The EarthScope Portal

Integrated Data, Data Products, and Tools

October 11, 2004

USArray Data Products Workshop, Oregon

Christel Hennet

EarthScope Analyst

Context for Portal

- The promise of EarthScope is to take a multi-disciplinary approach to studying the structure and evolution of the North American Continent and the physical properties that control earthquakes and volcanoes.

- Requires making different data types accessible to a broad range of scientists and educators.

EarthScope Portal

The most important legacy of the National Science Foundation's largest investment in solid-Earth Science, and a fundamental data base for the next generation of Geoscientists.

EarthScope Portal - Mission

1) Provide seamless, single-point access to all EarthScope data, data products and tools, independent of whether they are developed through the EarthScope facilities or through the RFP process.

"Proposals to this competition should include aspects of the following elements:… A clear commitment to work with the EarthScope Office to make data products and tools openly accessible through the EarthScope data and products portal."

-- NSF Program Announcement: EarthScope: Science, Education, and Related Activities, Program Solicitation: NSF 04-589

2) Provide access to other relevant data sets and associated information.

3) Provide first order capability to integrate diverse EarthScope data sets

Preliminary Development Plan

- **Internal EarthScope Portal Working Group**
  - *Membership (Hennet, Guillemot, Weiland, Ellsworth, Anderson, Eriksson, Ahern, Taber, Pieper)*
  - *Meetings*
  - *Schedule*

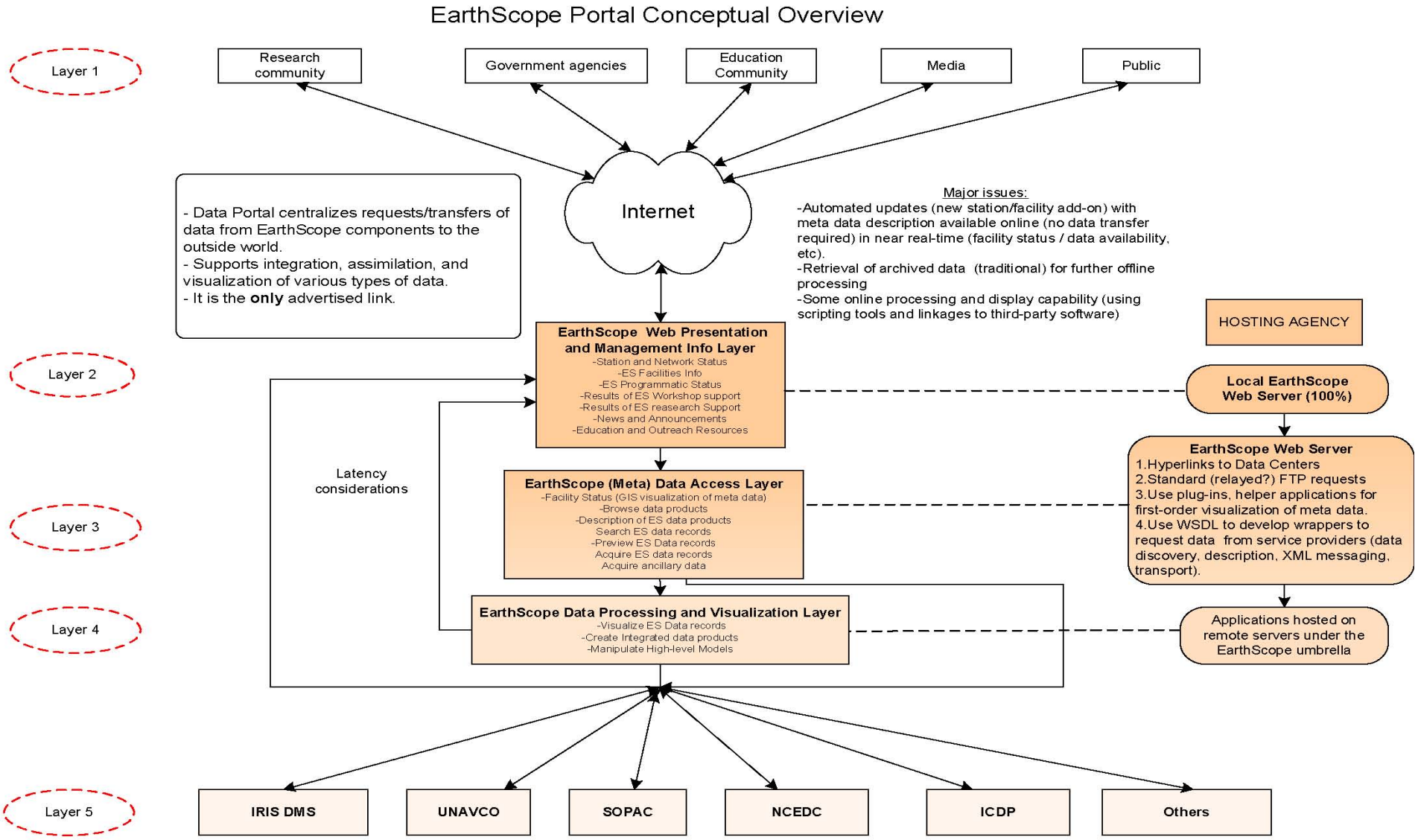
- **Near-Term Tasks**
  - *Make all data seamlessly accessible through EarthScope Portal*
  - *Coordinate efforts between EarthScope components to promote the sharing of resources and capabilities, and to avoid unnecessary duplication of effort.*
  - *Identify system for development of Level 2 data products and below for inclusion within the EarthScope O&M Proposal*

- **Time-Frame**
  - *EarthScope O&M Proposal may be submitted as early as the end of this year.*

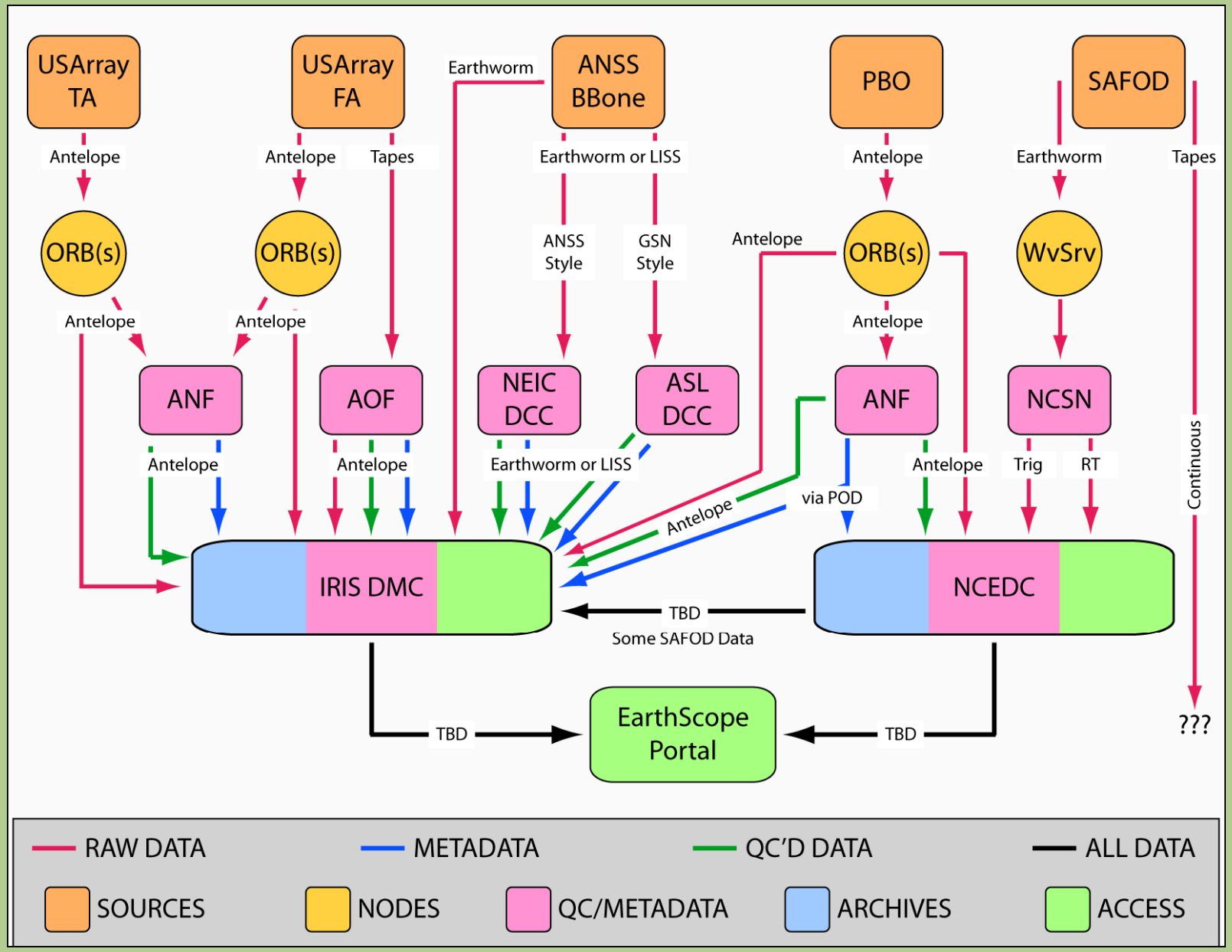
General Requirements

- All content is platform and browser independent
- All data are accessible, independent of who, or what EarthScope component collected it.
- All software are open-source
- Access to other relevant data sets available

Portal Conceptual Overview



EarthScope Seismic Data Flow



ES Data Product Level Definition

- *Based on ESEC-suggested data product levels*
- *EarthScope has developed a common terminology for describing EarthScope products*
  - *within the EarthScope facility*
  - *the broader EarthScope community*
- *The terminology will extend to data generated outside the facilities (incl. ES funded proposals, RFPs, workshops)*

Level 0

- Level 0: Raw data and Metadata:
  - Unprocessed data and data products that have not undergone quality control
  - May be available within minutes
  - Examples include: Physical samples, borehole logs, seismic waveforms, borehole

Level 1

- Level 1: Quality Controlled Data and Associated Metadata
  - Defined as data that has passed quality assurance procedures
  - May be available within 24 hours
  - Examples include: borehole logs, seismic waveforms, laser strainmeter data

Level 2

- Level 2: Low-level Derived Products and Associated Metadata:
  - Require facility scientific and technical interpretation and include multiple-sensor data
  - Examples include: borehole and laser time series, full geodetic velocity solutions, Earthquake locations and magnitudes

Level 3

- Level 3: Mid-level Integrated Products and Associated Metadata
  - Require researcher (PI) driven analysis and interpretation, model based interpretation using other data or strong prior assumptions
  - Examples include: 3-D crustal model of the SAFOD environment, phase-velocity maps, time-variable strain rates, tomograms

Level 4

- Level 4: High Integrated Products and Associated Metadata
  - Require PI driven scientific interpretation and multidisciplinary data integration and include model-based interpretation using other data and/or strong prior assumptions
  - Examples include: Large-scale 3-D and 4-D tectonic models, Earthquake slip distribution models

EarthScope Data Product Levels

- **Level 0 through Level 2** *data products - submitted as part of EarthScope O&M Proposal*.

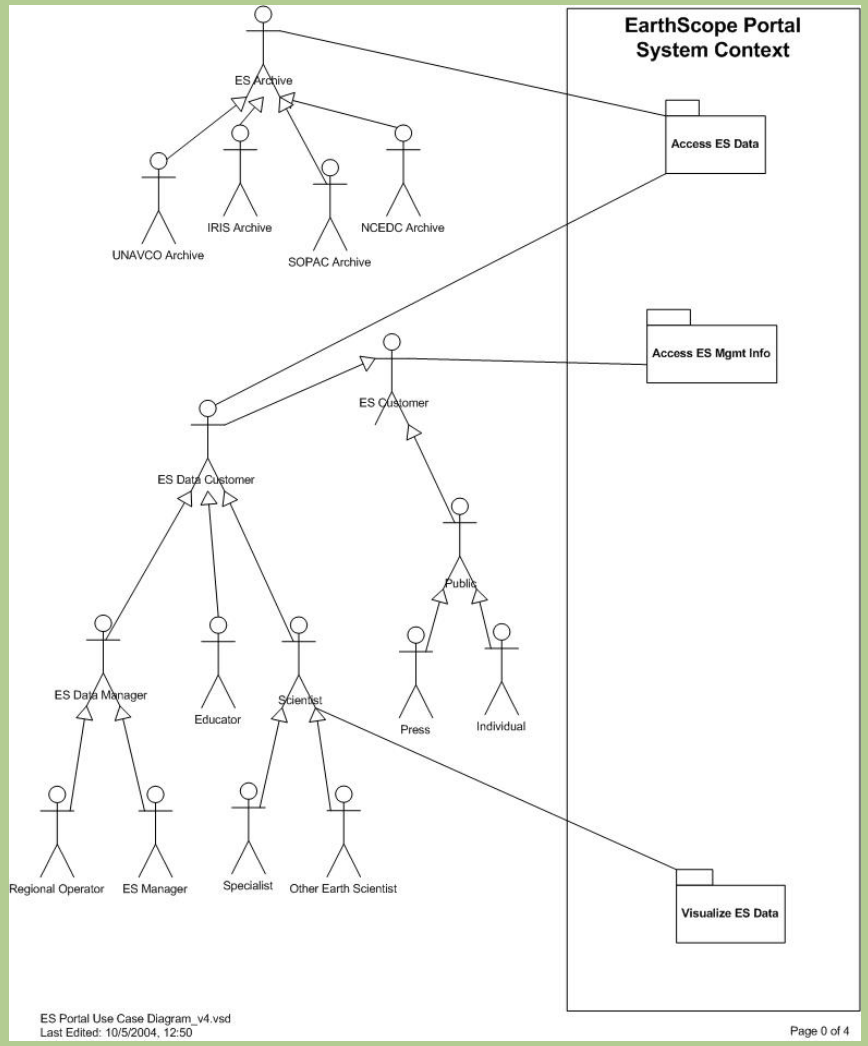
- **Level 3 and Level 4** *data products - funded through RFPs and the peer-review process, with commitment to make them available through the EarthScope Portal*.

Basic Portal Structure

I. Management and Information Layer

II. Data Access Layer

III. Data processing and Visualization Layer



I. Management and Info. System

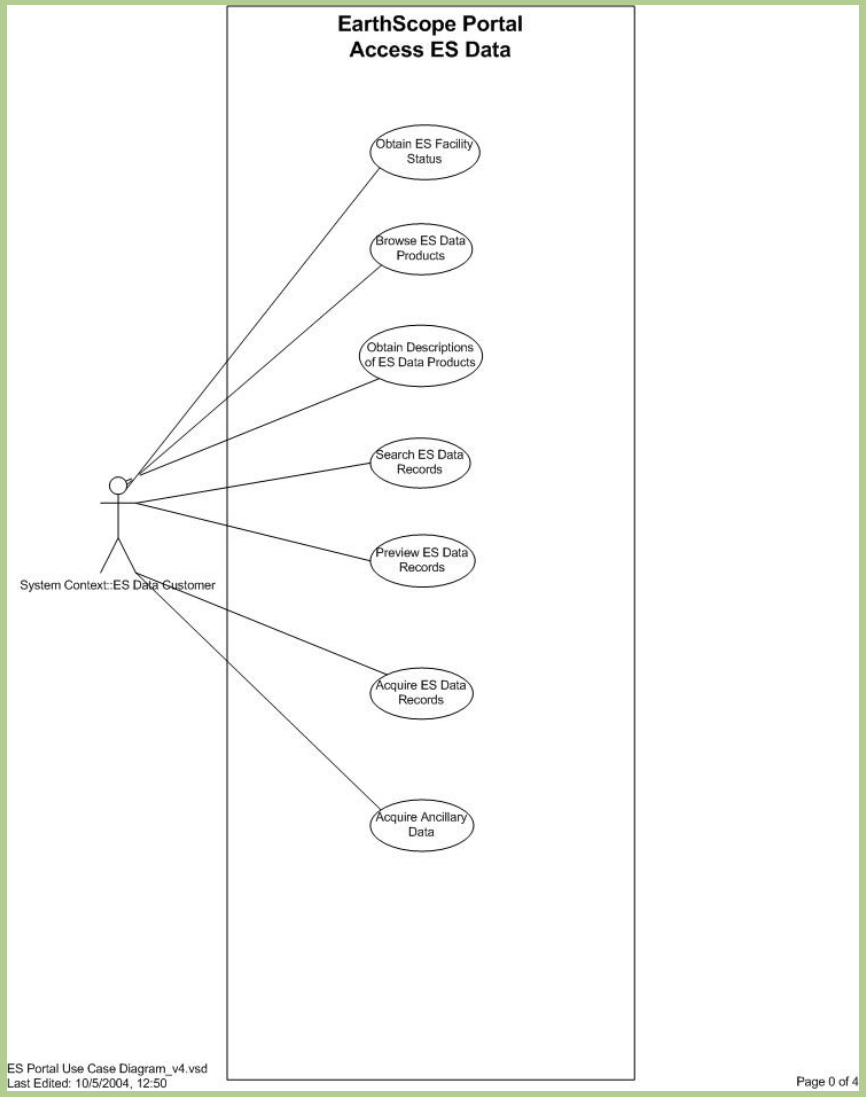
I. Management and Information System

- *Station and Network information*
- *Facilities Information*
- *Programmatic Status*
- *Results of Workshop support*
- *Results of research support*
- *News, Announcements*
- *E&O resources*

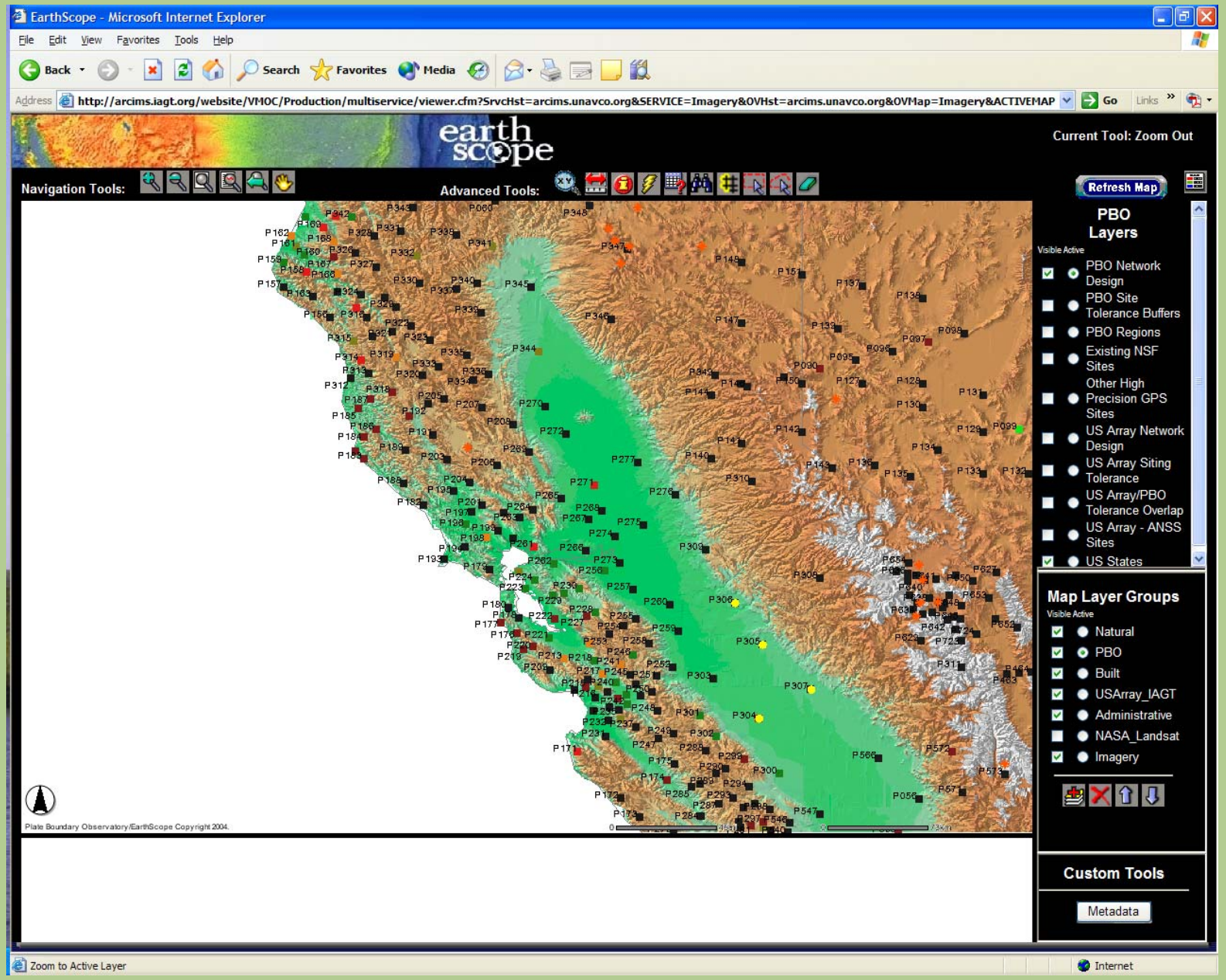
II. Integrated Data Access

II. Data Access Layer

- *Facility Status (GIS visualization of metadata)*
- *Browse data products*
- *Description of data products*
- *Search data records*
- *Preview data records*
- *Acquire data records*
- *Acquire ancillary data*



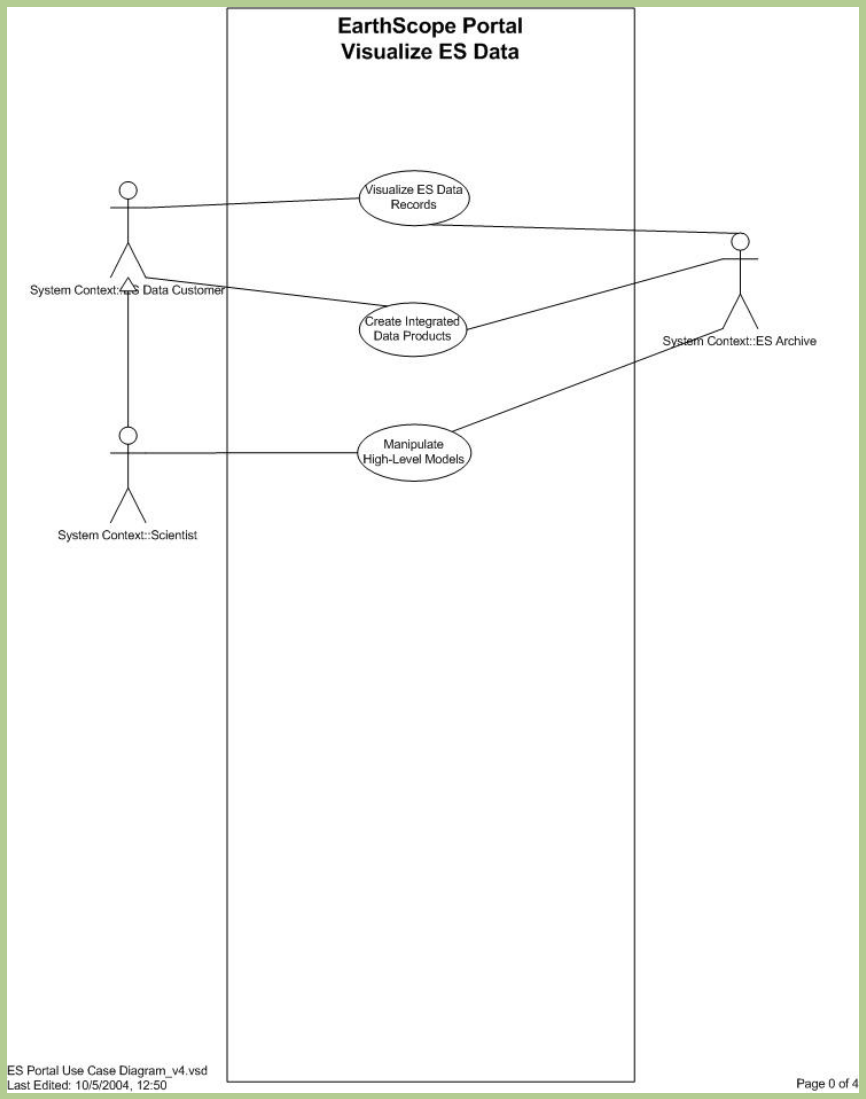
Possible Layer I & II Access Tool



III. Data Processing and Visualization

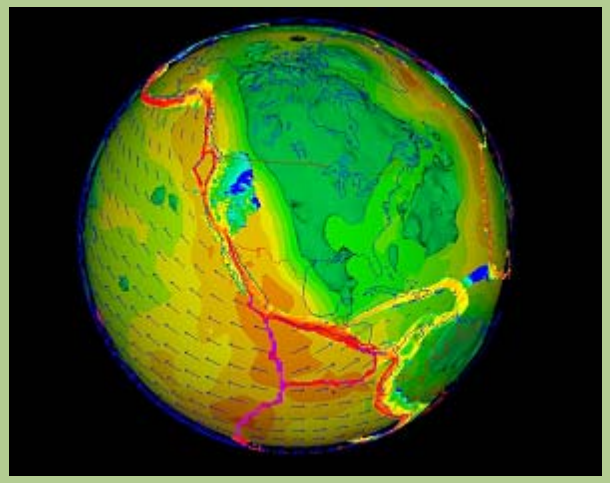
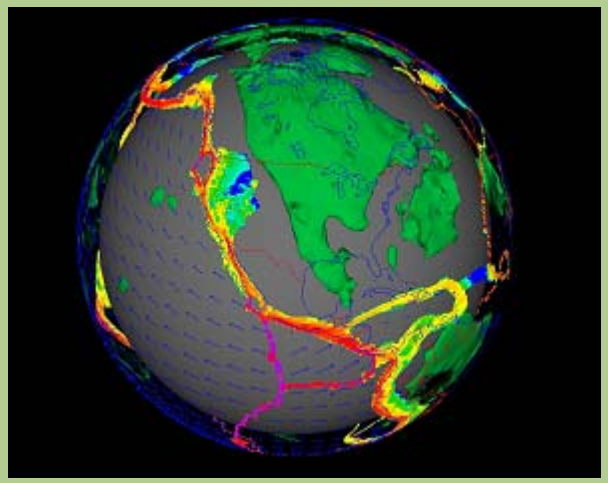
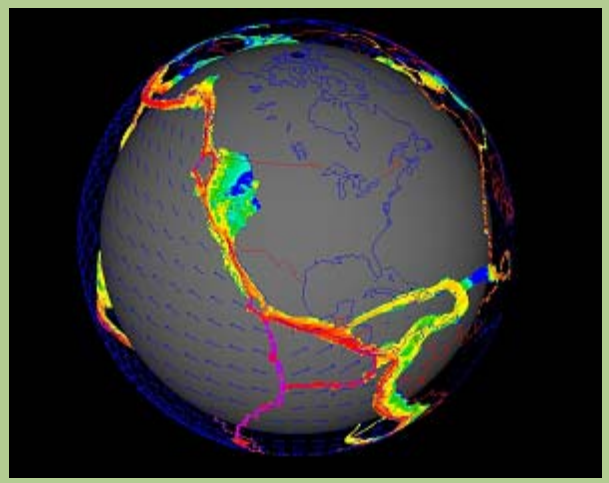
III. Data Processing and Visualization Layer

- *Visualize data records*
- *Create Integrated data products*
- *Manipulate high-level models*



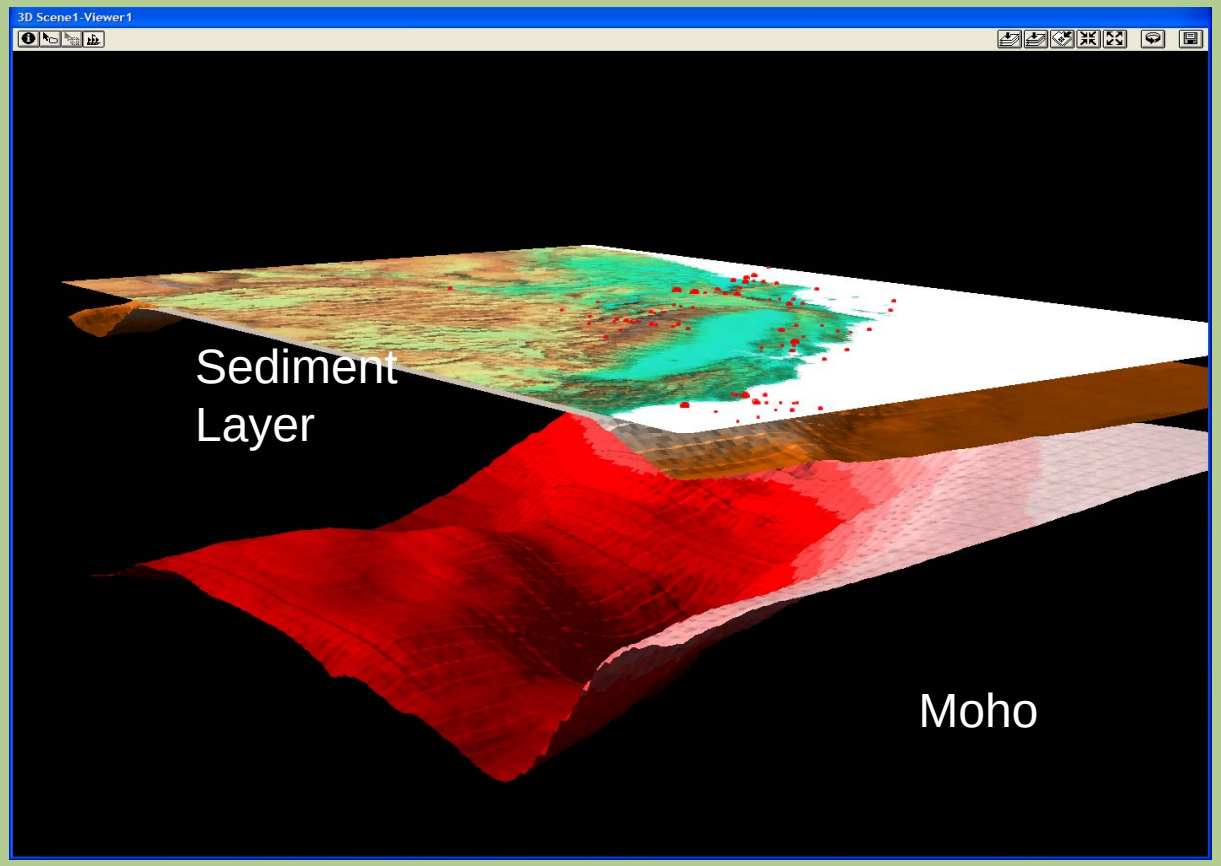
Integration w/ GEON developments?

- GEON –visualization models of integrated data sets

Synseis

Region-specific structure Model from GEON ArcIMS webservice



Slide from Dogan Seber, SDSC

EarthScope Seismic Data Products

Technical description of data products level 0 – level 2 as basis for portal developement

Type	Level	ES Component	Description	Product Type	Format	Frequency	Latency
Seismic	0	U-S-P	Raw Continuous	BUD	miniSEED	continuous	seconds-minutes
Seismic	0	U-S-P	Raw Segmented	SPYDER®	SEED	by event (Ml?=4.0)	2 hours
Seismic	0	U-S-P	QA Estimates		Oracle	continous	1 day
Seismic	0	S	Raw Continuous	Assembled	SEGD	continuous	
Seismic							
Seismic	1	U-S-P	QA'd segmented data	FARM	SEED	by event	5 weeks
Seismic	1	U-P	QA'd continuous data	Archive	SEED		
Seismic							
Seismic	2	U-S-P	Inst Corrected	NA	SEED	continuous	1 day
Seismic	2	U-S-P	Inst Corrected	NA	SEED	by event	2 hours
Seismic	2	U-S-P	Record Sections	Exist without a name	image	by event	2 hours
Seismic	2	U-S-P	Picks	NA	Oracle TBD		
Seismic	2	U-S-P	Hypocenters		Oracle		
Seismic	2	U-S-P	Station Metadata		dateless SEED	-	pre iinstallation
Seismic	2	U-P	CMTs		Oracle TBD		
Seismic							
Seismic	3						
GPS	Level		Description		Format	Frequency	Latency
GPS	0	P	15-sec raw data		BINEX/GHC RI	Daily	2 hours
GPS	0	P	5-sps raw data		BINEX/GHC RI	As needed	2 hours
GPS	0	P	SGPS Raw data		BINEX	As needed	Less than six mon
GPS	0	P	Station metadata		POD/XML-MD	Daily	Real-time

Next Steps

- Development of specific plans for EarthScope O&M Proposal

- Development of integrated collection, Q/C. archiving and distribution for all EarthScope seismic data regardless of origin (i.e. PBO's 175 Borehole seismic stations, SAFOD seismic arrays)