NASA EOSDIS* Evolution in the BigData Era

*Earth Observing System Data and Information System

Christopher Lynnes
EOSDIS Systems Architect
Code 586, NASA/GSFC
HPC Forum 2015
EOSDIS processes, archives and distributes data from Earth observing satellites

http://earthdata.nasa.gov
Example: Atmospheric Infrared Sounder (Aqua Satellite)

Total Column Carbon Monoxide, Night
11 Aug 2011
EOSDIS manages data from downlink to distribution
Data are archived and distributed by DAACs oriented around science disciplines.

*Distributed Active Archive Centers

- **ASF SDC**
  - SAR Products
  - Sea Ice
  - Polar Processes
  - Geophysics

- **LP DAAC**
  - Surface Reflectance
  - Land Cover
  - Vegetation Indices

- **NSIDC DAAC**
  - Snow and Ice
  - Cryosphere
  - Climate Interactions
  - Sea Ice

- **PO.DAAC**
  - Gravity
  - Sea Surface Temperature
  - Ocean Winds
  - Topography
  - Circulation & Currents

- **GES DISC**
  - Global Precipitation
  - Solar Irradiance
  - Atmospheric Composition and Dynamics
  - Global Modeling

- **SEDAC**
  - Human Interactions
  - Land Use
  - Environmental Sustainability
  - Geospatial Data

- **OB.DAAC**
  - Ocean Biology
  - Sea Surface Temperature

- **ORNL DAAC**
  - Biogeochemical Dynamics
  - Ecological Data
  - Environmental Processes

- **CDDIS**
  - Space Geodesy
  - Solid Earth

- **LaRC ASDC**
  - Radiation Budget
  - Clouds
  - Aerosols
  - Tropospheric Chemistry

- **GHRC DAAC**
  - Hydrologic Cycle
  - Severe Weather Interactions
  - Lightning
  - Atmospheric Convection

- **LAADS**
  - MODIS Level-1 and Atmosphere Data Products

http://earthdata.nasa.gov
DAACs and users are supported by EOSDIS Common Services

http://earthdata.nasa.gov
EOSDIS Evolves Continually

*Almost
**Thank you, HDF internal compression!

http://earthdata.nasa.gov
Big Data Volume Growth

Archive Volume (PB)

http://earthdata.nasa.gov
Big Data Distribution Growth
Big Data User Growth

Destination IPs of Data Users

- 2000
- 2002
- 2004
- 2006
- 2008
- 2010
- 2012
- 2014
- 2016

http://earthdata.nasa.gov
Big Data Variety Growth
EOSDIS in the Big Data epoch will enable more analysis closer to the data.
Let’s Break that Down...

“more analysis closer to the data”
“More Analysis”

More Complexity

Subset

Data Variable

Spatial Area

Quality Filter

Transform

Reprojection

Mosaicking

Analyze

Simple Stats

Complex Stats

End User’s Algorithm

Seasonal Time Series

http://earthdata.nasa.gov
“more analysis closer to the data”
“Close To” = At Archive
“Close To” = Near Archive

MapServer (WMS)
Giovanni Viz.

Cloud Server
Cloud Storage

Near-Archive Server

OPeNDAP

Archive Server
Archive

http://earthdata.nasa.gov
“Close To” = Near Processing

- Cloud Storage
- Google Earth Engine
- NASA Earth Exchange
- Archive Server
- Archive

http://earthdata.nasa.gov
“more analysis closer to the data”
## “The Data”

<table>
<thead>
<tr>
<th>Suitable Data Form</th>
<th>Processing Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Data as Archived</td>
<td>GrADS Data Server, ArcIMS for Server</td>
</tr>
<tr>
<td>Groomed Data (reformatted, annotated)</td>
<td>Earth System Grid, Giovanni</td>
</tr>
<tr>
<td>Reorganized Data</td>
<td>Google Earth Engine, SciDB, MapReduce</td>
</tr>
</tbody>
</table>
What's Next?

Prototype Cloud Analysis and Storage

Collect More Science Use Cases

Cloud Server

Near-Archive Server

Archive Server

Cloud Storage

Build Out Analysis Support Capabilities

Archive

http://earthdata.nasa.gov