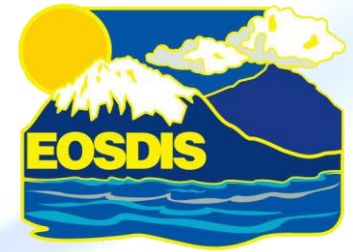# Development, Operation and Evolution of EOSDIS – NASA's major capability for managing Earth science data

Presented at CENDI/NFAIS Workshop on Repositories in Science & Technology: Preserving Access to the Record of Science
November 30, 2011

H. K. Ramapriyan
Assistant Project Manager
ESDIS Project, Code 423
Rama.Ramapriyan@nasa.gov

Topics

- **NASA Earth Science and EOSDIS**
- **Scope**
  - **Earth observing satellites and Earth science measurements**
  - **"Interesting" Images**
  - **EOSDIS Context**
- **EOSDIS – A System of Systems**
- **Search and Access**
  - **Interoperability**
  - **Near-real-time Data**
- **Metrics and customer feedback**
- **History, Development Issues and Evolution**
  - **Policy**
  - **Missions to Measurements**
  - **Technology**
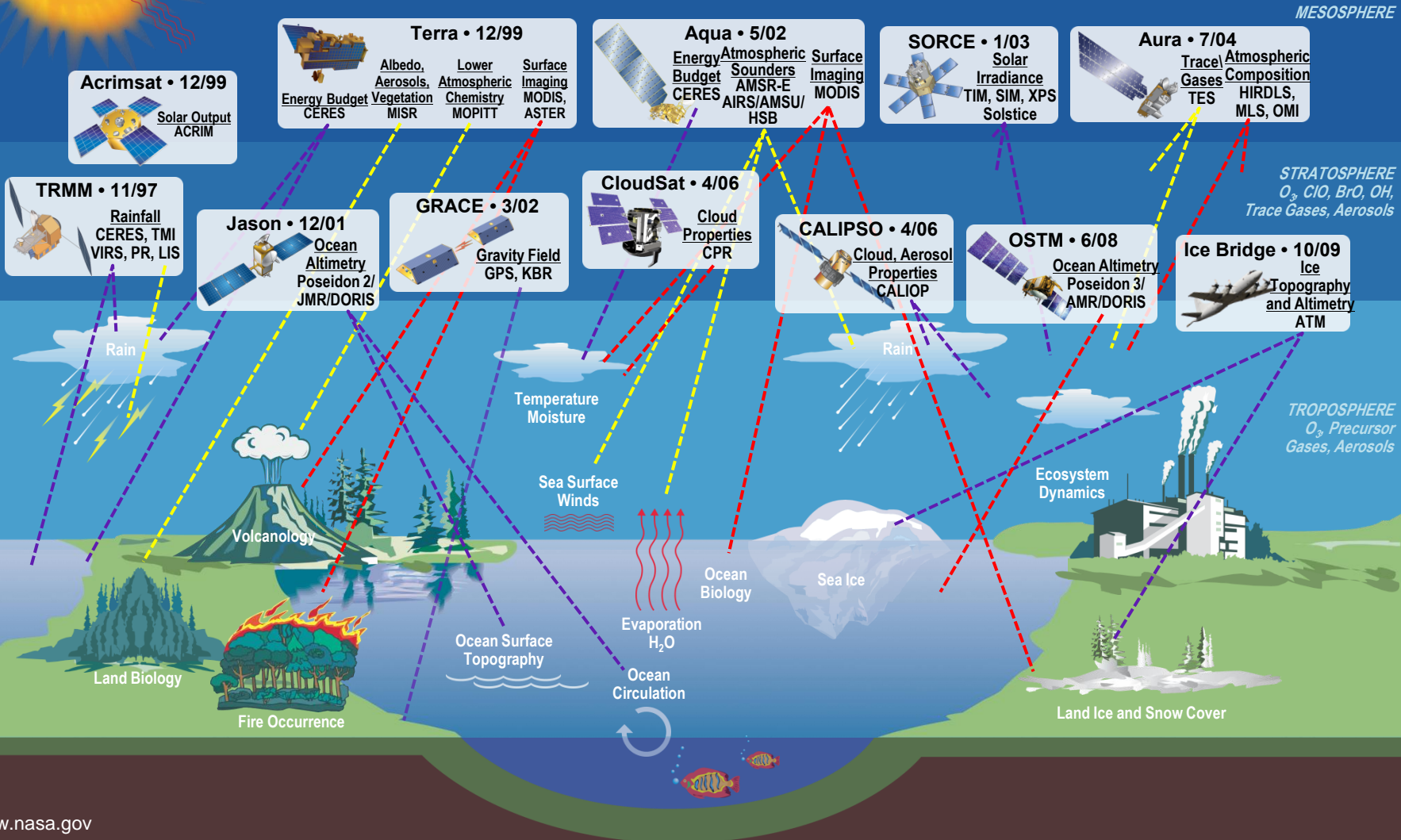- **Preservation**
- **Conclusion**

NASA Earth Science and EOSDIS

- **"Advance Earth system science to meet the challenges of climate and environmental change." --** *2011 NASA Strategic Plan*
- **NASA's Earth Science Data Systems directly support this objective by providing end-to-end capabilities to deliver data and information products to users**
- **NASA's Earth Science Data and Information Policy promotes usage of data by the community**
  - **No period of exclusive access - Data are available after initial checkout**
  - **Data available at no cost to all users on a non-discriminatory basis except where agreed upon with international partners**
- **EOSDIS provides:**
  - **Science Operations**
    - **Science data processing**
    - **Data management**
    - **Interoperable distributed Data archives**
    - **On-Line data access services**
    - **Earth science discipline-oriented user services**
  - **Network Data Transport to distributed system elements**
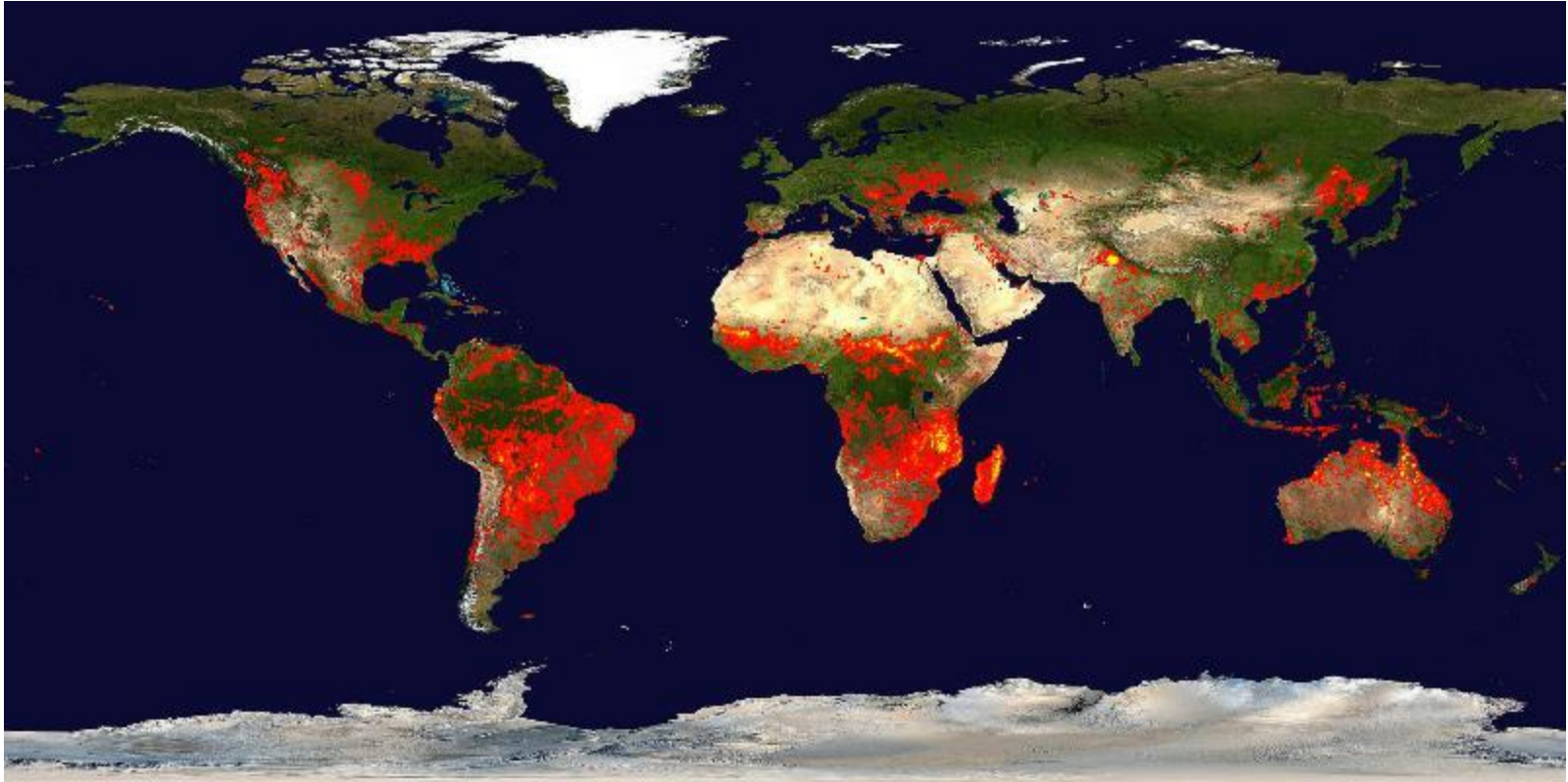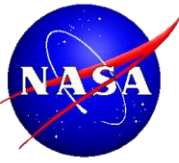
Earth Observing Satellites

Global Fire Map


Locations of the fires detected by MODIS on board EOS Terra and Aqua satellites over a 10-day period (10/28/11-11/06/11). Fire detection algorithm developed by Louis Giglio. Fire maps created by Jacques Descloitres, MODIS Rapid Response System at NASA/GSFC. Fire locations produced by the MODIS Rapid Response System since mid-2001. Image from http://rapidfire.sci.gsfc.nasa.gov/firemaps/

Ocean Chlorophyll Concentration
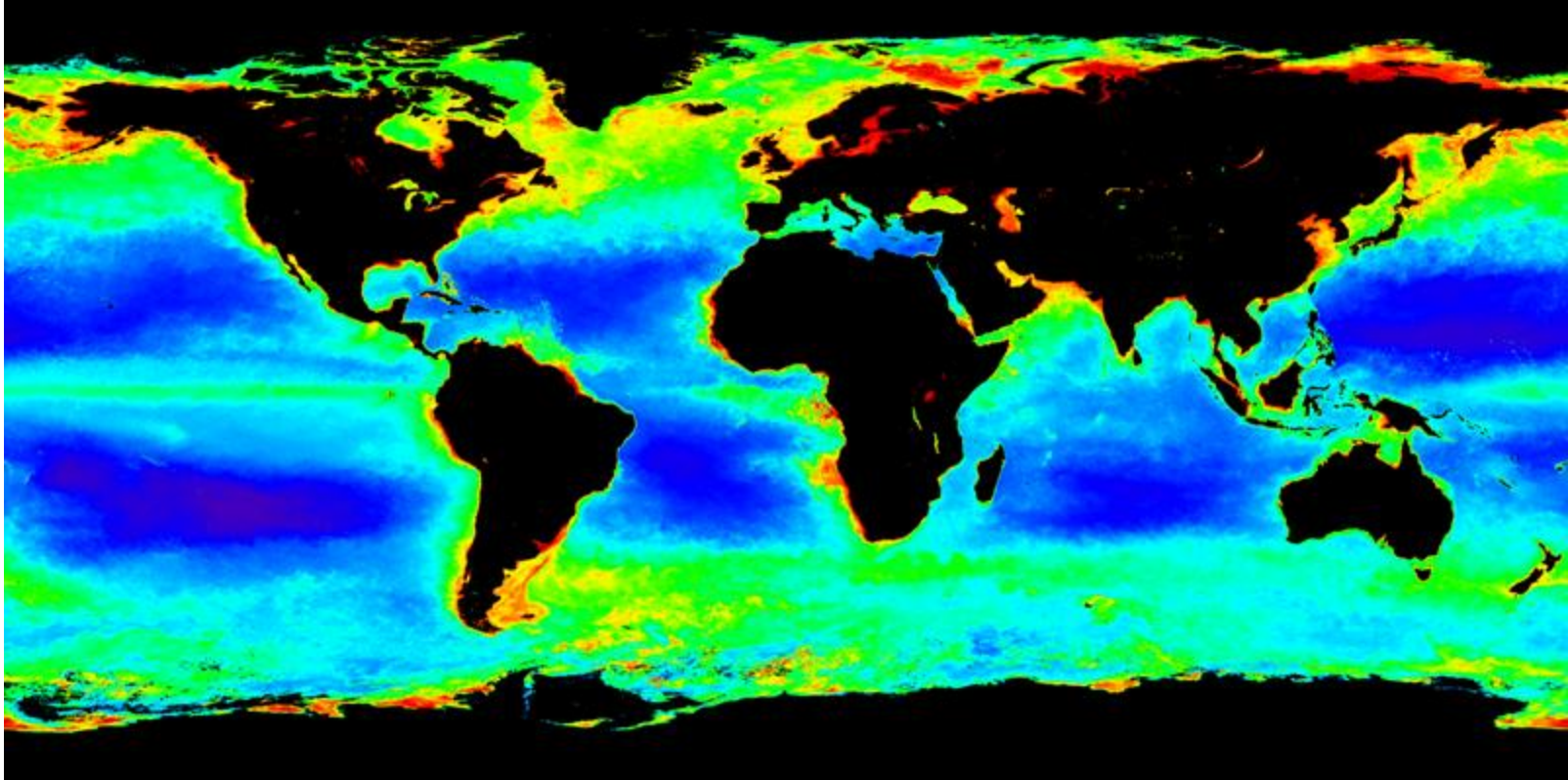
Annual composite Chlorophyll concentration generated with data from MODIS instrument on EOS Aqua for 2010. Image from http://oceancolor.gsfc.nasa.gov/.
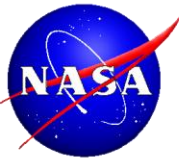
Winter Storm in Alaska
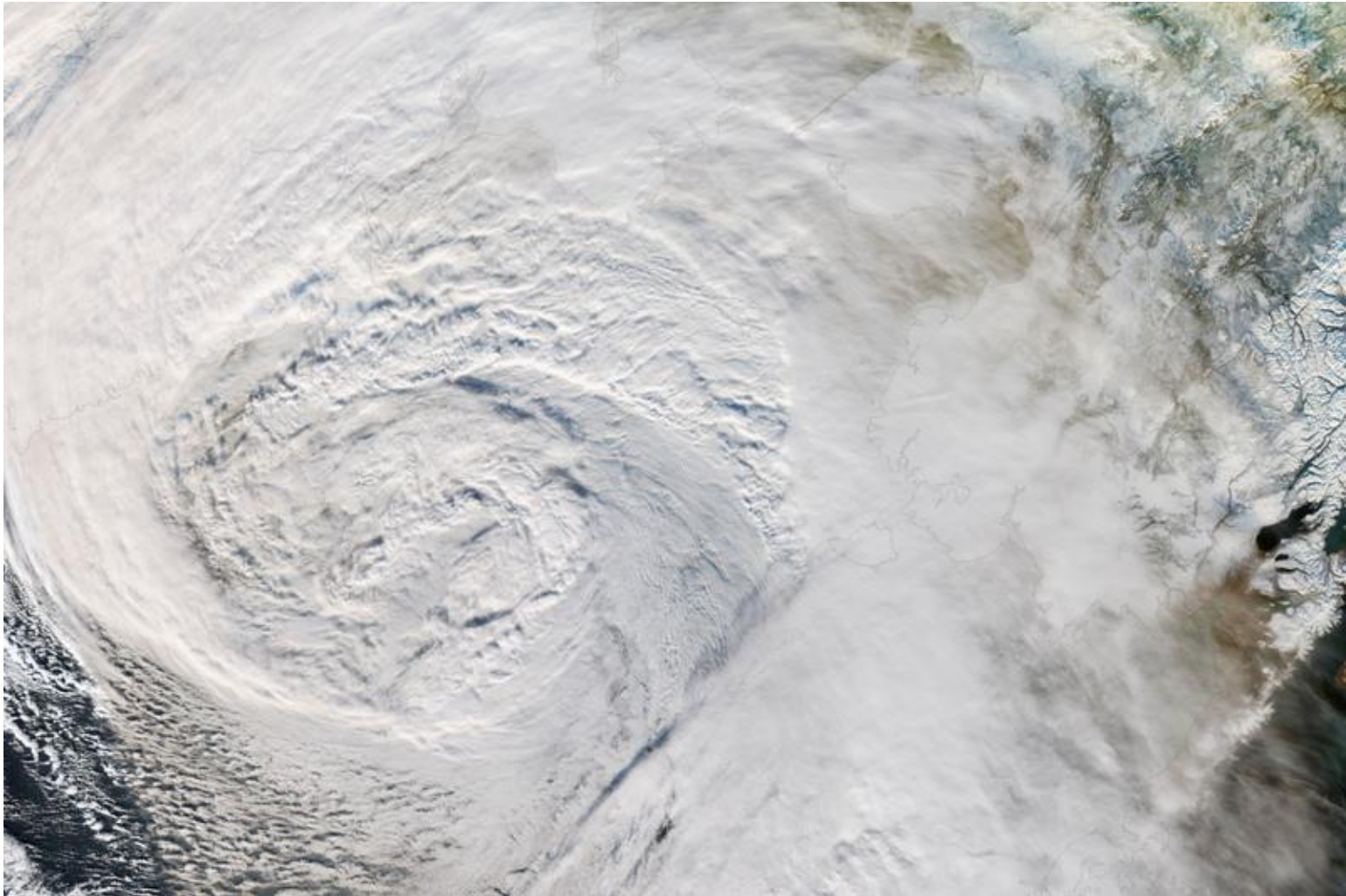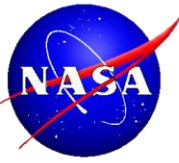
Winter storm over northwestern Alaska - November 8 and 9, 2011. NASA image courtesy LANCE/EOSDIS at NASA GSFC. Instrument - MODIS on Aqua satellite

Antarctic Ozone Hole
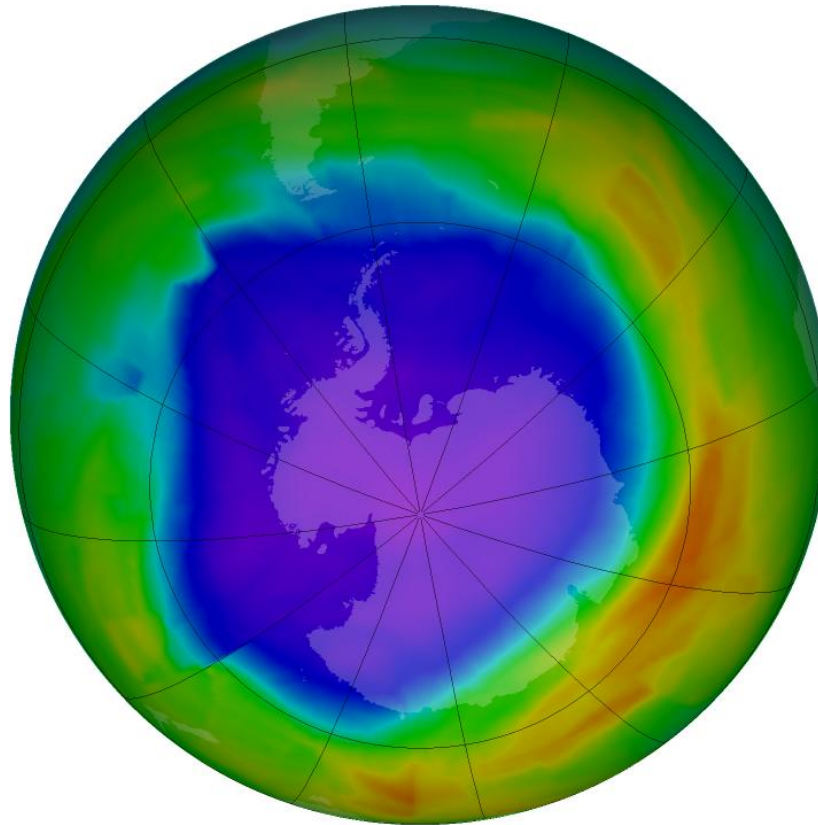
September - October time period is considered the "Ozone hole" season. The ozone data produced by OMI on EOS Aura spacecraft is used by the Ozone watch website (along with other instruments) to produce a map of the ozone over Antarctica.

Above is an image of the Ozone hole as seen on October 8, 2011 from the Ozone watch website, http://ozonewatch.gsfc.nasa.gov. Ozone reached its lowest value for the year, 95 Dobson Units, on this day.

EOSDIS – A System of Systems

- Instrument and Science Expertise needed to process data
  - Most EOS standard products are generated at Science Investigator-led Processing Systems (SIPSs) under supervision of PIs
- Earth science discipline knowledge needed to ensure data stewardship
  - Processed data are archived and distributed by discipline-specialized EOSDIS Science Data Centers
- Expertise in system interoperability needed to provide cross-system (interdisciplinary) data access
  - EOS Clearing House (ECHO) middleware and associated clients provide search and access to data across all EOSDIS Science Data Centers
- EOSDIS data collections are diverse:
  - Primary sources are NASA spacecraft
  - Ancillary, airborne, in-situ and socio-economic data
  - Data from international partners
  - Comprehensive approach to multi-discipline science
  - Feed growing need by models (e.g., climate models)
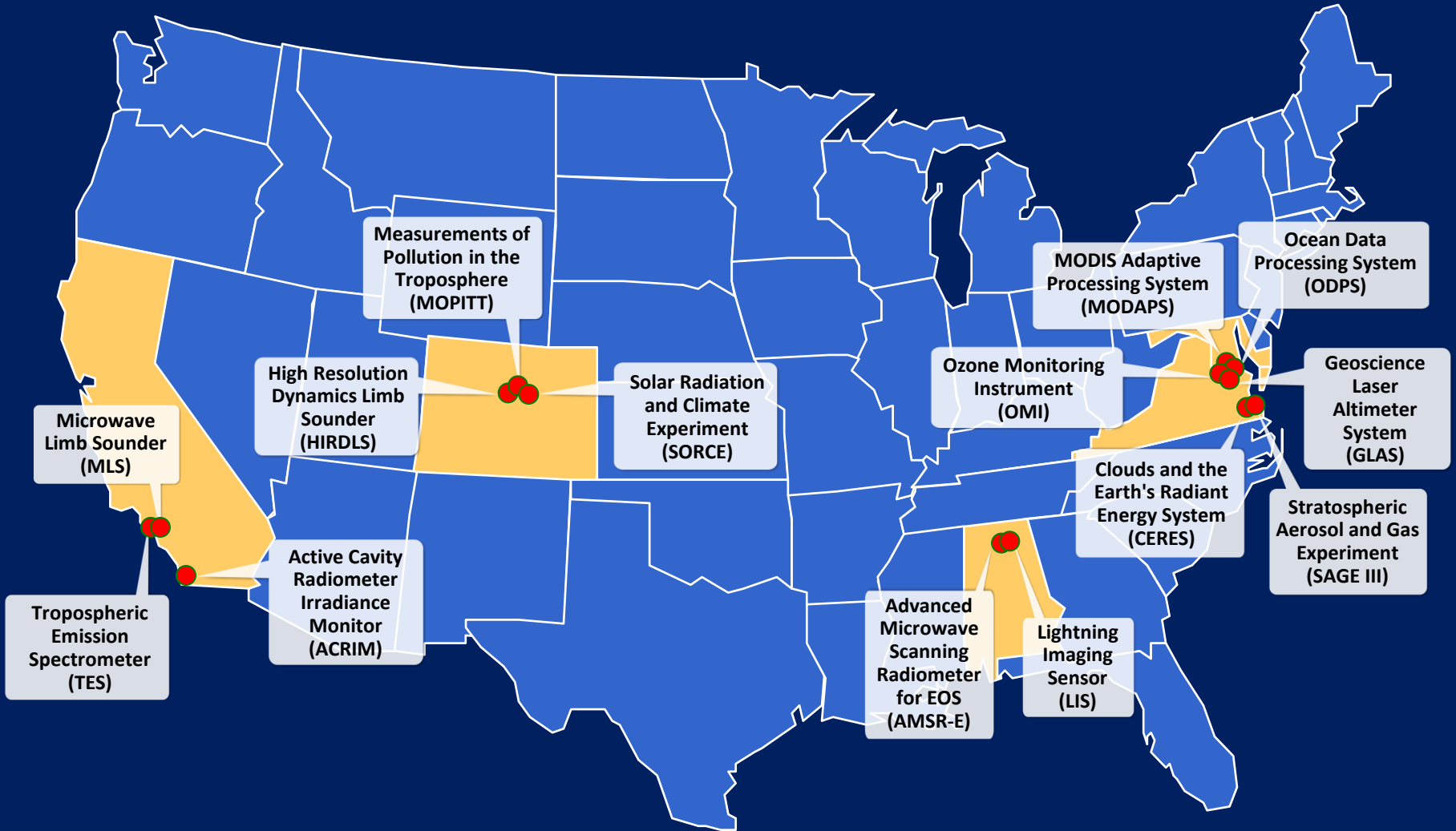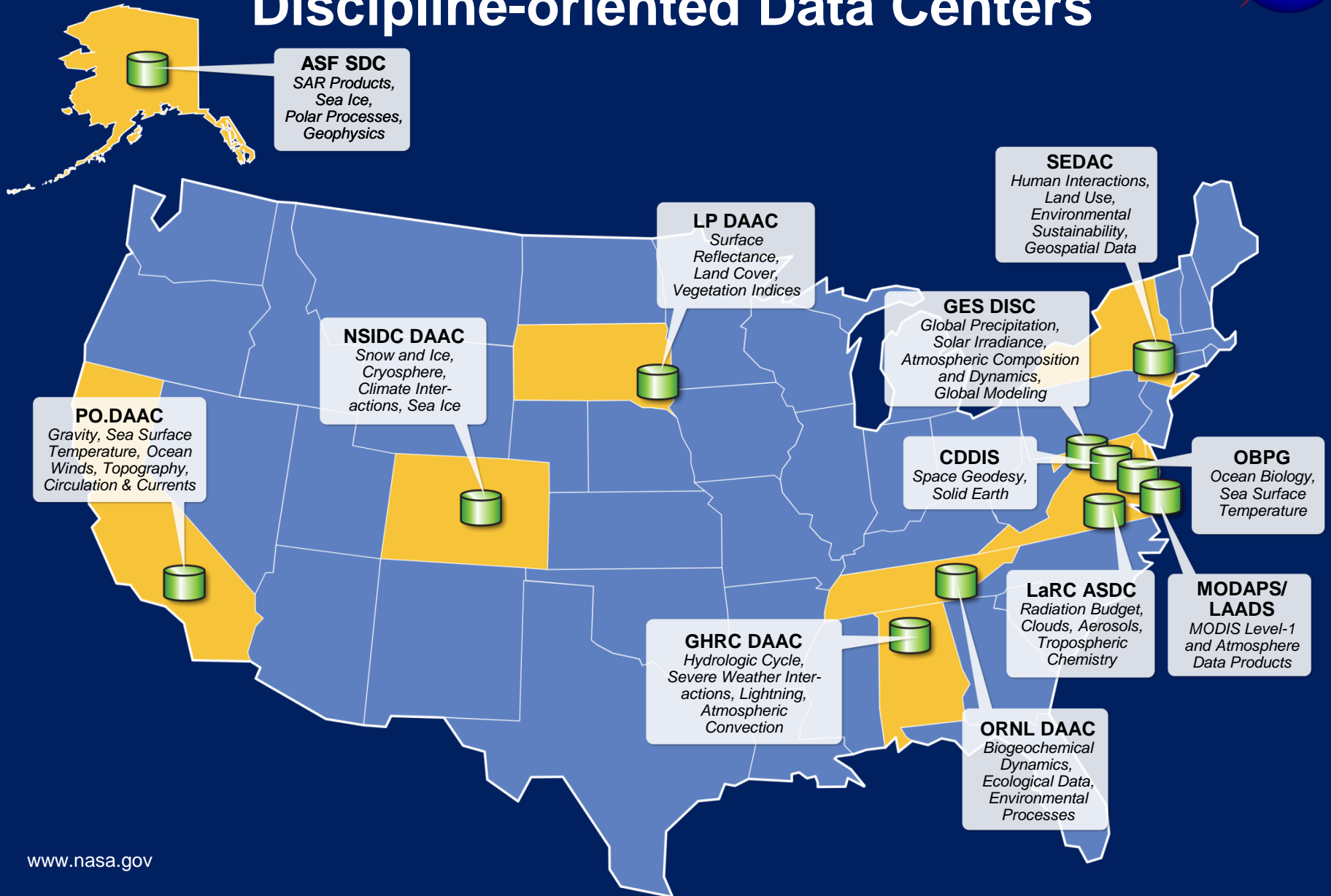
EOS Science Data Distribution Network Overview

Key

- NISN Backbone (10 gbps)
- NISN Tail Circuits
- Internet2 Backbone (20+ gbps)
- Internet2 Router Node
- Interconnect
- NISN CIEF
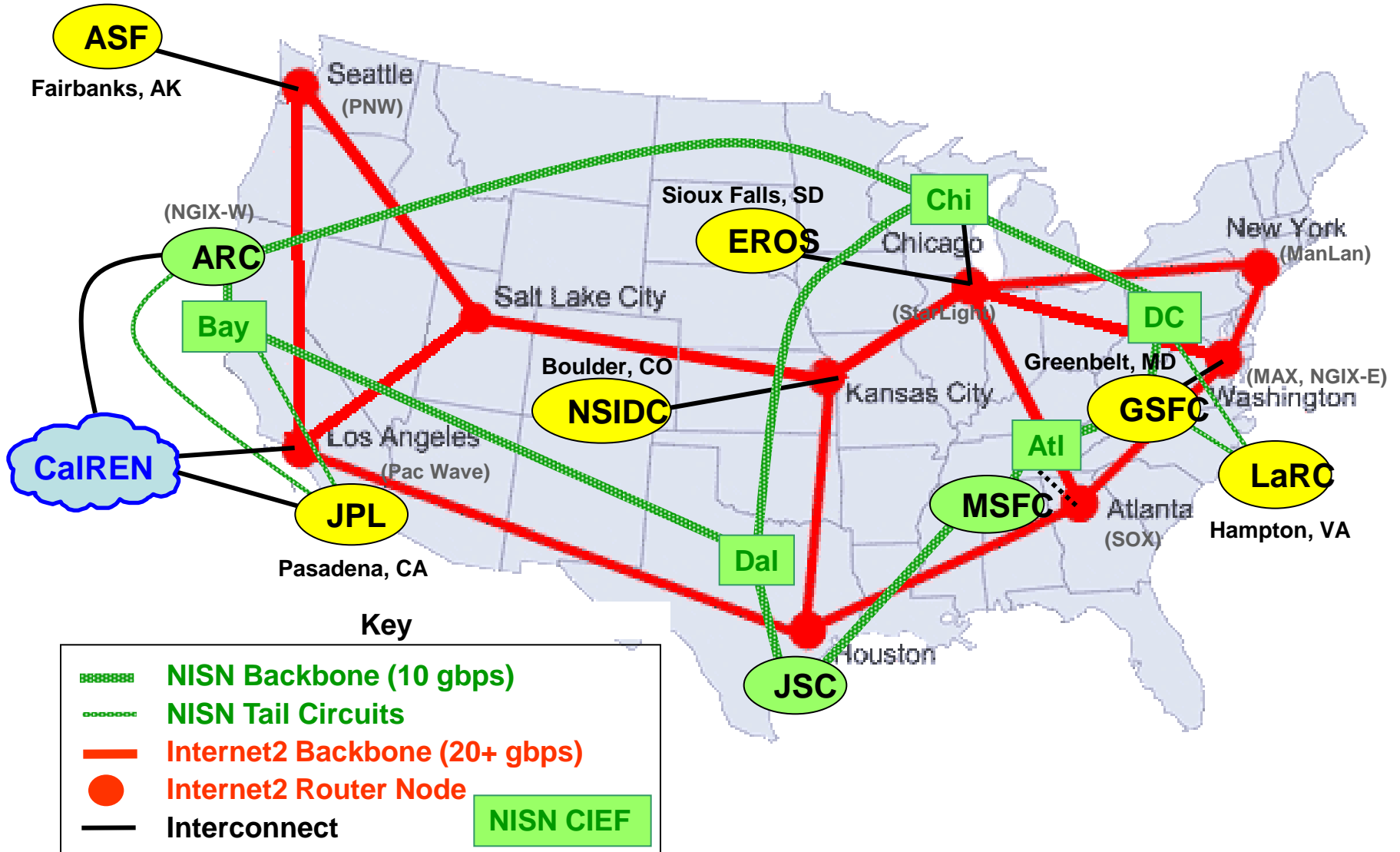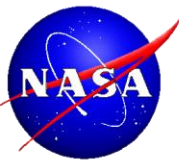
EOS International Connectivity

Key

NISN Backbone (10 gbps)	NISN CIEF
NISN Tail Circuits	
Internet2 Backbone (20+ gbps)	
Internet2 Router Node	
Interconnect (1 gbps)	
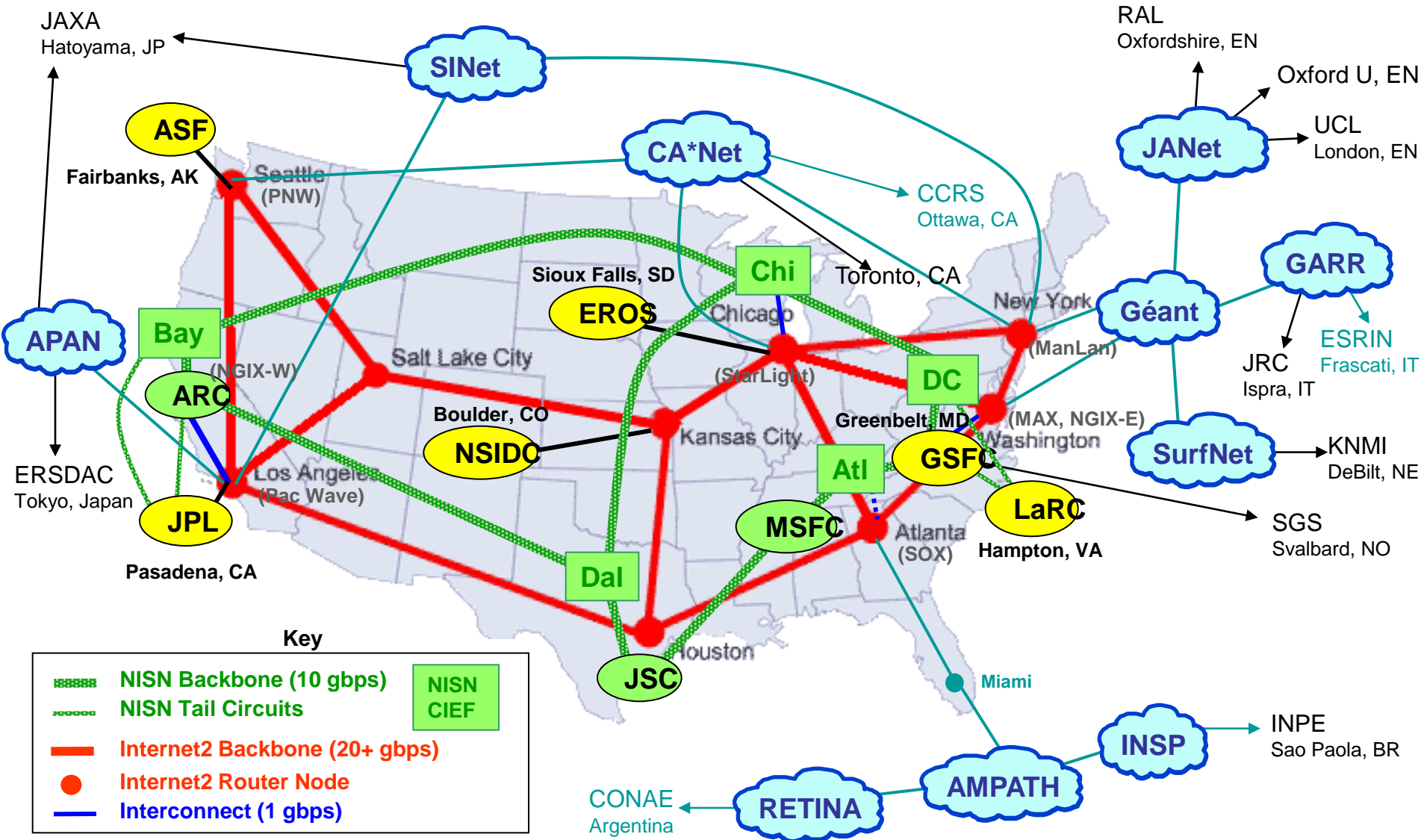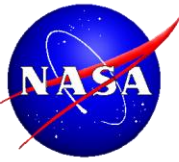
Variety of Search and Access Methods
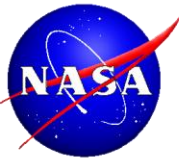
- There are several ways to search for data of interest
  - Directory level information from Global Change Master Directory
  - Cross-Data Center searches through REVERB– uses EOS Clearing House (ECHO) metadata repository
  - Tailored client software using ECHO metadata repository
  - Data Center-specific search tools
- Almost all data in EOSDIS are held on-line and accessed via ftp
  - A small part, still held in near-line robotic tape archives are being migrated to on-line storage
- On-line services are available
  - e.g., subsetting, reprojection, mosaicing, format conversion
- Several data visualization and analysis tools are available at EOSDIS Data Centers

Interoperability – Directory Level

- Used for locating data set (collection) of interest and go to data center holding the data set

- NASA's Global Change Master Directory (GCMD)

  - \>25,000 Earth science data set and service descriptions, which cover subject areas within the Earth and environmental sciences.

  - All of NASA's EOSDIS datasets are discoverable in the GCMD.

  - The CEOS International Directory Network (IDN) has a subset of the GCMD holdings.

  - The Geospatial One-Stop and the Global Earth Observation System of Systems (GEOSS) Clearinghouse harvest GCMD for all NASA contributed datasets (~4,000) on a weekly basis.

  - Data.gov has been harvesting the Geospatial One Stop for datasets from multiple agencies including NASA on a monthly basis.

Interoperability – Inventory Level

- Used for searching for and accessing files of data meeting specific spatio-temporal (and other) criteria
- ECHO connects users to NASA's Earth science data through a metadata middleware and web portals
  - Metadata middleware offers publicly available APIs accessible through SOAP, REST and OpenSearch web services for Search and Order functionality
  - General purpose web portal (REVERB client) allows users to submit cross discipline search and order capability across all EOSDIS data providers.
  - Several other agencies and groups are accessing EOSDIS data, e.g.,
    - With European Space Agency's (ESA) eoPortal Catalogue Client users can query Earth Science products from ESA, German Space Agency (DLR), Korea Aerospace Research Institute (KARI), and NASA (via ECHO)
- Committee on Earth Observation Satellites (CEOS) coordinates satellite Earth Observation programs of the world's government agencies, along with agencies that receive and process remote sensing data
  - EOSDIS provides data to CEOS WGISS (Working Group on Information Systems and Services) Integrated Catalog (CWIC) for search and access to satellite data within partner systems. CWIC data providers include AOE-China, CLASS-NOAA, ECHO-NASA, INPE-Brazil , and LANDSAT-USGS.

Land, Atmosphere Near-real-time Capability for EOS (LANCE)

- Building on existing EOSDIS elements provides data from MODIS, OMI, AIRS, MLS, and AMSR-E instruments in near real-time (< 3 hours from observation)

- Utilizes software for Standard Science Products, but relaxes requirements for ancillary data inputs

- High operational availability

- Applications of LANCE data include:
  - Numerical weather & climate prediction/forecasting
  - Monitoring of Natural Hazards
  - Disaster Relief
  - Agriculture
  - Air quality
  - Homeland Security
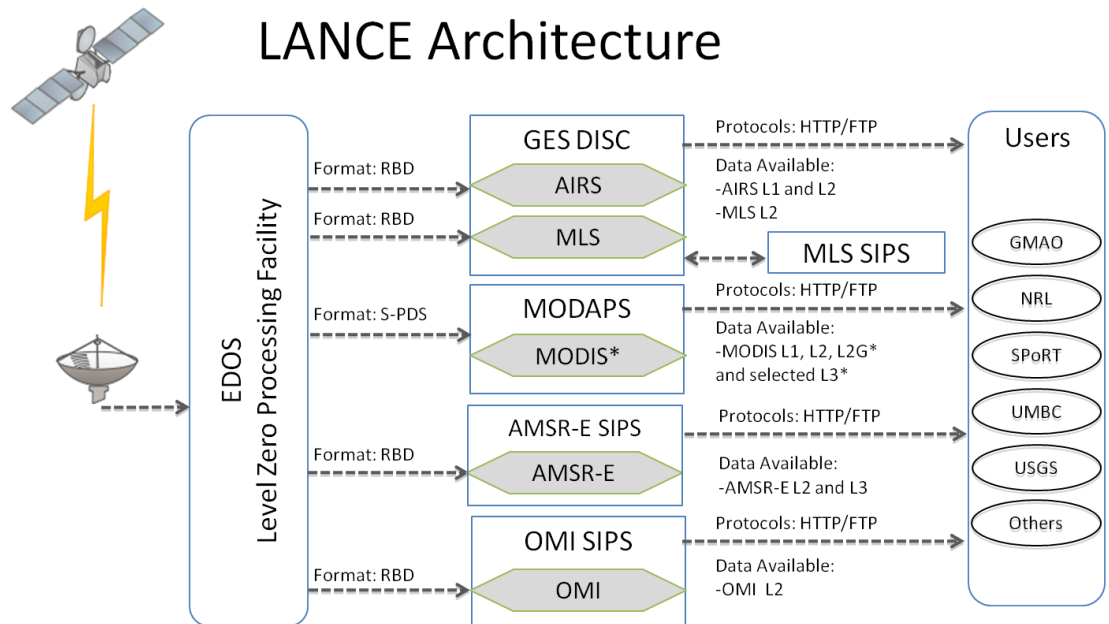
LANCE Architecture

RBD: Rate Buffered Data
S-PDS: Session Based Production Data Set
* L2G and L3 daily products have latency of 27-28 hours. The Climate Modeling Grid (CMG) is the only L3 MODIS product produced by LANCE.

EOSDIS Key Metrics
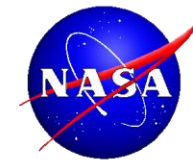
EOSDIS Metrics (Oct 1, 10 to Sept 30, 11)

Unique Data Products	> 5,000
Distinct Users of EOSDIS Data and Services	> 1.1 M
Average Daily Archive Growth	3.8 TB/day
Total Archive Volume	5.1 PB
End User Distribution Products	> 500 M
End User Average Daily Distribution Volume	13.9 TB/day

ESDIS Project Supports

Science System Elements	Data Centers	12
	SIPS	14
Interfaces	Interface Control Documents	22
Partnerships	US	10
	International	6
Missions	Science Data Processing	10
	Archiving and Distribution	37
	Instruments Supported	87

EOSDIS Products Delivered (Millions)

Millions — bar chart, FY00 through FY11

FY	Value (approx., Millions)
FY00	~5
FY01	~7
FY02	~18
FY03	~35
FY04	~47
FY05	~68
FY06	~90
FY07	~128
FY08	~157
FY09	~255
FY10	~413
FY11	~500
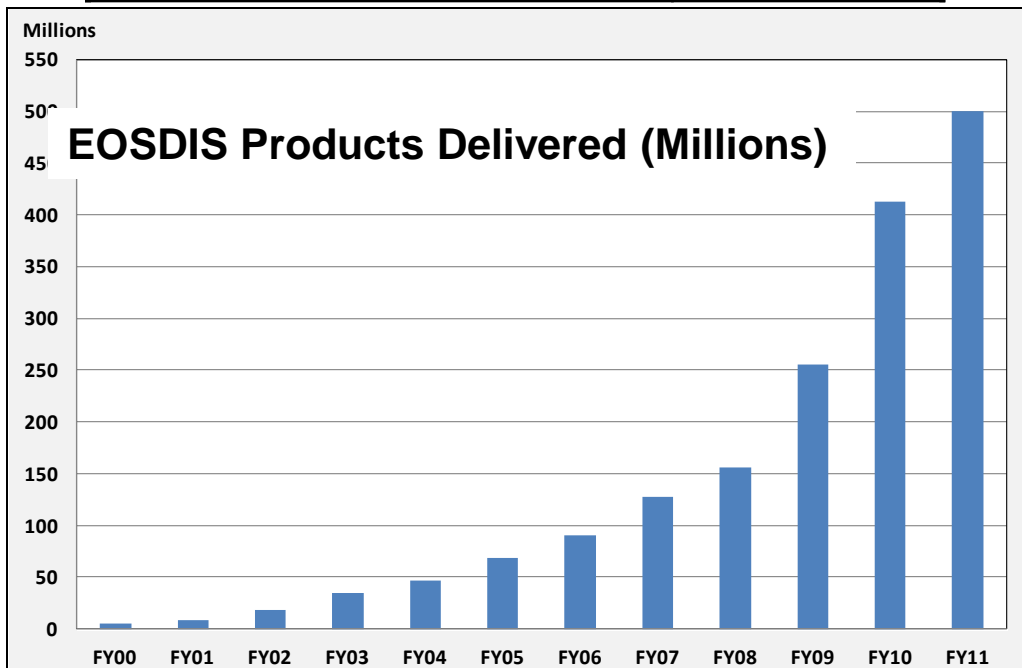
EOSDIS ACSI Customer Satisfaction Survey 2011: Relative Rankings

- EOSDIS sponsors an annual independent customer survey in conjunction with the American Customer Satisfaction Index (ACSI).
- EOSDIS consistently exceeds the Federal Government average
- Ratings in the mid-to-upper 70s are considered "very good" by the independent rating organization, the CFI Group
- 2011 Survey results based on 3996 responses (8%)
- Comments in surveys help define system improvements

Category	Score
Federal Government Overall (2010)	65
NASA EOSDIS	77
E-Business News & Information (Public Sectors)	75
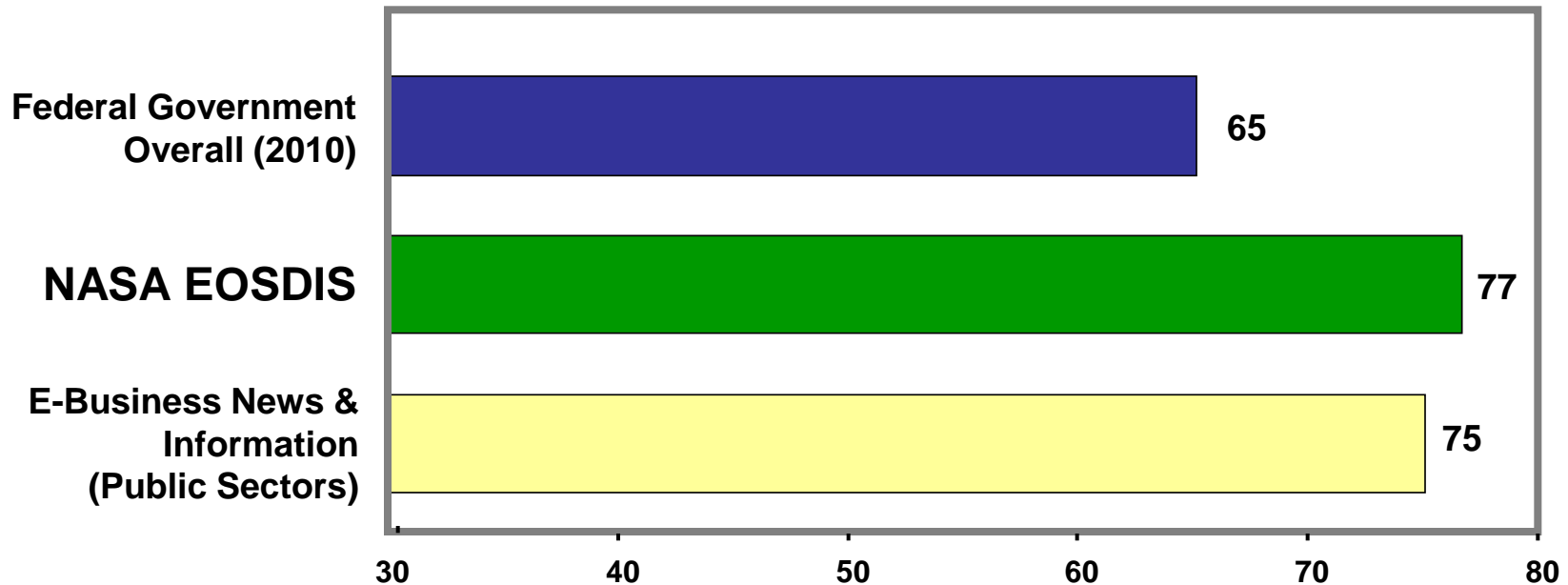
Evolution from Pre-EOS Era

- Data policy
  - Formerly PI' s had exclusive access periods
  - Now free and open
- Focus on data management
  - Formerly small parts of flight projects
  - Starting with EOS, separate data system program/project
- Missions to measurements
  - Discipline-based data centers
  - MEaSUREs Program
- Technology upgrades and system evolution

Data Policy

- ■ NASA Earth Science Data Policy
  - No period of exclusive access
  - Except where agreed upon with international partners, data and derived scientific products are available at no cost to all users
  - Any variation in access will result solely from user capability, equipment, and connectivity
  - All NASA-generated standard products are made available (upon request) along with the source code for algorithm software, coefficients, and ancillary data used to generate these products.
  - See http://science.nasa.gov/earth-science/earth-science-data/data-information-policy/ for full text of policy
- ■ Data are made available to all users promptly
  - After an initial checkout period
  - Appropriate caveats about data quality are provided in product documentation

NASA Mission Instrument Data and Measurement Earth System Data Records
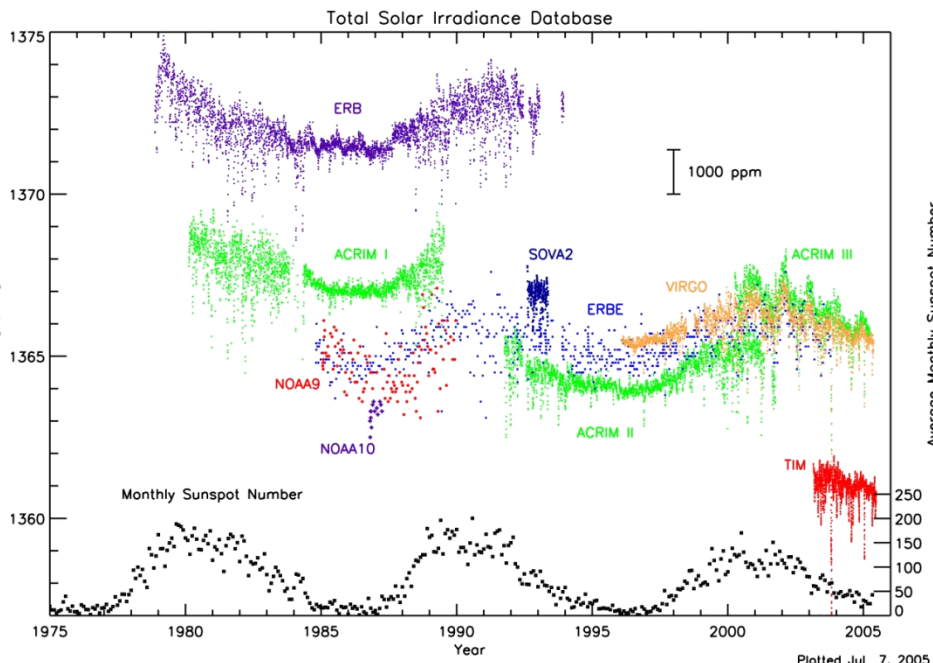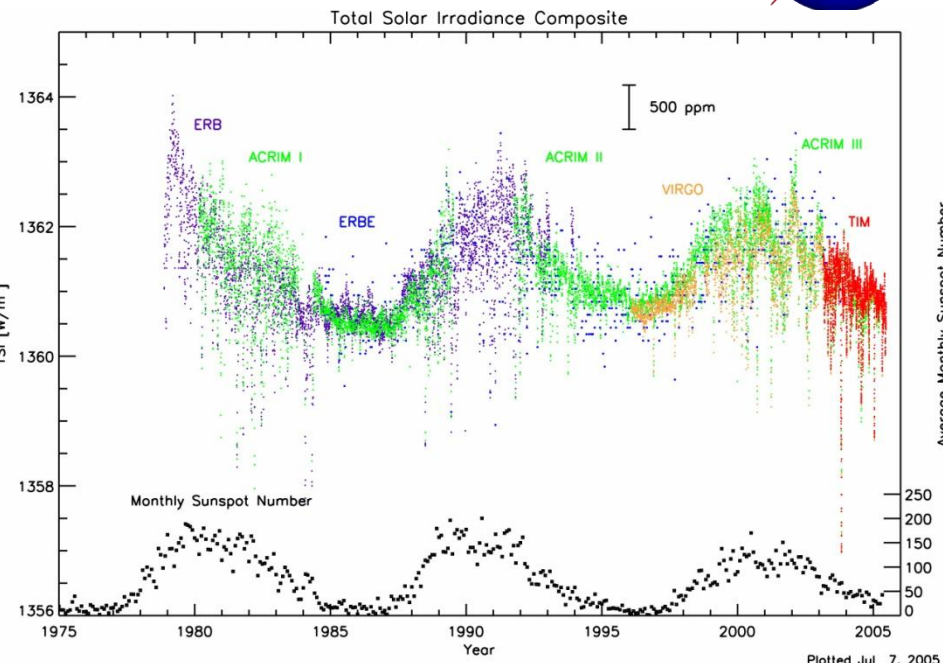
Total Solar Irradiance (TSI) measurements shown are existing 26 year record - precise data records from multiple sensors have been corrected for accuracy through data overlaps.
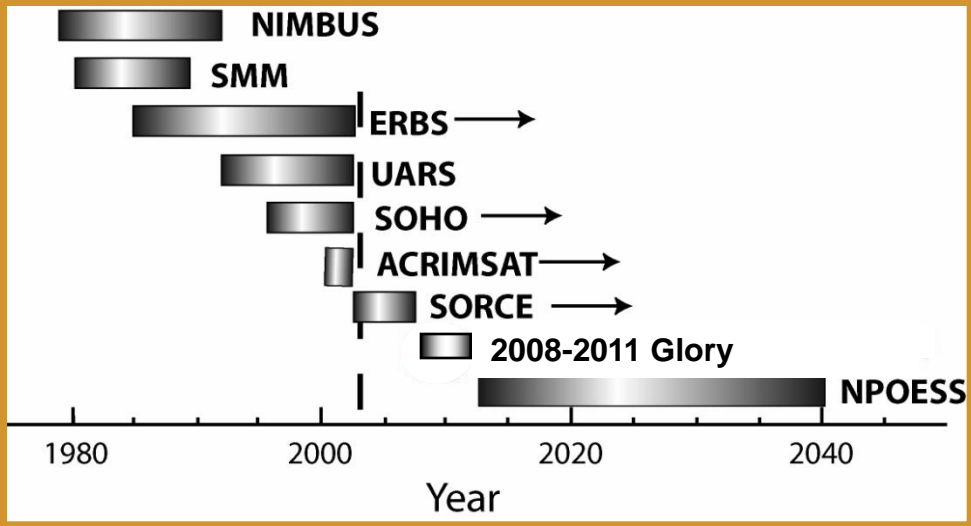
New data sets from MEaSUREs

- MEaSUREs (Making Earth System data records for Use in Research Environments)
  - Objective is to develop long-term, consistent and calibrated data and products that are valid across multiple missions and satellite sensors
  - Data are now becoming available through websites at EOSDIS Data Centers.
    - Cross-Calibrated Multi-Platform Ocean Surface Wind Velocity Data Set through PO.DAAC
    - Goddard Satellite-based Surface Turbulent Fluxes (GSSTF) Data Set for Global Water and Energy Cycle Research through GES DISC
    - Global Record of Daily Landscape Freeze/Thaw Status at NSIDC
    - Greenland Ice Sheet Velocity Map from InSAR data at NSIDC/ASF
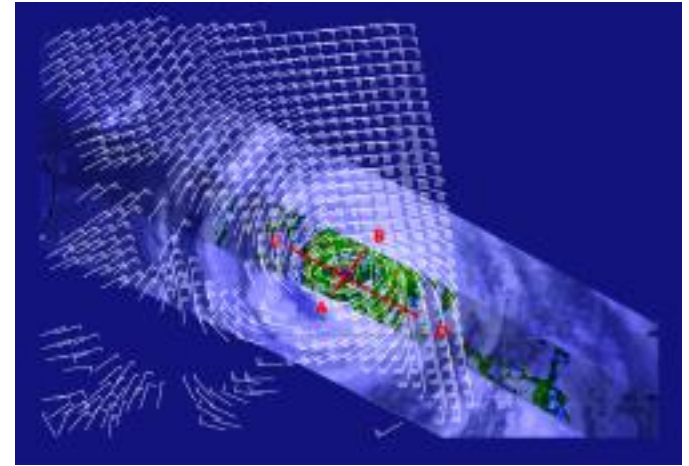
Ocean Surface Wind Velocity Dataset/PO.DAAC
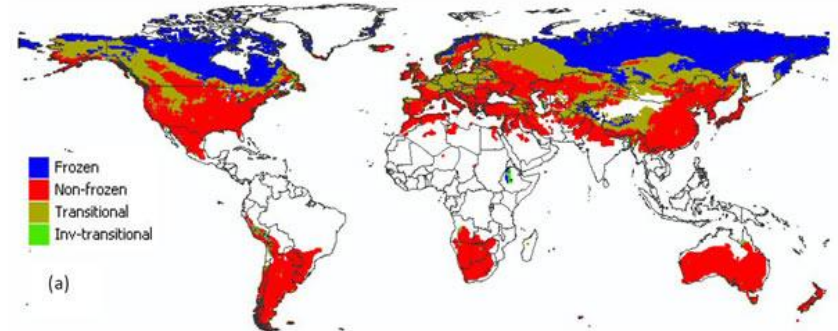
Freeze/Thaw Status/NSIDC

Technology Upgrades and System Evolution

Discipline/ mission specific data systems

Community-specific standards only

Data inter-use proved cumbersome

Improved access to heritage data

Cross-system search and order access via data interoperability model

Common distribution format (HDF); other formats also supported

Support for high data volumes

Integrated core plus coupled elements

Common data model

Expanded software tools and services

Options to support or interoperate with external data sources

Coexistence of hetero-geneous, distributed data providers / information partners

Minimal set of core standards; support for community-specific standards

Coherent Web – earthdata.nasa.gov

Preservation – content specifications

Reusable software

Service Oriented Architecture

On-line archives and cross-system service invocation

Near Real-Time access

Ease of innovation and technology infusion

<1990 **Mid-1990s** **Late 90s +** **Past decade to near future**

Lessons learned and information technology advances coupled with advice/comments from community supports a continuously evolving data system with growing capabilities

26

Preservation

- **NASA is not a "permanent archive" agency**
  - It has to maintain a "research archive" for as long as data are used for scientific research and/or transition responsibility to permanent archives
  - Research archive responsibilities persist well beyond lives of missions
  - NASA works with USGS and NARA for long-term preservation
  - NASA has to ensure data and other critical items are preserved and made available to permanent archival agencies
- **General requirements**
  - Don't lose bits
  - Maintain readability
  - Maintain understandability
- **NASA is working on Preservation Content Specifications for Earth Science Data**
- **NASA is participating in Earth Science Information Partners (ESIP) Cluster on Data Stewardship and Preservation, on an "emerging" Provenance and Context Content Standard**

Categories of Content to be Preserved

1. **Preflight/Pre-Operations:** Instrument/Sensor characteristics including pre-flight/pre-operations performance measurements; calibration method; radiometric and spectral response; noise characteristics; detector offsets

2. **Science Data Products:** Raw instrument data, Level 0 through Level 4 data products and associated metadata

3. **Science Data Product Documentation:** Structure and format with definitions of all parameters and metadata fields; algorithm theoretical basis; processing history and product version history; quality assessment information

4. **Mission Data Calibration:** Instrument/sensor calibration method (in operation) and data; calibration software used to generate lookup tables; instrument and platform events and maneuvers

5. **Science Data Product Software:** Product generation software and software documentation

6. **Science Data Product Algorithm Input:** Any ancillary data or other data sets used in generation or calibration of the data or derived product; ancillary data description and documentation

7. **Science Data Product Validation:** Records, publications and data sets

8. **Science Data Software Tools:** product access (reader) tools.

Conclusion

- EOSDIS success has been based on its ability to meet and adapt to needs of diverse Earth science communities
  - 17 years of diverse science data centers' operation to meet the needs of a growing user community
  - Over 11 years of support for EOS missions
  - Preparing for new missions (e.g., EV-1 aircraft investigations, NPOESS Preparatory Project (NPP), SMAP, ICESat-2, OCO-2)

- Managing and preserving Earth science data is a long-term ongoing responsibility to support Earth system science and to meet national and international needs for understanding global climate change

- Science discipline expertise is required for management of data

- Flexibility, ability to evolve, and responsiveness to community feedback are key to success

BACK-UP

EOSDIS Acronyms

Acronym	Definition	Acronym	Definition	Acronym	Definition
ACSI	American Customer Satisfaction Index	GES	Goddard Earth Sciences	NRL	Naval Research Laboratory
ACRIM	Active Cavity Radiometer Irradiance Monitor	GHRC	Global Hydrology Resource Center	NSIDC	National Snow and Ice Data Center
AIRS	Atmospheric Infrared Sounder	GLAS	Geoscience Laser Altimeter System	OBPG	Ocean Biology Processing Group
AMSR-E	Advanced Microwave Scanning for EOS	GMAO	Global Modeling and Assimilation Office	OGC	Open Geospatial Consortium
API	Application programming interface	GMU	George Mason University	OMI	Ozone Monitoring Instrument
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	GRACE	Gravity Recovery and Climate Experiment	ORNL	Oak Ridge National Laboratory
ASDC	Atmospheric Sciences Data Center	GSFC	Goddard Space Flight Center	PB	Peta Byte
ASF	Alaska Satellite Facility	HDF	Hierarchical Data Format	PO.DAAC	Physical Oceanography DAAC
AMSU	Advanced Microwave Sounding Unit	HIRDLS	High Resolution Dynamics Limb Sounder	RBD	Rate Buffered Data
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations	HSB	Humidity Sounder for Brazil	SAGE	Stratospheric Aerosol and Gas Experiment
CDDIS	Crustal Dynamics Data Information System	IWGDD	Interagency Working Group on Digital Data	SAR	Side Aperture Radar
CERES	Clouds and the Earth's Radiant Energy System	JPL	Jet Propulsion Laboratory	SEDAC	Socioeconomic Data and Applications Center
CEOS	Committee on Earth Observation Satellites	LAADS	Level 1 and Atmosphere Archive and Distribution System	SIM	Spectral Irradiance Monitor
CFI	Claes Fornell International	LANCE	Land, Atmosphere Near-real-time Capability for EOS	SIPS	Science Investigator-led Processing Systems
DAAC	Distribute Active Archive Center	LIS	Lightning Imaging Sensor	SNOW-I	Search 'N Order Web Interface
DISC	Data and Information Services Center	LP DAAC	Land Processes DAAC	SOLSTICE	Solar Stellar Comparison Experiment
ECHO	EOS ClearingHOuse	ManLan	Manhattan Landing (high performance exchange point in New York City)	SORCE	Solar Radiation and Climate Experiment
ECS	EOSDIS Core System	MISR	Multi-angle Imaging SpectroRadiometer	SPoRT	Short-term Prediction Research and Transition Center
EDOS	EOS Data and Operations System	MLS	Microwave Limb Sounder	TB	Tera Byte
EOC	EOS Operations Center	MODAPS	MODIS Data Processing System	TES	Tropospheric Emission Spectrometer
EOS	Earth Observing System	MODIS	Moderate Resolution Imaging Spectroradiometer	TIM	TRMM Microwave Imager
EOSDIS	EOS Data and Information System	MOPITT	Measurements of Pollution in the Troposphere	TRMM	Tropical Rainfall Measuring Mission
ESDIS	Earth Science Data and Information System	MSFC	Marshall Space Flight Center	UMBC	University of Maryland, Baltimore County
ESIP	Federation of Earth Science Information Partners	NASA	National Aeronautics and Space Administration	USGS	U.S. Geological Survey
ESSI	Earth and Space Science Informatics	NGIX	Next Generation Internet Exchange	WGISS	Working Group on Information Systems and Services
FGDC	Federal Geographic Data Committee	NISN	NASA Integrated Services Network	WIST	Warehouse Inventory Search Tool
FRGP	Front Range GigaPOP	NITRD	Networking and Information Technology Research and Development	XPS	XUV Photometer System
gbps	Giga bits per second	NPP	NPOESS Preparatory Project		
GCMD	Global Change Master Directory	NPOESS	National Polar-orbiting Operational Environmental Satellite System		