ECHO Services –

Foundational middleware for a Science Cyberinfrastructure

WGISS – March

2005
Goals of the presentation

- View of the future
- ECHO Overview
- Approach to Registries
- Interoperability
- Current status & activities
Supporting a comprehensive, and coordinated environment

- For science, education, governments, industry
- For system development and contribution
- Organic resource utilization
  - Resource capacity is established throughout the partner community
  - Partners offer what they wish

- Enabling Science that is...
  - Comprehensive
  - Coordinated
  - Sustained
New paradigm

- Service-oriented enterprise
- Net-centric computing
  - Power to the participants – producers and consumers
  - Enable with infrastructure of service middleware
- GEOSS momentum
Cyberinfrastructure - Underlying Fabric

Infrastructure
Services
Directory
Security
Collaboration
System Management
Storage
Discovery
...

© 2004 Blueprint Technologies, Inc. All Rights Reserved
• **Goals**
  - Support a Marketplace …
  - Enabling …

• **Approach**
  - Interoperable Registries
  - Not “the” solution/system
    • But integrated into a dynamic enterprise
    • Building systems, interoperating with systems
  - Net-centric
    • Power to the partners
**ECHO Is...**

- **Interoperability middleware solution** -
  - Open, XML-based APIs
  - Supporting net-centric architectures and solutions
  - Set of interoperable registries for both data (metadata) and services
  - Provides user accounts and common infrastructure for the registries
  - Built upon a layered architecture with extensible infrastructure for supporting community unique protocols
  - Operational since November 2002
  - Available as open source

---

**Data Registry**
*Representing data resources through metadata*

- **Publish Capability**
  - Supports Collection, Granules and Browse publication
  - Product Specific Attributes extending the standard data model

- **Discovery Capability**
  - Collection/Dataset
  - Granule/Inventory
  - Based on Z39.50 Standard
  - Full results Management

- **Access Capability**
  - Online Access
  - Legacy Order Mechanisms
  - Access Controls
    - Visibility of data resources
    - Access to data resources

---

**Service Registry**
*Service offerings leveraging Web Service Standards*

- **Publish Capability**
  - Advertisements
  - Service Interfaces
  - Service Implementations
  - Service GUIs
  - Linkage to Data Registry

- **Discovery Capability**
  - UDDI Standard Based
  - Data Registry View *(What services are appropriate for this collection/granules?)*

- **Future**
  - Service Brokering
  - Access Controls

---

**Metadata Subscriptions**
Power in the hands of the community

- **Data and Service Partners**
  - Use ECHO's Partner Interface tools to manage how their resources are represented in the registries
  - 2-level access control in Data Registry (metadata visibility, ordering)
  - In defining each dataset, partners can establish which spatial search mechanism (Cartesian, geodetic, orbital) is best suited for searching that dataset
  - Special data modules supported as needed. (ie. Orbital search module)
  - Extending the standard data model through product specific metadata

- **Client Partners**
  - Provide discipline specific, application specific or individual preferred views of the contents of the clearinghouse
  - Use ECHOs APIs to access the clearinghouse to convey a seamless, efficient view of data holdings

- **Service Partners**
  - Advertise their service offerings
    - General Advertisements, Service Interfaces (API specifications), Service Instances (web services) or GUIs
  - Classification of services through taxonomies
  - Service offerings describe in WSDL
  - Access to services through SOAP

© 2004 Blueprint Technologies, Inc. All Rights Reserved
Registries in the Enterprise

**Enterprise network solutions**
- point solutions
- open system solutions
- closed system solutions

Service Resource Registry

Data Resource Registry

Interoperability
Data Model
Taxonomies/Categorization

Brokerage Service
ECHO Service Registry

- Supporting SOA
  - Publish
  - Find
  - Bind

- Using Web Service Strategy
  - Service Description
    - WSDL (Interface & Implementation)
    - Parameter language (ECHO/ISO 19115 subset)
  - Service Binding
    - SOAP
    - WxS Issues
  - Registry
    - UDDI
Fundamental Use Cases

- Publishing
- Discovering
- Understanding
- Accessing
Publish
Resource Partners "advertise" their offerings to publicly available registries

**Data Registry**
- Public Interface: "IngestService"
  - Register, Ingest services
- Standards
  - Data Dictionary ISO 11179 based
  - Data Model based on EOS
  - ISO 19115 Compliant
  - GCMD mapping
  - FGDC mapping
- Technology - Oracle database

**Service Registry**
- Public Interface: "ExtendedServicesManagement"
  - Register service
- Standards
  - Service Description – WSDL
  - Registry – UDDI
  - OGC Service Taxonomy (ISO 19119)
- Technology – Systinet WASP UDDI Server

**Issues**
Data Model inconsistency
- Ingest adapters for data model translation and validation
Data Fidelity
- Update and delete interfaces available
- Time delays between registry and source
**Discovery**

Consumers participate in the marketplace by looking for resources of interest

---

**Data Registry**

- Public Interface: "CatalogService"
  - Query
- Inventory and Dataset level
- Standards
  - Z30.50 based queries and results management
- Technology – XML-based API
  - Oracle augmented with additional spatial query engine

---

**Service Registry**

- Public Interface:
  - UDDI: Inquiry
  - Registry (domain)
- Standards
  - UDDI
- Technology – Systinet WASP UDDI Server

---

**Issues**

Interoperability

- Data represented in a consistent data model
- Service – Data interoperability through Taxonomies
- Common Data Language for data definition and parameter specification
Brokering
Active Interoperability by facilitating the invocation of services upon data resources from different sources

Brokering Service

• Public Interface: "BrokerService"
  – Validate, Submit
• Optimize data access pattern, bind to service on behalf of consumer, manage status and results delivery
• Standards
  – TCP/IP based data access and delivery
• Technology – XML-based API

Issues

Interoperability – Matching services to data
  • Common Data Language for data definition and parameter specification
  • Validation operation (optional)
• Co-location
  • Moving data, not services (so far)
  • Temporary resources provided for facilitating data movement

Resource Efficiency
  • Validation operation (optional)
• Security
  • Registered Users only
  • Moving data, not services
Multiple paths for data movement in support of brokering or orchestrating services

Brokering Service optimizes data movement from path based on:
- Data partner access options
- Service Partner access and delivery options
- Consumer requests
Orchestration Service

- Public Interface: "Orchestrator"
  - Validate, Submit
- Ensure service invocation sequence is valid
- Translucent and Aggregate/Opaque capabilities
- Standards
  - OCG efforts (OWS 2.0) ISO 19119
- Technology – XML-based API

Issues

Interoperability –
- Data represented in a consistent data model
- Service – data interoperability through taxonomies
- Common data language for data definition and service parameter specification

© 2004 Blueprint Technologies, Inc. All Rights Reserved
Service model

Service Registry

Partner
- organization
- reference web page
- provider contacts
- description of holdings
- description of services
- provider type

Web Service Interface
- interface name
- service description
- reference web page
- data types
- messages

Service
- reference web page
- service name
- description
- access point

Taxonomy
- name
- description
- reference web page

Taxonomy Node
- name
- value
- reference web page
- description

Web Service Implementation
- binding protocol

Web Service GUI

Advertisement

comprised of

classified by

realizes

is a type of

© 2004 Blueprint Technologies, Inc. All Rights Reserved
Interoperability

• Between registries
  - Service Registry through Taxonomies
  - Data Registry through Catalog Service
    • <details>

• External Interoperability
  - Adaptors
    • Protocol interoperability with Partners for order and ingest
    • Data model interoperability through Ingest Adapters and data set registration
    • Can establish "Views" of results

  - API's enable external other interoperability through adaptors
    • Ostensibly, ECHO clients
Classification

- **Taxonomies (so far)**
  - Data Set
    - Data Partner → Dataset → Version
  - Data Format
    - Similar to mime-type
  - Service Type
    - OCG/OSE model
    - Looking for others

- **Other Classification schemes possible**

- **Potential for user defined taxonomies**
  - Dynamically defined and leveraged
  - Profile entities
  - Semantic web integration
Service Semantic understanding

- **Ensure that services can work together**
  - Enabling binding is easy for a registry
  - Flexibility always has its price
    - Understanding relevance of context and parameters in an open, dynamic and loosely coupled network enterprise is the challenge
  - Developing a Common Data Language for parameter specification

- **Service Metadata**
  - Common way of describing
    - Partner
    - Service Type
    - Interface
    - Source
    - Capability
    - Understanding
    - Applicability
UDDI v3

• Upcoming Technology evolution

• Key Features
  – Multi-registry environments
    • Allows for federation of service registries
  – Publisher specified keys
    • Beyond UUIDs
    • Flexible and powerful namespace management
  – Security enhancements
    • Adds digital signing to core data types
Rules of Engagement for Service Partners

• Register

• Provide Reference Web Page

• Active Services
  – Provide WSDL
  – SOAP binding point

• Optionally:
  – Categorize resource

• Brokered
  – Declare Input access support
  – Declare Output support
# Current Data Partner Status

## 5 Operational:

<table>
<thead>
<tr>
<th>Data Partner</th>
<th>Department</th>
<th>Collections</th>
<th>Granules</th>
<th>Browse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oak Ridge National Laboratory (ORNL)</td>
<td>Department of Energy, ORNL, Oak Ridge, Tennessee</td>
<td>679</td>
<td>105,706</td>
<td>0</td>
</tr>
<tr>
<td>Distributed Active Archive Center (DAAC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land Processes EOS Core System (ECS) DAAC</td>
<td>USGS Eros Data Center, Sioux Falls, SD</td>
<td>87</td>
<td>11,587,961</td>
<td>2,349,125</td>
</tr>
<tr>
<td>Goddard ECS DAAC</td>
<td>NASA/GSFC, Greenbelt, MD</td>
<td>357</td>
<td>18,071,296</td>
<td>1,890,246</td>
</tr>
<tr>
<td>Socioeconomic Data and Applications Center (SEDAC)</td>
<td>Columbia University, New York, NY</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alaska SAR Facility</td>
<td>University of Alaska, Fairbanks, AK</td>
<td>14</td>
<td>1,872,361</td>
<td>0</td>
</tr>
</tbody>
</table>

## 3 Development/Test:

<table>
<thead>
<tr>
<th>Data Partner</th>
<th>Department</th>
<th>Collections</th>
<th>Granules</th>
<th>Browse</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Snow and Ice Data Center (NSIDC) DAAC</td>
<td>University of Colorado, Boulder CO</td>
<td>35</td>
<td>62,066</td>
<td>0</td>
</tr>
<tr>
<td>Atmospheric Sciences Data Center (ASDC) DAAC</td>
<td>NASA/LaRC, Hampton, VA</td>
<td>93</td>
<td>2,954,549</td>
<td>1,507,711</td>
</tr>
<tr>
<td>Stennis Space Center Data Purchase Project</td>
<td>NASA/SSC, Stennis Space Center, MS</td>
<td>22</td>
<td>17,766</td>
<td>15,892</td>
</tr>
</tbody>
</table>

Total data items in ECHO = 34,671,705
## Current Client Partner Status

### 2 Operational:

<table>
<thead>
<tr>
<th>Project</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercury EOS</td>
<td>Oak Ridge National Laboratory (ORNL) DAAC, Department of Energy, ORNL, Oak Ridge, Tennessee</td>
</tr>
<tr>
<td>Power User Interface (script based)</td>
<td>Earth Science Data and Information System (ESDIS) Project, NASA/GSFC, Greenbelt, MD</td>
</tr>
</tbody>
</table>

### 9 Development/Test:

<table>
<thead>
<tr>
<th>Project</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASF ECHO Client</td>
<td>University of Alaska, Fairbanks, AK</td>
</tr>
<tr>
<td>Data Validation User Interface (DVUI)</td>
<td>ESDIS Project, NASA/GSFC, Greenbelt, MD</td>
</tr>
<tr>
<td>MODIS Website</td>
<td>MODIS Project, NASA/GSFC, Greenbelt MD</td>
</tr>
<tr>
<td>SNOWI-E</td>
<td>NSIDC, University of Colorado, Boulder, CO</td>
</tr>
<tr>
<td>WISRD</td>
<td>NSIDC, University of Colorado, Boulder, CO</td>
</tr>
<tr>
<td>WIST (Warehouse Inventory Search Tool)</td>
<td>ESDIS Project, NASA/GSFC, Greenbelt, MD</td>
</tr>
<tr>
<td>New Earth Observer (NEO)</td>
<td>Earth Science Directorate, NASA/GSFC, Greenbelt MD</td>
</tr>
<tr>
<td>Simple MODIS ECHO Client (SIMECC)</td>
<td>MODIS Rapidfire Project, NASA/GSFC, Greenbelt MD</td>
</tr>
<tr>
<td>Invasive Species</td>
<td>NASA and USGS</td>
</tr>
</tbody>
</table>
Current Activities

- Brokering Services
- Service Orchestration – exploration within the ESIP community
ESIP Participation

- **Strategic Direction**
  - Promote Services that can participate in SOA systems
  - Descriptions in WSDL
  - SOAP access/interface

- **Establish a testbed of a service registry**
  - Using ECHO
  - Semi-public

- **Explore techniques and technologies of orchestration**