【コミュニティ動向】
e-Science コミュニティからの話題

Grid と GIS: OGF と OGC* の連携模索について

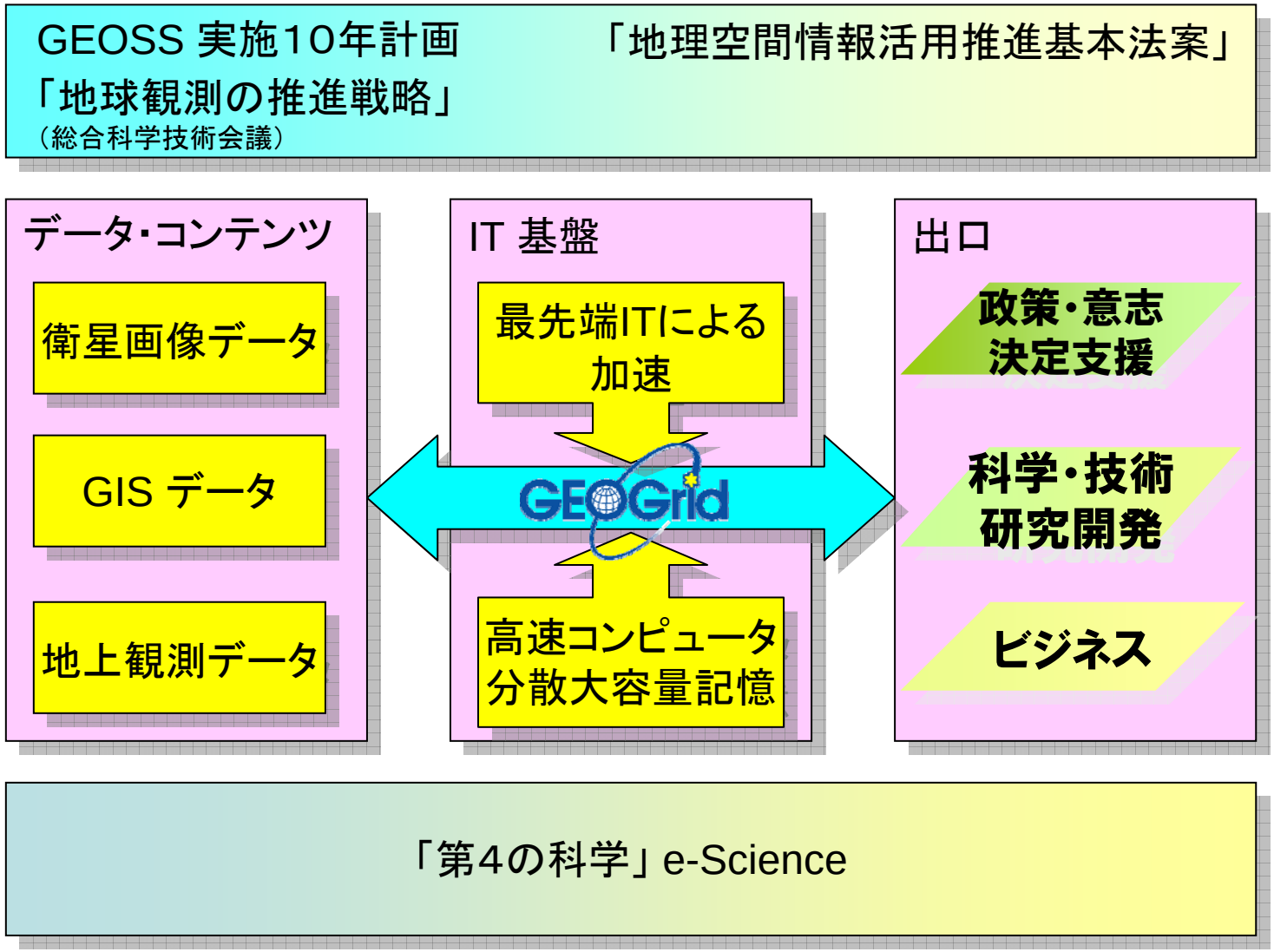
関口智嗣（産総研）

OpenGridForum
OPEN FORUM | OPEN STANDARDS

(*Open Geospatial Consortium)

Grid
Consortium Japan
グリッド協議会

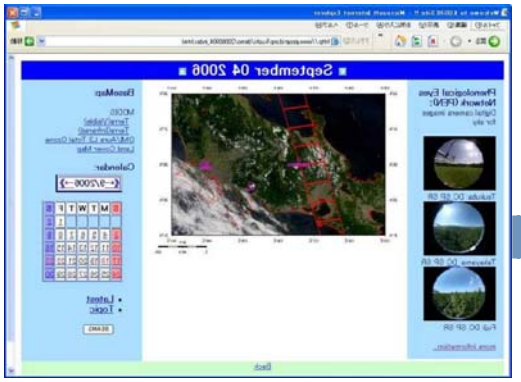
GEO Grid の枠組み



GEO Grid Challenges in system design

- ASTER Complete Data Online
  - Handling large (>200TB) data archive for satellite imagery
- Data Integration
  - Handling wide diversification of data types, associated metadata, products and services
- Data Access Policy
  - Retaining data owner's publication policies without increasing management cost.
- On-Demand data processing capability

Large Data Archive at AIST



AIST

Data Grid
gfarm

GRID DATA FARM

Computing Grid

ASTER

MODIS

PALSAR

Geometric, Radiometric &
Atmospheric Correction
Mosaic & Composite,
Image Processing

210TB /P4 3.5TB(raid5)/node x 60C

AIST Supercluster
(11TFlops)

ASTER & MODIS

Tsukuba
(Japan)

AIT
(Thailand)

NASA
(USA)

MODIS

L0 or L1

ASTER

APAN/JGN II

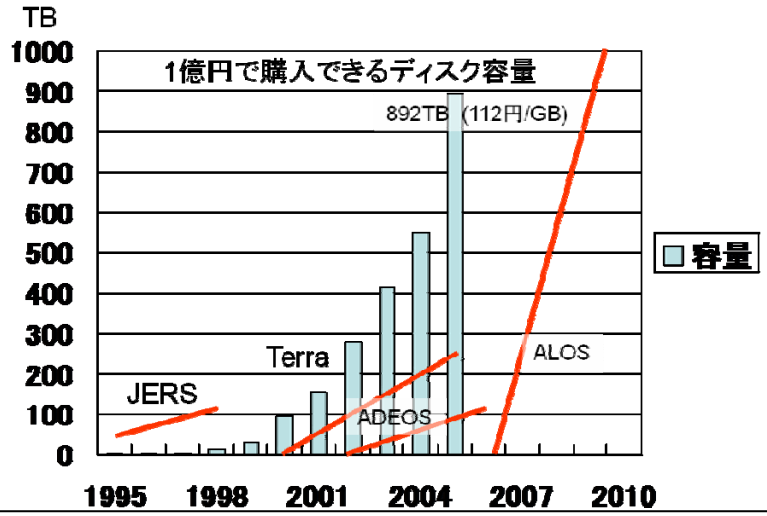
PALSAR

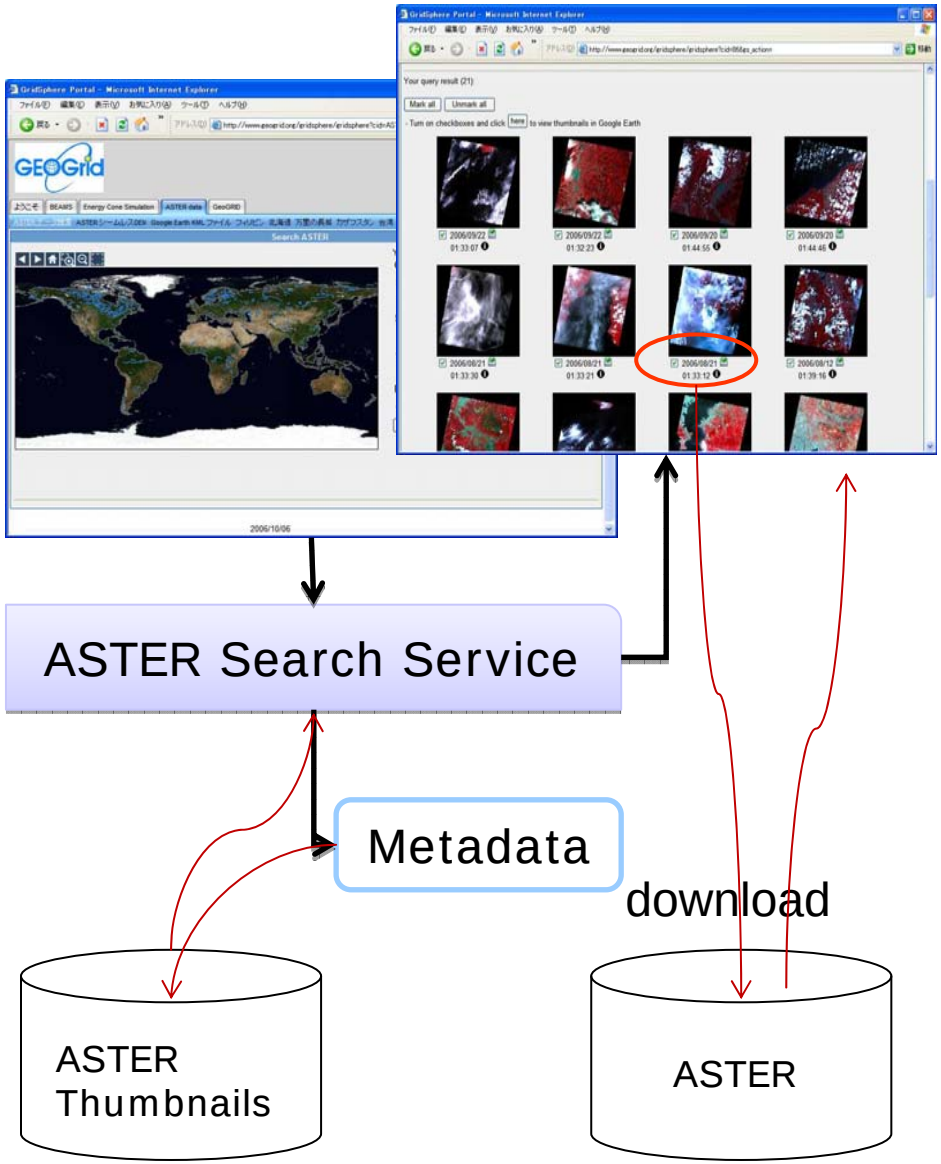
HATOYAMA (Japan)

L1A data will not be in public

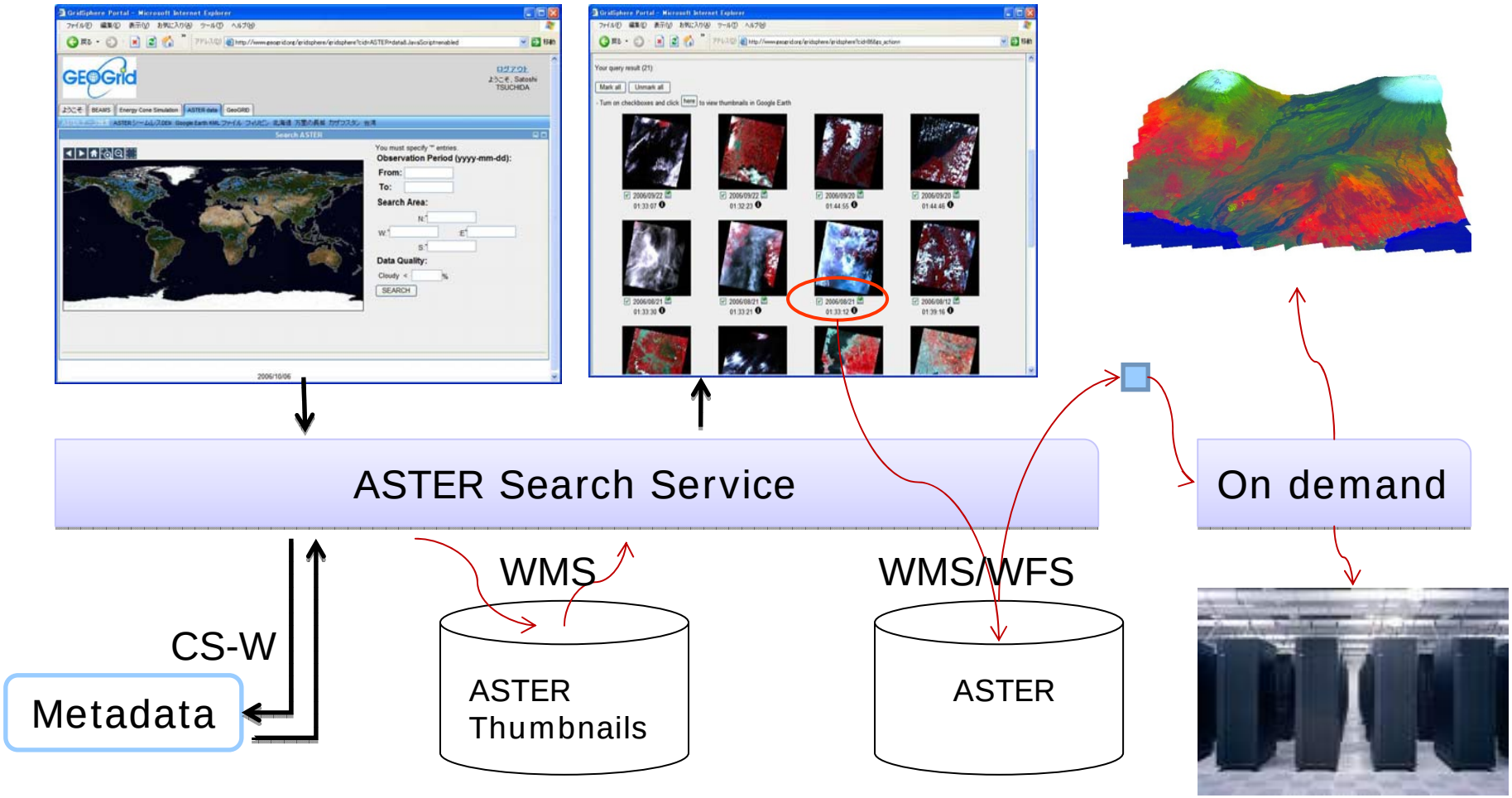
ASTER Complete Data Set – On Line

- Believe me!
  - Grand Stn designed in late 90's
  - HDD cost 1/200
- Benefits
  - On-demand producing higher level products
    - Less storage volume
    - Choose algorithms interactively – parameter, calibration, projection, etc.
  - Easy to overlay other GIS data
    - Put in meta data registries
    - Low latency to get data from disk

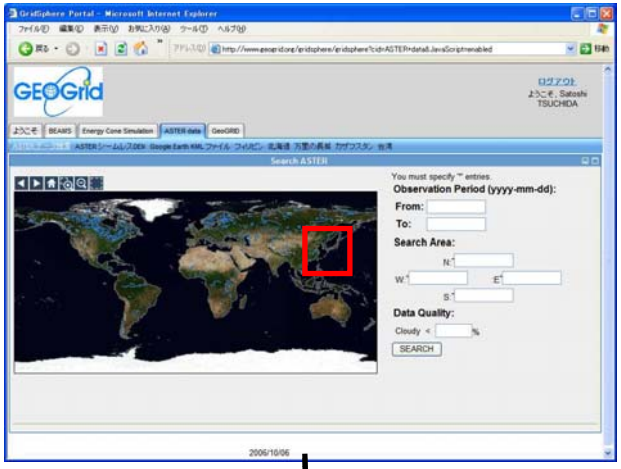




More sophisticated, more standard approach

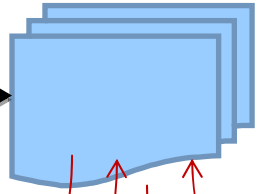


Data Integration and Access Policy



List of metadata

profile	WMS	WFS
ASTER	○	○
Formosat-2	○	



It works only if a user has been adopted a unique access policy for all data set

Search Service

CS-W

Metadata

WMS?
WFS ?

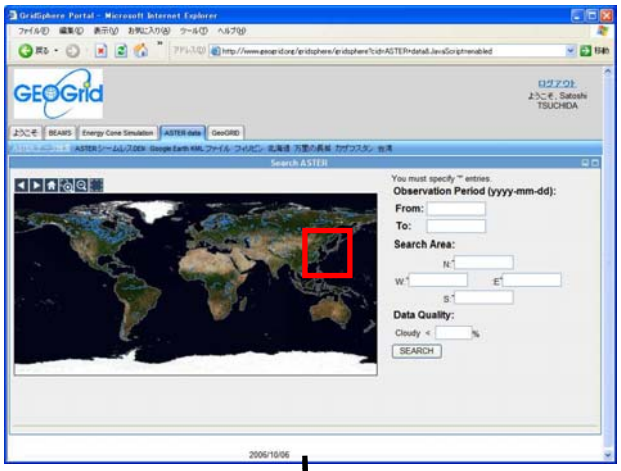
WMS/WFS

FORMOSAT-2

ASTER

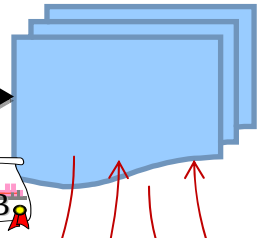
Each Access Policy is naturally diverse – GEO Grid should deal with otherwise data owner can't contribute.

From individual account to Virtual Organization



profile	WMS	WFS
ASTER	○	○
Formosat-2	○	

List of metadata



It works! but obviously data owner don't want to manage for over a million of account.

Search Service

WMS

WFS

CS-W

Metadata

FORMOSAT-2

ASTER

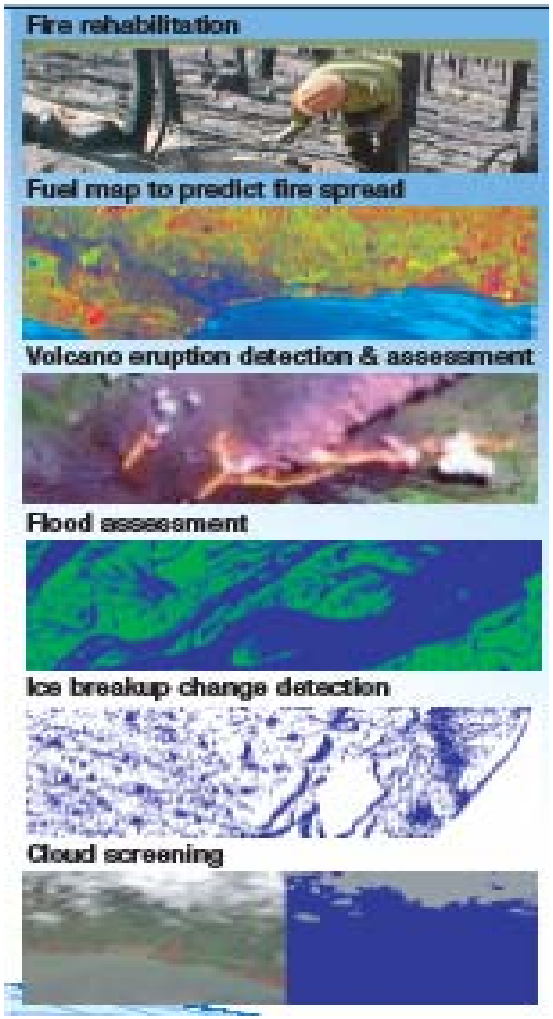
Introduces management group or tier – GEO Grid implement this as VO (virtual organization).

Earth Observation Grids, OGC and OGF

November 14, 2007

Dr. Craig A. Lee
The Aerospace Corporation
(a non-profit, federally funded R&D center)

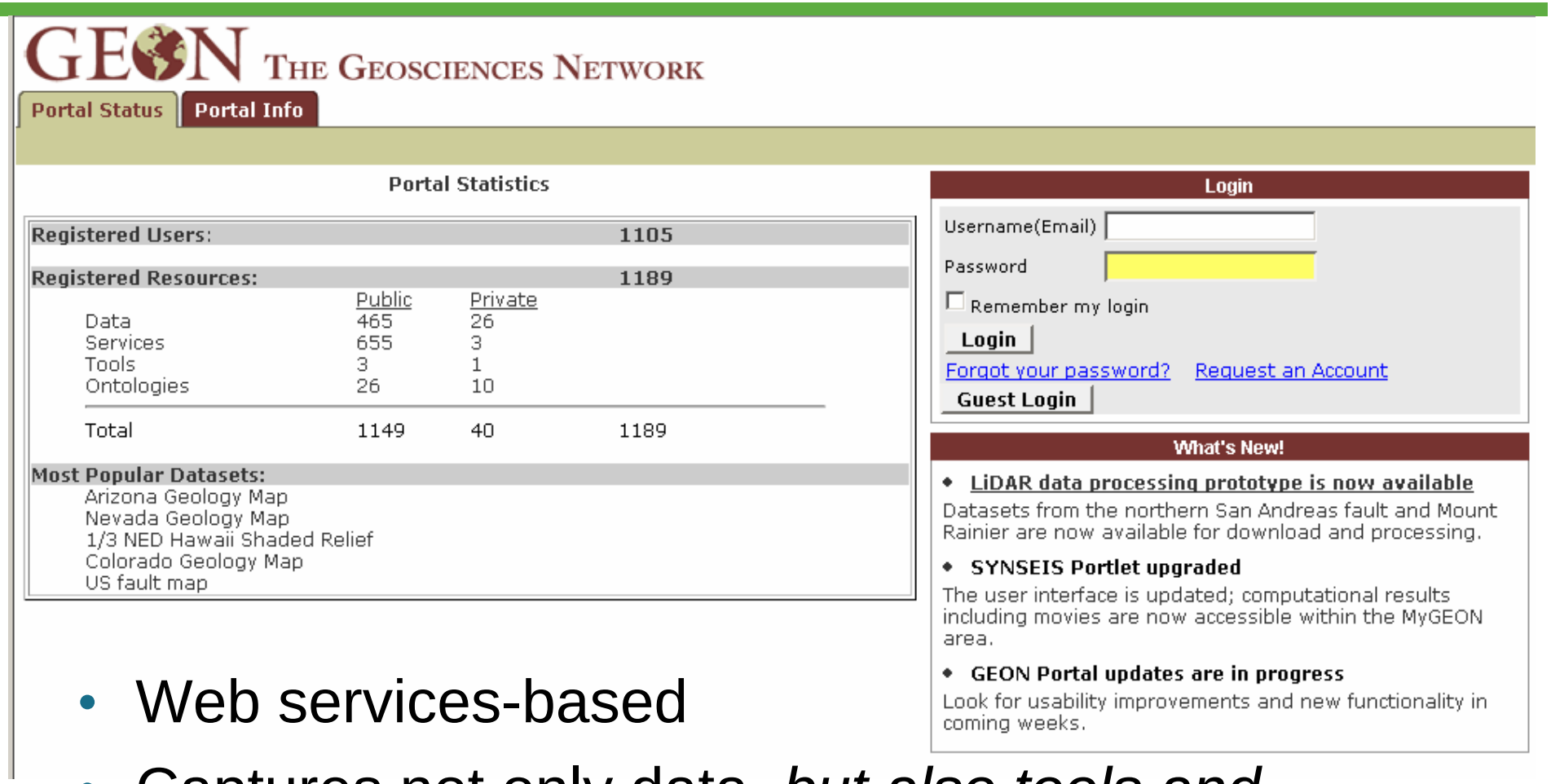
Why Do We Need EO Grids?



- EO can be applied to many important fields:
  - Weather
  - Fire prediction & rehabilitation
  - Flooding
  - Carbon cycle
  - Climatology
  - Earthquake prediction
  - Volcano detection & assessment
  - … and others

Example:

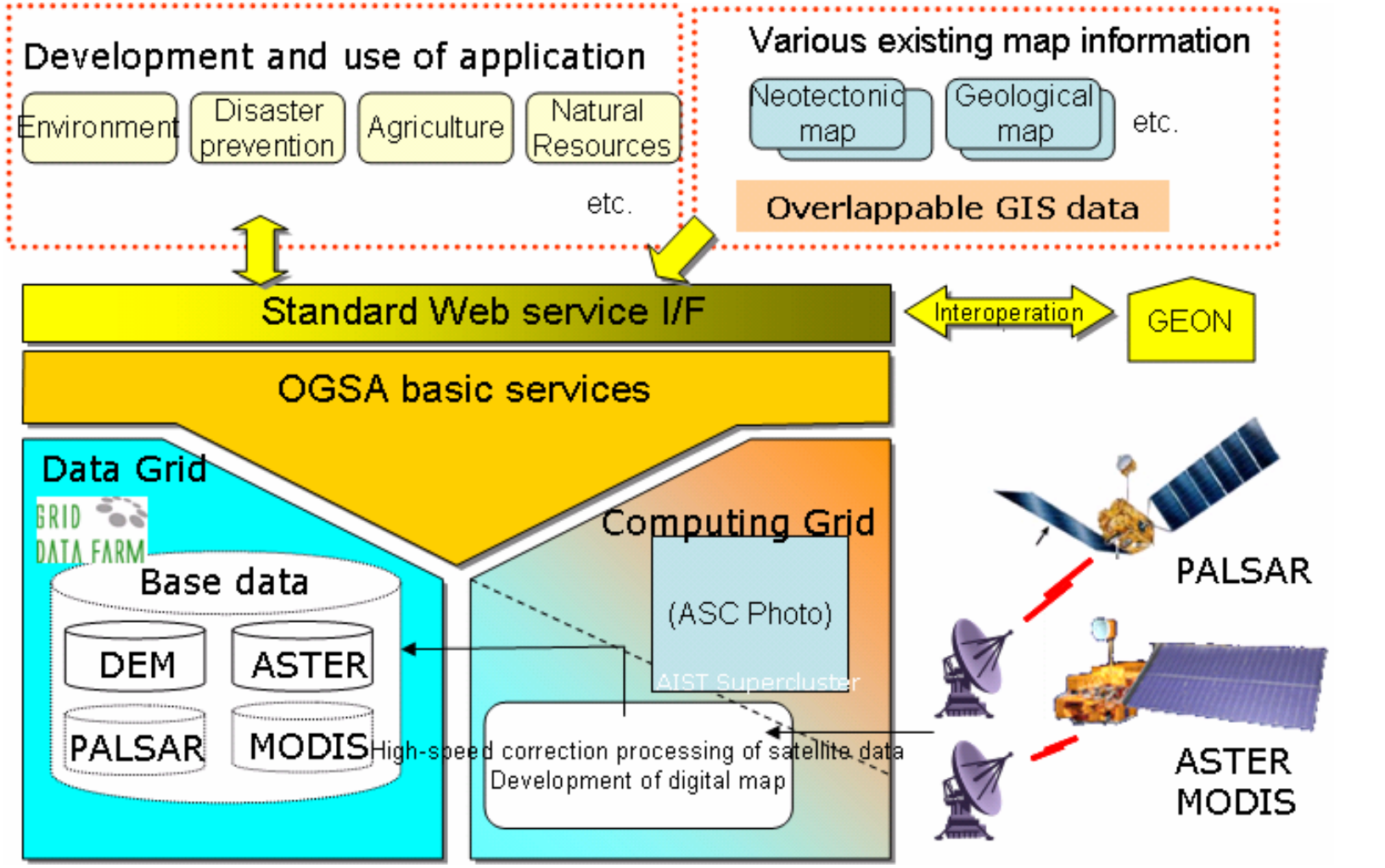
GEON – The Geoscience Network



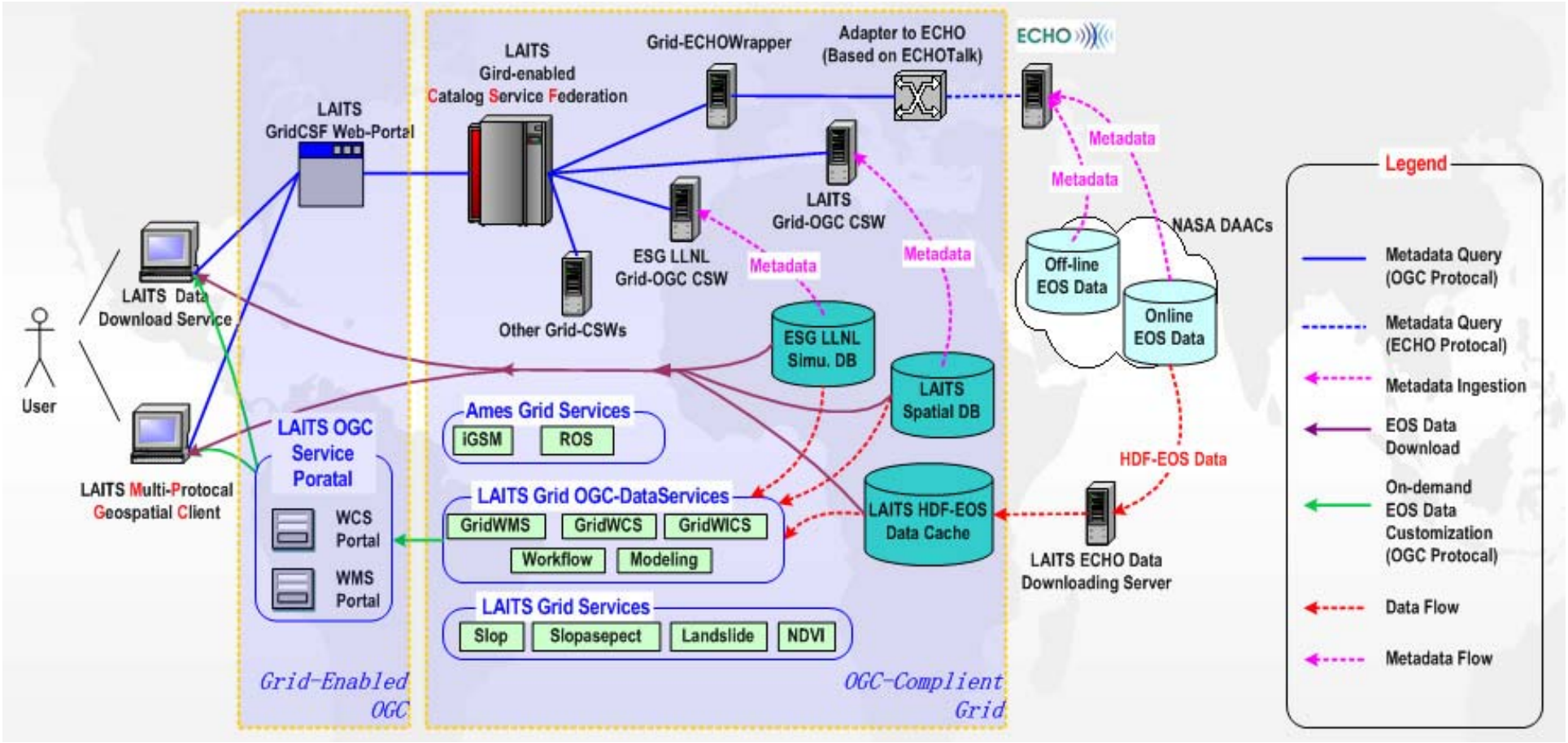
- Web services-based
- Captures not only data, *but also tools and knowledge representation*

Example:

GeoGrid – www.geogrid.org



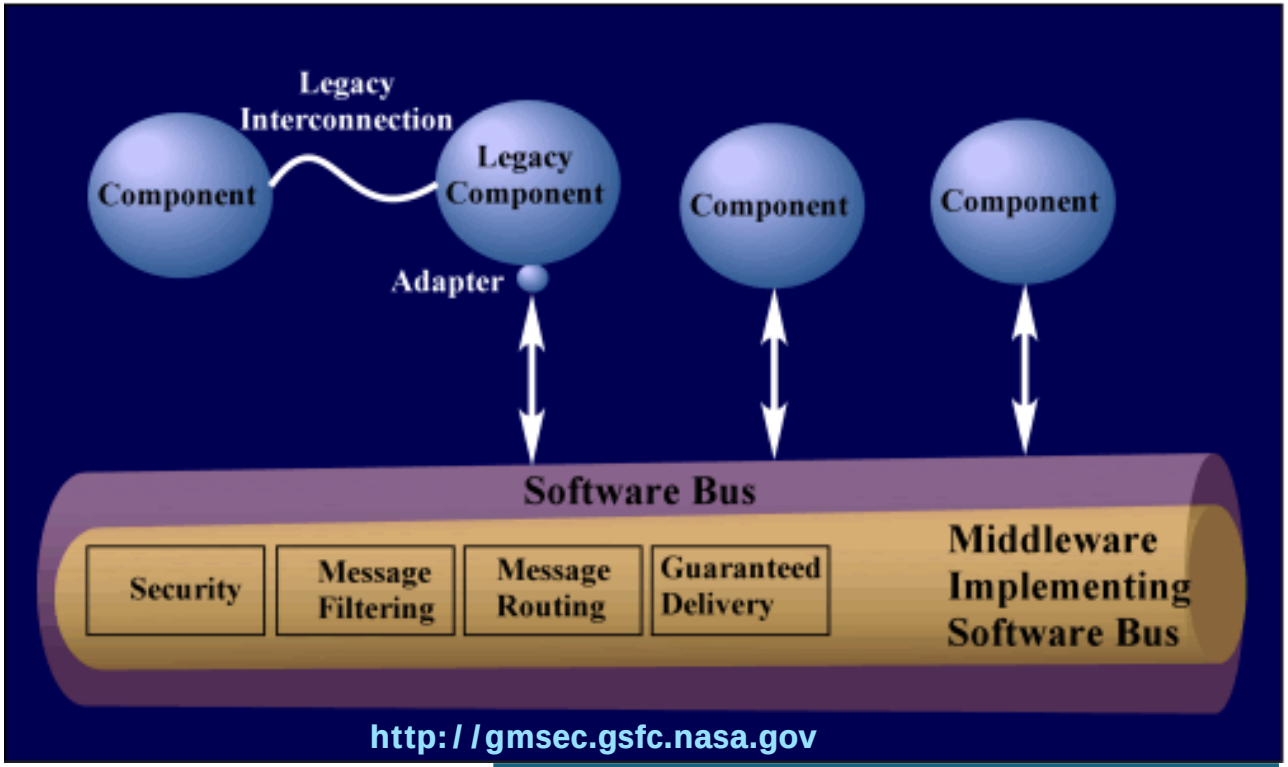
GMU Geospatial Grid



Di, Chen, Wei, Yang @ OGF-18

The Need for Common Tools

- Goddard Mission Services Evolution Center (GMSEC)
  - A *Software Message Bus framework*
- Integrates ~30 COTS/GOTS products
- Increase reuse -- Reduce cost and risk



COTS：市販コンポーネント/パッケージ
GOTS：政府・自治体コンポーネント/パッケージ

Vision & Goals

- *Achieve integrated, distributed resource management for geospatial tools*
  - Enable a geospatial service architecture
  - (Grid-enabling geospatial tools w/o using the "g-word")
- *Collaboration between OGC and OGF!*
  - Identify concrete steps/plan to accomplish this integration
  - Get buy-in from the OGC and OGF leaderships
  - Sign Memorandum of Understanding
- Execute work plan defined by MOU
  - *Working group-level collaboration!*

What is OGC?

- Open Geospatial Consortium
  - www.opengeospatial.org
- “Helping the World to Communicate Geographically”
  - Any type of geospatial data
  - Anything that goes on a map
- A Few Current OGC Standards
  - Web Map Server (WMS)
  - Web Feature Server (WFS)
  - Web Coverage Server (WCS)
  - Catalog Service for Web (CSW)
- Commercialization of these standards
  - ESRI (www.esri.com)
  - IONIC (www.**ionic**soft.com)

OGC-OGF Common Objective

- Promote international standardization with the goal of providing distributed data processing capabilities for geospatial data users in a way that is:
  - *Transparent* -- the users do not have to be aware of the exact data and computing resources they are using or the details about doing so,
  - *Interoperable* -- the data and computing resources can come from different sites, and
  - *Scalable* -- the same user model can service small computing tasks that can be done locally, as well as large computing tasks that require massive remote platforms.

Basic Terms of the MOU

- Renewable one-year periods
  - Automatically renewed unless one party provided written notification three months in advance of non-renewal
- OGC and OGF exchange memberships for key collaboration members
  - Key collaboration members have full access to both organization's documents
  - OGC has members-only web site areas

Concrete Goal

- Provide *data processing capabilities* for geospatial data users in a way that is:
  - *Transparent* -- the users do not have to be aware of the exact computing resources they are using or the details about doing so
  - *Interoperable* -- the data and computing resources can come from different sites
  - *Scaleable* -- the same user model can service small computing tasks that can be done locally, as well as large computing tasksthat require massive remote platforms
  - *Simple* -- See *Transparent*

OGC Web Processing Service

- WPS defines three operations:
  - GetCapabilities returns service-level metadata for those services supported by a particular server (but not including things like input and output parameters)
  - DescribeProcess returns a full description of one or more "processes", including its input and output parameters
  - Execute sends input arguments to the service, and up on service completion, returns the results
- The WPS specification does not address issues of
  - Data or service discovery
  - Data management, including data lifetime
  - Security, e.g., who is authorized to use a particular service
  - Service statefulness

Approach

- Integrate the WPS with a range of "back-end" processing environments
  - Generic web servers
  - Web 2.0 "mash-ups" based on AJAX (Asynchronous Javascript and XML)
  - Complete, distributed service architectures, represented by many local and national grid infrastructures that offer large parallel machines
- Be Use Case-driven
  - Focus our efforts to address high-priority issues for geospatial users
- Participate as an OGC Web Services (OWS) testbed thread

Collaboration: Some Use Case Drivers (and Possible Collaborators!)

- GEO Grid
  - Global Earth Observation Grid
  - AIST, Japan
  - Environment conservation, resource exploration, natural disaster prevention, risk management
- ADMIRE
  - Advanced Data Mining and Integration Research for Europe
  - UK e-Science Institute with EU FP 7
  - Flood and pollution-spill transfer modeling of international river systems
- CYCLOPS
  - CYber-Infrastructure for CiviL protection Operative ProcedureS
  - EU
  - Prevention and management of forest fires, flash floods

WPS Platform-Neutrality

- Current and Possible Implementations:
  - PyWPS – A Python implementation of WPS is reported at http://pywps.wald.intevation.org/index.psp
  - GMU/LAITS implementation of WPS?
  - The Grid OGC Collision Programme is investigating issues of security and workflow that are highly relevant to WPS
  - Web 2.0
    - Ad hoc way of scripting web-enabled operations that have simple requirements concerning discovery, security, reliability, etc.
  - Grid/Web Service Architectures
    - WS-Interoperability Basic Profile (WS-I)
    - Web Services Resource Framework (WSRF)
    - Open Grid Services Architecture (OGSA)
    - OGF HPC Basic Profile
  - Network-Enabled Services (NES)
    - NetSolve/GridSolve, Ninf/Ninf-G, and DIET all use GridRPC (Remote Procedure Call) standard developed at OGF

Possible Processing Scenarios to Demonstrate as on OWS Thread

Start Simple. Incrementally add and demonstrate key capabilities.

1. No service/resource discovery, just local data selection, stage to server, job submission/mgmt, data retrieval.
2. Data discovery. Selection data from any location (local or remote -- data virtualization). Stage from transparent location to server. Submit/Manage job. Retrieve result.
3. (2) with result data stored back to some virtual location via registry
4. (2) with result data left on server for subsequent use
5. A single client selects data from a virtual location and submits jobs on two different grids transparently, e.g., TeraGrid and NGS. Demonstrates both data and service virtualization
6. (5) with result data left on server for subsequent use

Some Immediate Goals

- Demonstration of a grid-enabled WPS that is easy to use
- Transparent integrated access to different data stores
- Transparent access to resources across different grids
- Understanding of how users typically want to use a capability like WPS
- Provide support to enable users to easily move from a lightweight mash-up environment to an industrial-strength grid when necessary
  - Need for stronger security
  - Need for proper discovery
  - Need for reliability, etc.
- Wider recognition and acceptance of WPS, across multiple grid and geospatial communities, as a way of doing real work across a spectrum of compute resources

Longer Term Goals

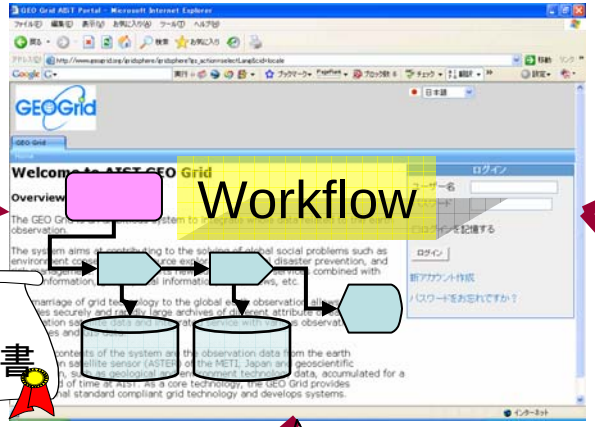
- What is the Future of GIS and Distributed Computing?
  - Google Earth mash-ups?‼
- Generalization of domain-specific data services
- Domain-agnostic infrastructure w/ services that have domain-specific interfaces
- Loose coupling between end-user map tools and system producing geospatial data
- Integration of geospatial metadata standards, federated geospatial catalogs with workflow management
- While WPS may be primary identified integration approach, survey users to determine (re-validate) what it is they really need and want to do!

GEO Grid Security とデータアクセス

ユーザ

(1) ログイン
（認証）

GEO Grid Portal



代理
証明書

(1*) portalへのログインにより
webサーバ上にユーザの
代理証明書が生成される

(2) データ検索

(2*) 代理証明書を用いた
シングルサインオン

代理
証明書

(3) データ補正
サービス等起動

(3*) 委譲によりユーザの代理
証明書が計算機上に生成
される。

代理
証明書

(4) クラスタがデータにアクセス

(4*) 委譲された代理証明書を
用いたデータベースへの
アクセス認証

ASTER
(PALSAR)

WMS WFS
GSJ
(geology map)

MODIS

ユーザ毎にデータ提供方法の選択

Summary

- OGCとOGFの協力課題の整理
  - WPS – Web Processing Service
  - Federated Data Access w/ Policy
- Workshop 開催予定
  - OGF-22 in Cambridge, Massachusetts, February 25-29, 2008,
  - on the topic of grid-enabling standard geospatial processing tools
  - 乞うご期待

OpenGridForum
OPEN FORUM | OPEN STANDARDS

Grid Consortium Japan
グリッド協議会