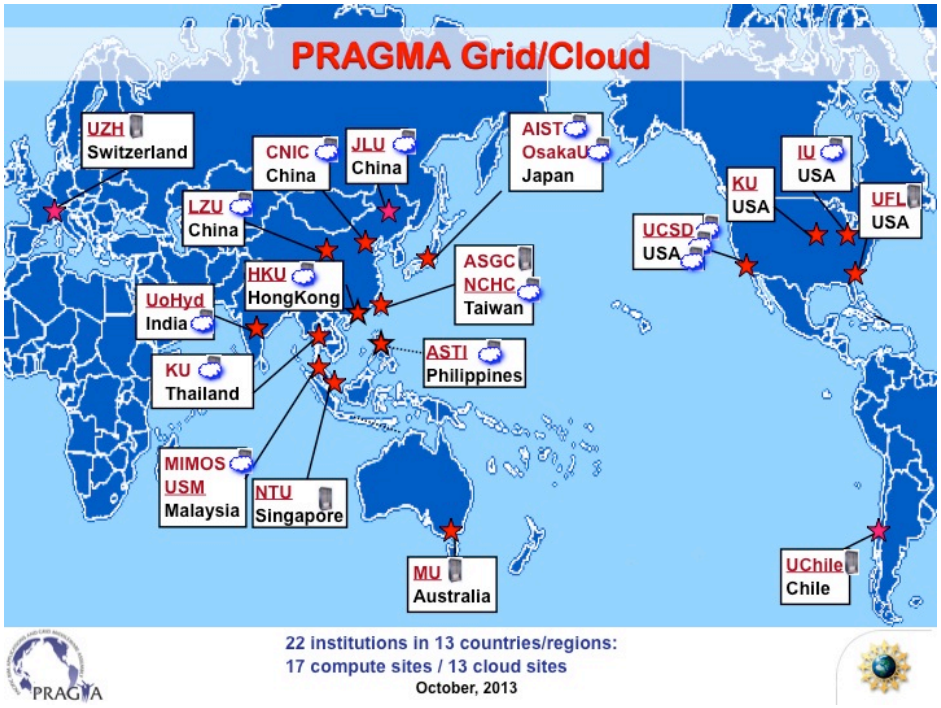
Resources Working Group Report

PRAGMA 27

This Map is out of date



- As a group we have effectively stopped using Globus
- A number of sites have changed personnel
- We have been effectively working on component technologies.
- Time for the next step

Rebooting the Persistent PRAGMA Testbed

- PRAGMA-wide persistent testbed resources
  - Various technologies are now "good" enough that we can profitably revisit.
    - PRAGMA Boot, Cloud Controller, Overlay Networks
  - Discussion was about
    - What, How, Who
  - Two types of natural testbed resources
    - Pure VM and Virtual Cluster Hosting.
    - VM and Virtual Cluster Hosting w/ programmable networking.
  - Need to be able to integrate data resources (particularly important in Biodiversity expedition)

Areas of Discussions

- **What cloud deployment technologies are people using today?**
  - Rocks (UCSD), OpenStack, OpenNebula (NCHC),CloudStack (AIST, UFL, KU (thai),
- **How should we handle accounts?**
  - something simple. Recognize that we don't need to scale to 1000s of users.
    - Central place for usernames, contacts, public ssh keys?
    - Eduroam? Leveraging FutureGrid? OpenID? (weicheng to investigate)
    - Centralized SSH pubkey list via private GitHub Repository (Via an academic)
- **Access? How do we determine who/when a remote user can spin up a virtual cluster for an experiment**.
  - Can we steal from HPC schedulers? What about Condor Scheduler?
  - Components from INCA that can be used to detect collisions.
  - Central Place for Reservations/Availability? ORCA from GENI/Planetlab?
  - Leases? Shava + Jose to investigate
- Record keeping/performance
  - Tracking.
  - perfSONAR --- mesh config (John Hicks to assist in PRAGMA mesh config).

What's an endpoint look like?

- a Node to start/stop VMs/VC's
- N nodes to host VMs (N >= 4)
  - Support KVM.
- perfSONAR host in the Rack (perhaps a VM on the head/ submit node? – Public IP)
  - Probably need 8 available Public IP addresses
  - v6 addressing (unlimited!)
- Sites:
  - UCSD, UF, NAIST, Internet2, CNIC?, Osaka, NCHC, IU (when Staffing issues are sorted)
    - UCSD (Luca, Phil, Nadya) to coordinate

Goals by Next PRAGMA

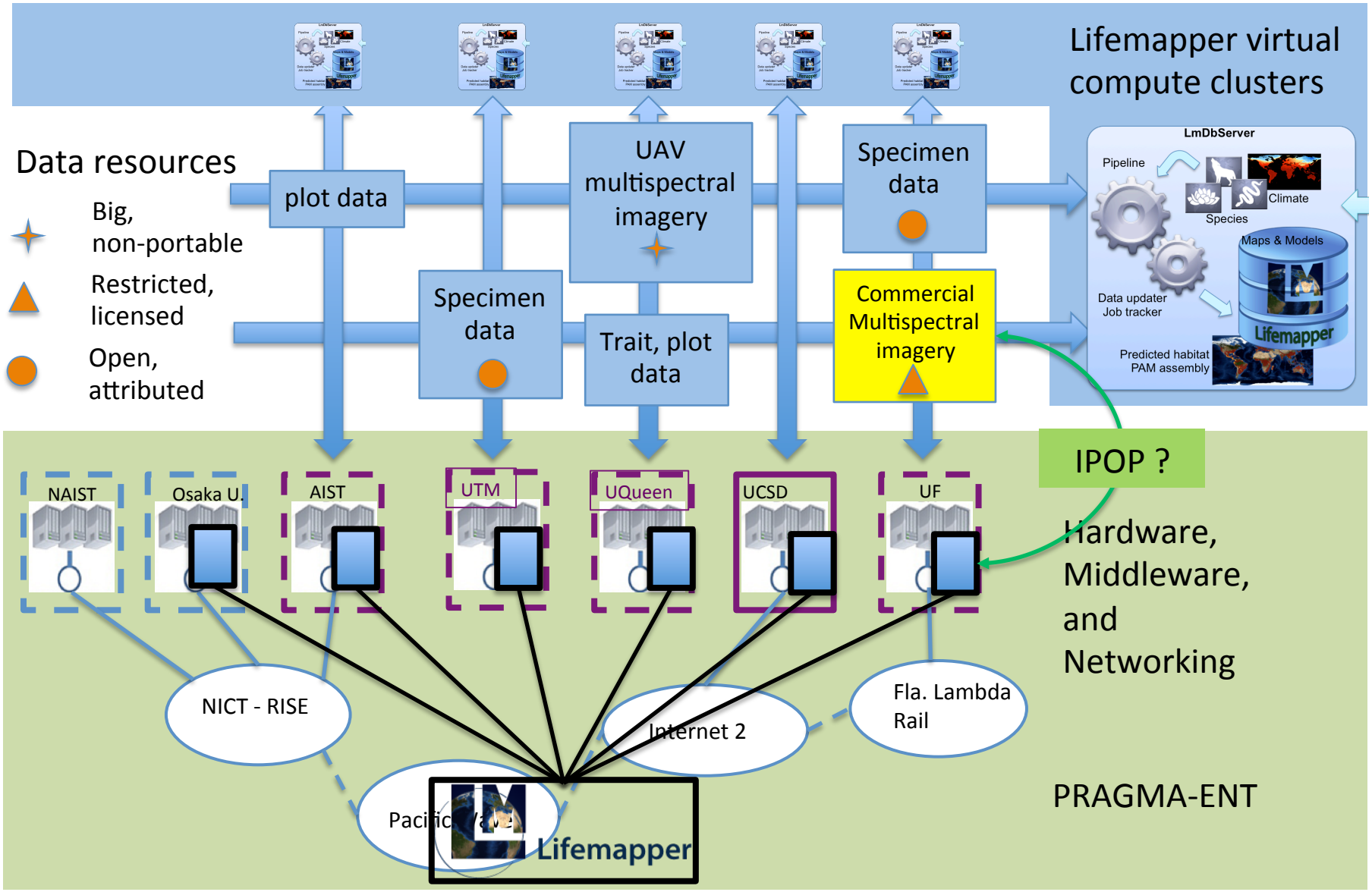
- Resources Defined/Deployed
- PRAGMA_BOOT at all sites
- VM Images stored in Amazon S3 (Possibly Available through CloudFront
- GitHub (Private) of users + ssh public keys
- perfSONAR PRAGMA Mesh Deployed at all sites
- Simple (one email?) request for resource allocation at sites.

Lifemapper + ENT + Specialized Data (Overnight Hike)

Overall Goal: Run Lifemapper Cluster on PRAGMA ENT w/access to private satellite data

- Data in Florida (Data is NOT on ENT Resources, but close to Florida Resources)
- Have many of the component technologies but, the integration will drive activity and uncover specific issues
  - Need to extend the ENT openflow overlay to the data server
    - IPOP integration? To be worked out
  - ENT needs to have slicing so that we can have a (Semi) stable experiment env.
  - Some work to be done in Lifemapper itself to integrate the new Satellite Data

- Demonstrated capability by Pragma 28 (2)
  - Sciene output: run/insight gained by using the data (Reed)
  - Technology Demo, all the pieces needed to make this work (the rest of us)
- People
  - ENT: Kohei, Mauricio, Pong? Chawa? (for controller expertise)
  - IPOP: Renato
  - Cluster deployment: Aimee and Nadya
  - domain science: Reed


Lifemapper virtual compute clusters
Data resources
Big, non-portable
Restricted, licensed
Open, attributed
plot data
UAV multispectral imagery
Specimen data
Specimen data
Trait, plot data
Commercial Multispectral imagery
LmDbServer
Pipeline
Climate
Species
Maps & Models
Data updater
Job tracker
Lifemapper
Predicted habitat
PAM assembly
IPOP ?
NAIST
Osaka U.
AIST
UTM
UQueen
UCSD
UF
Hardware, Middleware, and Networking
NICT - RISE
Internet 2
Fla. Lambda Rail
Lifemapper
PRAGMA-ENT

Discussion about Resources WG and Other working groups

- Possible (Probable) resources WG ½ day meeting before usual 2-day meeting
- For at least one working group session during meeting, Resources will NOT meet
  - members of resources will participate directly in other working groups
    - E.g Telescience/Geosciences
    - Biosciences
    - Cyber-learning
  - Goal: learn more of what others in PRAGMA need/ desire from resources.