



LP DAAC Development

Land Processes DAAC (LP DAAC) User Working Group August 22-23, 2007

Brian Sauer EOS Project Architect SAIC Contractor to the USGS



Agenda

- Major Technologies
- Architectural Goals
- Development Methodology
- Access & Delivery
 - Seamless, MRTWeb, WMS/WCS, Google
- Archive & Processing
 LODADS, E-MODIS
- Things that didn't work...
- Future Plans



What are the Major Technologies?

EOSDIS Core System (ECS)

- Heterogeneous Storage Infrastructure (SAN, Tape Libraries)
- Open Systems (Linux, Intel)
- Web Services Interfaces
 - ECHO WSDL Ordering Component (EWOC)

DAAC Unique Extensions

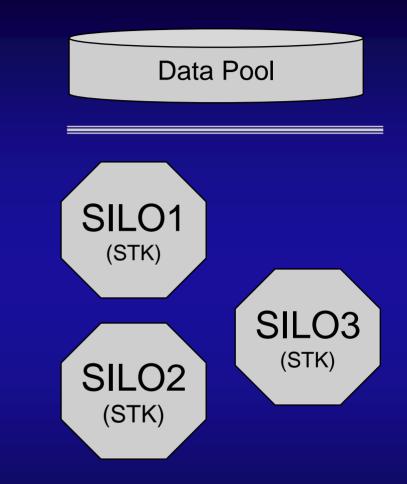
- Web Services Interfaces
 - Machine to Machine Gateway (ECHO / EWOC Interfaces)
 - USGS Billing and Accounting
 - MRTWeb
- Message Service Bus
 - Used behind MRTWeb
- OpenLaszlo Flash Interface
 - ASTER Data Acquisition Request (DAR) Tool
- Extensive use of PERL
 - Customization of S4PM
- VMWare





LP DAAC Archive

- Nearly 3PB of capacity when fully loaded with 9940B tapes
- Proven ingest into the deep archive of over 3TB of day continuous while accepting more than 1TB/day of orders for distribution
- Data pool archive is separately managed
- Data is not considered "ingested" until data is stored in the deep archive





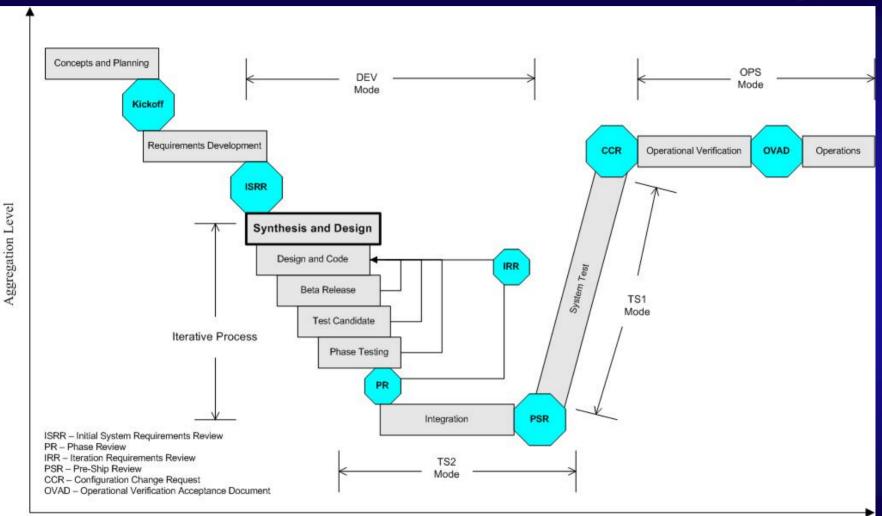


Architecture Goals

- Customer Data Access Driven
 - Fast, powerful, spatial / simple
 - Access to Data Services ie. Subset, mosaic, reproject, reformat
- Reliable
 - As an archive center, data stewardship holds significant value
- Scaleable
 - Tradition of large orders and mass distribution of data
 - Allow for inclusion of multiple data sets
- Simple
 - Integrate evolved ESDIS Core System (ECS)
 - Publish data to ECHO
 - Physical reduction in footprint
- Interoperable
 - Interoperable with ECHO, EROS Billing and Accounting, GDS, MODAPS, etc



LP DAAC Development Methodology









Iterative Development

- Utilize an Iterative Development Process
 - Involve key stakeholders heavily early in the process
 - Several development iterations
 - Stakeholders perform initial tests on early adopted iteration code; gather stakeholder feedback
 - Formal system test period





Access and Delivery

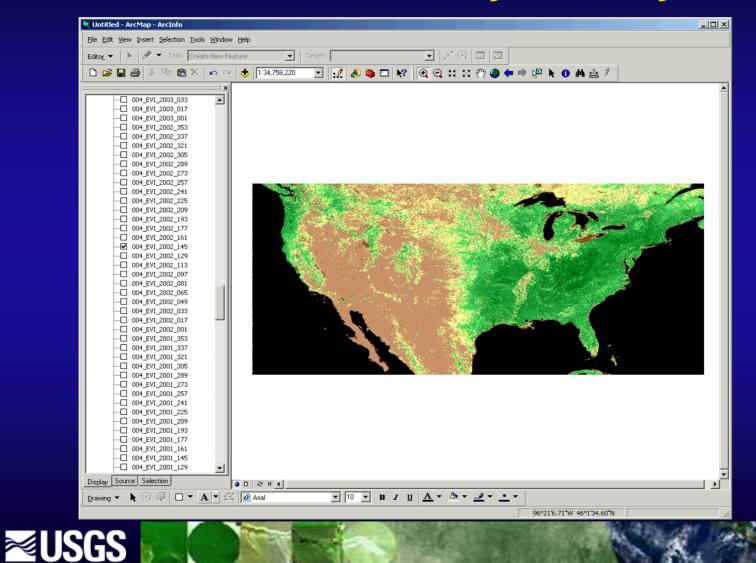
USGS Seamless WMS Capability

- MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid
- MRTWeb
- NASA GIO OGC Prototype Project
- Google Discussion





USGS Seamless Project OGC WMS MOD13A2 EVI - 2002 Day 145 Layer

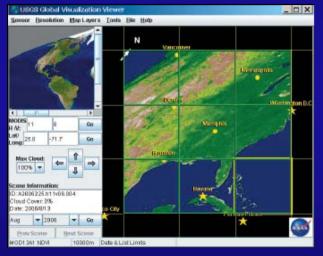




MRTWeb – Software Re-use

LP DAAC adapted and integrated two existing technologies to create MRTWeb

GloVis (Selection Interface)



MRT (Processing Tool)

			-
Be actor	jermus Hen		
Nexo:			Theodie of Law
t put lites			Special Internet Re
UMOREICA	a Andreast State and Andreast		
		Ginni ngati me	Column Ric
		end application	Colgoi Bin 1990
-		Were Introduct The	
		10	Plander gilling Types
Regard Marinela			Feranet Netgiber
Number	Ner (1418, 580), 2808, 5	461 3,060 3 462 3 463 5 60 406 3000 3 88 9400 3480 3 8 3400 368 9406 3600 768	Ell Degester Deventers -
No and an of a	Iver (1410, Strot, Strot, Strot), S armster, (2010, 3400, 341 Listen, of Carver, 1437 - 1	406, 2000, 3188, 2400, 3490, 2 8, 3400, 2688, 2406, 3600, 218	Edd Despection The resident - Ontari Piert Ster
history of Astronomy Attractions at	over (1410, Stritt, Sector, Sector, S remotiver, (2010, 3400, 340 Library, of Carrier, 1437–1 radia 8	100, 5100, 3181, 5100, 3400, 3 8, 5100, 3181, 3106, 3100, 310 311571674711 (#1	Outer theo Ser
Automotive for Land and the Automation for Star Land Land Star Land Land	nave (1410, 5400, 5400, 540) Internet (2410, 5400, 541) Internet (470, error (4716) 2.1 4 2.1 4 3.1 30	106, 5100, 5188, 5100, 5400, 5 6, 5100, 5188, 5176, 5100, 518 518, 5176, 5176, 5100, 518 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	and the second se
Automotive of the second secon	Inter (14) 400 300 5 Interior (2003)0340 Interior (2003)0340 Interior (2003) Interior (2003) I	TR. 2000 3188 2000 3483 20 5 300 3485 200 300 20 115275301 1 0 0 0 0 0 0 0 0 0 0 0 0 0	Outer theo Ser
Australia Co	International Activities of the second and the second and the second activities of the second activitities of the second activities of the second	TR. 2000 3188 2000 3483 20 5 300 3485 200 300 20 115275301 1 0 0 0 0 0 0 0 0 0 0 0 0 0	Carenada
Automotive of Landson of Landson of Landson of Landso	neer 148, 4m, 3m, 5m renter 121 C380, 30 team of Cross 1411 2,1 8,1 8,1 8,1 1, as	TR. 2000 3188 2000 3483 20 5 300 3485 200 300 20 115275301 1 0 0 0 0 0 0 0 0 0 0 0 0 0	Carenada
Automotive of Landson of Landson of Landson of Landso	bar (14) 400 500 50 meter (- 24) 530 50 1000 40 50 1000 40 50 1000 40 50 1000 40 1000 4	TR. 2000 3188 2000 3483 20 5 300 3485 200 300 20 115275301 1 0 0 0 0 0 0 0 0 0 0 0 0 0	Connect Pres See
Annual of a second seco	have cital in strong banks and marketing (2010) 2010, 2010 Jacobia of Cartoner, 1823 of Arabia o	100.000.0149.020.0149.020 100.0149.020.0149.020.014 100.0149.010.000.01 100.0140.000.01 100.0140.000.00 100.000.000 100.000.000 100.0000 100.0000 100.00000 100.0000 100.0000 100.00000 100.00000 100.000000	Connection Community Land Personale File - Sans Forced on File - Parandom File
Automotive of Landson	new chail dam jam 3 martie (24) (300,24) lateral of Carone (43) 2,1 8,1 8,1 8,1 8,1 8,1 8,1 8,1 8	en and and an and an and an and an and an and an	Connection Community Land Personale File - Sans Forced on File - Parandom File
Annual of a second seco	have cital in strong banks and marketing (2010) 2010, 2010 Jacobia of Cartoner, 1823 of Arabia o	100.000.0149.020.0149.020 100.0149.020.0149.020.014 100.0149.010.000.01 100.0140.000.01 100.0140.000.00 100.000.000 100.000.000 100.0000 100.0000 100.00000 100.0000 100.0000 100.00000 100.00000 100.000000	Connection Community Land Personale File - Sans Forced on File - Parandom File

Mosaic tiles Subset an area from a tile, mosaic, or time series Eliminate unwanted bands or layers Define projection (n=14) Set resampling options Choose file format (HDF, GeoTIFF, binary)

Rapidly visualize tiles within map context Navigate through time and space Select tiles of interest for processing





MRTWeb Architecture

MRTWeb Interface (GloVIS Servelets)

Business Process and Orchestration Layer (MRTControl, Status)

LP DAAC SERVICE BUS

(Advertised Services - Reprojection, Subset, Reformat, DPOS)

Production "Brick" * MRT Backend

Production "Brick" * MRT Backend Production "Brick" * MRT Backend

Storage Layer ECS Infrastructure - Data Pool and Near-line Data



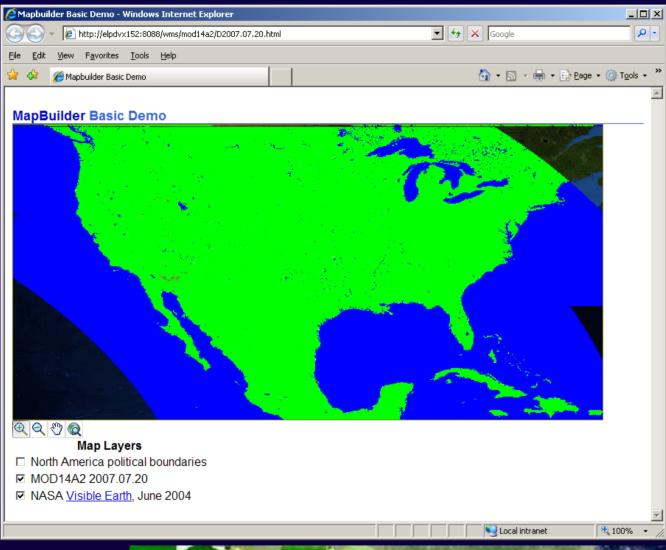


NASA GIO OGC Prototype

- Demonstrate the functionality of OGC WCS and WMS on MOD14A1 and MOD14A2 data sets
- Prototype (not operational)
- Available through key stakeholders
- Understand potential impacts of running these services on the data pool
- Delivery of a white paper including architecture outline of performance, societal benefits, interoperability and next steps towards an operational environment



Mapserver MOD14A2 Layer





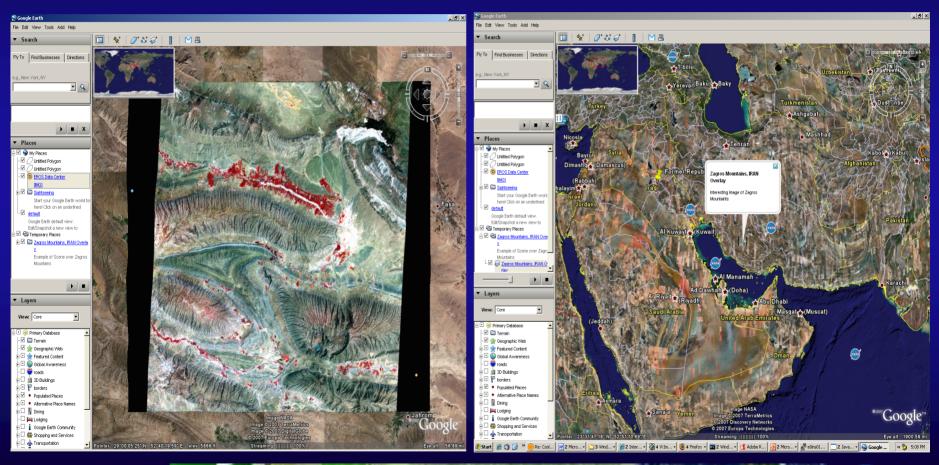


LP DAAC and Google

- Currently the LP DAAC does not utilize Google Earth or other Google tools for data discovery
- The LP DAAC has interest in understanding how Google Earth (or other Google tools) may be utilized for enhanced data discovery
- The LP DAAC plans to do some internal prototyping using Google as a data discovery tool
 - Google Mini Crawl and Search Metadata
 - How can the LP DAAC leverage Google tools for data discovery and other uses?
 - Understand Layers / Utilization of the NASA Layer?
- Is Google Earth of interest to the user community for discovery of science data?



Google Earth - Discussion







Archive and Processing

- Low Cost Data Archive and Distribution System (LODADS) Prototype Work
- USGS eMODIS





LODADS USGS LTA Prototype

- Demonstrate the Capability to Ingest, Archive and Produce of both ASTER and MODIS Data
 - USGS EOS Project (LRS) with USGS Long-term Archive
 - Produce ASTER L1B's
 - Produce MODIS Surface Reflectance Swath Products (MOD09)
 - Ingest, Archive, and Distribute Data
- Commodity Hardware
- Software Reuse (S4PM)
- Interoperable (USGS EBAS, Data Providers, etc)
- Apply Engineering Principals



FY06 USGS LTA Prototype 'DATA GRID' System Concept

Data and Service Discovery

Business Process and Orchestration Layer

SERVICE BUS (http://services.usgs.gov)

Production "Brick"

Production "Brick"

* Distribution

* Production

* Ingest

* Distribution * Production

* Production * Ingest Production "Brick"

- * Distribution
- * Production * Ingest

Storage Layer





LTA MODIS Processing Prototype Lessons Learned

- MODIS Processing from Level 0 to Surface Reflectance can be Challenging but Manageable
 - Ancillary Products
 - Complex Data Flow
- The Project Exposed Ingest, Distribution and Production as a Service and Clustered It
 - Very Scaleable
 - Low Cost Hardware
- Reuse
 - Service Bus Reuse for MRTWeb
 - Potential for Expanded Distribution Capabilities Built on the ECS Infrastructure
- Work Continues in Support of the USGS LTA and the eMODIS Project

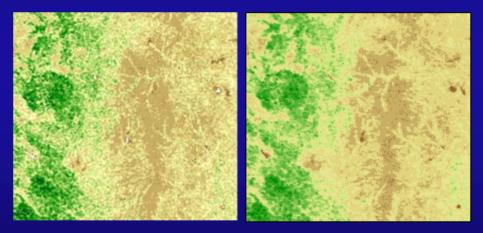




eMODIS: Enhanced MODIS Processing

- USGS-funded activity to prototype the system and evaluate products
- Specifications developed to satisfy U.S. monitoring requirements:
 - Products generated within ~24 hours of satellite data acquisition (using inputs obtained from the NOAA MODIS Near-Real-Time (NRT) System)
 - Core output product: Rolling 7-day NDVI composite (i.e., a new 7-day composite every day)
 - Continental U.S. extent, Lambert Equal Area Azimuthal projection, Geotiff format
- Leverages existing code (e.g., from discontinued MODIS DB system)





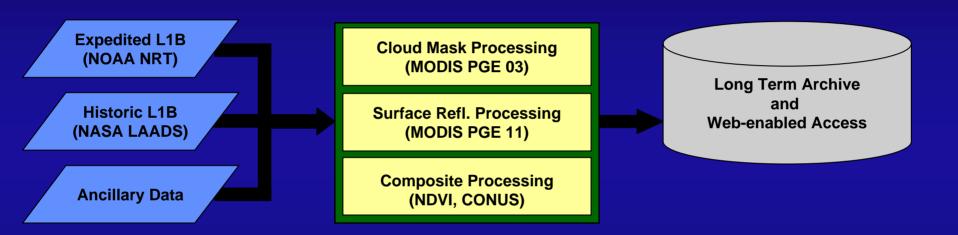
A prototype Terra eMODIS NDVI composite for the period August 2-8, 2006

The increased spatial resolution of MODIS (left) improves upon AVHRR (right) for characterizing and monitoring greenness





eMODIS Processing Flow







Were there technologies that didn't pan out? Why?

- DOWS Prototype
 - Implementation Woes
 - Data management activities (create themes of data in GeoTiff)
 - Slow response
 - Internal knowledge
- SOA
 - Difficult to Build a Robust Service Bus Using HTTP as the Transport Layer
 - Need a robust message queue
 - HTTP can be prone to error and timeouts
- Grid Computing (Globus Toolkit)
 - Heavy framework given the scope of work
 - Trade ruled this out quickly



What changes do you see coming?

- Decreasing storage cost enables more data to be accessed on-line
- Increasing computational power allows more sophisticated data delivery services
- Increasing user expectation for fast, visual interfaces ups the design ante
- Increasing emphasis on societal benefits underscore challenges in enabling the use of disparate data sets as well as communicating the relevance and impact of the data



Data Access Vision

- Data Pool
 - Enable more online data
- Web Services Prototype
 - NASA GSFC GIO and USGS Infrastructure
 - Potential delivery mechanism for ASTER Global DEM
- MRTWeb
 - Go Live (as new Data Pool client?)
 - Expand concept to ASTER
- Data Visualization
 - Google Earth / Google Tools / Enabling Technologies
 - Clients ?



Questions...



