Outline

• CERES User Metrics
• ASDC EOSWEB Update
• Data Products Online (DPO)
• New IBM Hardware
• Advanced Architecture & Engineering (AAE)
CERES USER METRICS
CERES and FLASHFlux Archive Volume
By Data Date Through March 2014

Total Volume Archived ~ 880 TB
NPP Volume Archived ~ 5.14 TB
CERES Data Archived
(September 2008 – March 2014)

CERES Production

CERES Production

April 22, 2014
ASDC Update for CERES STM
Page: 5
CERES Data Orders
(June 2010 – March 2014)

Total Orders: 18,111
CERES Data Distribution
(June 2010 – March 2014)

Total data distributed
• ~289 TB
• 968,979 data months
• ~2,428 users
Number of Users by Product
(June 2010 – March 2014)

Total: 3,491
User Affiliations

- Other: 29%
- Commercial: 17%
- Educational: 21%
- Government: 33%
Users by Country (June 2010 – March 2014)
ASDC EOSWEB UPDATE

eosweb.larc.nasa.gov
EOSWEB Update

- ASDC deployed a redesigned EOSWEB website April 8, 2013
- Since then, the team has made several improvements to the site
  - The main improvement that affects CERES is a change in the way EOSWEB displays Tier 1 and Tier 2.

- Process notes:
  - The EOSWEB team updates the CERES End Dates each week, based on data ingested the previous week.
  - Having one point of contact email (eosweb-content@lists.nasa.gov) has really helped in response time and assured that when the main point of contact is out of the office, a back up person is there to respond in a timely manner.
  - Over a hundred CERES related EOSWEB tickets have been resolved in the last year.
EOSWEB Update

Status

• Ongoing effort to incorporate enhancements to website
• Significant interaction with CERES team (D. Doelling, P. Mlynczak, N. Loeb) since initial deployment to tailor data informational pages

CERES Updates

• CERES data info pages reflect 4 tier design
  • Tier 1 – Level (1, 2, 3, 3b)
  • Tier 2 – Processing Stream Name
  • Tier 3 – Product Name
  • Tier 4 – Filename

Has been updated

Has been updated
EOSWEB Update

Example - Tier 2 – Stream
Each stream contains data products. Products organized by temporal and spatial resolution vs by filenames.

Before:

Tier 2 now a separate page, instead of inside an accordion

Tier 1 and 2 changes together results in fewer clicks for the user to get to the data

<table>
<thead>
<tr>
<th>Products</th>
<th>Temporal Resolution</th>
<th>Spatial Resolution</th>
<th>Temporal Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYN1deg-Month Ed3A</td>
<td>Monthly</td>
<td>Regional, Zonal, Global</td>
<td></td>
</tr>
<tr>
<td>SYN1deg-Day Ed3A</td>
<td>Daily</td>
<td>Regional</td>
<td></td>
</tr>
<tr>
<td>SYN1deg-M3Hour Ed3A</td>
<td>Monthly 3-Hourly</td>
<td>Regional, Zonal, Global</td>
<td></td>
</tr>
<tr>
<td>SYN1deg-3Hour Ed3A</td>
<td>3-Hourly</td>
<td>Regional, Zonal, Global</td>
<td></td>
</tr>
<tr>
<td>Legacy Data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DATA PRODUCTS ONLINE (DPO)
Data Products Online (DPO)

- Access by ASDC Data Production systems & SCF Interactive & Compute servers
- Available to all AMI Users desktop systems over LaRC campus network
- Possible to process very long data streams without staging data from tape archive
Data Products Online (DPO)

- Original DPO (~ 900 TB) deployed 2008
- Early 2013 added 524 TB new storage to DPO and migrated ~200 TB MODIS data off older disks
- Late 2013 added 1 PB new storage to replace original disk systems purchased in 2008
- In January 2014 encountered major disk problem (old DPO) while migrating CERES data to new storage
  - Most of CERES data on DPO was not accessible or corrupted.
  - Worked closely with CERES DMT to identify highest priority CERES data to restored to new DPO storage as soon as possible.
  - Sets identified as category1, category2, and category3-STM subset have been restored and verified. Remainder of category3 in progress. Expect to work category4 next.
- CERES data in DPO = ~250 TB (10.3 million files)
NEW IBM HARDWARE
New IBM Power 7+ and Intel Processors

IBM PureFlex System
• Eight Power 7 Compute Nodes (128 cores)
  • 16-core 4.1 GHz POWER7+ Processor
  • Targeted for running Clouds PGEs
• Four Intel Xeon Compute Nodes (64 cores)
  • 2 x Intel Xeon 8-Core Processor Model E5-2690 2.9GHz
• Testing and integration to begin early summer
• New hardware will support increased levels of CERES production processing facilitated by running more PGEs via CATALYST
ADVANCED ARCHITECTURE & ENGINEERING (AAE)
Advanced Architecture & Engineering (AAE)

iRODS (integrated Rule Oriented Data System) Pilot

iRODS is an open-source data management software capability in use at research centers and at many gov agencies. It provides a mechanism for federating data bases for mash-ups and fusion between multiple products and DAACs, policy-based data management for "smart" stewardship, subsetting, conversion of file formats, and even some authentication.

Accomplished:

• Initial federation between GSFC & ASDC stood up
  • Sharing CERES data products (FLASH/ISCCP/SYN/SSF/EBAF)

• Bi-Weekly meetings are held w/ NCCS to ensure project stays on schedule and any issues are resolved immediately

In Progress:

• Continuing to add data products
• Updating to new release
• Integration and testing with EOSWEB
Geospatial Analytics Capability
Esri ArcGIS provides geospatial technologies to enhance new user access and analysis

Accomplished:
• Drafted a guide for loading NetCDF files in ArcGIS for Desktop
• Server setup and configured for testing access to OGC WMS/WCS using ASDC datasets

In Progress:
• Coordination with science teams on grants/proposals (A36 win with Paul Stackhouse)
• Working with Esri to address issues when loading HDF files, which uses GDAL library
• Working with Esri to support multidimensional WMS/WCS
• Integration and testing with EOSWEB
ODISEES provides a capability to highlight the different contexts that numeric data values were derived and can help explain potential differences in values using an ontology.

**Accomplished:**
- Generated RDF metadata for CERES parameters (includes all currently available CERES data sets)
- Completed prototype for the parameter based search application

**In Progress:**
- Ingest tool to extract metadata
- Including more datasets
- Integration and testing with EOSWEB
OPeNDAP
OPeNDAP provides subsetting, dataset aggregations, and conversion to/from various data formats

Accomplished:
• Server setup and configured with access to both ANGe and ECS data holdings

In Progress:
• Addressing performance issue with public vs. private data
• Integration and testing with EOSWEB

Cloud Computing
Provides a capability for scalable, on demand computational power to store and process data

Accomplished:
• Cloud bursting proposal with AMES and JPL started
• Participation in Cloud Computing ESDSWG
• Supplied Initial draft of System Security Plan to OCIO for incorporation into Center protective measures

In Progress:
• Identifying architecture requirements for implementing Hadoop