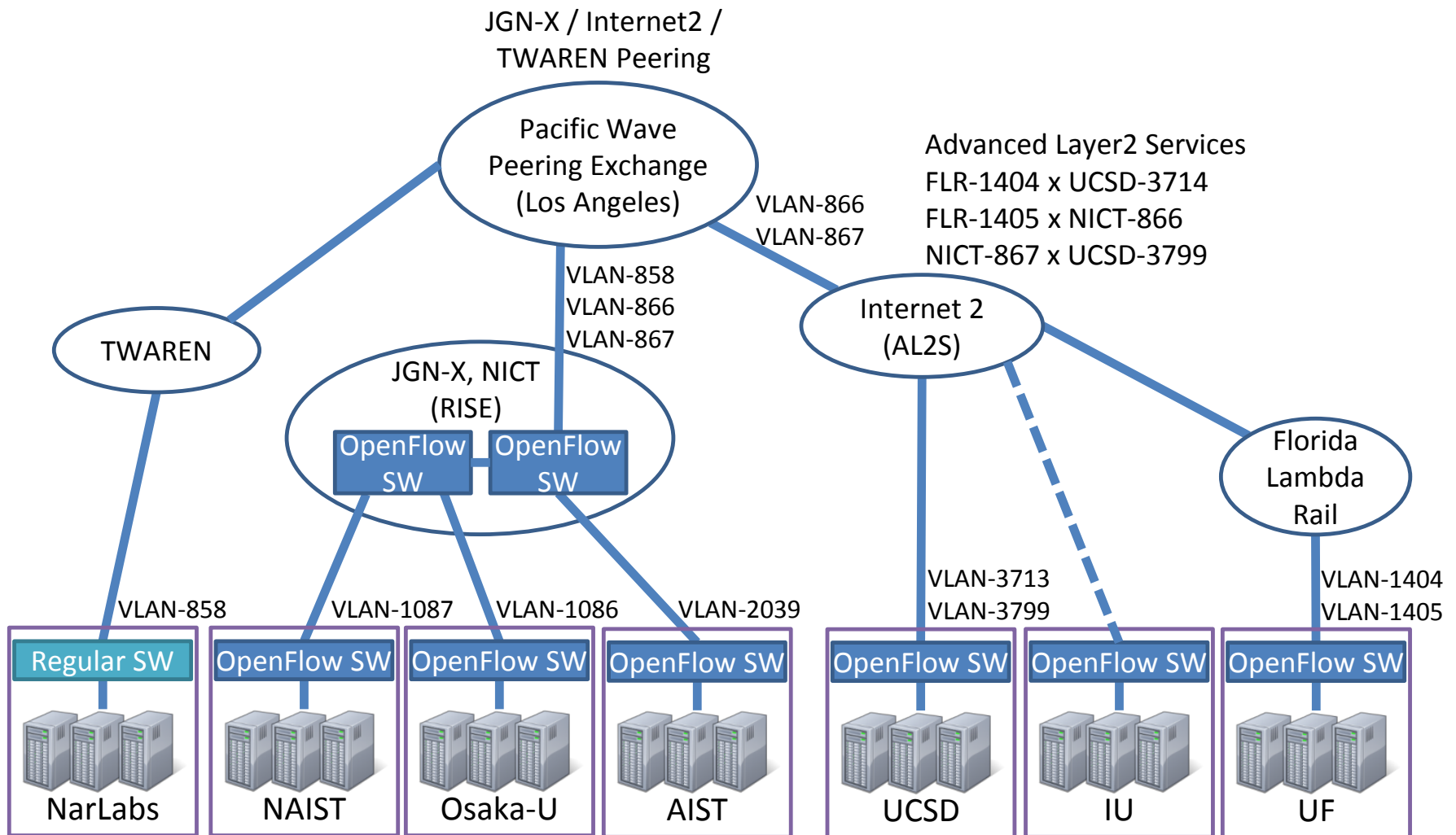# PRAGMA-ENT (Experimental Network Testbed)

An international SDN/OpenFlow testbed for use by PRAGMA researchers and collaborators

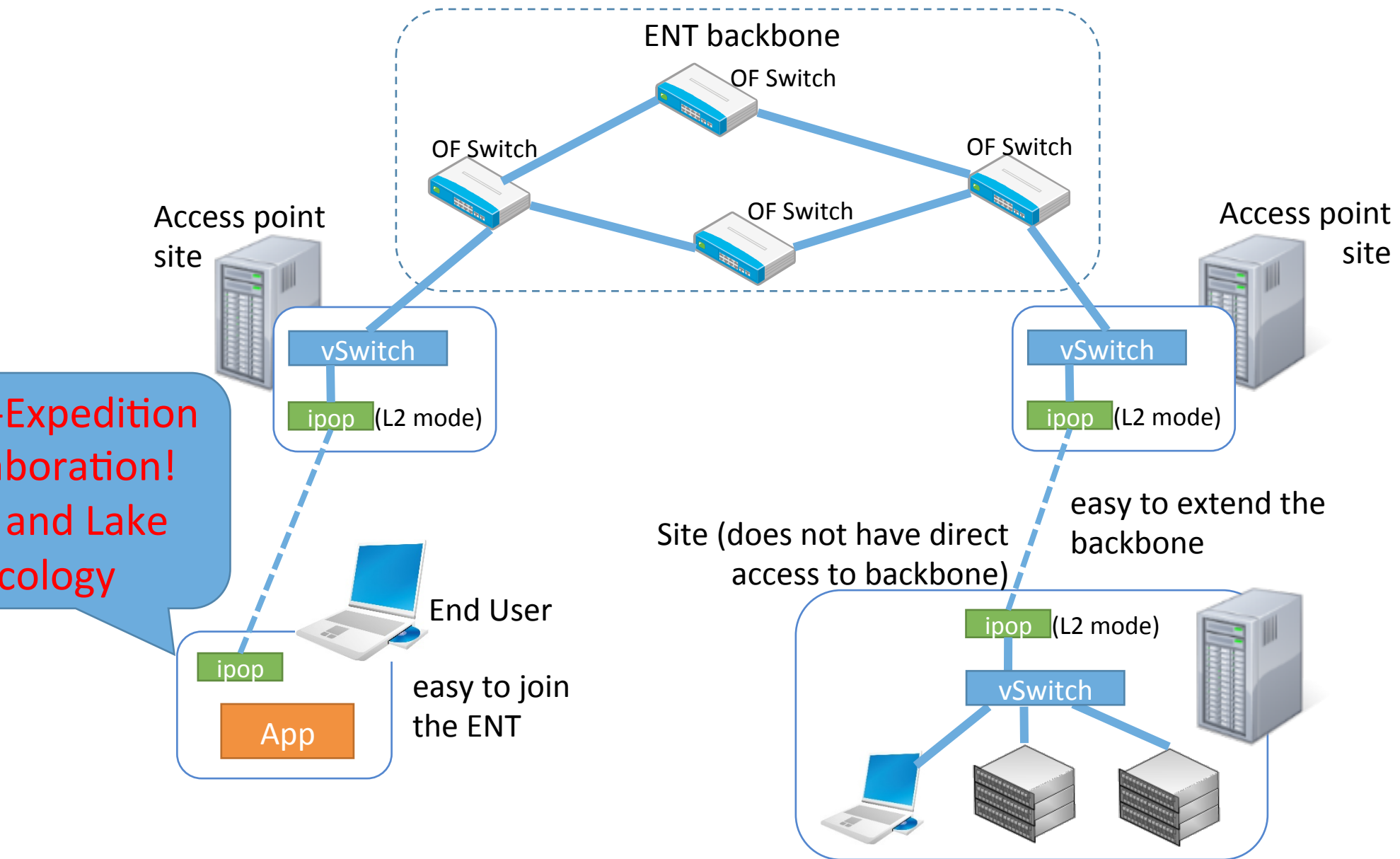- provides complete freedom to access and configure network resources

Backbone of PRAGMA-ENT

Currently, seven institutions are involved in the project. Their resources are connected each other through academic network backbones

JGN-X / Internet2 / TWAREN Peering

Pacific Wave Peering Exchange (Los Angeles)

Advanced Layer2 Services
FLR-1404 x UCSD-3714
FLR-1405 x NICT-866
NICT-867 x UCSD-3799

VLAN-866
VLAN-867

VLAN-858
VLAN-866
VLAN-867

TWAREN

JGN-X, NICT (RISE)

OpenFlow SW — OpenFlow SW

Internet 2 (AL2S)

Florida Lambda Rail

VLAN-858

VLAN-1087

VLAN-1086

VLAN-2039

VLAN-3713
VLAN-3799

VLAN-1404
VLAN-1405

Regular SW	OpenFlow SW	OpenFlow SW	OpenFlow SW	OpenFlow SW	OpenFlow SW	OpenFlow SW
NarLabs	NAIST	Osaka-U	AIST	UCSD	IU	UF

Activities: Extending backbone through IPOP (demo)

Activities: Usability study of ENT

Based on usability study, we have been re-organizing the PRAGMA-ENT documents on github wiki

- https://github.com/pragmagrid/pragma_ent/wiki

Documentation Then

Official PicOS documents – 34 pages

vSwitch documents – 257 pages

Single user thread

Documentation Now

PicOS documents for ENT – 4 pages

vSwitch documents for ENT – 5 pages

Multiple user thread
- Application user
- Infrastructure user

RDA-PRAGMA Sprint -- PRAGMA 30 @ Philippines

Quan (Gabriel) Zhou, Nadya Williams, Aimee Stewart

Jason Haga, Beth Plale

2/4/16

Objectives

- Assess recently released tools and best practices from RDA for contribution to PRAGMA services. Carry out assessment through 2 phase demo.

- Demo: verify lineage of projection data objects, and enable rerun when new data exists

- Enhancements to PRAGMA testbed: Provide common persistent identifiers and landing pages to VMs and datasets of Lifemapper

- Feed results back to RDA (Research Data Alliance)

Demo Phases

Phase I: Use static GBIF subset for Southeast Asia as input to Lifemapper,
- Input datasets bundled into VM.
- User has ID of two projection result datasets (both result sets have same internal ID (e.g., 317)), and uses RDA services to determine whether they came from the same VM or from the primary VM and its clone

Phase II:
- Input datasets ingested dynamically into VM (workflow dynamically accesses iDigBio.)
- After seeing change to iDigBio input dataset, use new PRAGMA data infrastructure to identify, download, and faithfully replay run with new iDigBio input dataset to visually compare before and after.

Adding a Significant Reproducibility Dimension to Lifemapper Experiments

New architectural components

PRAGMA-RDA Data Service

- Stores metadata, objects, and landing pages
- Maintains metadata about both data sets and VMs
- Assigns unique handle to incoming objects
- Displays landing page for each object
- Interacts with RDA PIT/DTR service
  - PID – Persisent Identifier, PIT = PID Information Types, DTR – Data Type Registry

RDA PIT/DTR service: stores type information about minimal metadata that allows interpretation of the metadata.

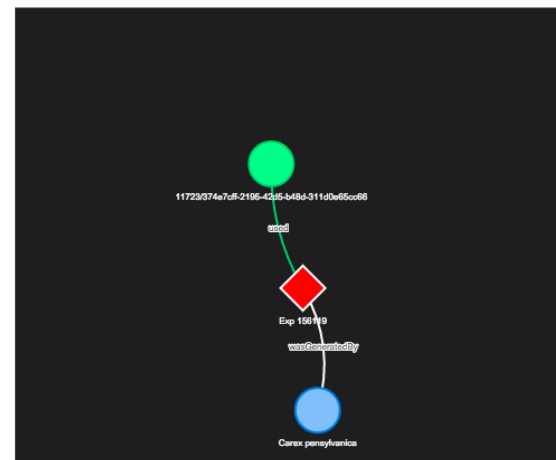Handle service : obtain handle PID

Landing Page

URL:

ndl.handle.net/11723/
[6e9-54f8-4c5c-9e95-054a26c](6e9-54f8-4c5c-9e95-054a26c)

ase example:

[/github.com/Gabriel-Zhou/](github.com/Gabriel-Zhou/)
dingpage

RDA-PRAGMA Landing Page for Lifemapper

ProjectionSet ID	7575077
Display Name	Carex pensylvanica
Scenario Code	WC-10min
Bounding Box	-180.0, -60.0, 180.0, 90.0
Resolution	0.16667
Last Modified	2015-12-16 20:32:54
OccrrenceSet PID	11723/374e7cff-2195-42d5-b48d-311d0e65cc66
Experiment ID	156119
Checksum	f540c8cc528596967fde3c9925e140c9

Download   Go to Occurrence Set

Projection Set Provenance

CheckSum Information Type

PID: 11723/377739b4-14df-441a-b219-15881cf6ae52

U Data Node Resources – Pathfinder for Data-centric nodes in PRAGMA

- RDA PIT/DataType Registry Service
- PRAGMA-RDA Data Service
- Handle V8 service (generates PIDs, runs at CNRI)
- PRAGMA-ENT Mesh
- Open HathiTrust Corpus

Cloud Scheduler Updates

- **Goal: Provide a lightweight cloud scheduler for sharing of PRAGMA resources**

- New: Migration from prototype to (very) early users

  - Setup Fiji (UCSD/SDSC) as our production cloud scheduler server

  - Initial integration of NCHC Ezilla deployment using OpenNebula pragma_boot driver

  - Added automated shutdown of virtual clusters
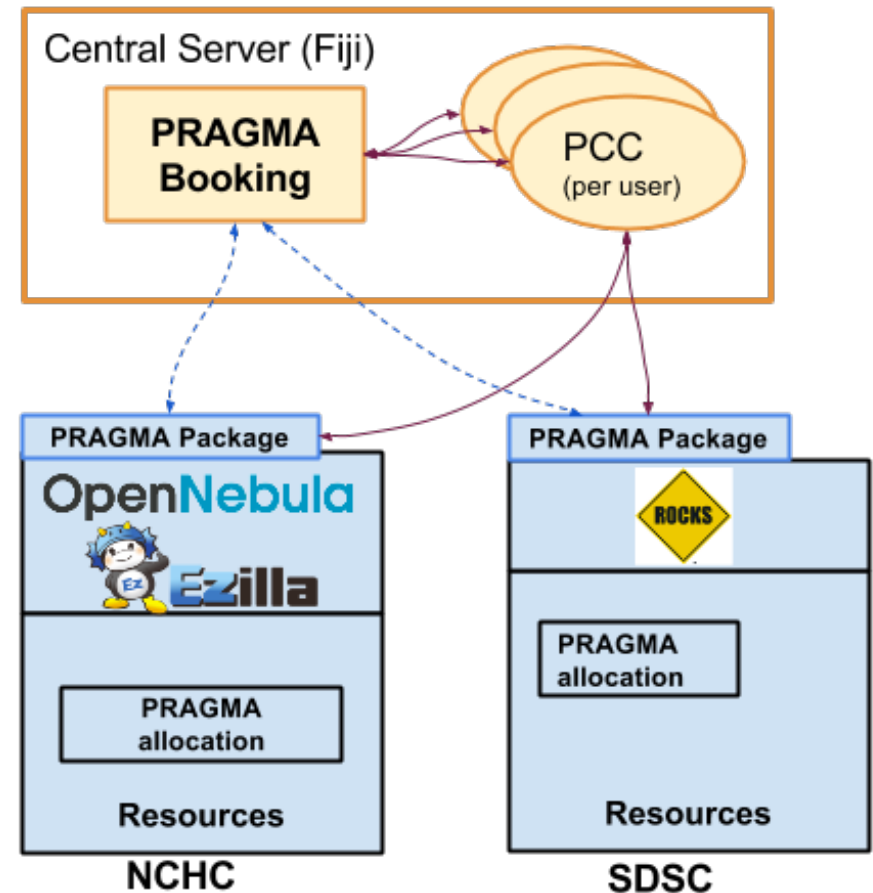
Pragma_boot updates

- Goal: Boot virtual clusters for users across PRAGMA institutions using local VM provisioner

  - **New**: Rewrote pragma_boot framework Python scripts and kvm_rocks driver (previously bash scripts)

  - **New:** Enhanced vc-out-parser to work for Ubuntu

  - **New**: Made virtual cluster images of the hku_biolinux.img from Biosciences WG (frontend and compute images) and leveraged Clonezilla to create sparsified ZFS volumes

Cross – working Group interaction

Using Tech from NCHC Taiwan

ractically dealing with Large disk images

- evious versions of PRAGMA boot had to fully copy Frontend and
- ompute Node images to remote hosts BEFORE booting
  - → Could take 30 – 60 minutes to boot a virtual cluster

- SC Developed tech –
  - Boot a virtual machine with its system disk homed on a central disk server
  - Migrate image to remote. Then ever iSCSI link → VM disk access become local

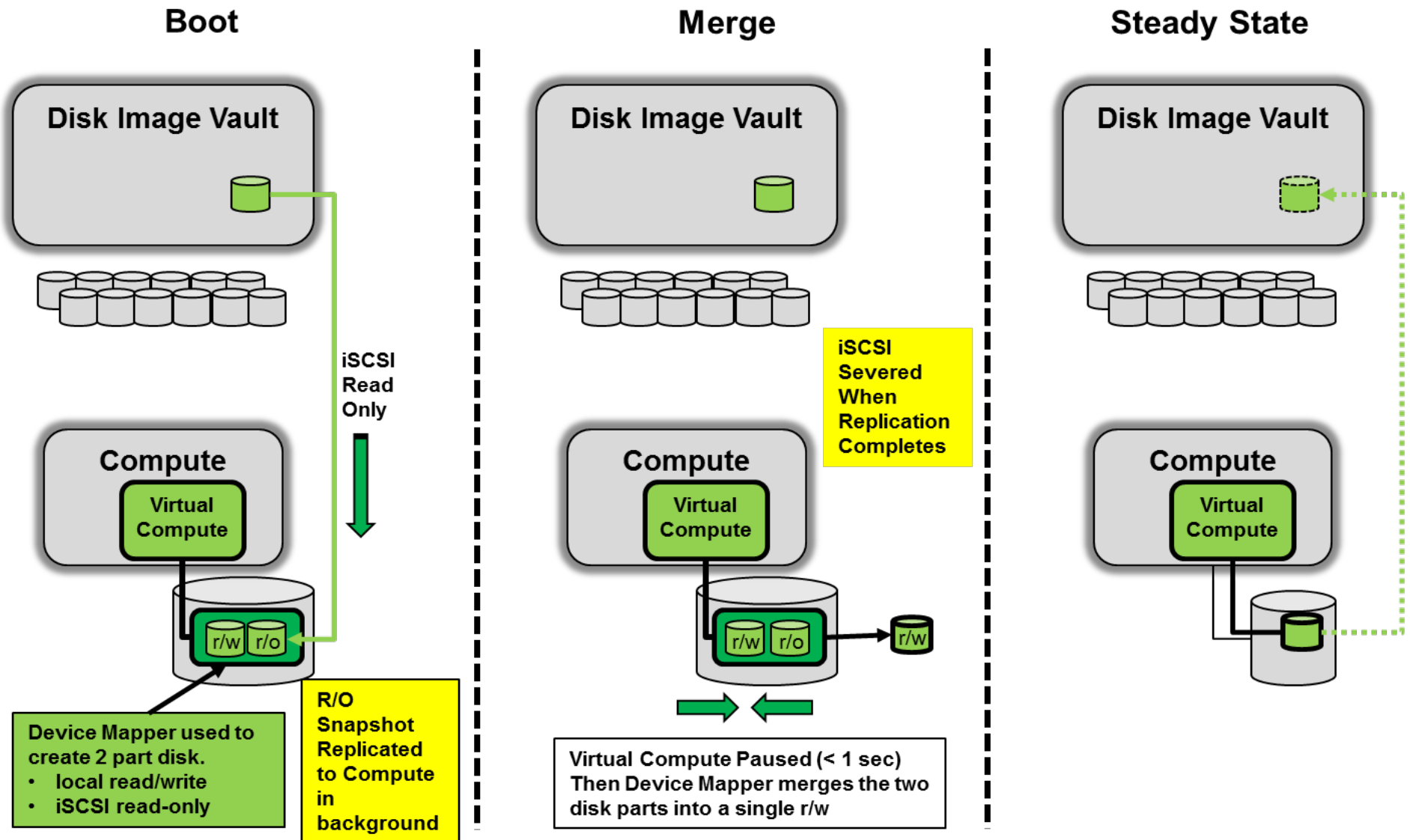- dapted for PRAGMA to support cloning of compute node images
  - → **Frontend** + 2 **Compute** node boot in less than 90 seconds!

```
-01-27 03:08:03,026 INFO Loading driver pragma.drivers.kvm_rocks.Driver

-01-27 03:08:30,320 DEBUG Executing command: '/opt/rocks/bin/rocks start
ocks-52'
-01-27 03:09:25,615 DEBUG Executing command: '/opt/rocks/bin/rocks start
m-rocks-52-1'
```

everaging and expanding tech from SDSC

Boot

Disk Image Vault

iSCSI
Read
Only

Compute

Virtual
Compute

r/w r/o

Device Mapper used to
create 2 part disk.
• local read/write
• iSCSI read-only

R/O
Snapshot
Replicated
to Compute
in
background

Merge

Disk Image Vault

iSCSI
Severed
When
Replication
Completes

Compute

Virtual
Compute

r/w r/o → r/w

Virtual Compute Paused (< 1 sec)
Then Device Mapper merges the two
disk parts into a single r/w

Steady State

Disk Image Vault

Compute

Virtual
Compute

Advantages

Virtual Machines boot very quickly
- Disk Image on Storage Vault is replicated in background

On Shutdown, final disk state is returned to vault

All Disks are lazily replicated
- If physical compute node fails, state of virtual disk is close to "up to date"

Utilizes the parallel I/O busses of all the compute nodes
- Disk vault can be built using commodity components.

→ More Details in Working Group

Planned Topics for PRAGMA 30

Beth - RDA, where are we, what are the near term plans, mid-term goals, how others can get involved

Shava - How PRAGMA boot works, thoughts on using Clonezilla within the workflow.

Renato/Paul/Cayelan -- How to get others involved in the Lake Eco expedition. Contribute resources, help find domain collaborators

Kohei - PRAGMA ENT and work done to make it easier to add users. Detailed status.

Yoshio -- concrete steps to rebuild the PRAGMA cloud

Phil - Update on image storage used on comet (reducing dependency on Rocks)