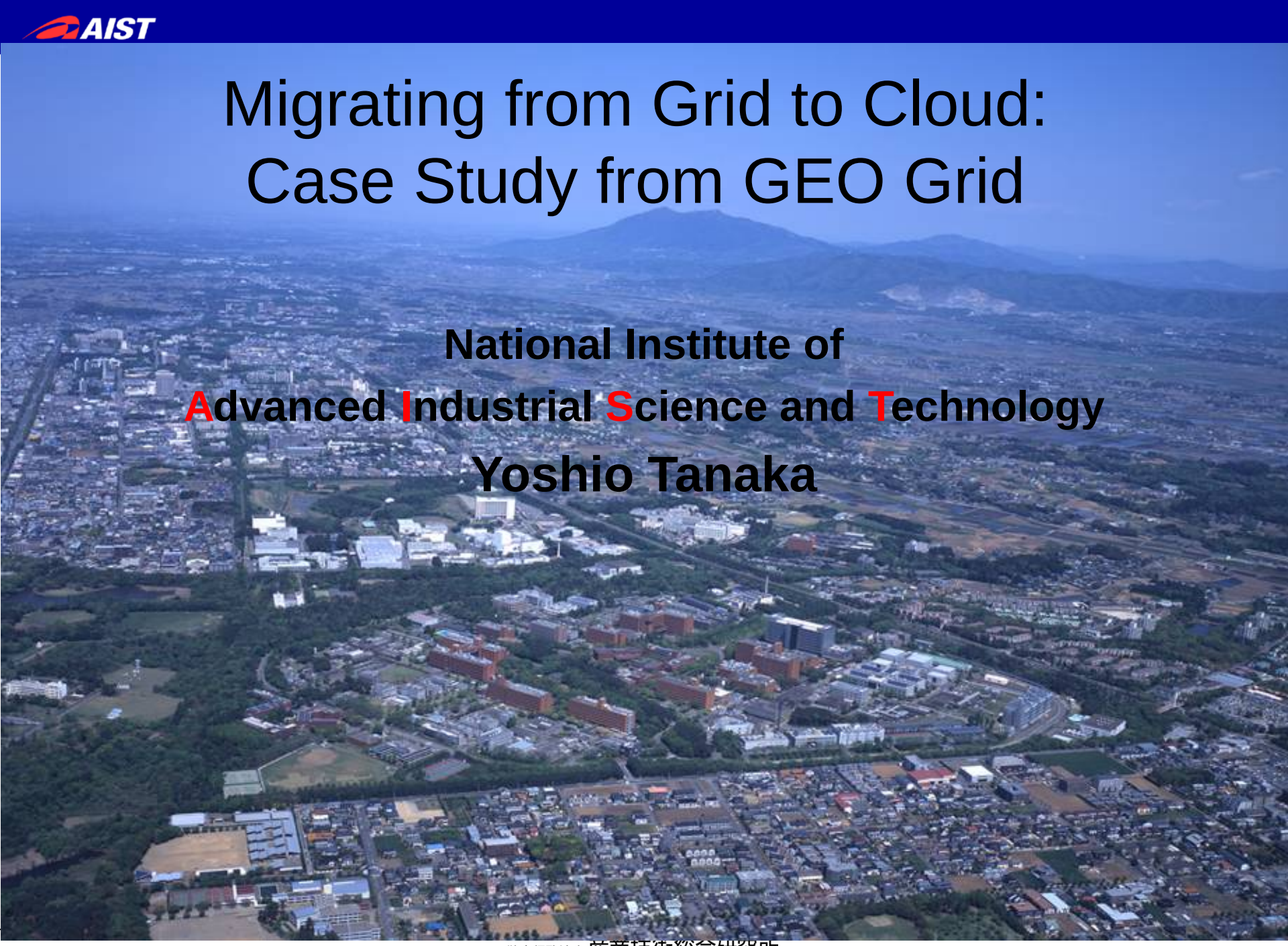
Migrating from Grid to Cloud: Case Study from GEO Grid



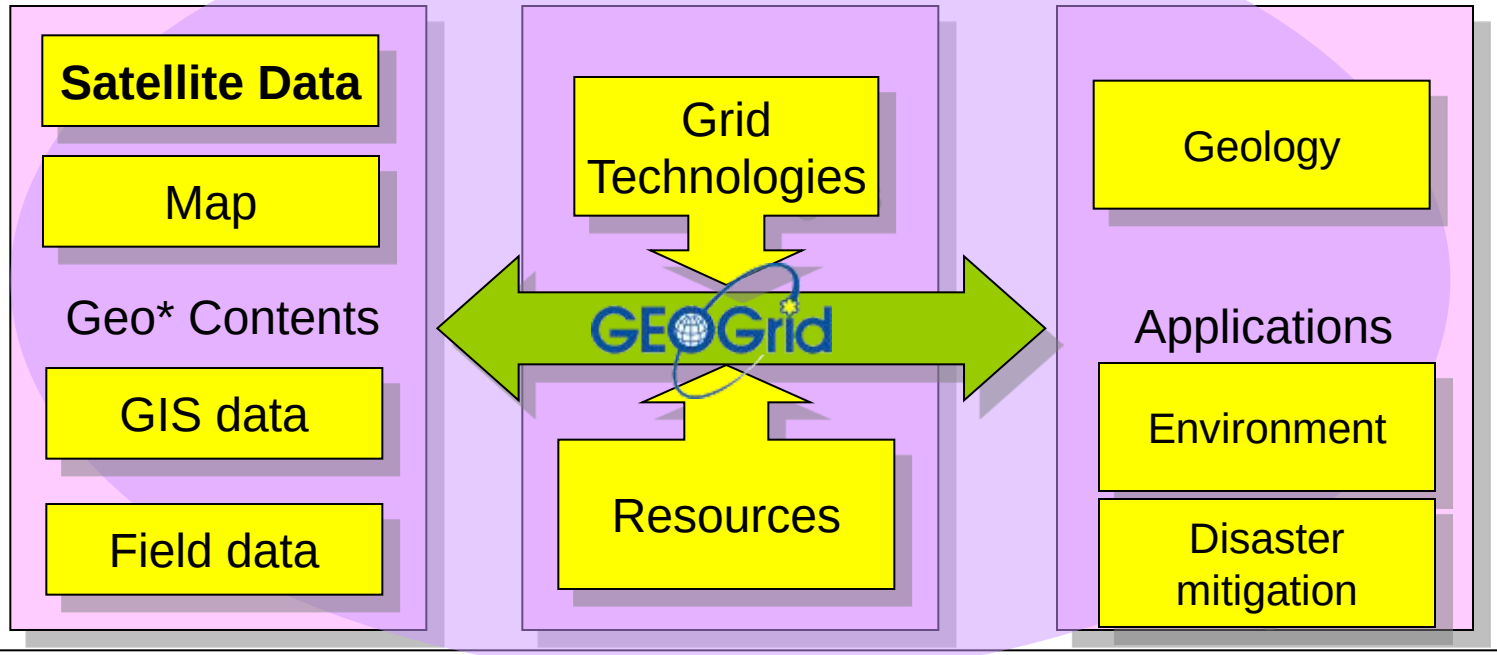
National Institute of
Advanced Industrial Science and Technology
Yoshio Tanaka

What is the GEO Grid ?

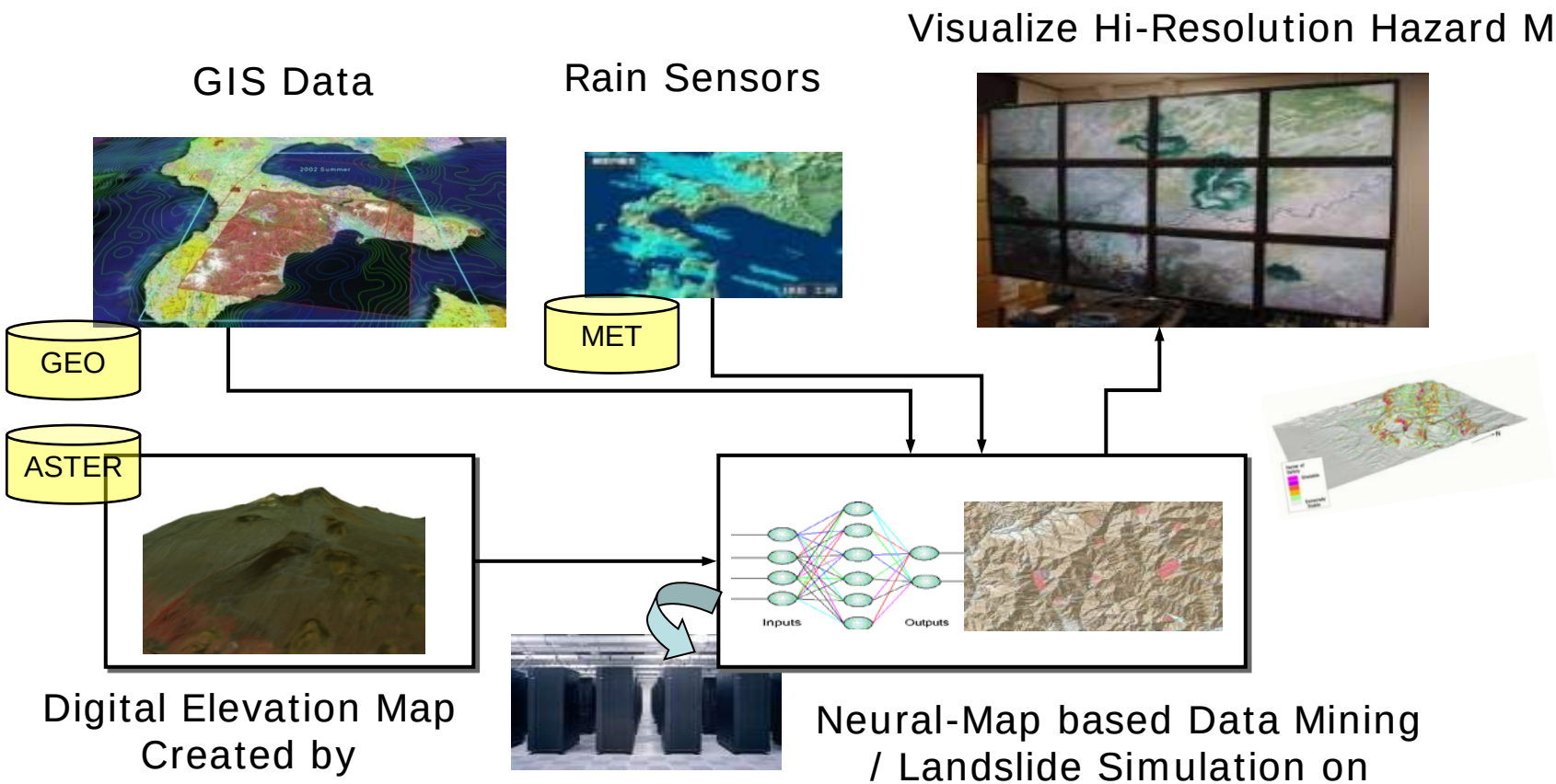
- The GEO (Global Earth Observation) Grid is aiming at providing a *Cyber Infrastructure* for worldwide Earth Sciences communities to accelerate GEO sciences based on the concept that relevant data and computation are *virtually integrated* with a certain *access control* and ease-of-use interface those are enabled by a set of Grid and Web service technologies.

AIST: OGF Gold sponsor (a founding member)

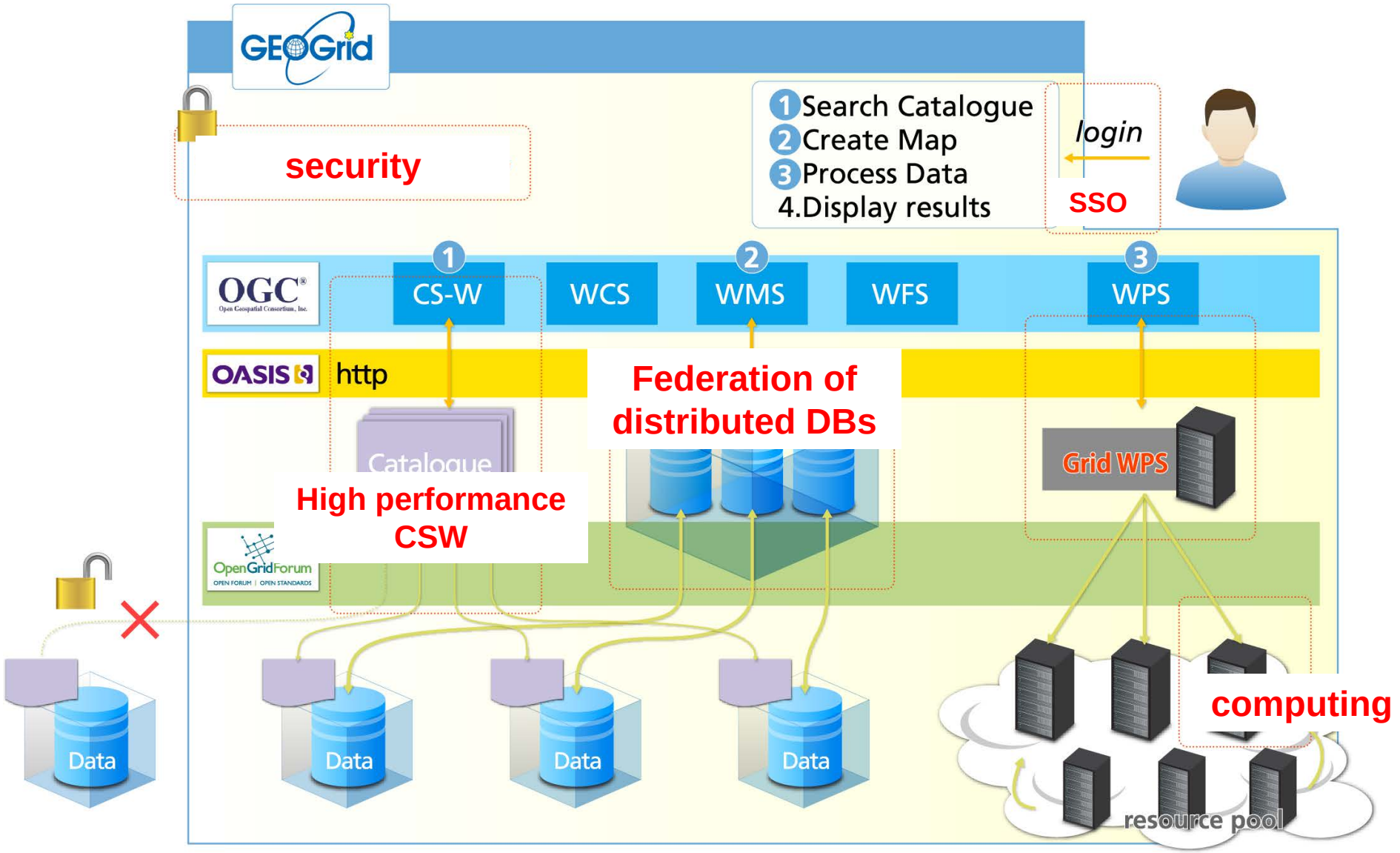
AIST: OGC Associate member (since 2007)



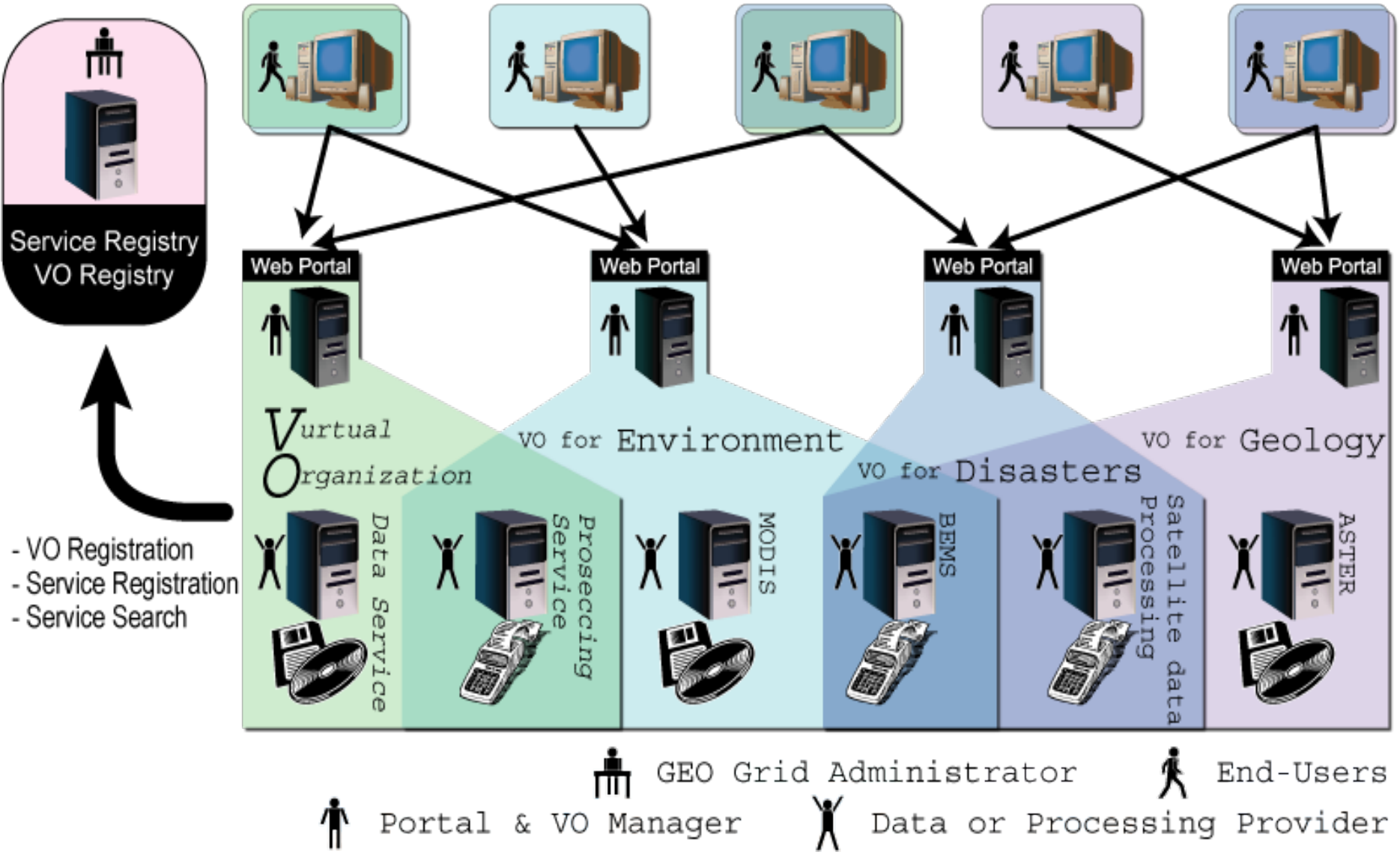
Example: Flood simulation



Why Grid? – federation of distributed resources -

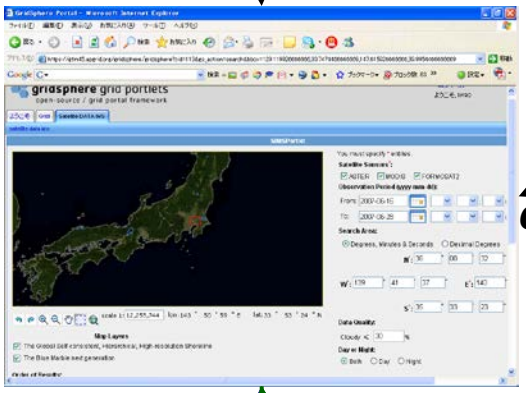


GEO Grid Security: GSI + VOMS



login

user



portal server

credential

account (GAMA) server

Account DB

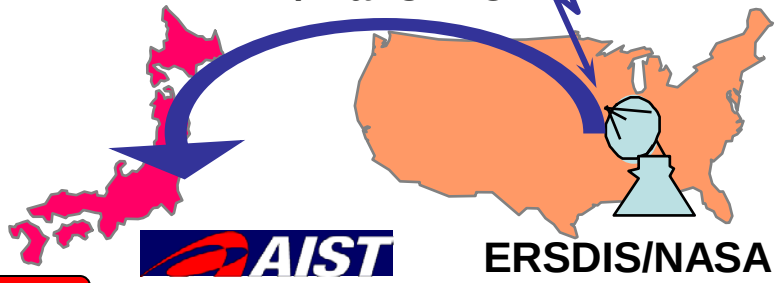
VO (VOMS) server

VO DB

Terra/ASTER

TDRS

APAN/TransPAC



AIST

ERSDIS/NASA

GET

exec

query

GSI + VOMS

GSI + VOMS

GSI + VOMS

WFS

WCS

WMS

CSW

OGSA DAI

GRAM

GridFTP

GIS server

map server

catalogue/ metadata server

gateway server

GEO Grid Cluster

GRID DATA FARM

L0

L0

L0

L0

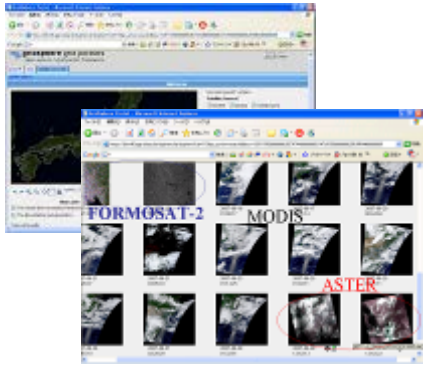
Data

Maps

Meta data

Storage (DEM)

Demo Environments in 2007 - SIMS (ASTER+MODIS+Formsat2)



Java Program

Integration Framework with OGSA-DAI

OGSA-DAI Client

SIMS portlet
- **query data**
- **create web page which shows thumbnail images**

SQL

SQL

SQL

SQL

SQL

Application Server

Globus

OGSA-DAI

VOMS

VOMS

Globus

OGSA-DAI

OGSA-DAI

SQL w/ JDBC

Database Server (Sybase)

FORMOSAT-2

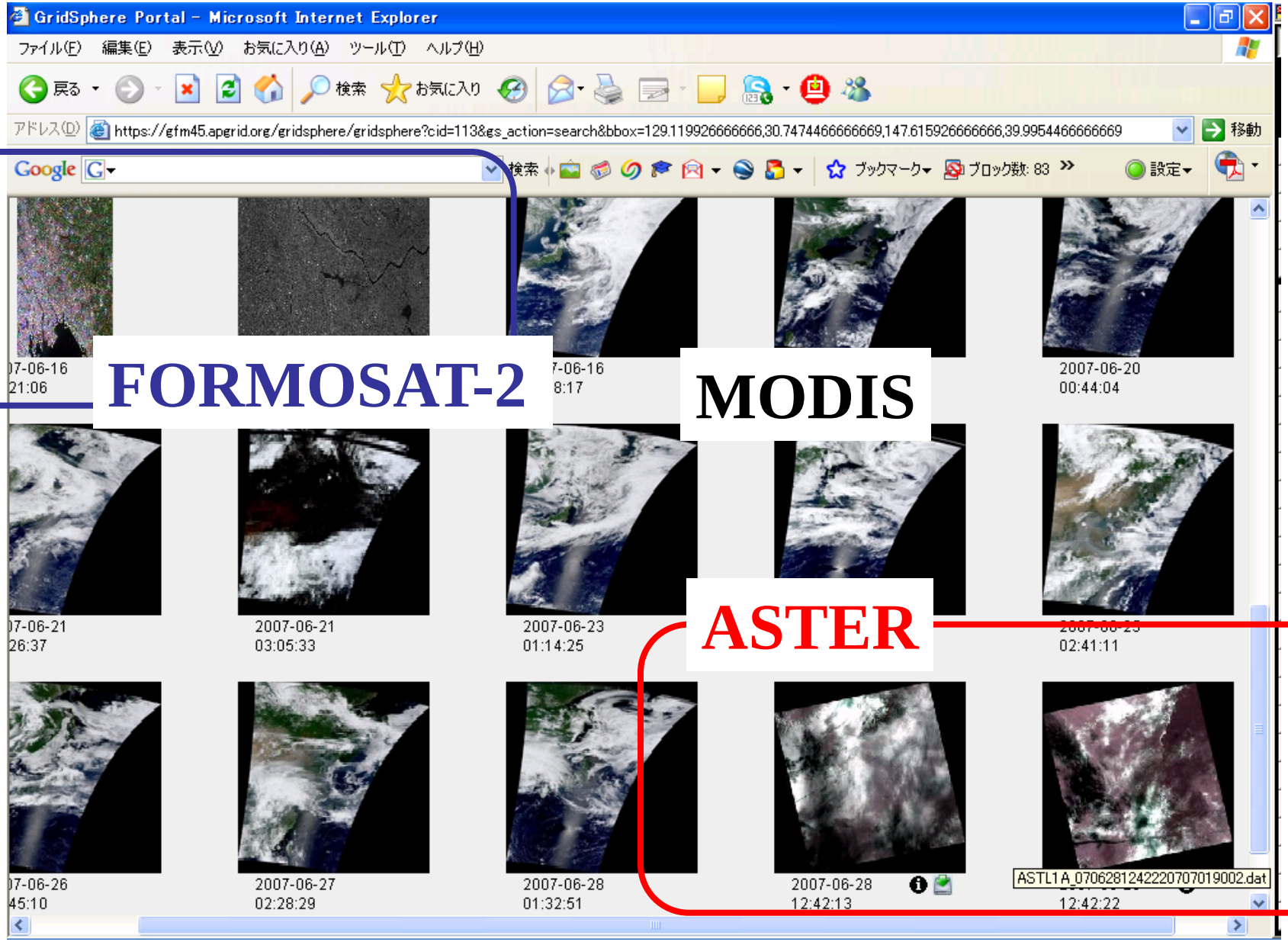
SQL w/ JDBC

Database Server (PostgreSQL)

ASTER

MODIS

NSPO@TW AIST@JP


GridSphere Portal - Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
戻る 検索 お気に入り
アドレス(D) https://gfm45.apgrid.org/gridsphere/gridsphere?cid=113&gs_action=search&bbox=129.119926666666,30.7474466666669,147.615926666666,39.9954466666669
移動
Google 検索 ブックマーク ブロック数: 83 設定
2007-06-20
00:44:04
2007-06-21
03:05:33
2007-06-23
01:14:25
2007-06-26
45:10
2007-06-27
02:28:29
2007-06-28
01:32:51
2007-06-28
12:42:13
ASTL1A_0706281242220707019002.dat
12:42:22
FORMOSAT-2
MODIS
ASTER

GEO Grid Service Examples

- Satellite data archive and processing
  - ASTER, PALSAR, MODIS, etc.
- Satellite data application
  - Application of Satellite-Field data Integrator (SFI) for aerosol monitoring Description http://fon.geogrid.org/aerosol/
  - SDCP (Science Degree Confluence Project) –Community validation tool for global land-cover & digital elevation models http://eco.geogrid.org/sdcp/
- Hazard information
  - QuiQuake (Quick Estimation System for Earthquake Maps Triggered by Observation Records) http://qq.ghz.geogrid.org/QuakeMap/index.en.html
  - Volcanic Gravity Flow Simulations on Volcanic Area http://volcano.geogrid.org/applications/EnergyCone/
- Geoscience data
  - Geological maps, Active fault data, etc.

Migration from Grid to Cloud

Motivation for migrating to Cloud

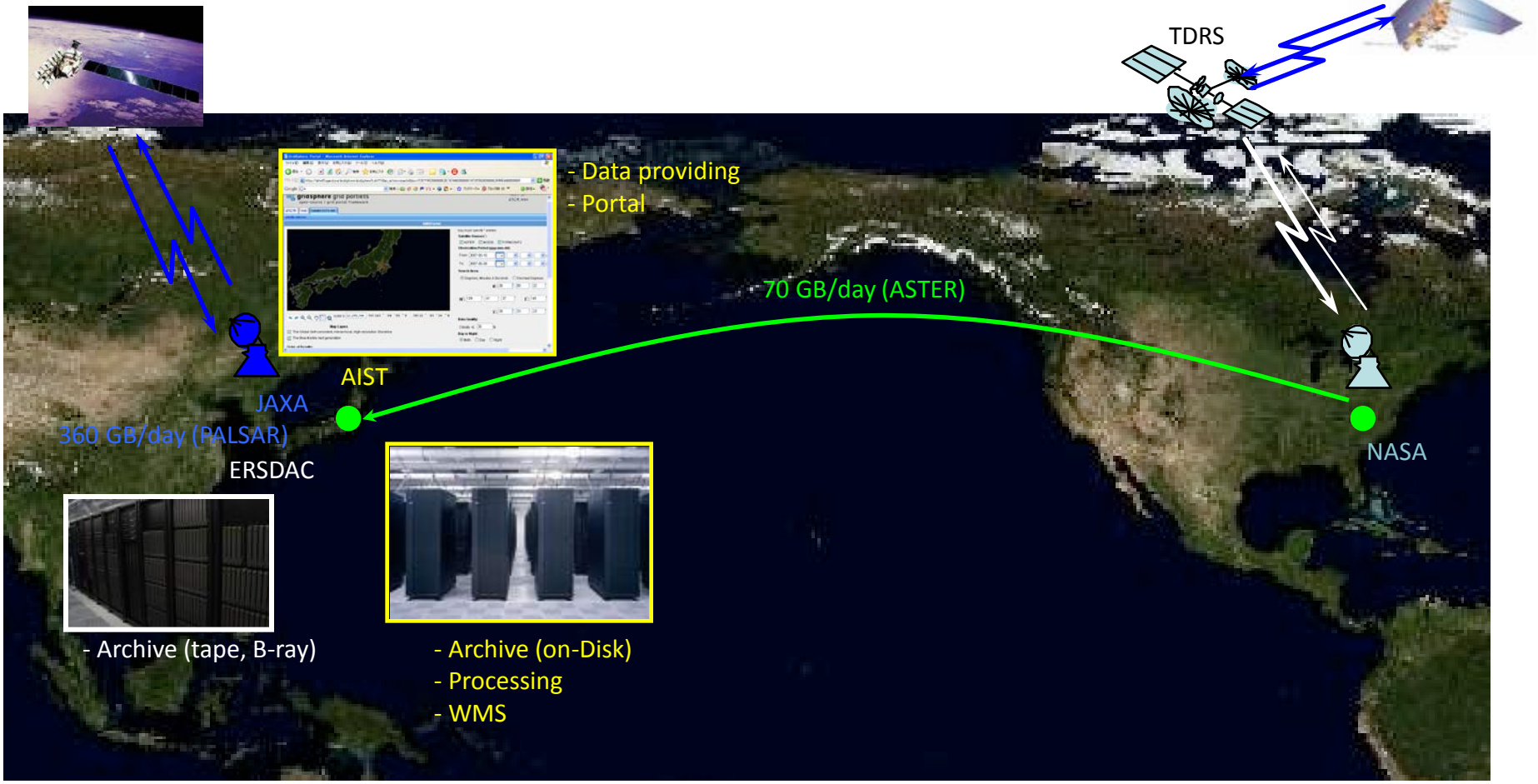
- Deployment of applications is not easy
  - Procedures for including new resources (deployments of applications) are troublesome.
    - Need easy-to-use.
    - Write once, run everywhere!
- Do we need Grid protocols?
  - Do we need Grid Security?
    - Delegation is necessary for third-party file transfer.
    - But key management is burden for end users.
    - Installation/configuration of VOMS is not easy.
  - Do we need Grid protocol (e.g. GRAM)?
    - GEO Grid applications use not Grid middleware/protocol but the other standards (e.g. OGC).
- Need to adapt the direction for wider use
  - GEO Grid system is stably in operation, but not extendable (elastic).
    - Data server and computing server are tightly coupled.
    - It's hard to use resources outside organization.
  - Is GEO Grid Design appropriate for use by business partners?
  - Japanese government has a plan of promoting use of satellite data for wide use.

Goals of and approaches by PRAGMA

- Enable Specialized Applications to run easily on distributed resources
  - Build once, run everywhere!!
- Investigate Virtualization as a practical mechanism
  - Supporting Multiple VM Infrastructures (Xen, KVM, OpenNebula, Rocks, WebOS, EC2)
- Share VM images in PRAGMA VM repository so that we can boot our application VMs at any site by any PRAGMA colleagues.
  - Discussed in PRAGMA 20 workshop @ HK, March 3rd and 4th, 2011, 1 week before the big earthquake in Japan…

2011 Tohoku Earthquake changed our R&D environments

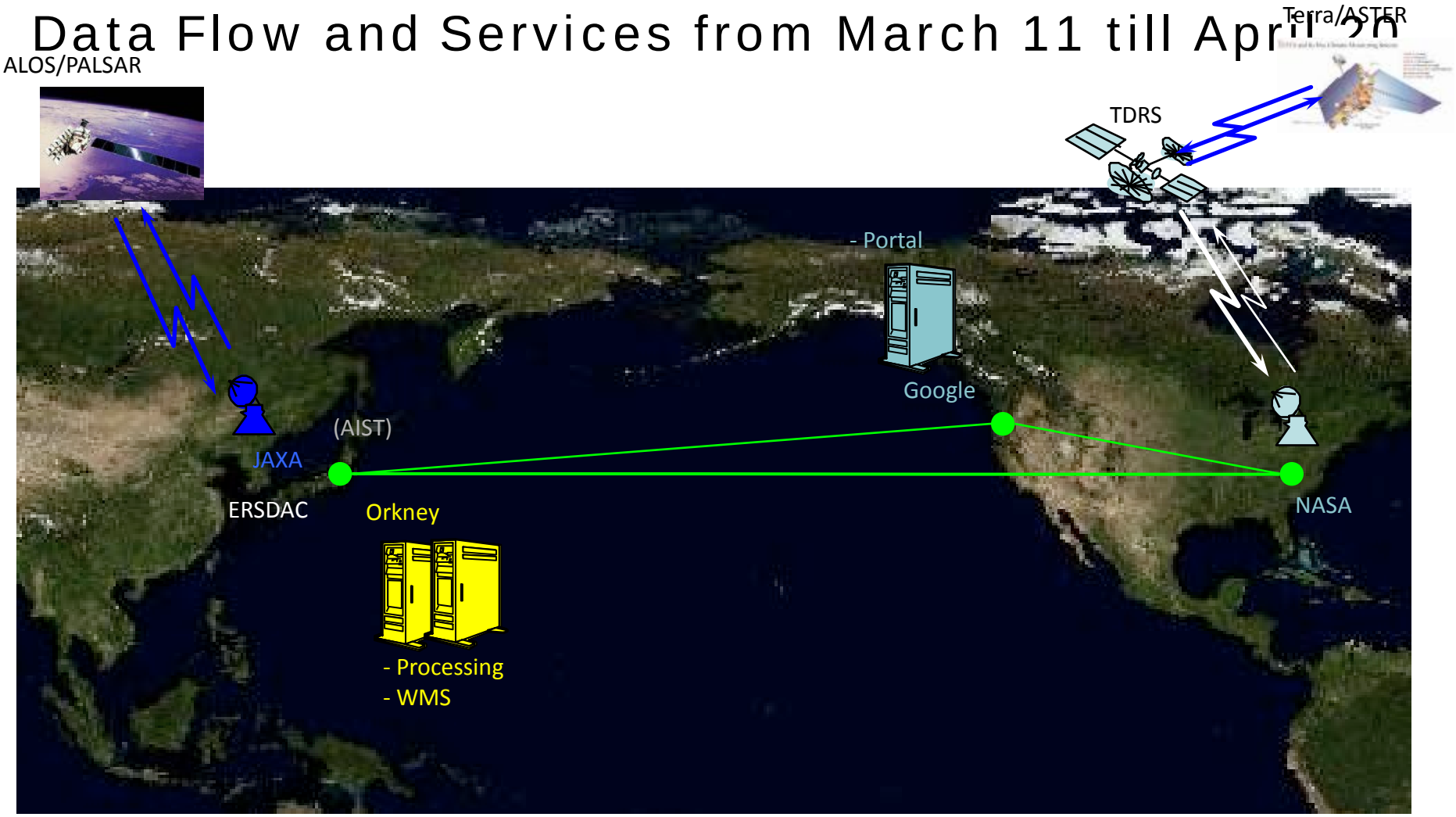
Satellite Data Flow and Services Prior to March 11



- **ASTER data: NASA→ERSDAC→AIST**
- **PALSAR data: JAXA→ERSDAC→AIST**

(processing, WMS, portal site, and data providing by AIST)

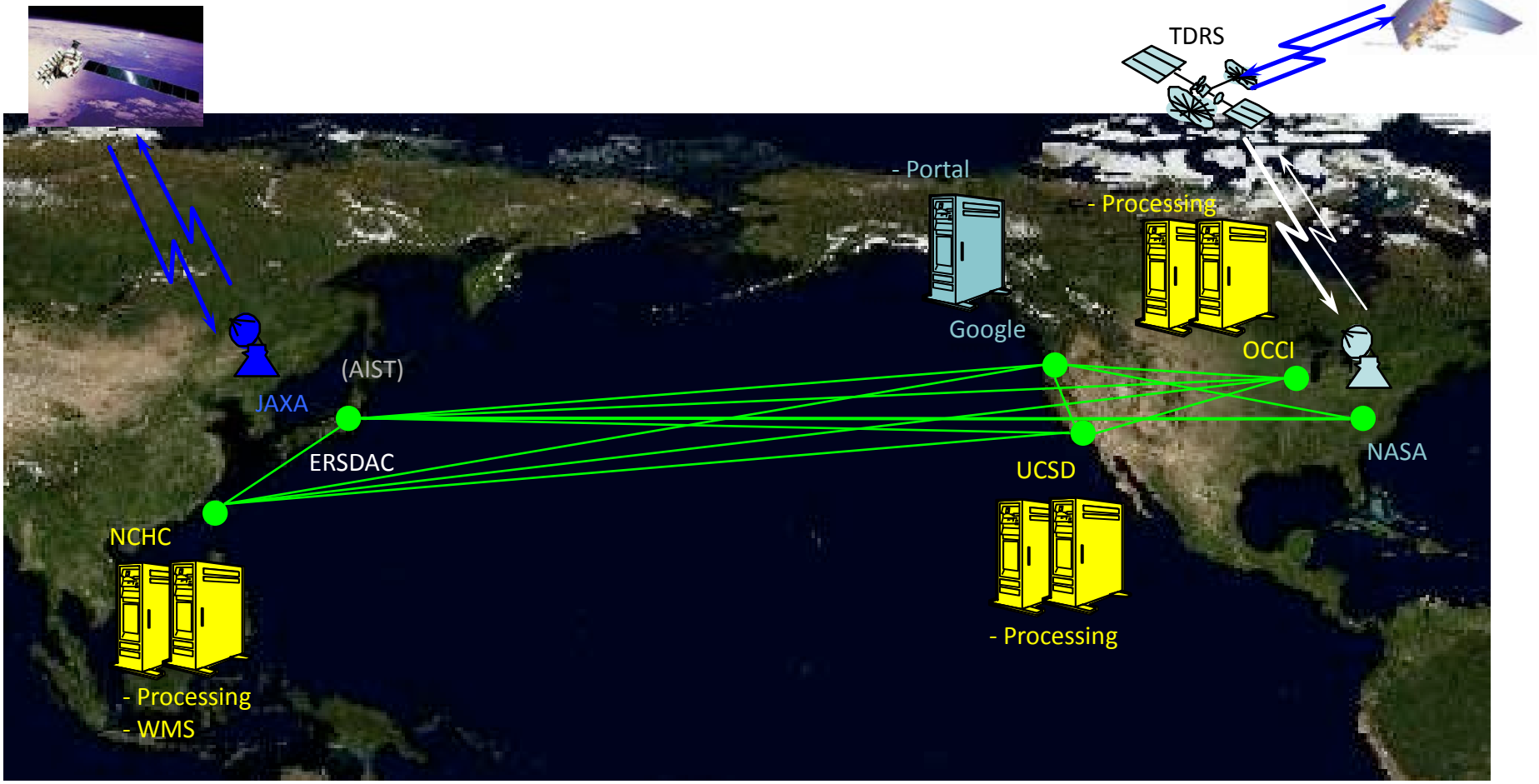
Data Flow and Services from March 11 till April 5th



- **ASTER data: NASA→ERSDAC→(AIST)→**
- **PALSAR data: JAXA→ERSDAC→(AIST)→**

(processing and WMS by Orkney, portal site by Google)

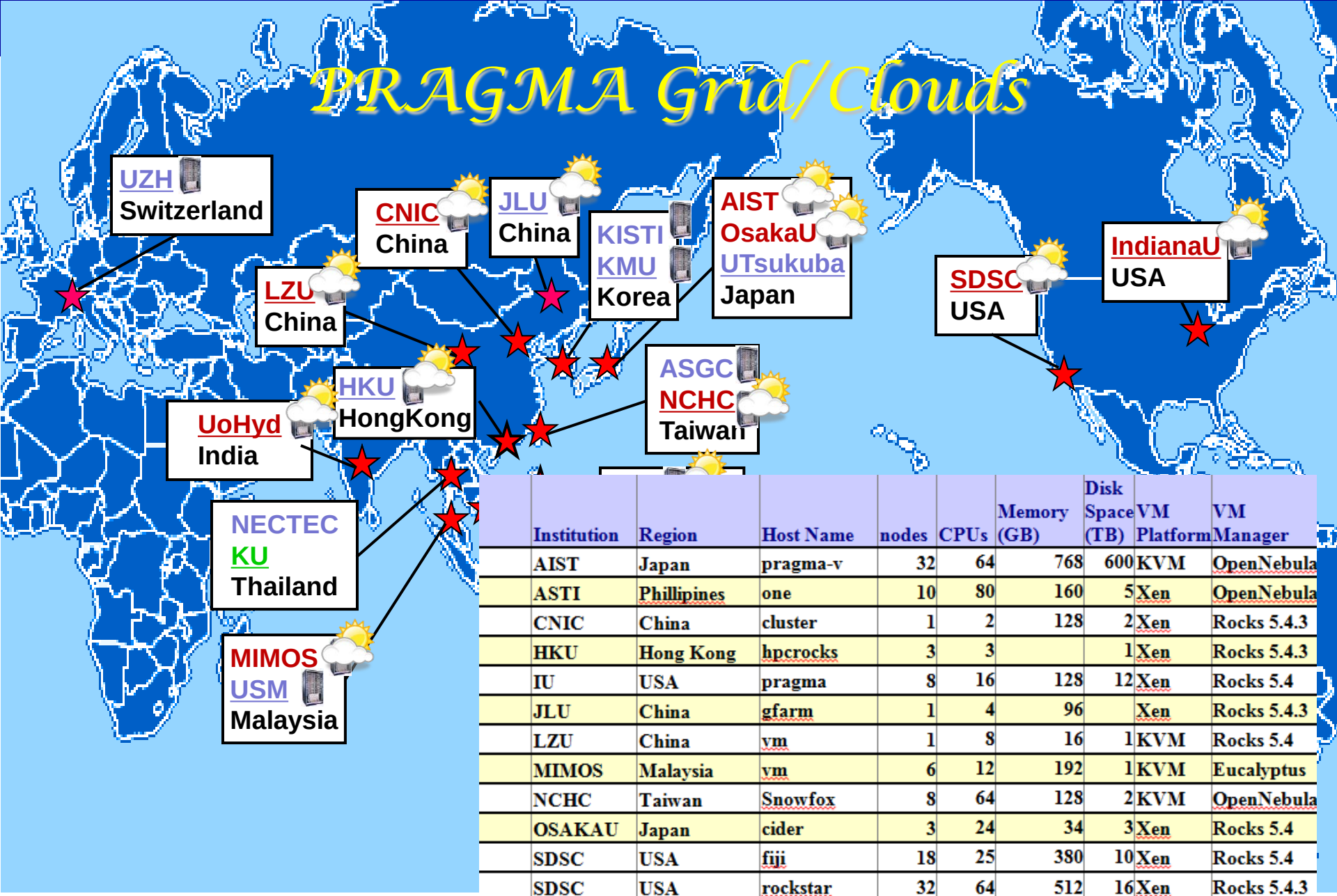
Data Flow and Services from April 21



- **ASTER data: NASA→ERSDAC→(AIST)→**
- **PALSAR data: JAXA→ERSDAC→(AIST)→** (processing by NCHC, SDSC, and OCCI, WMS by NCHC, portal site by Google)

Insights

- Fortunately, we already had VM images for satellite data processing.
  - We have prepared for using cloud.
- Need to make it routine use!
- PRAGMA members had disasters/accidents.
  - Japan earthquake
  - Thailand flooding
  - California power outage
- PRAGMA members has common interests/needs to build a sustainable infrastructure which could be used to support each other in case of emergency.
  - We accelerated the development/deployment of PRAGMA Cloud.



	Institution	Region	Host Name	nodes	CPUs	Memory (GB)	Disk Space (TB)	VM Platform	VM Manager
	AIST	Japan	pragma-v	32	64	768	600	KVM	OpenNebula
	ASTI	Phillipines	one	10	80	160	5	Xen	OpenNebula
	CNIC	China	cluster	1	2	128	2	Xen	Rocks 5.4.3
	HKU	Hong Kong	hpcrocks	3	3		1	Xen	Rocks 5.4.3
	IU	USA	pragma	8	16	128	12	Xen	Rocks 5.4
	JLU	China	gfarm	1	4	96		Xen	Rocks 5.4.3
	LZU	China	vm	1	8	16	1	KVM	Rocks 5.4
	MIMOS	Malaysia	vm	6	12	192	1	KVM	Eucalyptus
	NCHC	Taiwan	Snowfox	8	64	128	2	KVM	OpenNebula
	OSAKAU	Japan	cider	3	24	34	3	Xen	Rocks 5.4
	SDSC	USA	fiji	18	25	380	10	Xen	Rocks 5.4
	SDSC	USA	rockstar	32	64	512	16	Xen	Rocks 5.4.3
	SDSC	USA	pragma-kvm	1	4	32	1	KVM	Rocks 5.4
	UoHyd	India	venus	4	4	16	4	Xen	Eucalyptus
Total	12	8	14	128	374	2590	658		

26 institutions in

Slide by courtesy of PRAGMA

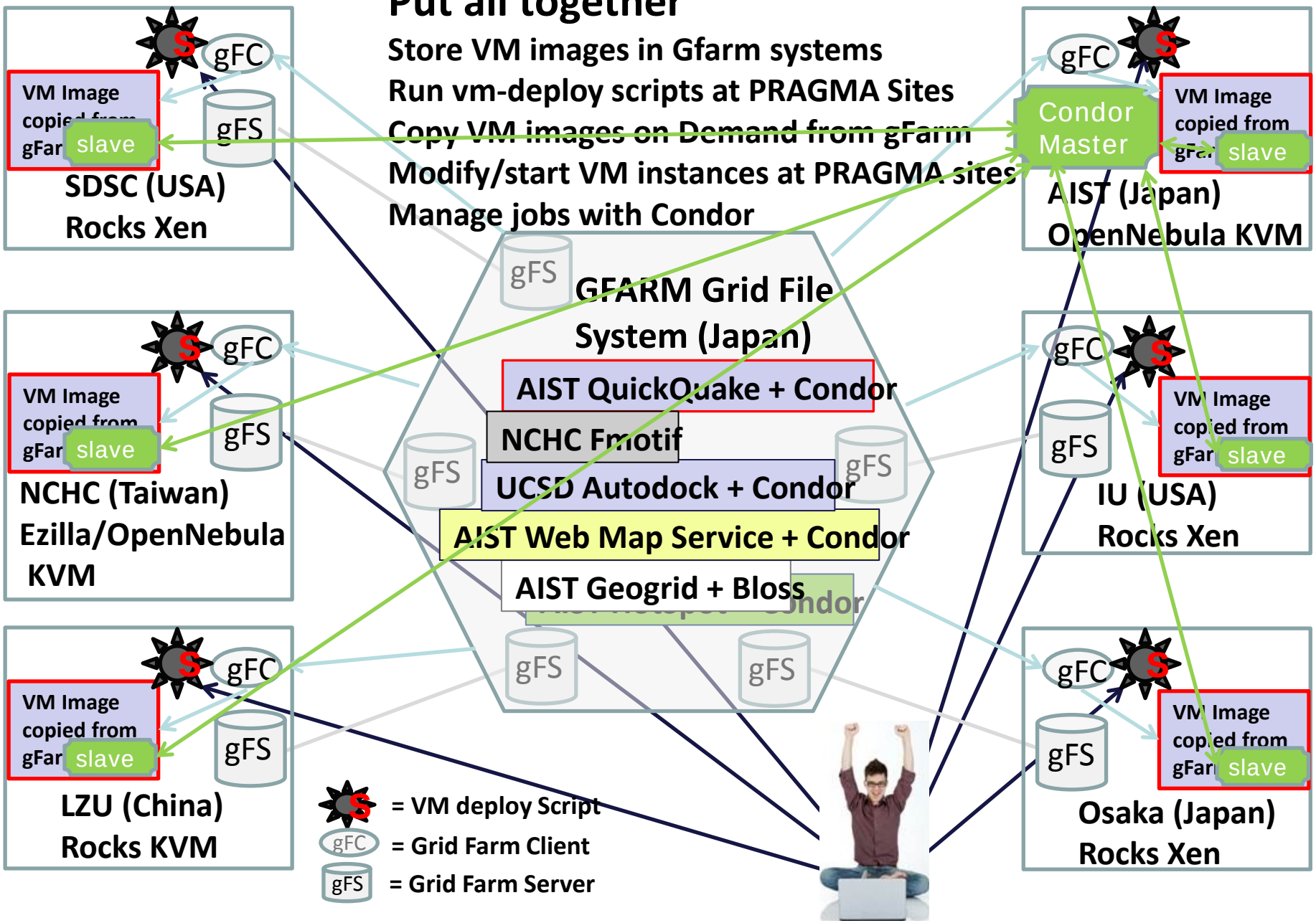
Deploy Three Different Software Stacks on the PRAGMA Cloud

- QuiQuake
  - Simulator of ground motion map when earthquake occurs
  - Invoked when big earthquake occurs
- HotSpot
  - Find high temperature area from Satellite
  - Run daily basis (when ASTER data arrives from NASA)
- WMS server
  - Provides satellite images via WMS protocol
  - Run daily basis, but the number of requests is not stable.

All these applications run as Condor workers

Put all together

Store VM images in Gfarm systems
Run vm-deploy scripts at PRAGMA Sites
Copy VM images on Demand from gFarm
Modify/start VM instances at PRAGMA sites
Manage jobs with Condor



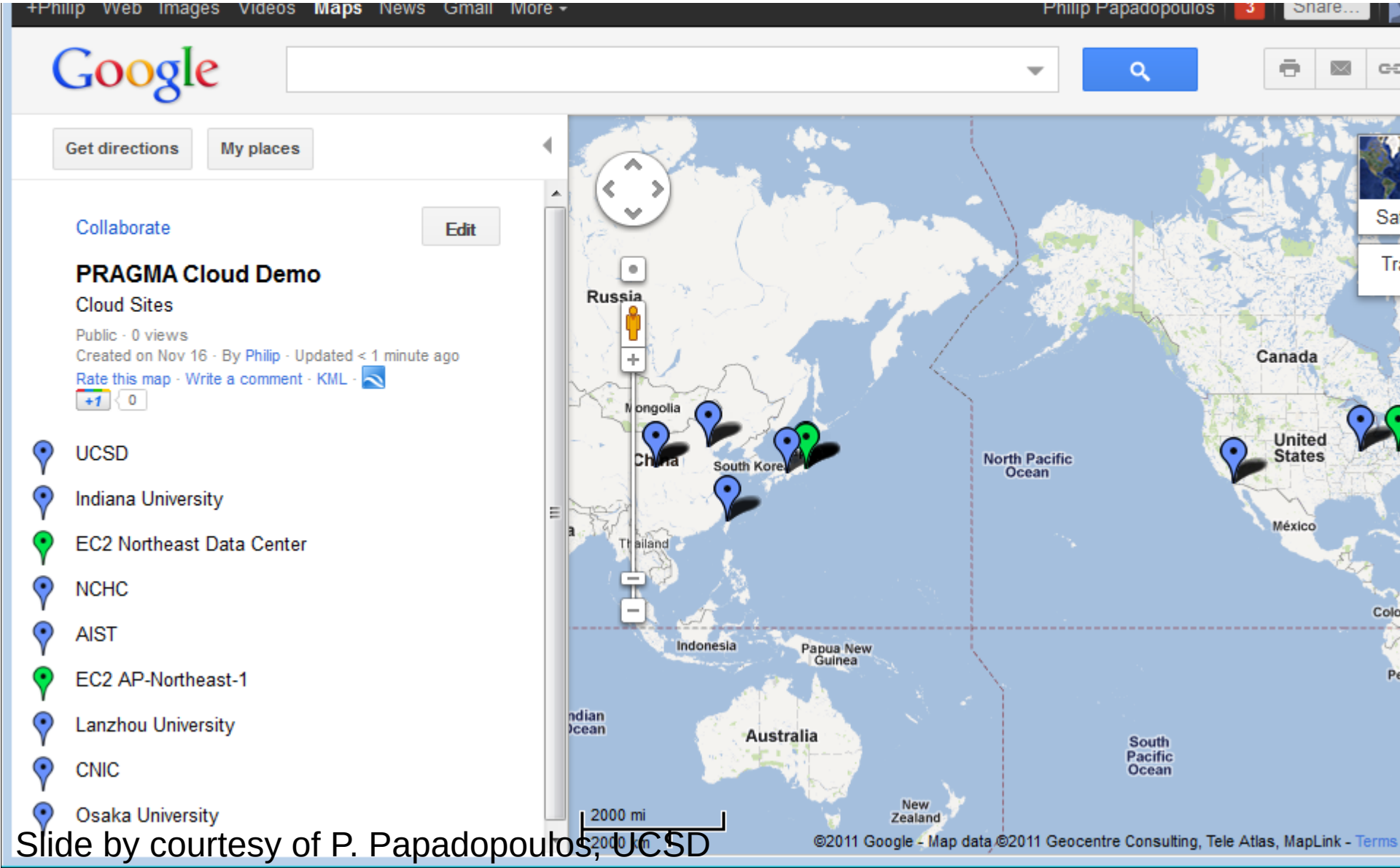
Slide by courtesy of PRAGMA

Essential Steps

1. AIST/GEO Grid creates their VM image
2. Image made available in “centralized” storage (currently Gfarm is used)
3. PRAGMA sites copy GEO Grid images to local clouds
    1. Assign IP addresses
    2. What happens if image is in KVM and site is Xen?
4. Modified images are booted
5. GEO Grid infrastructure now ready to use

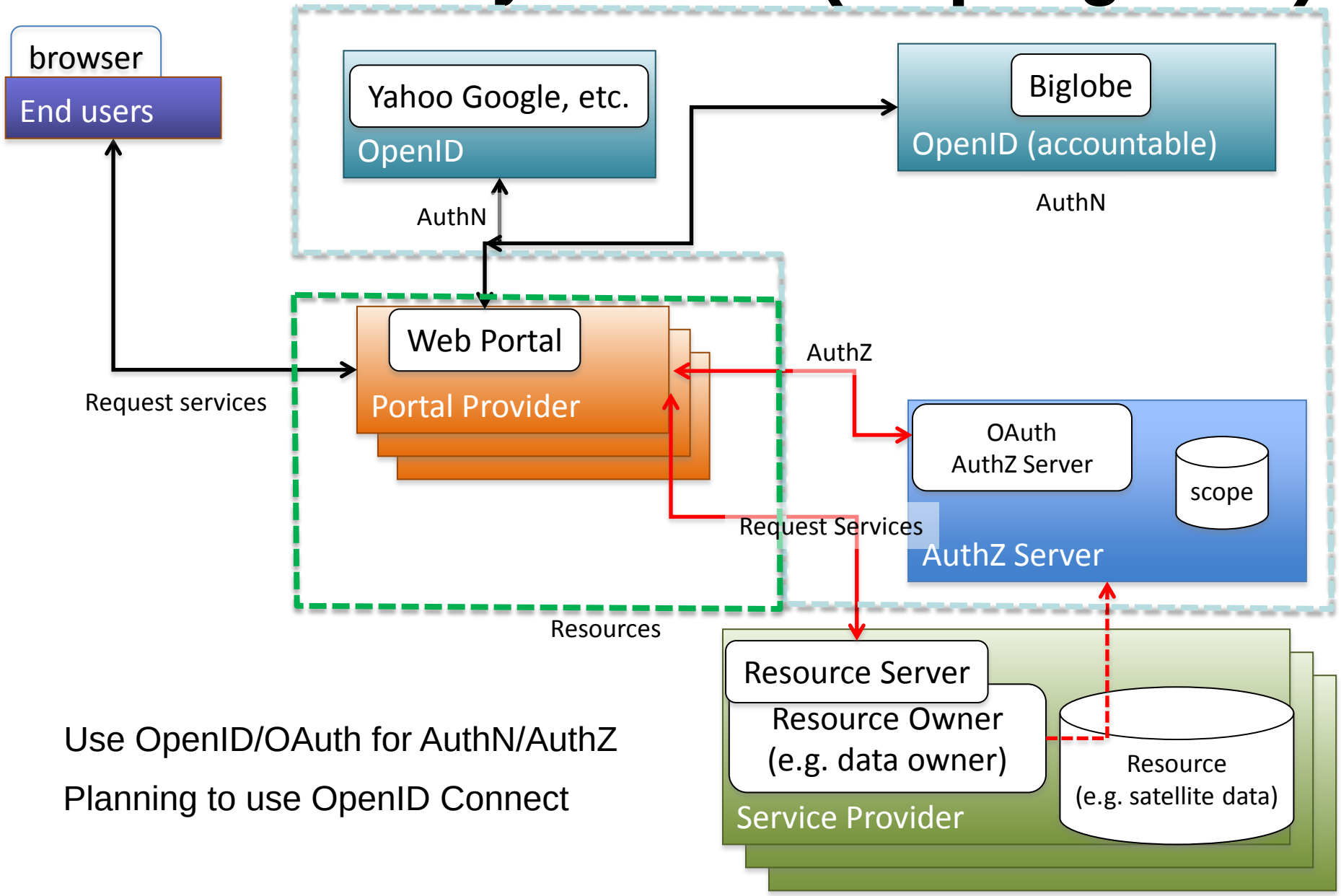
Slide by courtesy of P. Papadopoulos, UCSD

Cloud Sites Integrated in GEO Grid Execution Pool



Slide by courtesy of P. Papadopoulos, UCSD

New Security Model (in progress)

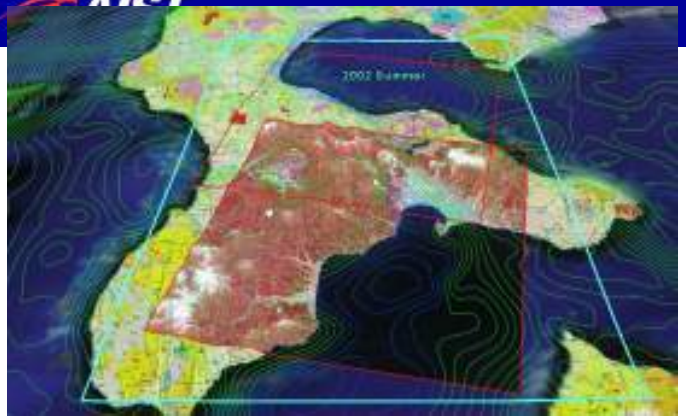


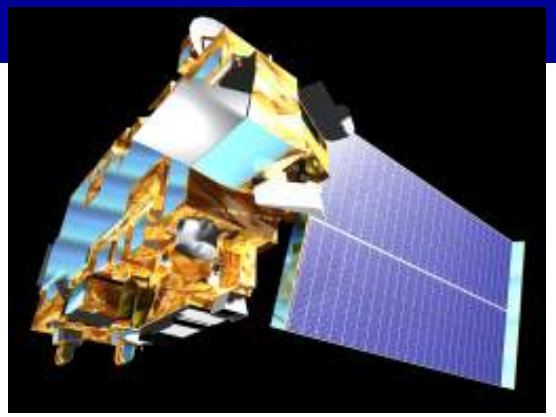
Use OpenID/OAuth for AuthN/AuthZ

Planning to use OpenID Connect

Summary

- We learned a lot through Grid experiments.
- Migrating from Grid to Cloud
  - Virtualization technologies is useful for making distributed infrastructure easy to use.
  - Better for business use.
- Still have many research issues.
  - Data
  - Network virtualization
  - Resource managements
  - Security
  - Making it routine-use





Thank you very much for your attention !

GEOGrid

Global Earth Observation Grid

http://www.geogrid.org/