High Performance GIS and GEO Grid

Satoshi Sekiguchi

(AIST, Japan)

HPC for Societal Benefit Area

- Understanding the Earth become to be more important.
  - The 10-Year Implementation Plan for GEOSS (Global Earth Observation System of Systems) agreed upon at the Earth Observation Summit
  - The Earth Observation Promotion Strategy of the Council for Technology and Science Policy in Japan.
- Necessary to integrate Global Earth Observation data for understanding the Earth.
- The GEO Grid is proposed as an E-Science infrastructure for understanding the Earth.

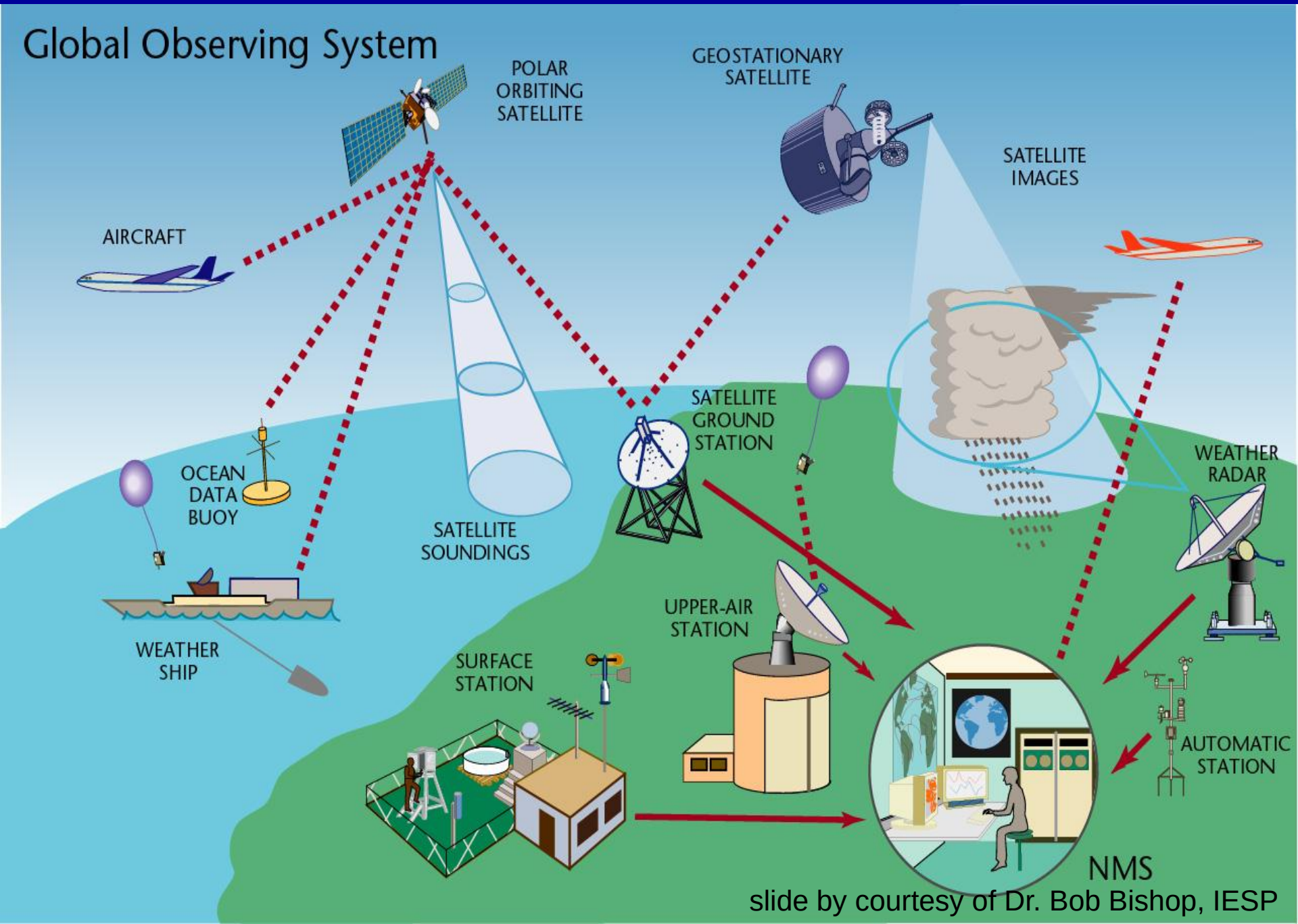
© GEO

Federation of distributed and heterogeneous Earth observation data is the key to implement the Global Earth Observation System of Systems.

Work like a dog
- before dying like a dog

Icelandic Ash Cloud: Mantle-Crust-Glacier-Rivers Weather-Climate-Agriculture-Economy-Society

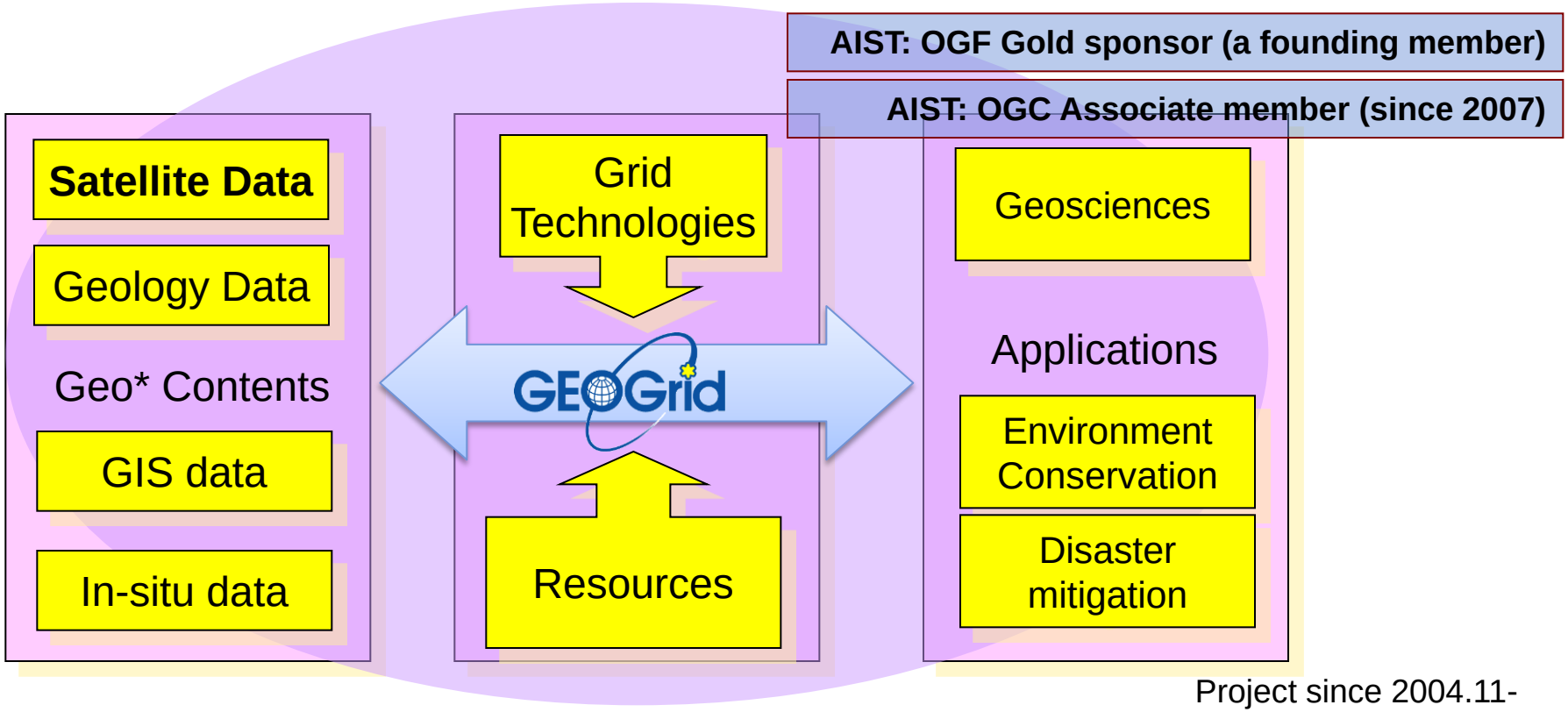
slide by courtesy of Dr. Bob Bishop, IESP



slide by courtesy of Dr. Bob Bishop, IESP

Motivation: How can HPC-Grid-Cloud help ?

- The GEO (Global Earth Observation) Grid is aiming at providing an *E-Science 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.

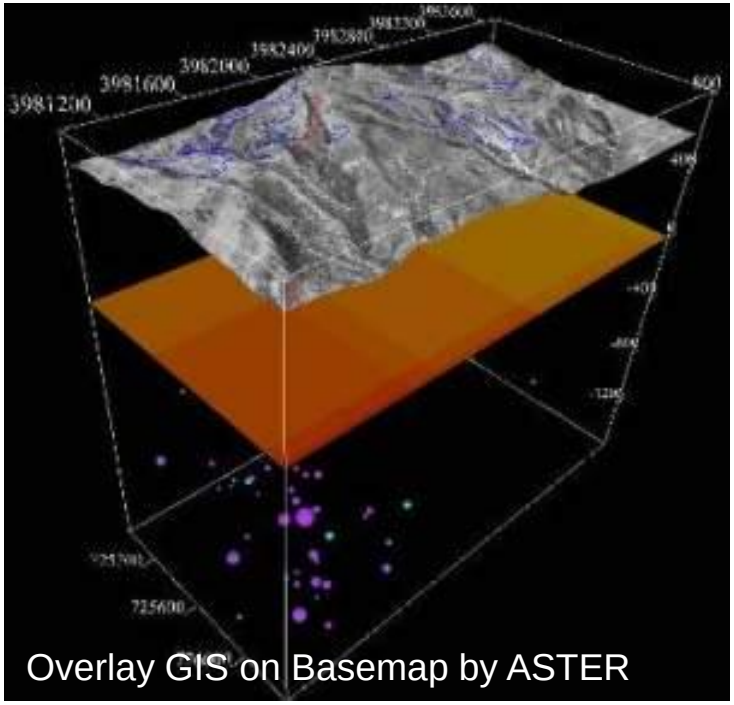


Project since 2004.11-

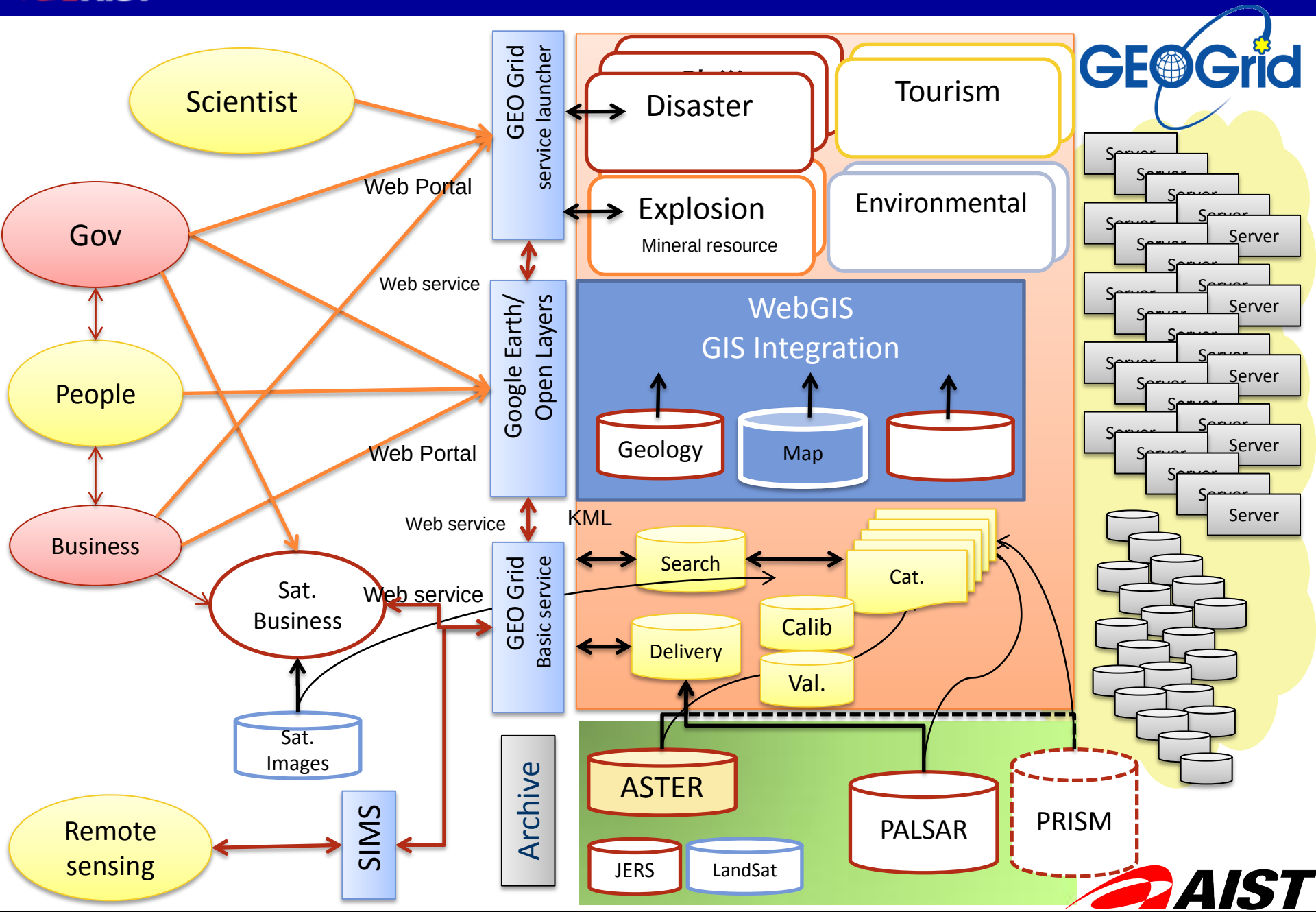
(Google Earth/Map 2005 Feb/Jun)

Objectives of GEO-* Grid

- Help Geo-* scientists to understand
  - Global warming, inventory of carbon dioxide
    - Kyoto protocol, environmental burden
  - Alternate energy
    - Biomass
    - Wind-power generator network
  - Harvest yield prediction/estimation
    - Weather, Soil, temperature, humidity, sunshine, etc.
- Help decision makers to plan
  - Hazard mitigation
    - Earthquake, Landslide, Flood, Volcano eruption, Tsunami
  - Exploration of natural resources
    - Oil, natural gas, mineral
- Unbeknown applications
  - Games, Amusements, Personal geo-record/history, etc.
  - Social science apps
- Multi funded project
  - National Data Archive
  - METI – 100M/JPY (x5yrs)
  - MEXT- 30M/JPY (x3yr)
  - AIST – 75M/JPY(x3yr)



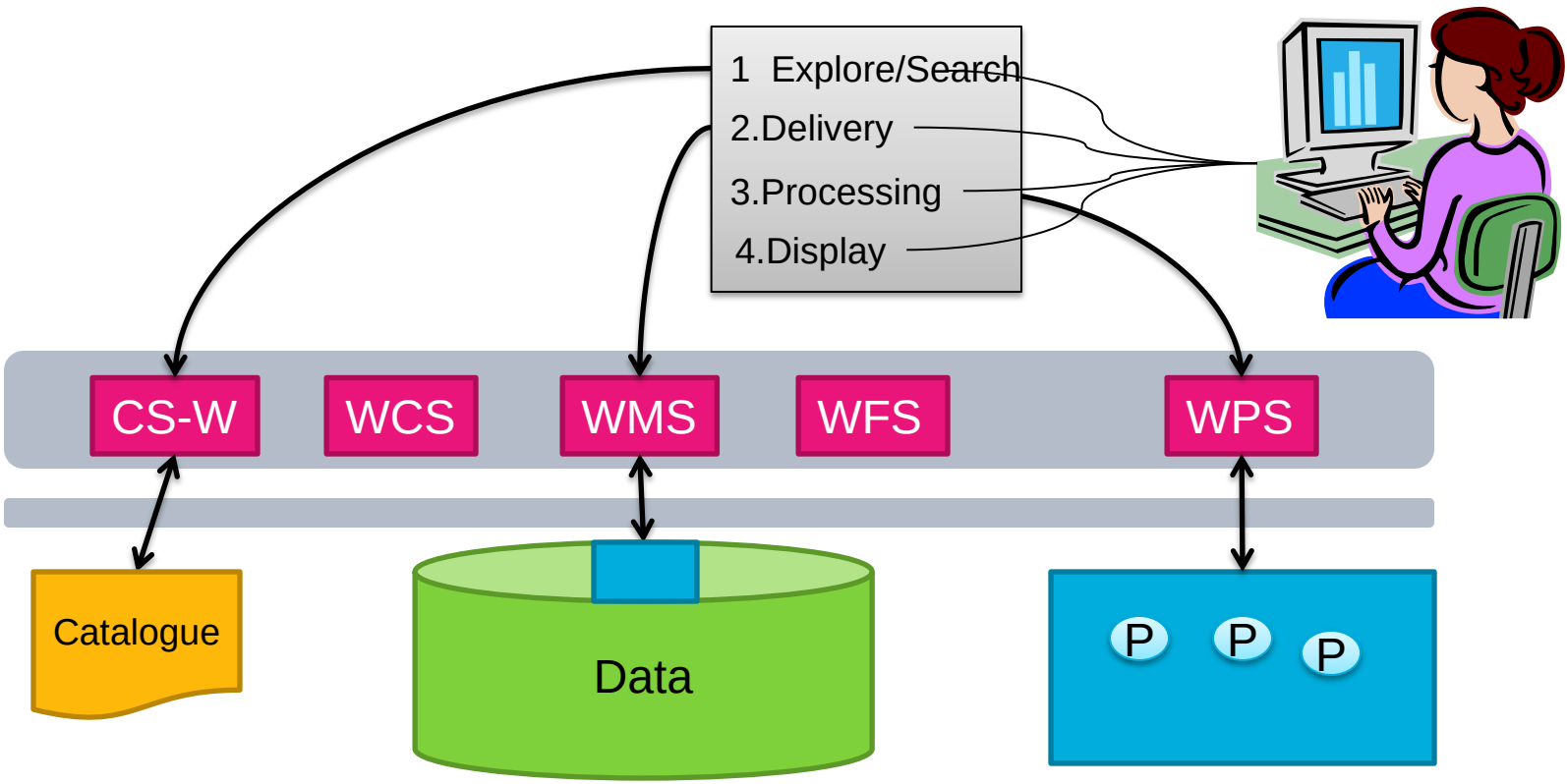
Overlay GIS on Basemap by ASTER


Scientist
Gov
People
Business
Remote sensing
Web Portal
Web service
Web Portal
Web service
Web service
KML
Sat. Business
Sat. Images
SIMS
GEO Grid service launcher
Google Earth/ Open Layers
GEO Grid Basic service
Archive
Disaster
Tourism
Explosion
Mineral resource
Environmental
WebGIS
GIS Integration
Geology
Map
Search
Cat.
Calib
Delivery
Val.
ASTER
JERS
LandSat
PALSAR
PRISM
GEOGrid
Server

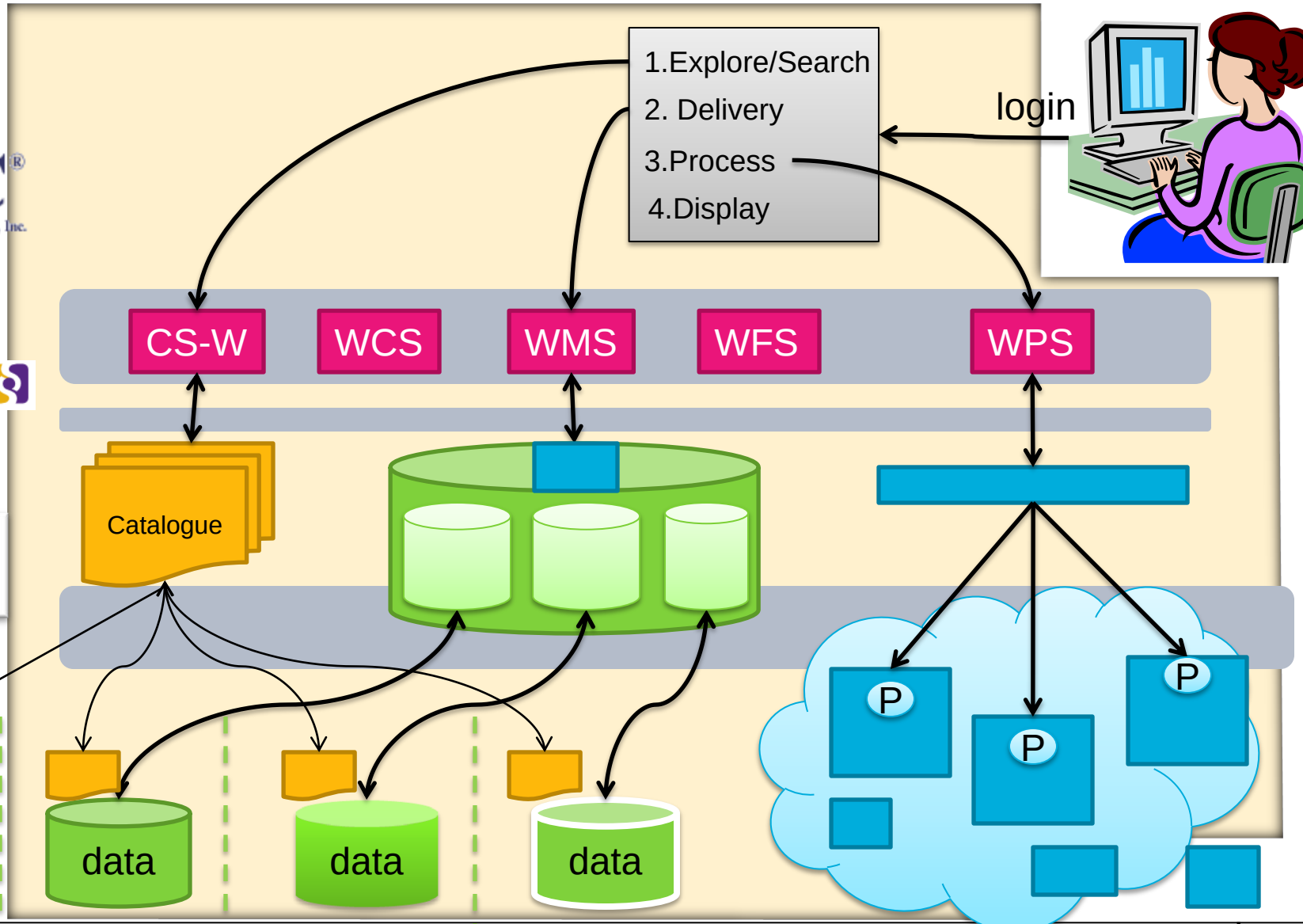
OGC Standard Architecture

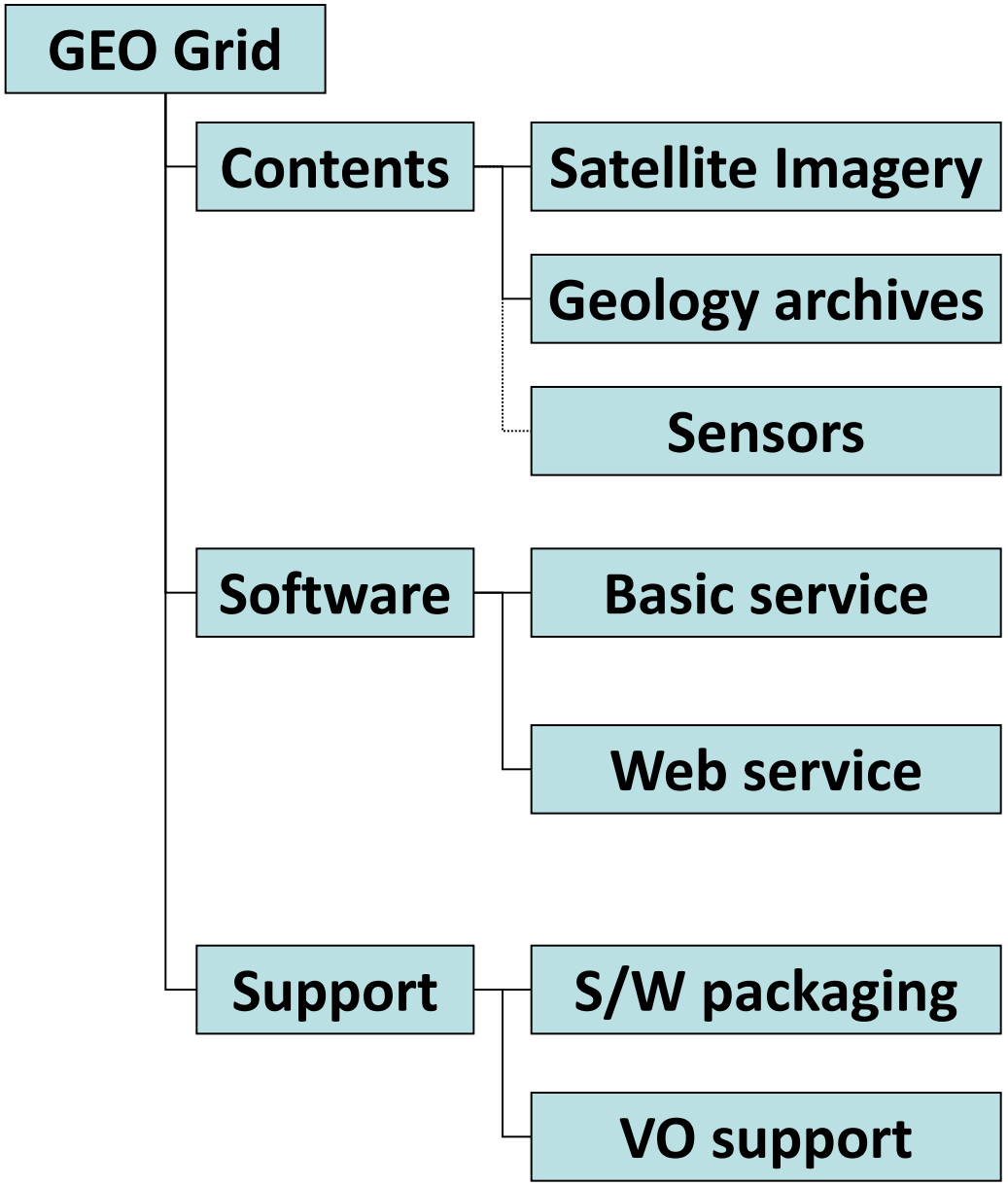
OGC®
Open Geospatial Consortium, Inc.

OASIS



GEO Grid Enhanced Architecture





Full L0 ASTER on disk
MODIS on disk
(Japan & Thailand path)

Japan, SE Asia

Asian Flux, Field server

Security, Resource management, Job, data/storage, information

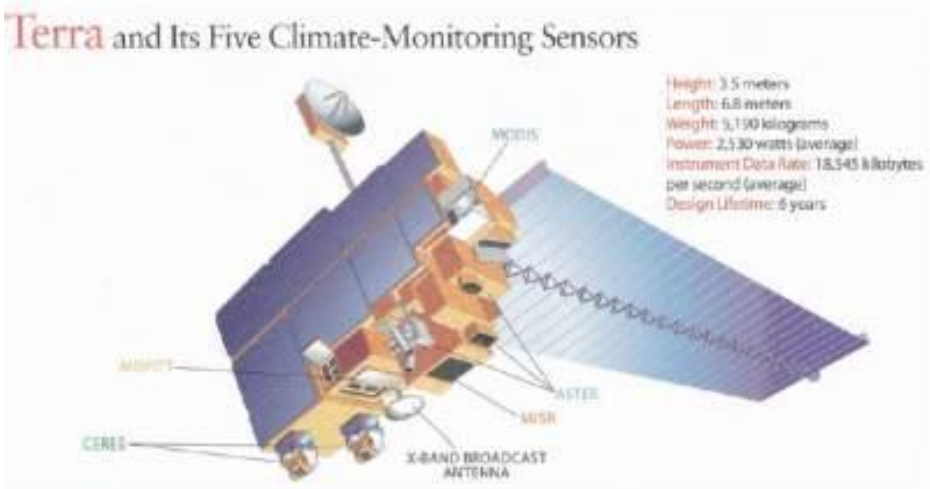
Web GIS, Workflow,
U/I Portal

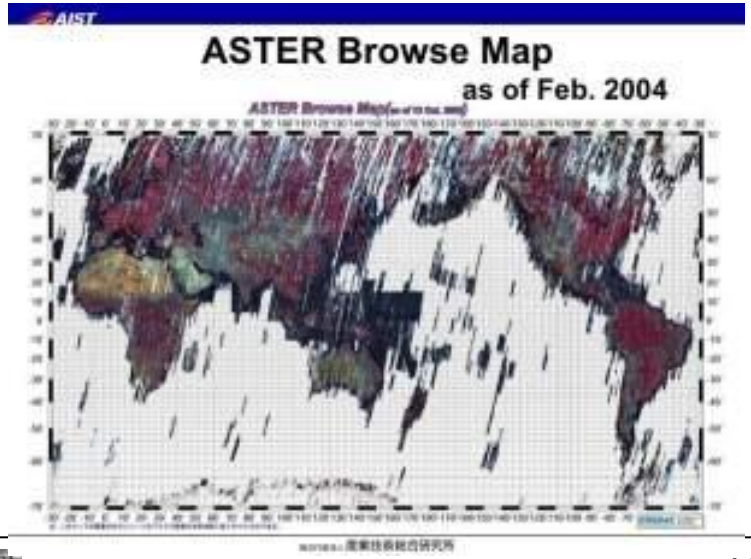
GEO Grid full/light

User account,
Operation support

ASTER : Main content of AIST GEO Grid

- ASTER was developed by METI with AIST scientific & engineering supports
- Global land coverage of Digital Elevation Model of 15m spatial resolution
- Excellent geo-location accuracy
  - Easy to mosaic (or make a seamless image/DEM)
  - Easy to overlay to GIS data
- Powerful spectral analysis
  - VNIR 3 bands (+backward 1band)/SWIR 6 bands/TIR 5 bands
- 2000~
- Attached to Terra
  MODIS has attached to the same satellite

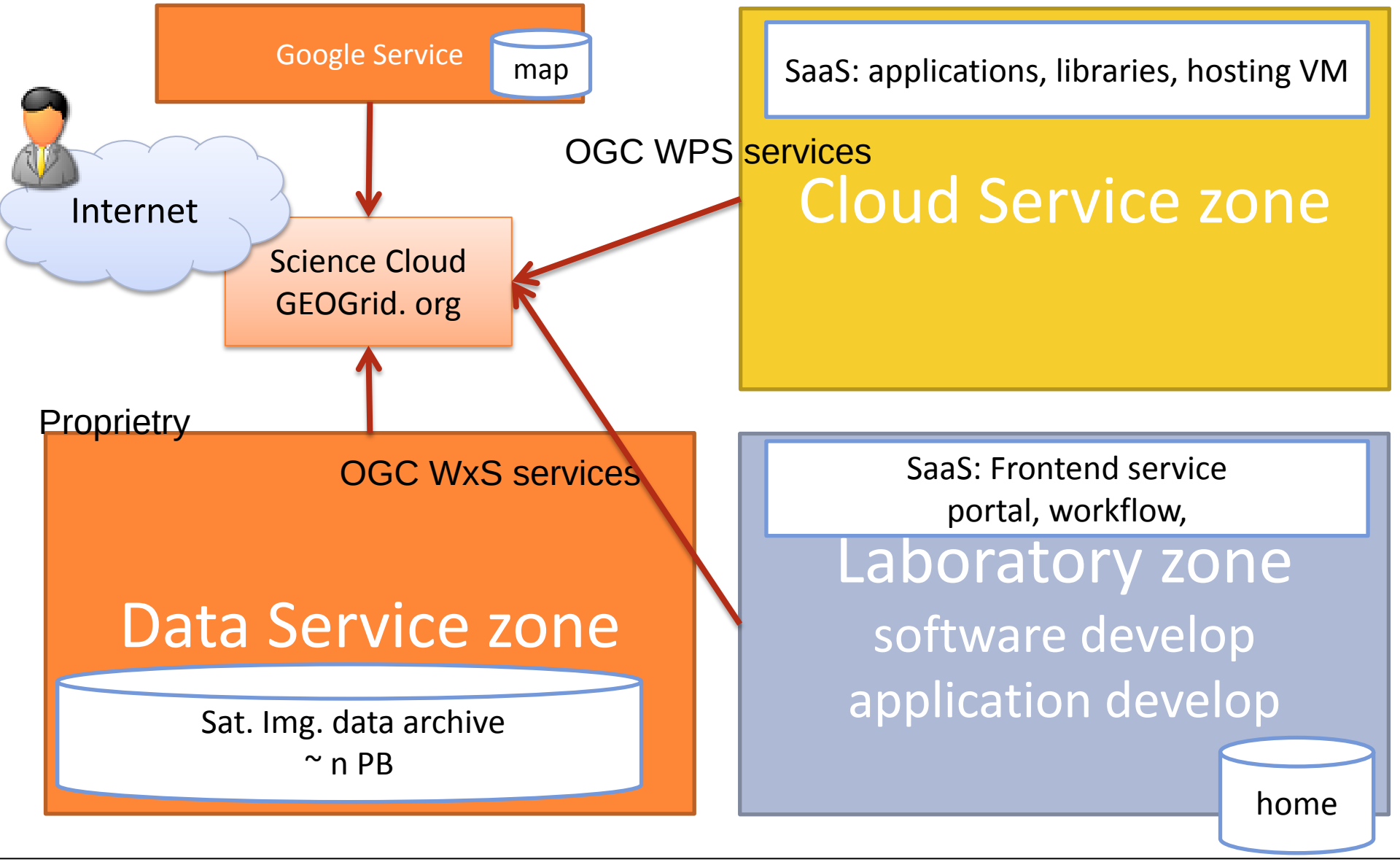




Why “GRID” ?

- Data Grid capability
  - large (>100TB) satellite imagery data
    - storage design, networking design
    - Next gen sensors require 10PB-
  - loosely couple of a wide variety of geographically distributed data
    - meta data (access method, server location,), ontology,
- Computing Grid capability
  - on-demand generation of high level data products
    - adopt the most accurate geometric-, radiometric- and atmospheric-correction methods on-the-fly
  - simulation jobs may consume computing resources
    - a “common” requirement of computing grid
- Grid Basic Service
  - compliance with owners’ access control policy of data/service
    - Grid Security Infrastructure – AuthN, AuthZ, Accounting
  - complex workflow support in portals incl. data access, simulation execution, visualization, etc.

GEO Grid System Zones

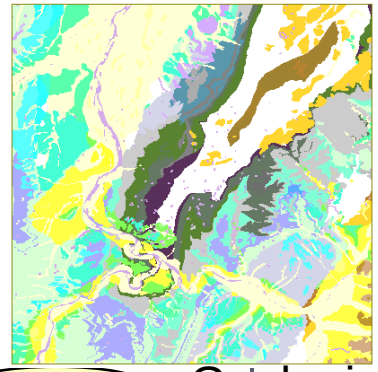


A Workflow example

“Disaster prevention and mitigation (Landslide)”

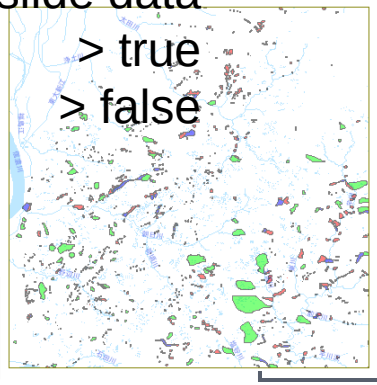
Early warning system based on Susceptibility map

Geology map (GSJ)



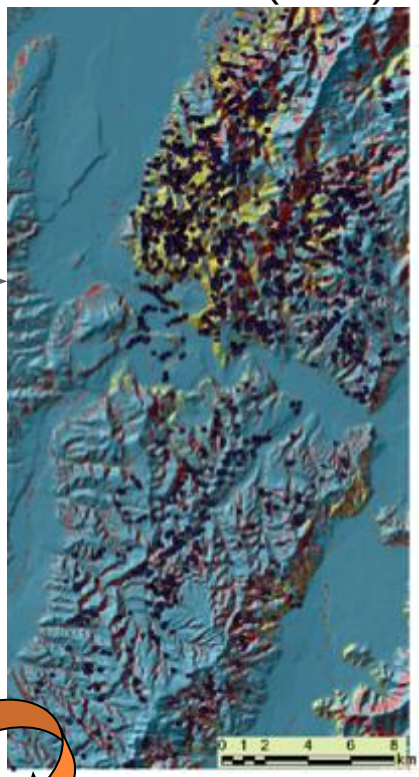
Geology
> rock type
> buffer

Landslide data
> true
> false



Landslide susceptibility index map and its legend
XML(GML)

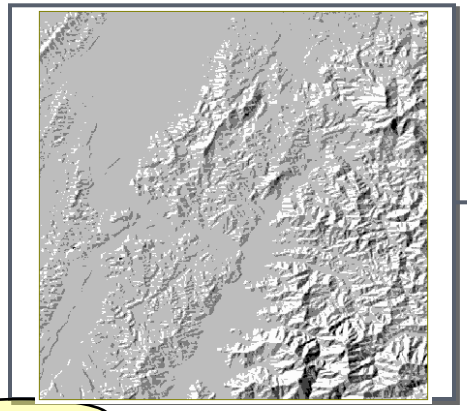
WCS



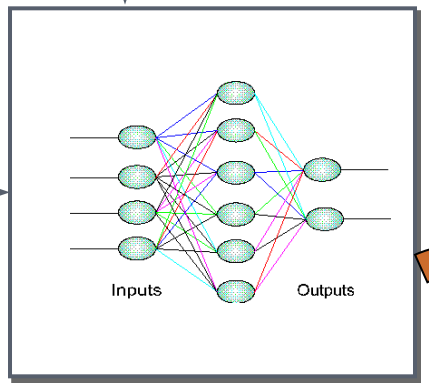
NIED

GSJ

Geologic data (GML) via WFS of Geology DB



DEM (GeoTiff) via WCS of ASTER DB (OnDEM)

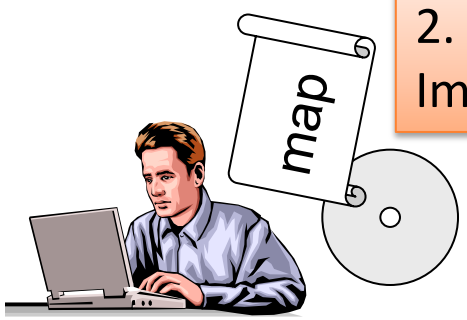


ASTER

High resolution DEM provided from ASTER
Elevation, slope, aspect

Large scale Computer simulation using actual landslide DB

Conventional Approach (Landslide)



2. Input data preparation using Image Processing & GIS software

In most case, commercial

- Convert Format
- Convert Projection
- Extract the Area of Interest

3. Number crunching Calculation

1. Data acquisition

Simulation Program

Request

Request

Request

Upload

If you need a large computing resource…

Application Server

Download

Result

DEM

CD-ROM
FTP

CD-ROM
FTP
Hard copy

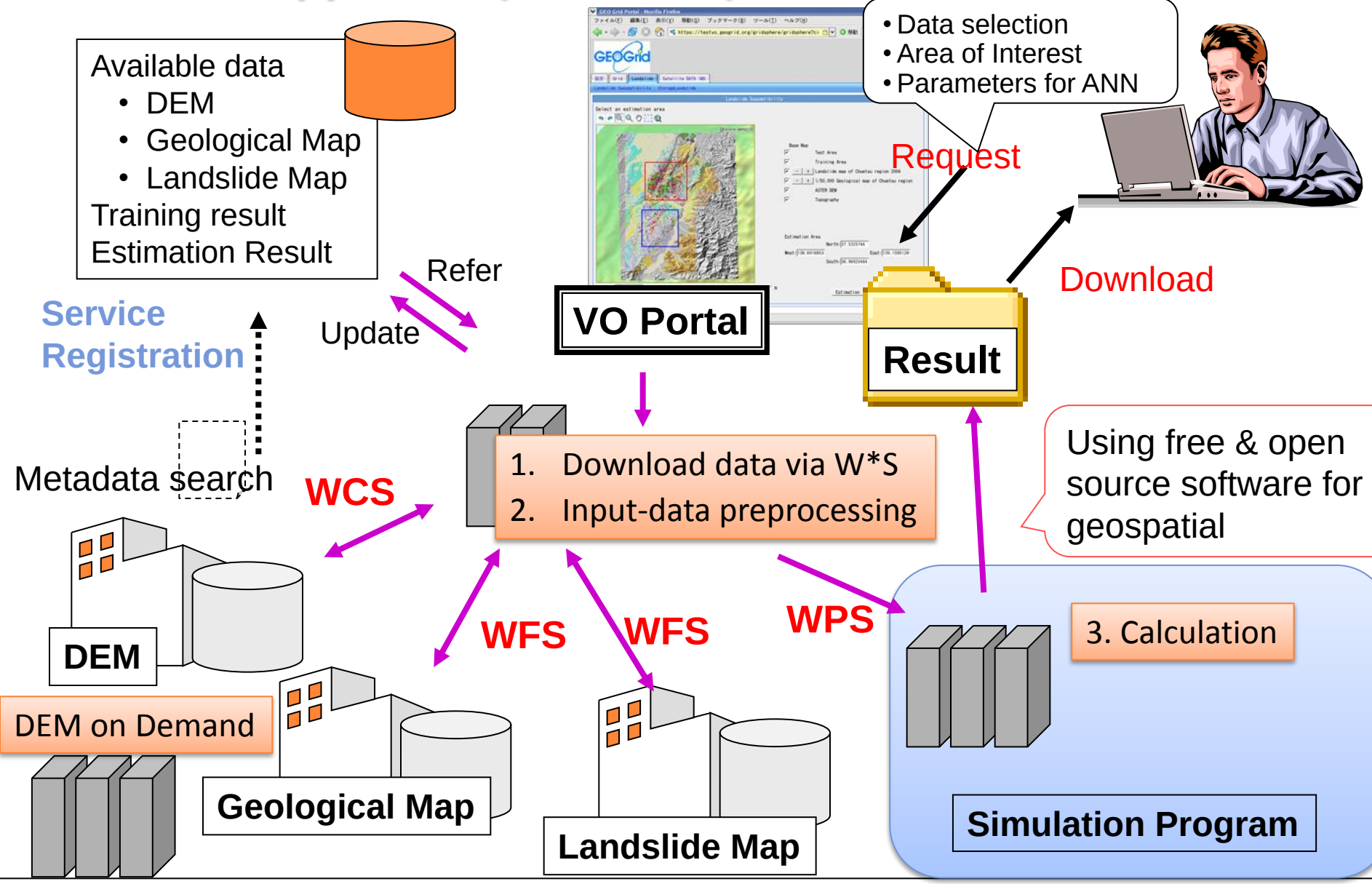
Geological Map

FTP
Hard copy

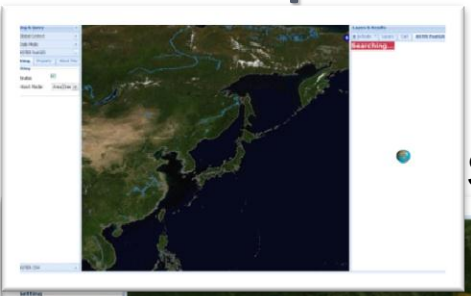
Landslide Map

Many manual operations!
(e.g. data preprocessing)
multi-data policies

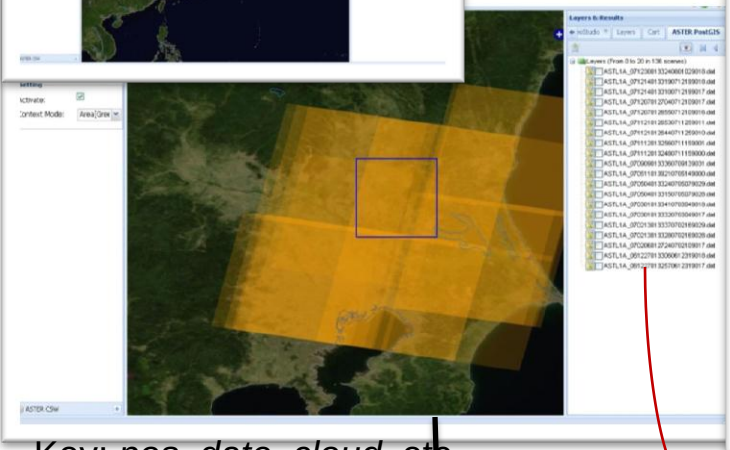
GEO Grid Approach (Landslide)

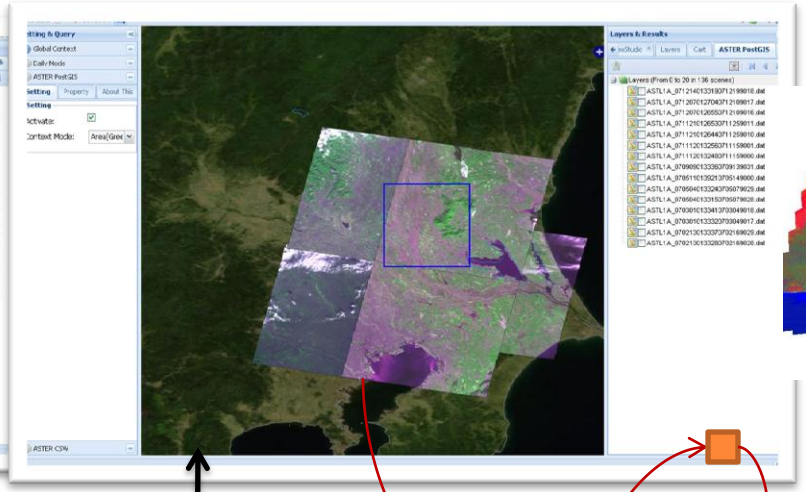


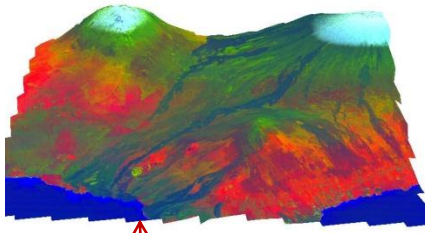
More sophisticated, more standard approach



Search







Key: *pos, date, cloud*, etc

GEO Grid Search Service

On demand

WMS

WMS/(WCS)

CS-W

Metadata

Sat. Data

WMS: Web Map Service – jpeg, png
WCS: Web Coverage Service - Raster

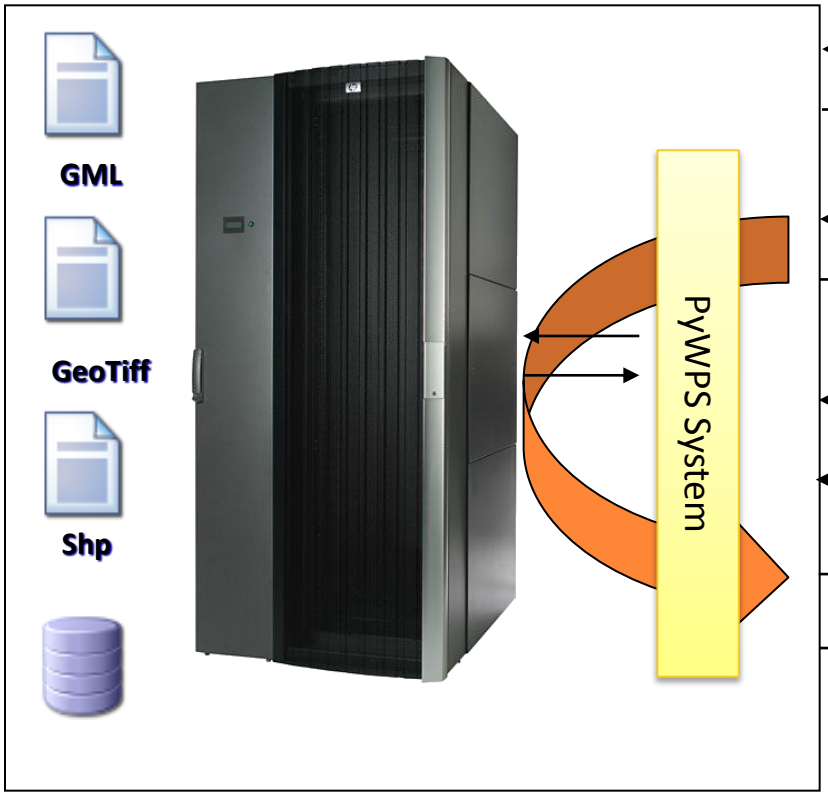
Web Processing Service – hooking GIS and HPC

- Recently, Open Geospatial Consortium (OGC) launches a draft specification of Web Processing Service (WPS) , originally named Geoprocessing Service.

- A kind of the Remote Procedure Call model

- The specified Web Processing Service provides client access to pre-programmed calculations and/or computation models that operate on spatially referenced data.
  - The result of request process are available to download for further analysis at user's machine.

Cloud enbaled Web Processing Service Server



Provision in a cloud

WPS url

GetCapability

Execute Request

send files

Notification

get files

Application
VM container

GeoGrid.org

data in cloud

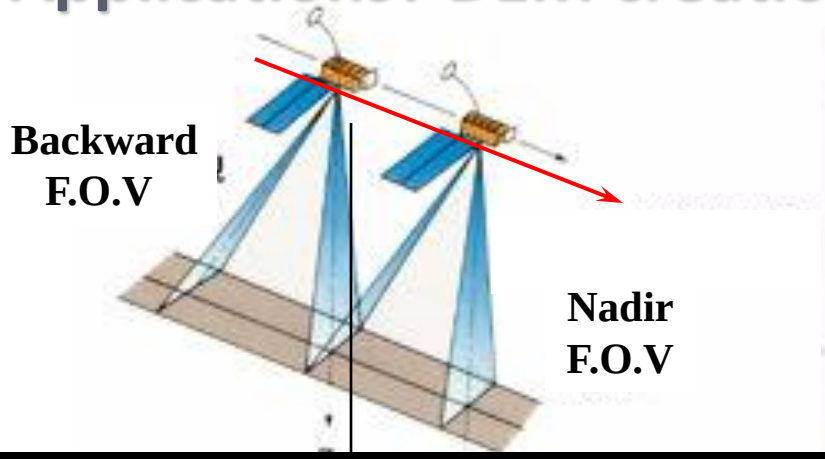
local data

User

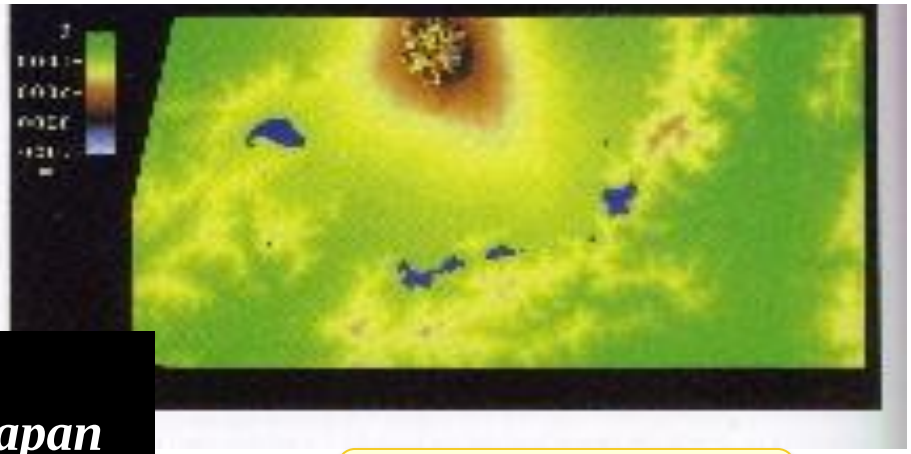
**Execute Request

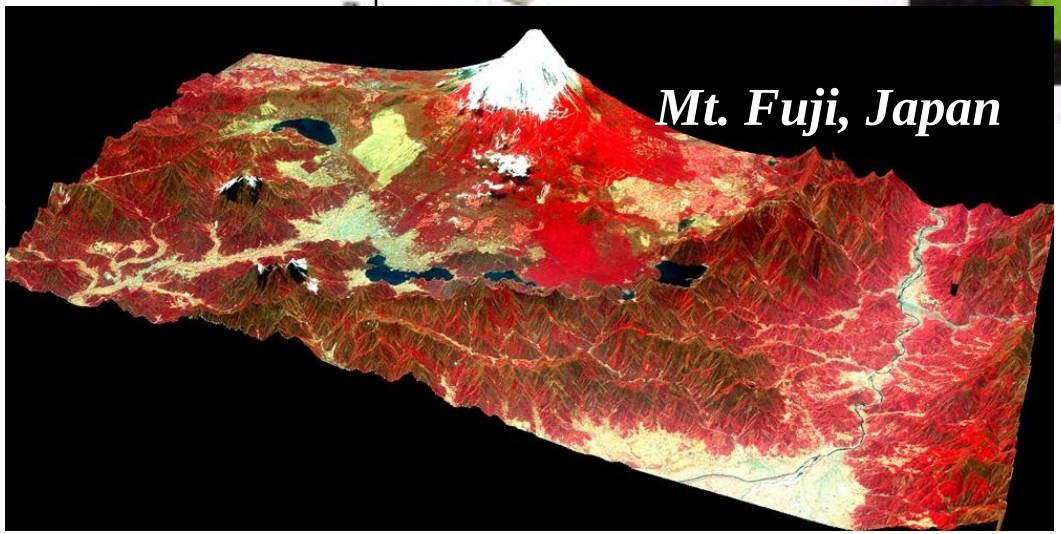
http://wps.org/cgibin/wps.py?service=wps&
version=0.4.0&request=execute&
Identifier=shortestpath&Datainputs=cost,0,x1,596527,y1,4921298,
x2,598173,y2,4923383&store=true&status=true

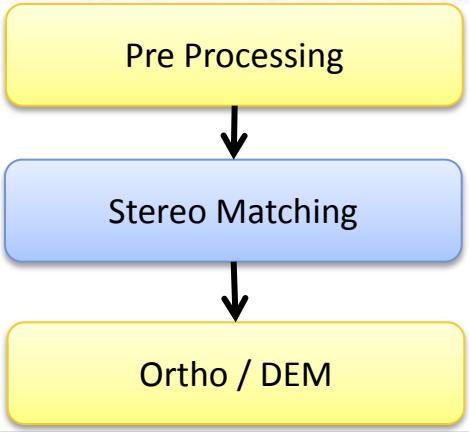
Applications? DEM creation and Stereo-Matching



DEM (Digital Elevation Model)



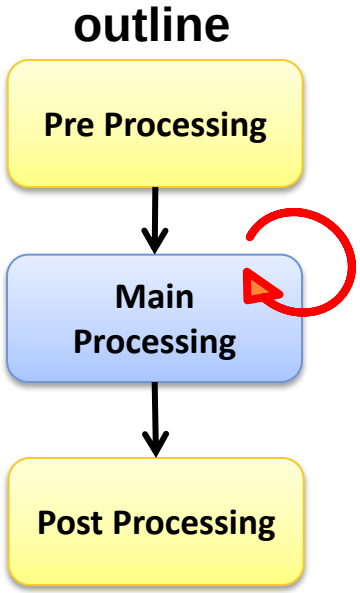


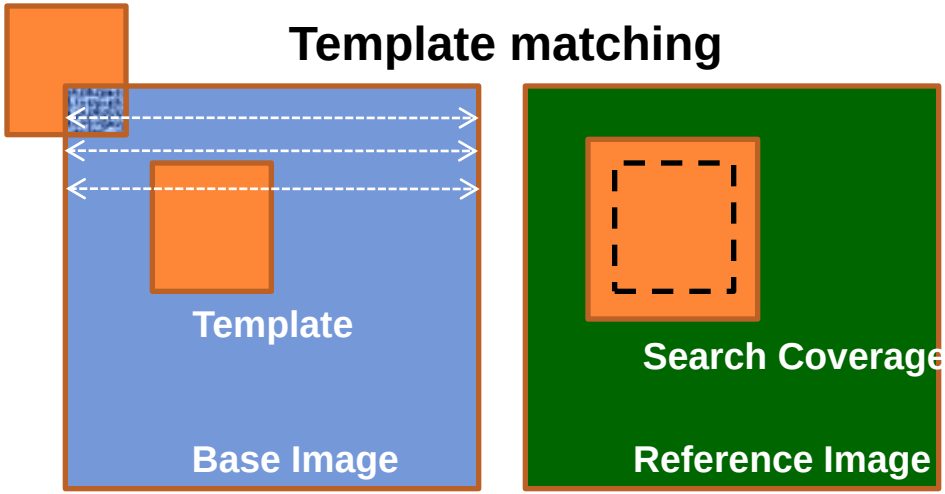


Stereo-matching software has often been used in generating a Digital Elevation Model (DEM) from a pair of satellite imagery data sets to compute height from parallax views using two photographic images.

Analysis of the program（outline） - core of DEM

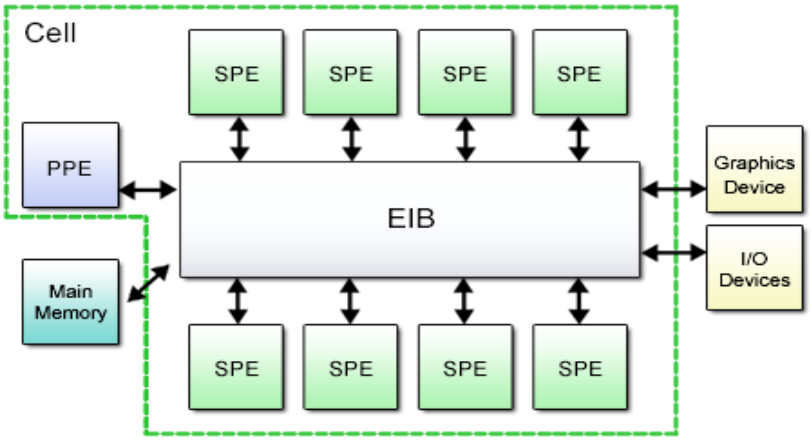
- Outline
  - Compare image data from different sensors
    - Calculate correlative coefficient and identify spots.
    - Complement missing data and generate altitude.
- Pre processing
  - Input data
  - Initialize structures
- Main processing
  - Template Matching
    - Compare two images and identify spots.
  - Interpolation
    - Complement missing data
  - Median Filtering
    - Remove noise
  - Other filtering
  - Output data
- Post processing
  - Free buffers





What we have done

- Optimized Stereo-matching software on multicore processors
- Target architecture
  - Heterogeneous (Cell)
  - Homogeneous (Nehalem)
- How to optimize
  - Manual optimization (Cell)
  - OpenMP + Manual optimization (Nehalem)
- Platforms

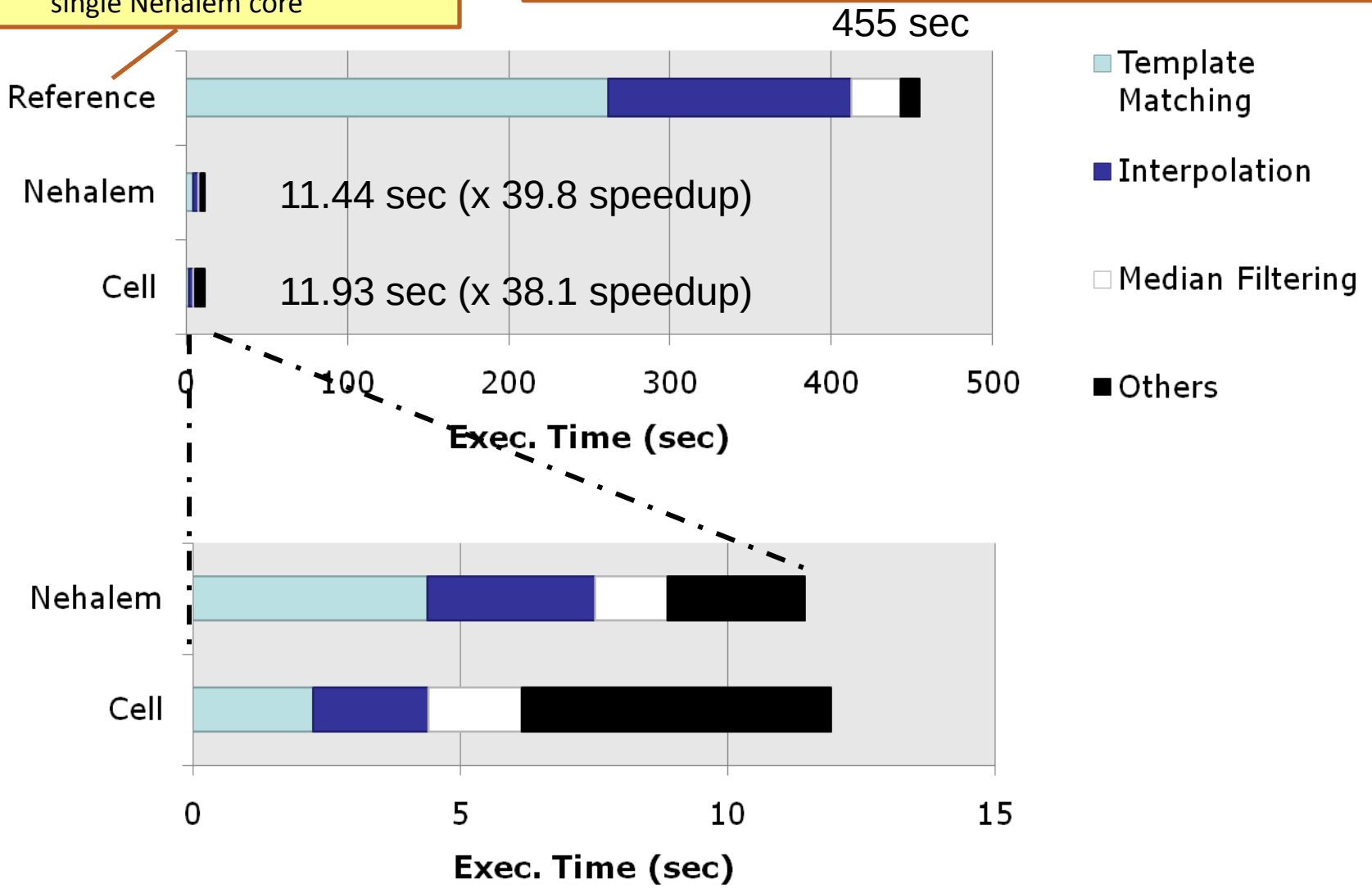


Platform	Processor	#CPU	memory	#core
IBM BladeCenter QS22	IBM PowerXCell 8i (3.2 GHz)	2	8GB	2 PPE + 16 SPE
HP Z800 Workstation	Intel Xeon X5500 (2.66 GHz)	2	4GB	8

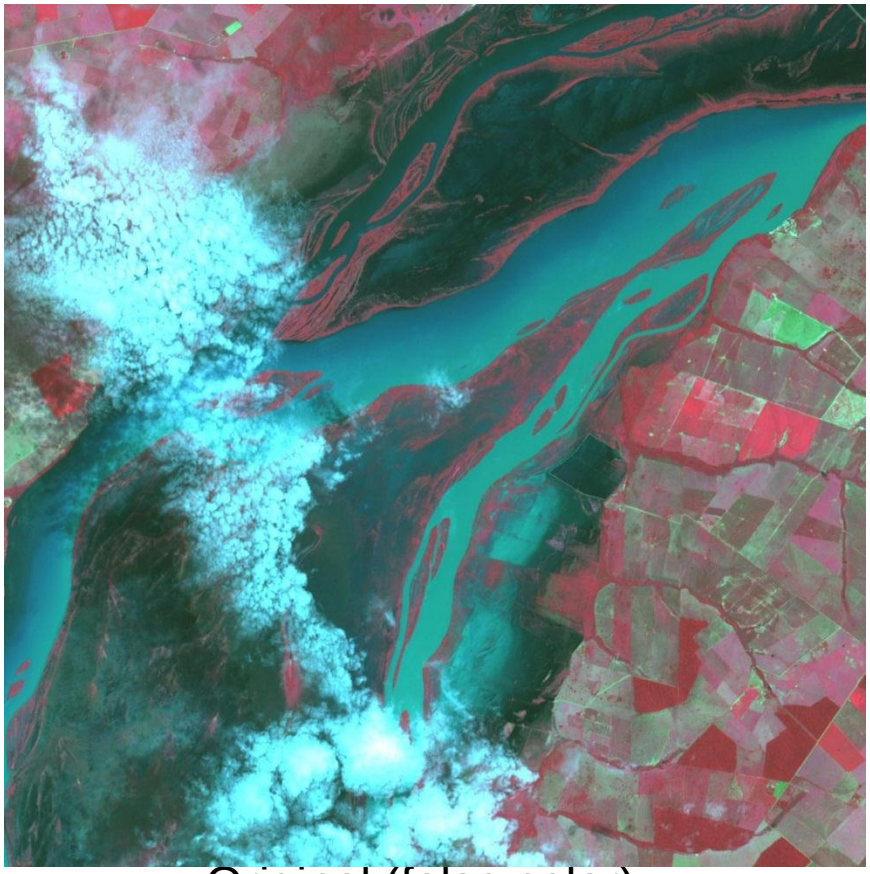
Results of the optimization – possible acceleration in clouds

Exec. time of the original program on a single Nehalem core

Many opportunities to improve performance by HPC



Color Modification – ASTER – high throughput computing

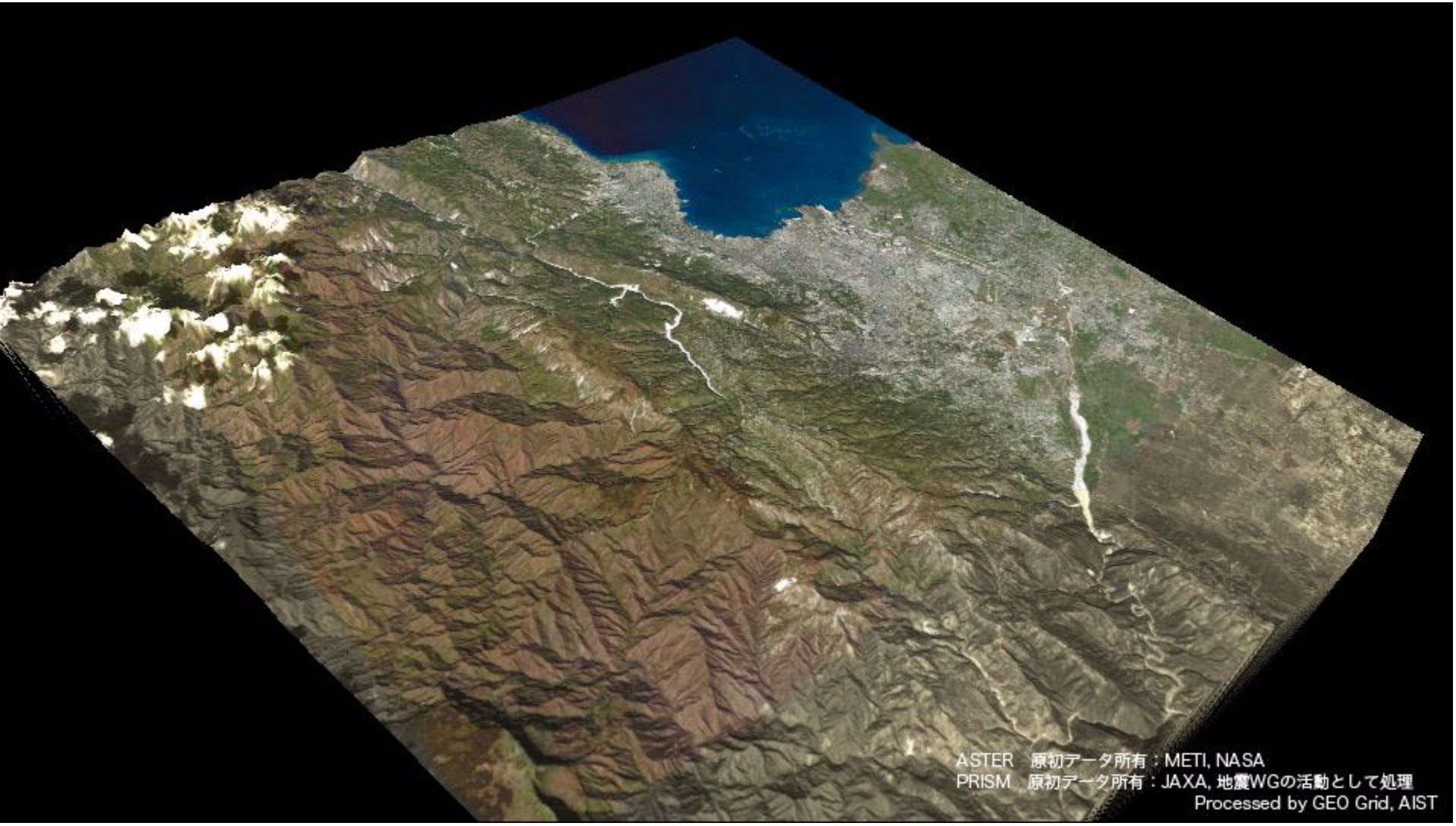


Original (false color)

Pseudo natural color

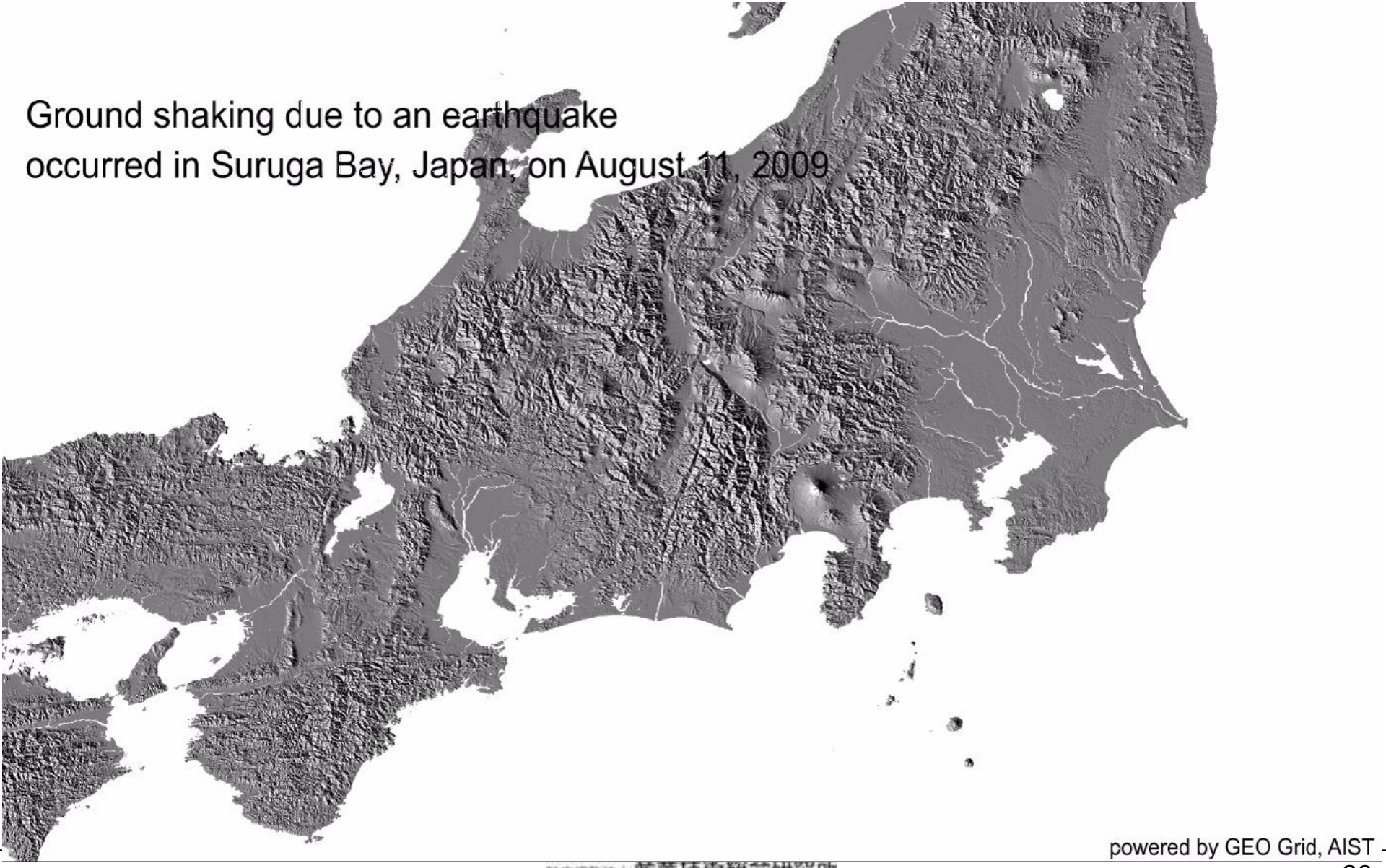
- For interpreting objects in the earth surface easily by human eyes, natural color composite image is sometime needed.
  - ASTER VNIR does not contain Blue band.
  - Color and atmospheric correction technique developed by Lille University, France was implemented in GEO Grid to generate pseudo natural color ASTER images.

Haiti landscape after shake – created by PRISM + ASTER



Ground shaking sensor data interpolation (250m mesh)

100ms



powered by GEO Grid, AIST

Summary

GEO Grid has done in

- Data Archives/services
  - ASTER all scenes in L1
  - PALSAR (part)
- AuthN, Group-AuthZ
  - VOMS, Tsukuba-GAMA
  - less management cost
- CS-W for Multiple data set
  - AIST CS-W, scalable
- Cloud based service
  - DEM, color coding
  - Earthquake, landslide, volcano

In progress

- Workflows – engine, language, tools
- Widgets design to build your own applications simply
- Digital Rights Management – water mark
- AuthZ in WMS, WFS, WCS, WPS
  - no clear strategic idea – whom to work with
- Full Cloud Service
  - easy adaptation
  - packaging
  - computing resource provider
- Google Earth integration – coming soon

GEO Cloud

See you again in SC10!