NASA's Evolution of ESE Data and Information Systems

Martha Maiden,

Program Executive for Data Systems
Program Planning and Development Division
NASA HQ

Land Processes DAAC Science Advisory Panel Meeting
September 10, 2003

Presentation Overview

- NASA's current satellite capabilities and future missions
- The current elements of the ESE data and information system
- The challenges to our current data and information systems
- Steps toward an evolution of the ESE data and information systems
- NASA's policies and plans for the long term archive

ESE Current Measurement Capabilities

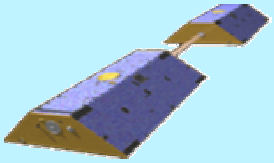
The Earth Observing System -- systematic measurement of interactions among land, oceans, atmosphere, ice & life

Exploratory missions to probe key Earth system processes globally for the first time


Landsat

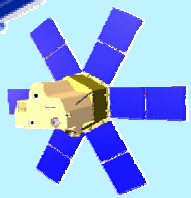

Aqua


TRMM

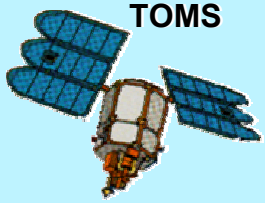

GRACE


Jason

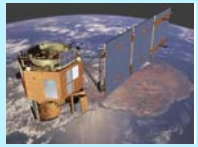

Terra

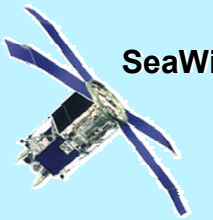

SORCE

TOPEX/ Poseidon


TOMS

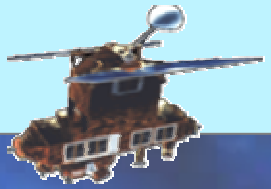

ICESat


EO-1: ALI & Hyperion


SeaWiFS

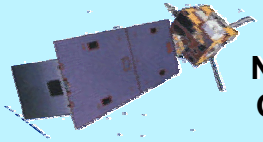
QuikScat

ERBS




ACRIMSAT

Operational precursor / Technology demos


NOAA/ GOES

Operational weather services missions with NOAA

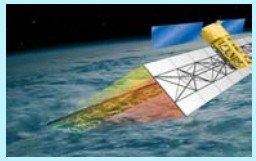

NOAA/ POES

ESE Next Generation Missions

Next generation systematic measurement missions to extend / enhance the record of science-quality global change data.

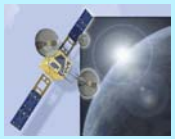
Future research measurements: Soil moisture, Advanced gravity, Ocean carbon, Cold climate processes, Vegetation recovery

Synthetic Aperture Radar

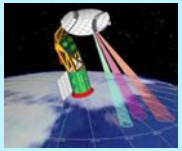


Aerosol Mission

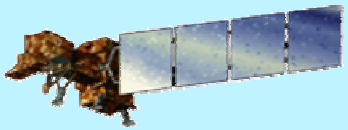
Global Precipitation Measurement



EO-3: GIFTS

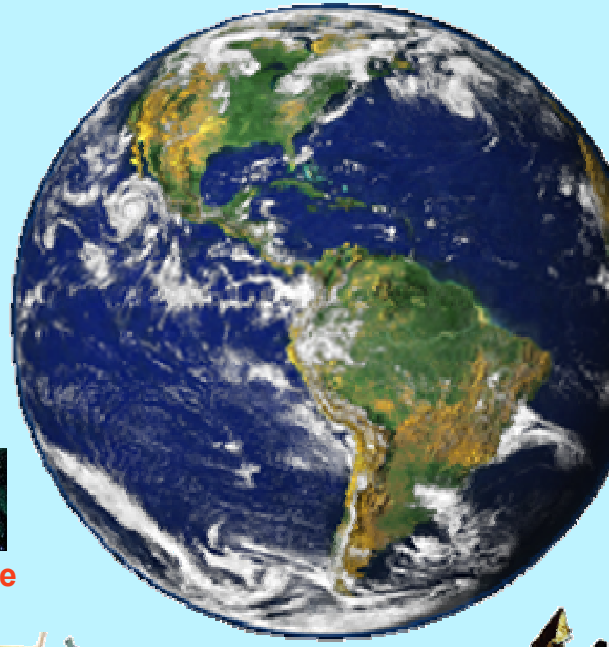


Aquarius

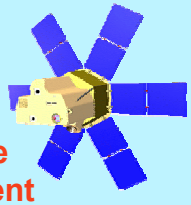


Landsat Data Continuity Mission

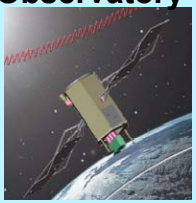
Chemistry / Climate Mission



Solar Irradiance Measurement

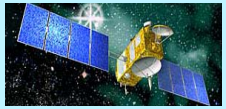


Orbiting Carbon Observatory



Calipso

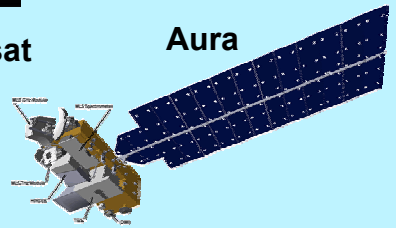
Total Column Ozone



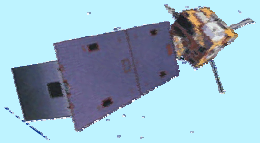
Ocean Surface Topography Mission

Cloudsat

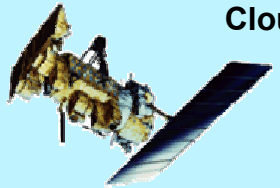
Aura



NPOESS Preparatory Project



NOAA/GOES-R



NOAA / NPOESS

Research missions to probe key Earth system processes globally for the first time

In Formulation

Candidate Future Missions

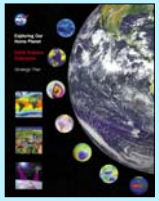
Operational weather services missions with NOAA/DOD

Cross-Cutting Science Observations to Support ESE Science Roadmaps

Science Roadmaps	Missions in Operation																									Missions in Implementation									Missions in Formulation								
	ERBS	UARS	TOPEX Poseidon	ERS-2	RadarSat I	IRS-P3 SIMBIOS	TOMS-EP	TRMM	QuikScat	Terra	ACRIMSAT	CHAMP	SAC-C	Jason-1	SAGE-III Meteor 3M-N1	Aqua	ADEOS-II (SeaWinds)	GOES I - M	POES	Landsat	EO-1	GRACE	ICESat	SORCE	Aura	SciSat-1	OrbView/SeaStar	GOES N - P	POES N/N'	NPOESS	CALIPSO	CloudSat	DSCVR	EO-3 GIFTS	NPP	Aquarius	LDCM	GOES R	OCO	Hydros	GPM	OSTM
Climate Variability and Change								S		S	S				S	S							P			S					P	S	P					S				
Atmospheric Composition		P		S			P			S					P	P								P	P													P				
Carbon, Ecosystems and Biogeochemistry				P	S					P			P			S				S	P						P						S				S					
Weather								P	P	S						S	P	P	P						S			P	P	P		P	S	P	P		P			P		
Water and Energy Cycles	P		P					P		S	P			P		P							S	P										P				P	P	P	
Earth's Surface and Interior Structure				S	P	P				S		P								P	P	P	S													P					

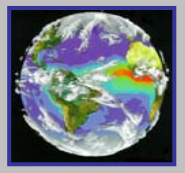
Shared mission between NASA and another agency or country
P = Primary Science theme
S = Support science or secondary science theme

Strategic Planning to Implementation through Science Roadmaps



ESE Strategy

Research Strategy



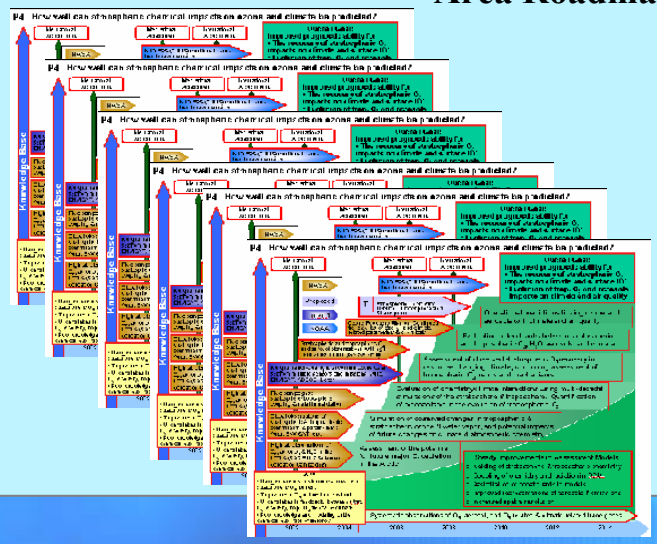
Education Strategy

Applications Strategy

Technology Strategy

Data Management Strategy

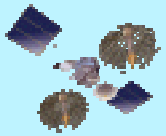
Six Science Focus Area Roadmaps



Measurement and Technology Requirements

Observation System Implementation

The Earth Science Enterprise



Spacecraft

Tracking & Data Relay Satellite (TDRS)

Ground Stations

Polar Ground Stations

Data Acquisition

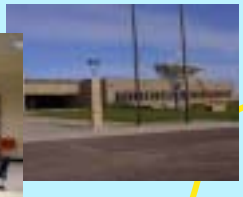
Flight Operations, Data Capture, Initial Processing, Backup Archive

Data Processing & Mission Control

Data Transport to DAACs

NASA Integrated Services Network (NISN) Mission Services

Science Data Processing, Info Mgmt, Data Archive, & Distribution



Distributed Active Archive Centers

Science Teams

Distribution, Access, Interoperability, Reuse

WWW

REASoNs

Research Users

Education Users

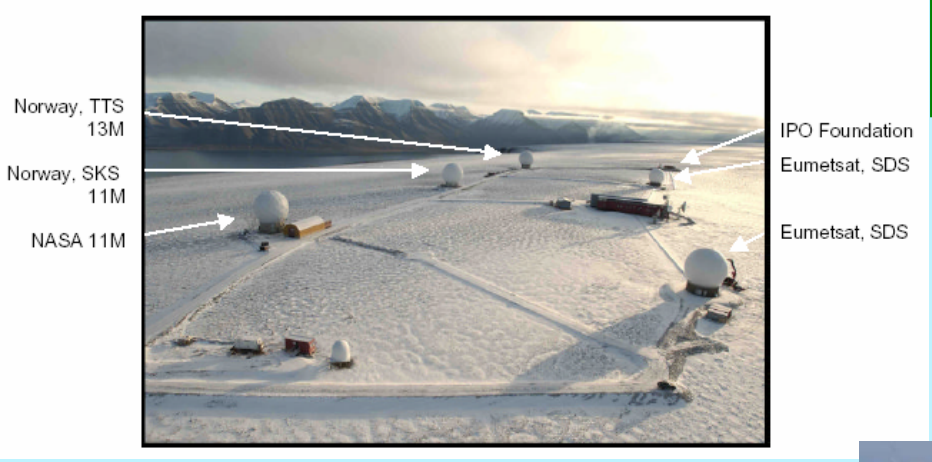
Value-Added Providers

Int'l Partners & Data Centers

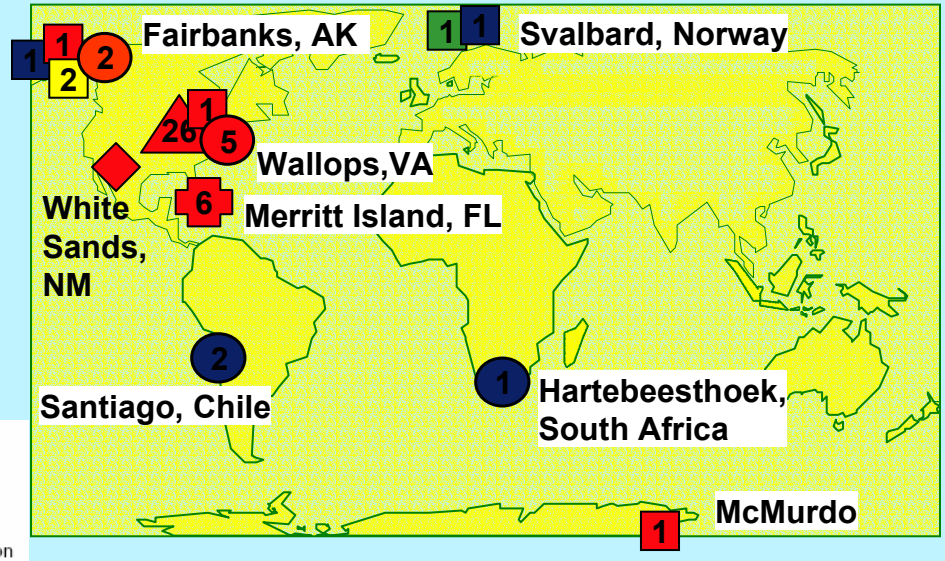
Interagency Data Centers

NASA's Ground Network

- 50 ground station antennas;
  7 geographic antenna locations
- 4 different owner/operator models
  - 14 contracts for commercial services



GN Antenna Map (Number of Antennas)



Primary Support Category

- Orbital S-Band @ 3 mbps
- Orbital X-Band @ 150 mbps
- Range
- Shuttle
- Scheduling

Owner/Operator Model

- # NASA/CSOC
- # NASA/university
- # NASA/commercial
- # Commercial

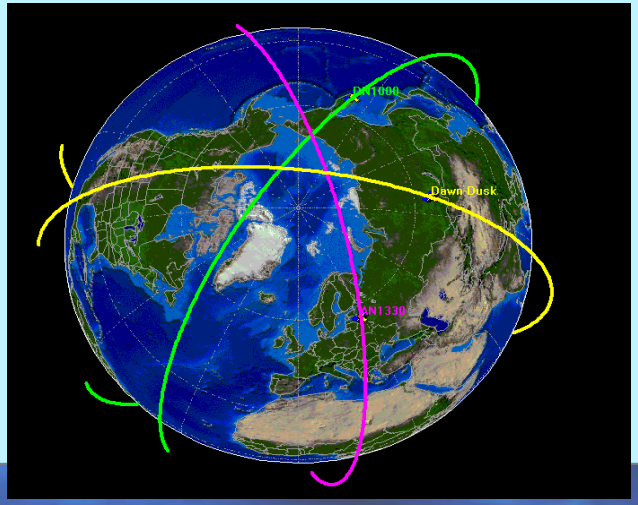
Shuttle Launch

Ultra Long Duration Balloons

Information Systems: EOS Mission Control

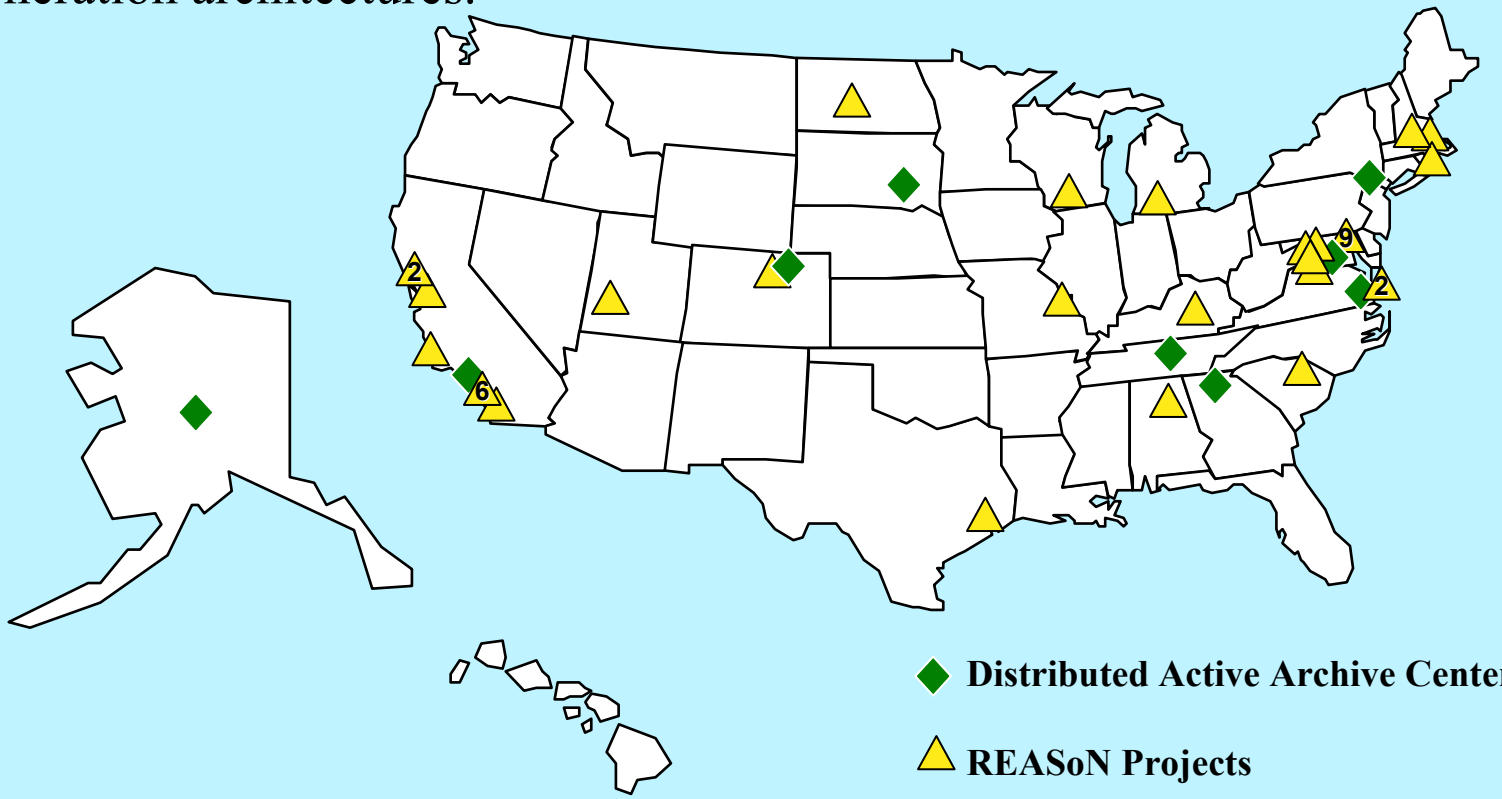
Earth Science Mission Operations

- Manage command and control spacecraft power, dynamics,
- Monitor instrument operations, build composite observation schedule
- Manage spacecraft recorder dumps and schedule retrieval with ground stations and space network
- House Landsat operations for USGS, and instrument operations for SeaWinds/ADEOS II



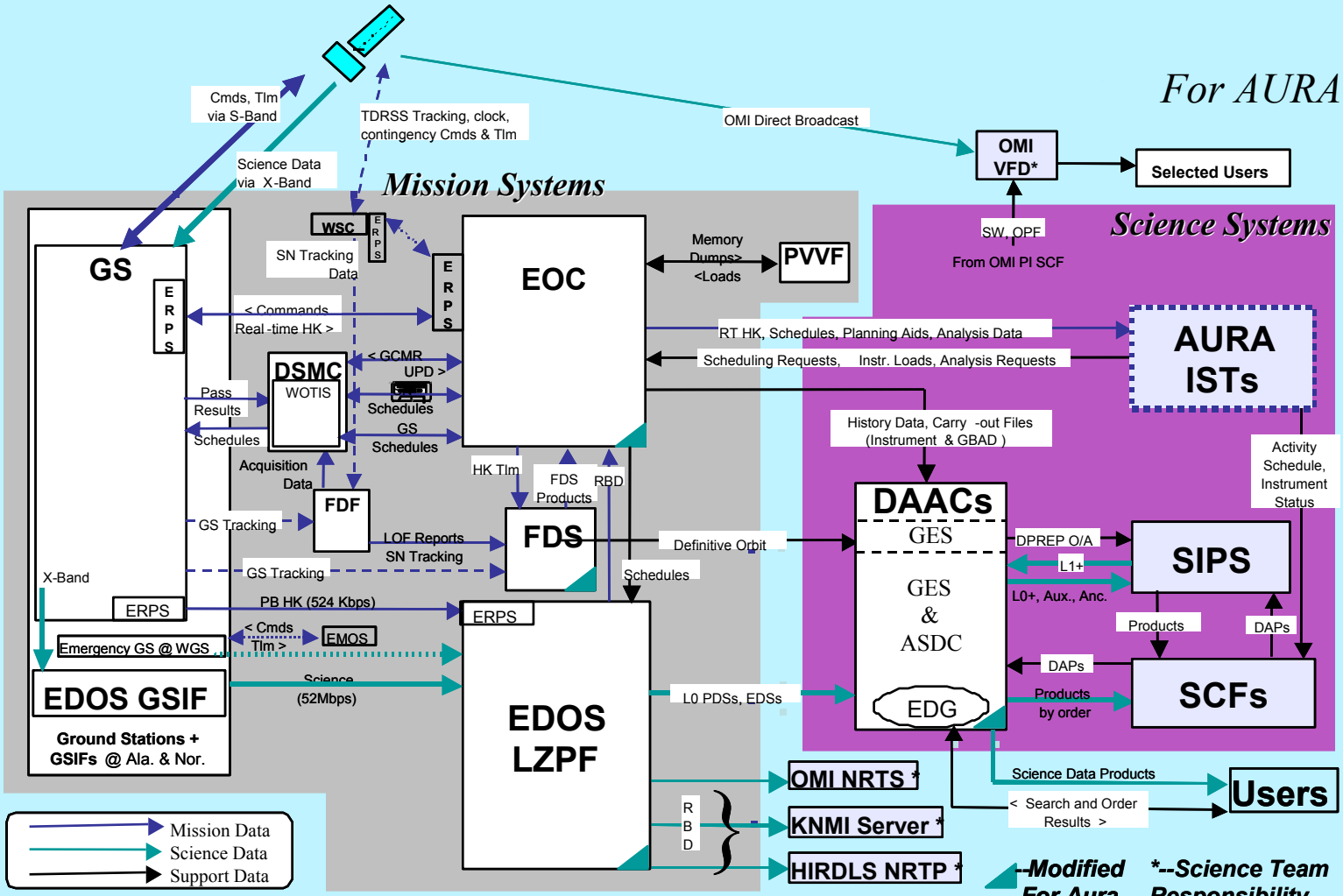
ESE Data Center Locations

- A total of 50 widely distributed data centers (some of which are at the same location).
- ESE has recently updated our peer reviewed data and information producing centers through the Research, Education and Applications, Solutions Network Cooperative Agreement Notice (REASoN CAN) for development of next-generation architectures.



EOSDIS Architectural Data Flow

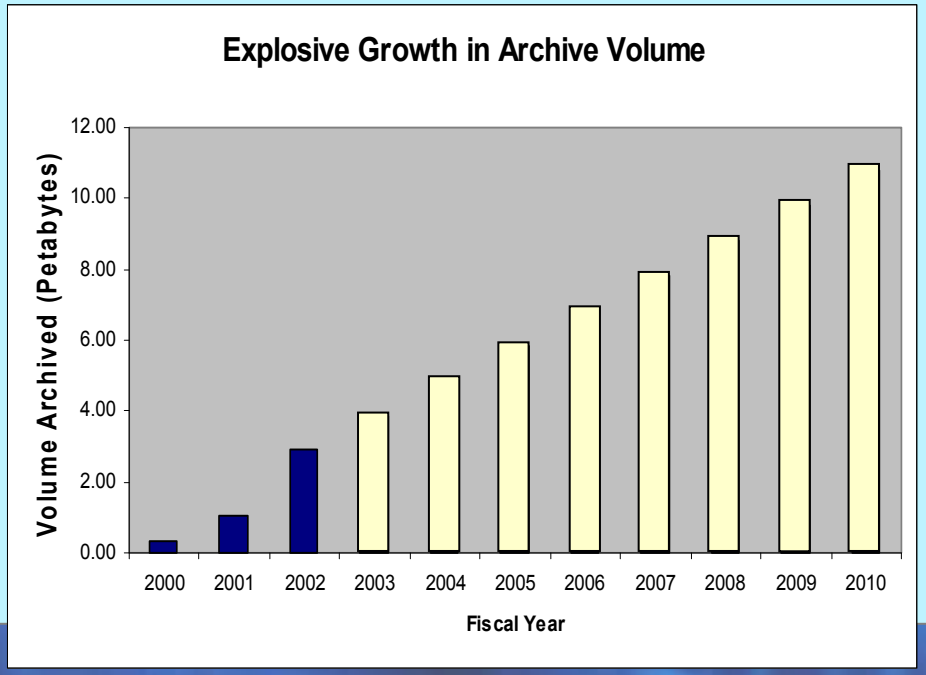
The end-to-end mission data handling responsibility must handle the information from photons in to science and operational products out.

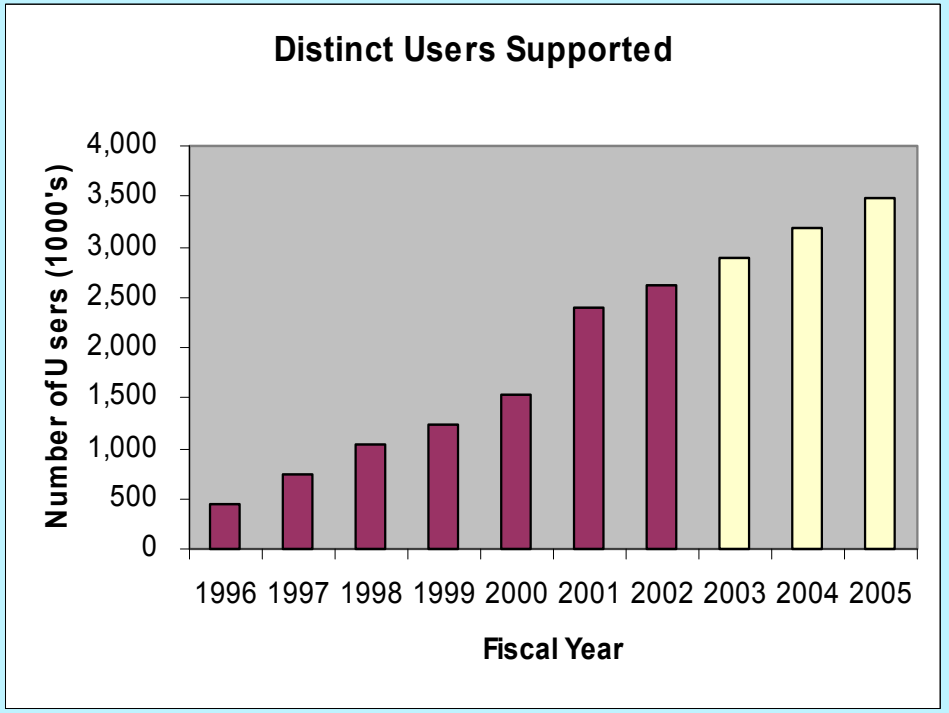


Information Systems: A Growing Demand for Earth Science Data and Information

Ingesting, processing, and archiving an unprecedented volume of climate and Earth science data (currently ~2TB/day).

We are benchmarking capabilities and processes for handling the capacities for future operational needs (e.g., NPOESS).





NASA provides access to Earth system science data, information, and services to millions of unique users.

Over the next decade, ESE will ensure the timely delivery of Earth Science information at an affordable cost by evolving to a more open, distributed set of data systems and service providers.

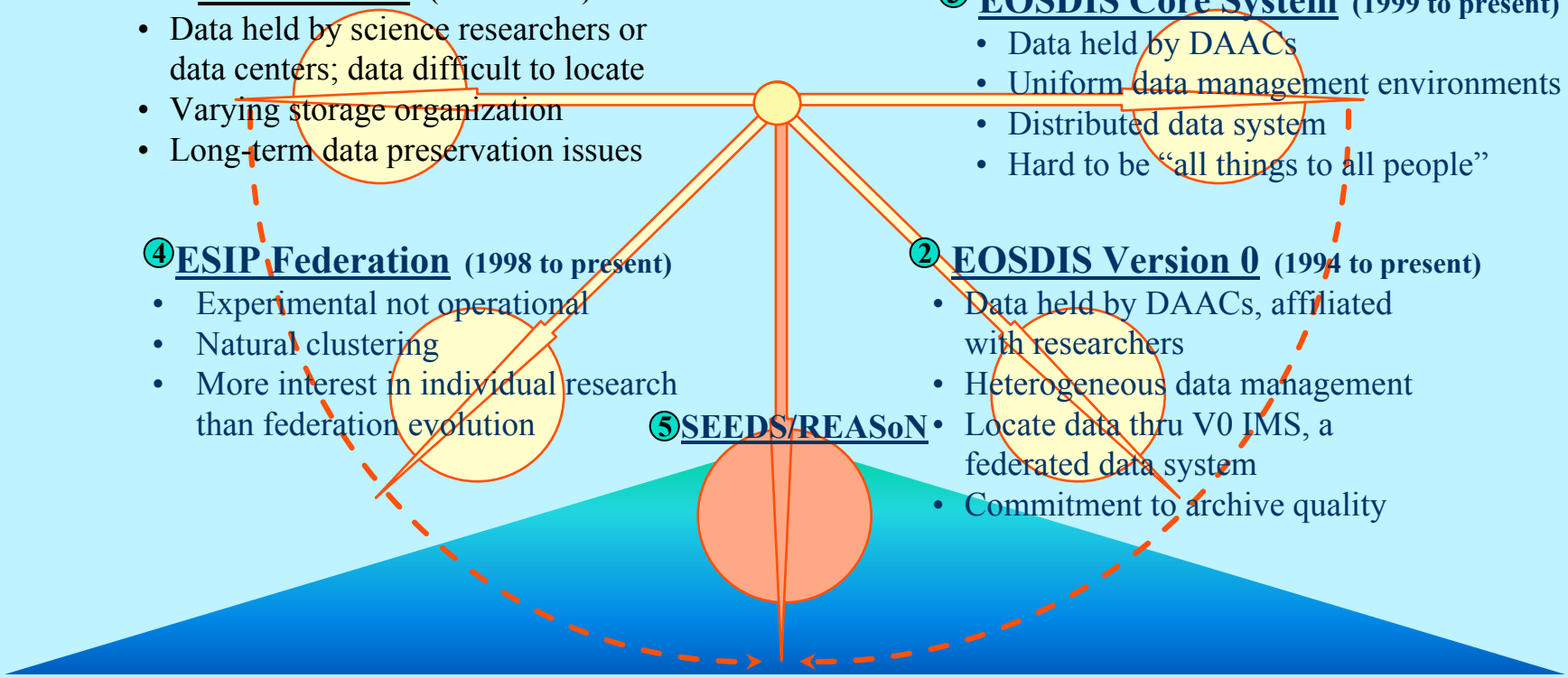
Challenges to NASA's ESE Data Systems

- Enable flexibility within data systems to adapt to new data stream(s) or to changes in current processing streams.
- Assure products flowing from processing stream meet the needs of the focus area science teams - scientists working in conjunction with data systems experts, and in consultation with their communities.
- Identify and create interfaces that facilitate the flow of data to modeling efforts (e.g. carbon assimilation) - "one size does not fit all". Enable seamless 'hooks' into data mining and high performance computing environments.
- Evolve from past instrument-focused processing systems to *measurement* oriented data systems within an interoperable framework that will help guide the flow of information and services, improve performance and access, and measurement focused and can be distributed geographically and logically.
- Meet the challenges of transitioning EOSDIS elements to a new DIS paradigm. What parts do we evolve? How to evolve these pieces without breakage or disruption?

The Evolution in ESE Data Management

Less NASA control in data management

More NASA control in data management



① **Pre-EOSDIS** (before 1994)
- Data held by science researchers or data centers; data difficult to locate
- Varying storage organization
- Long-term data preservation issues

③ **EOSDIS Core System** (1999 to present)
- Data held by DAACs
- Uniform data management environments
- Distributed data system
- Hard to be "all things to all people"

④ **ESIP Federation** (1998 to present)
- Experimental not operational
- Natural clustering
- More interest in individual research than federation evolution

② **EOSDIS Version 0** (1994 to present)
- Data held by DAACs, affiliated with researchers
- Heterogeneous data management
- Locate data thru V0 IMS, a federated data system
- Commitment to archive quality

⑤ **SEEDS/REASoN**

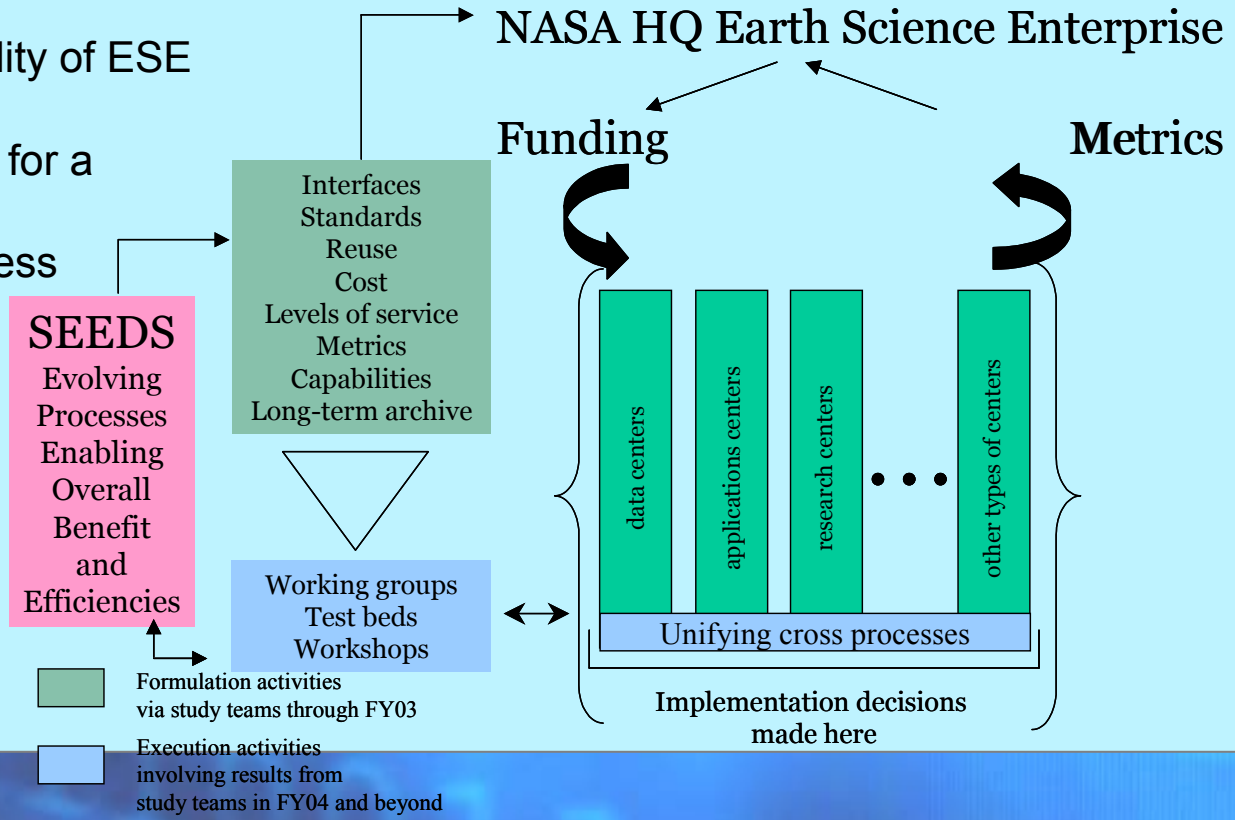
- **The intent of the SEEDS study and the REASoN CAN is to focus ESE data system evolution to:**
  - increase NASA's flexibility to adapt the network of data systems & service providers;
  - enable access for NASA's Applications program and its educational programs;
  - improve cost effectiveness throughout the data system development and operational life cycle;
  - leverage the capabilities, expertise, and lessons learned from existing data systems; and
  - assure long-term data stewardship and continuity of services.

Next Generation Data Processing Architecture

The Strategic Evolution of ESE Data Systems (SEEDS) study developed a framework for the evolution of data systems and service providers in the 2004 – 2015 timeframe - a move from centralized systems to distributed, heterogeneous systems.

SEEDS Study Objectives:

- Ensure timely delivery of ES information at an affordable cost
- Maximize availability and utility of ESE products
- Define a unifying framework for a more distributed network
- Assure continued effectiveness of an increasingly distributed and heterogeneous network of systems and services
- Engage the community on data management objectives and solutions



The Strategic Evolution of ESE Data Systems

The REASoN projects add to ESE next generation of data systems, while the evolution of EOSDIS elements continues.

The SEEDS study has concluded with recommendations and a framework for developing the next generation of distributed and heterogenous data systems.

SIPs, ESIPs, NewDISS and RESACs leverage heterogeneous activities and expand processing, archive and distribution capabilities through rapid prototyping.

DAACS - tightly coupled data system based on logically centralized, geographically distributed centers.

PIs collect and store data sets.

Disparate researchers working independently - difficult to exchange data and information.

Computer networking is difficult due to the variety of standards specific to highly technical IT research communities. Cost of computing infrastructure is high and only available to narrow range of users.

Emerging interoperability based on a set of minimal standards of interoperability - communication increases.

The merging of many networks into one common naming space - the World Wide Web. HTTP, URL and other protocols are widely adopted.

Communities form around common themes or interests - specialized methods for advanced information sharing are developed.

Rapid development in internet-based services and functions. Groups develop custom tools for expanding access to on-line resources. Ease of use increases the number of users.

Heterogeneous communities are connected via a robust set of standards and interfaces. Innovation proceeds rapidly as ideas from one community are shared among all.

# of Participants & Users

Computing $$

Difficulty to Interoperate

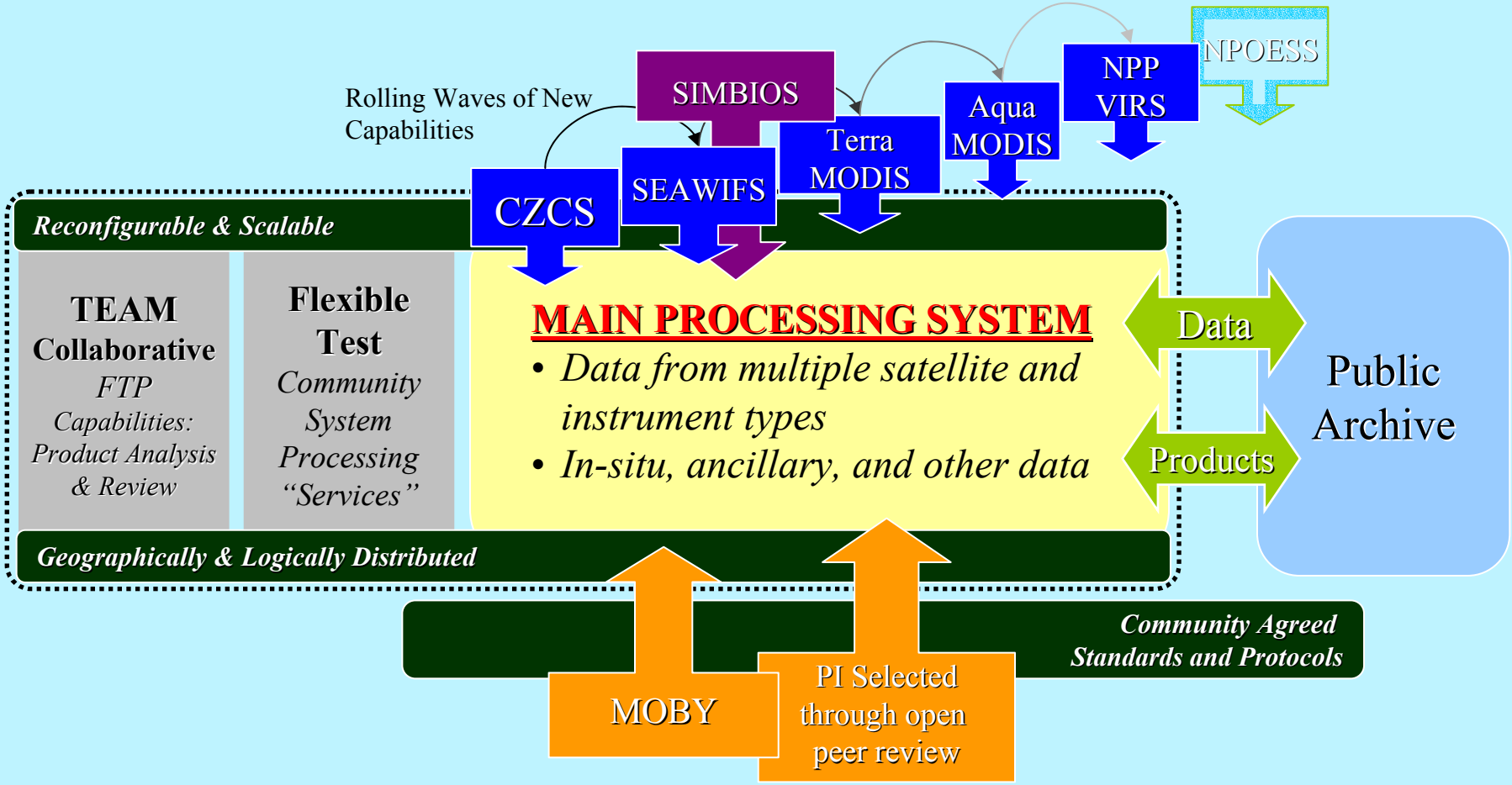
Summary of Near Term Data System Challenges

- **Missions to Measurements** will require data systems to support data coming from multiple missions and instruments. The emerging data systems should be capable of rapidly adapting to rolling waves of capabilities derived from new instruments and measurements.
- Foster **discipline/measurement focused science communities** to develop consensus views on important data and data system issues - broad community adoption and buy-in of algorithms, selection of products, sharing of expertise, etc.
- **SEEDS** activities in support of data system evolution will produce new opportunities to share and exchange solutions across communities. Information sharing from among theme communities will assure that data system solutions from one theme area can inform and possibly be adopted by others, including applications.
- Make initial steps in the development of **sensor web technologies**.

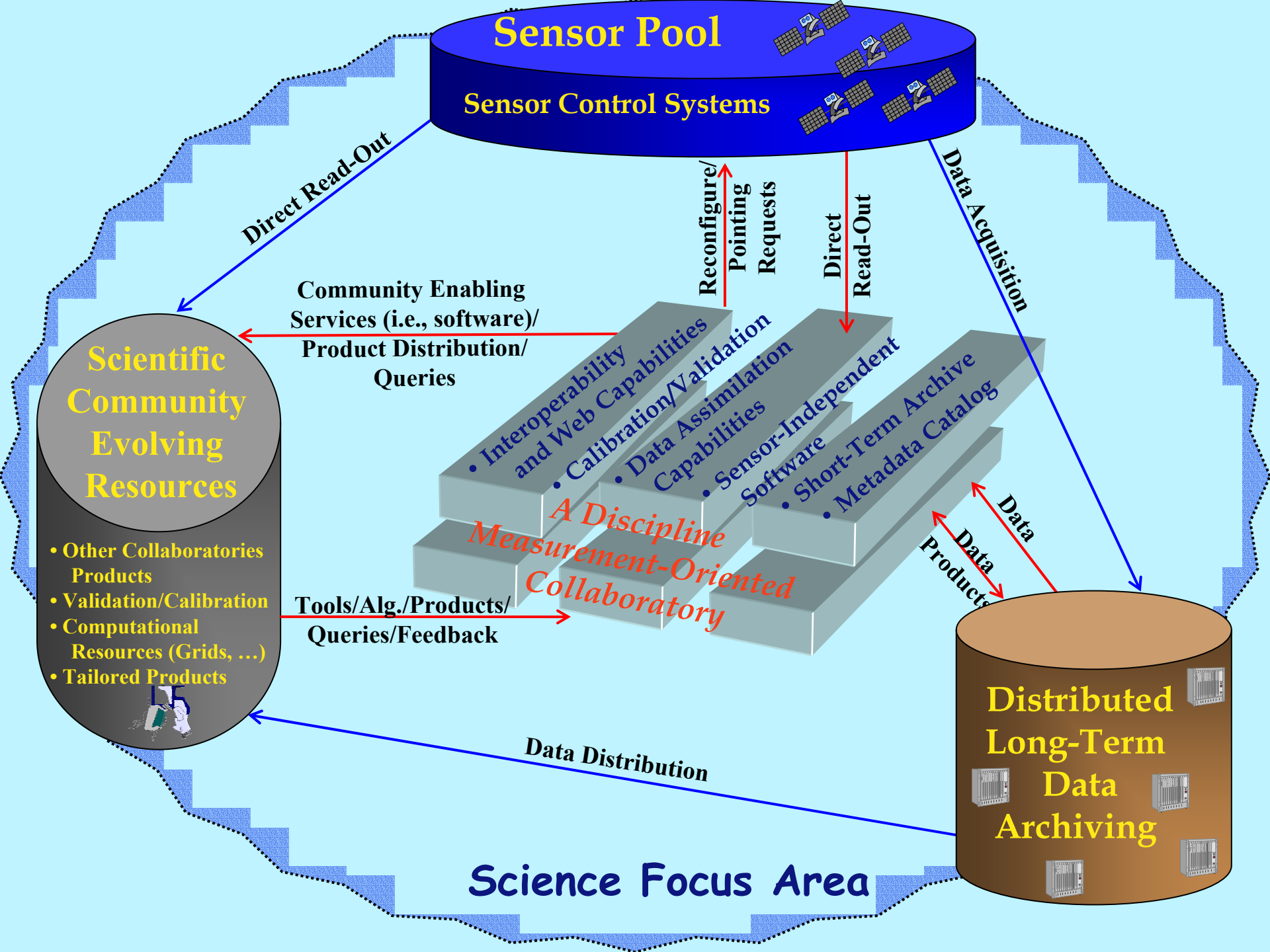
Response to challenges: An Earth Science Knowledge Environment

Evolve Earth Science driven data management from missions to measurements to support 'Data to Outcomes to Impacts' by infusing science community expertise and knowledge into transparent and seamless data and information systems.

A Data System Evolution Prototype: The Ocean Color Measurement Processing System



- Processing embedded within the science focus areas
- Distributed capabilities capturing "rolling waves"
- Community, participation, consensus and community services
- Move processing capabilities closer to the science teams
- Engage expertise through peer review selection


Sensor Pool
Sensor Control Systems
Direct Read-Out
Reconfigure/ Pointing Requests
Direct Read-Out
Data Acquisition
Community Enabling Services (i.e., software)/ Product Distribution/ Queries
Scientific Community Evolving Resources
• Other Collaboratories Products
• Validation/Calibration
• Computational Resources (Grids, …)
• Tailored Products
• Interoperability and Web Capabilities
• Calibration/Validation
• Data Assimilation Capabilities
• Sensor-Independent Software
• Short-Term Archive
• Metadata Catalog
A Discipline Measurement-Oriented Collaboratory
Tools/Alg./Products/ Queries/Feedback
Data
Data Products
Distributed Long-Term Data Archiving
Data Distribution
Science Focus Area

Managing the End-to-End Information Flow of the Future

Petabytes 10^{15}

Multi-platform, multi-parameter, high spatial and temporal resolution, remote & in-situ sensing

Calibration, Transformation To Characterized Geophysical Parameters

Terabytes 10^{12}

Interaction Between Modeling/Forecasting and Observation Systems

Gigabytes 10^{9}

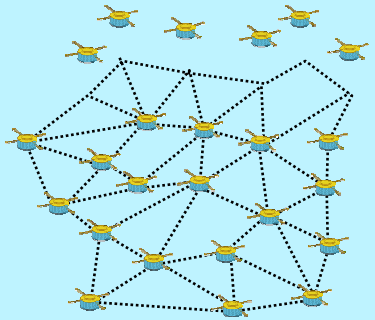
Interactive Dissemination

Predictions

Megabytes 10^{6}

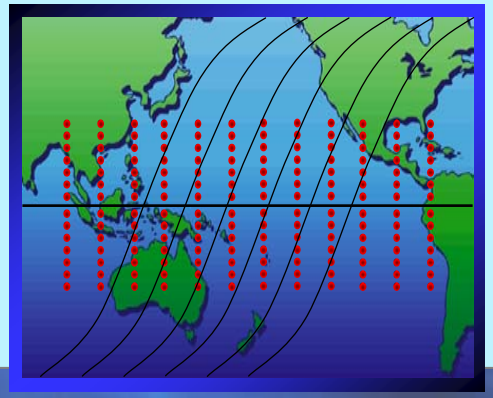
Advanced Sensors

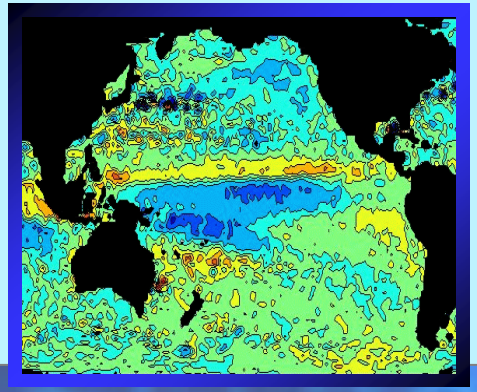
Data Processing & Analysis

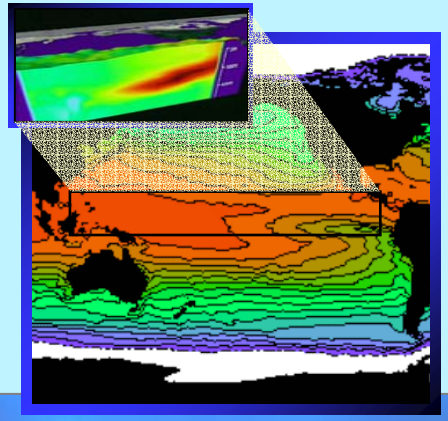


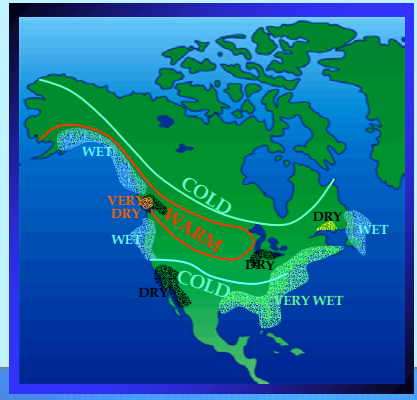
Information Synthesis

Access to Knowledge





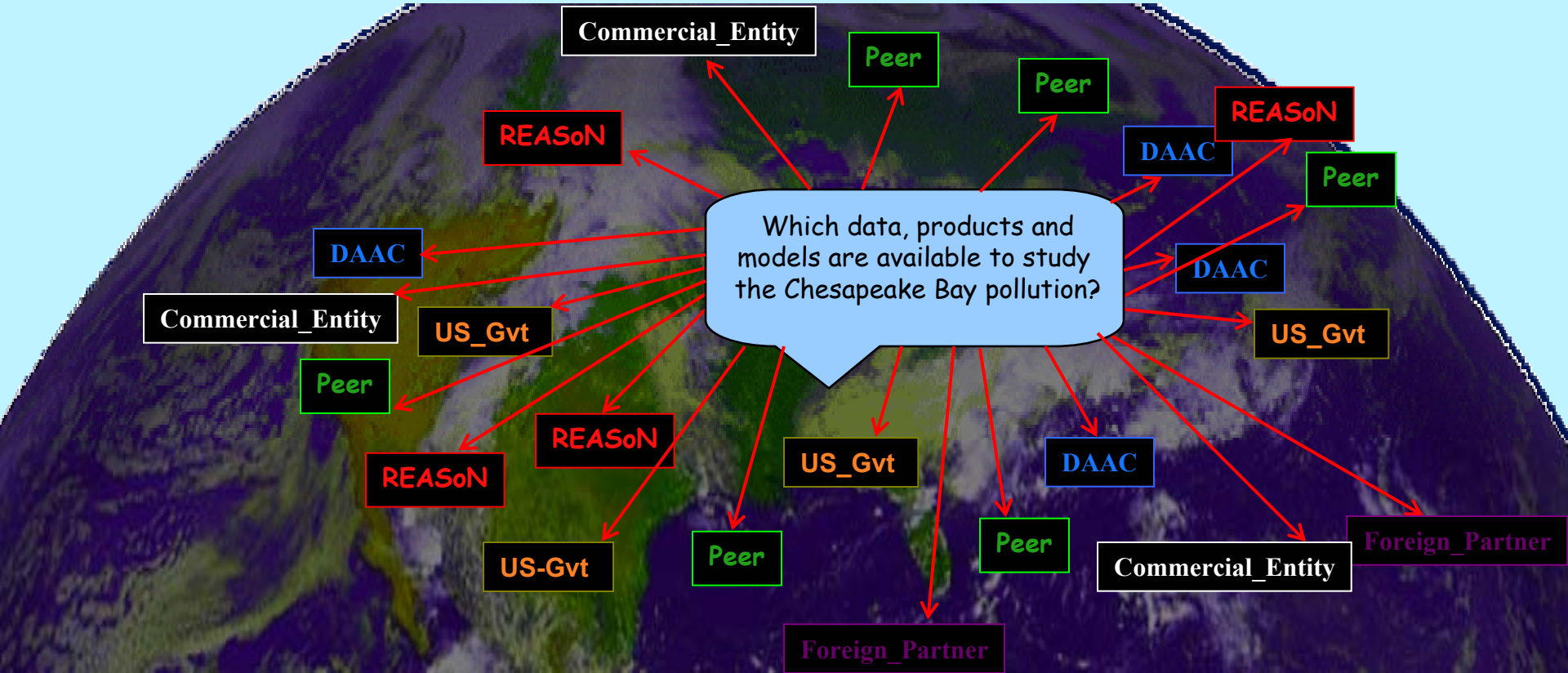




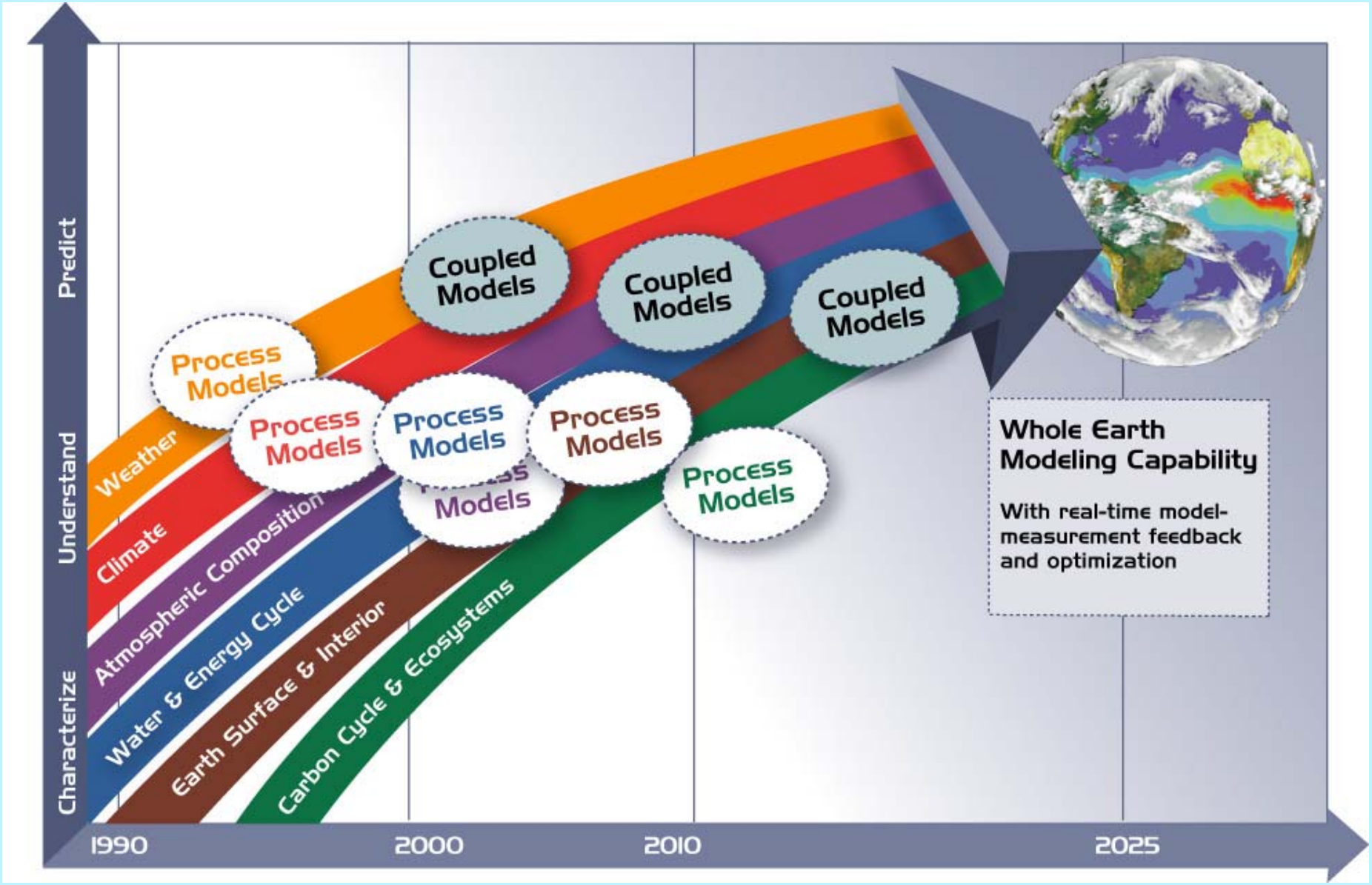
The Future of Knowledge and Data Sharing

Integration of:

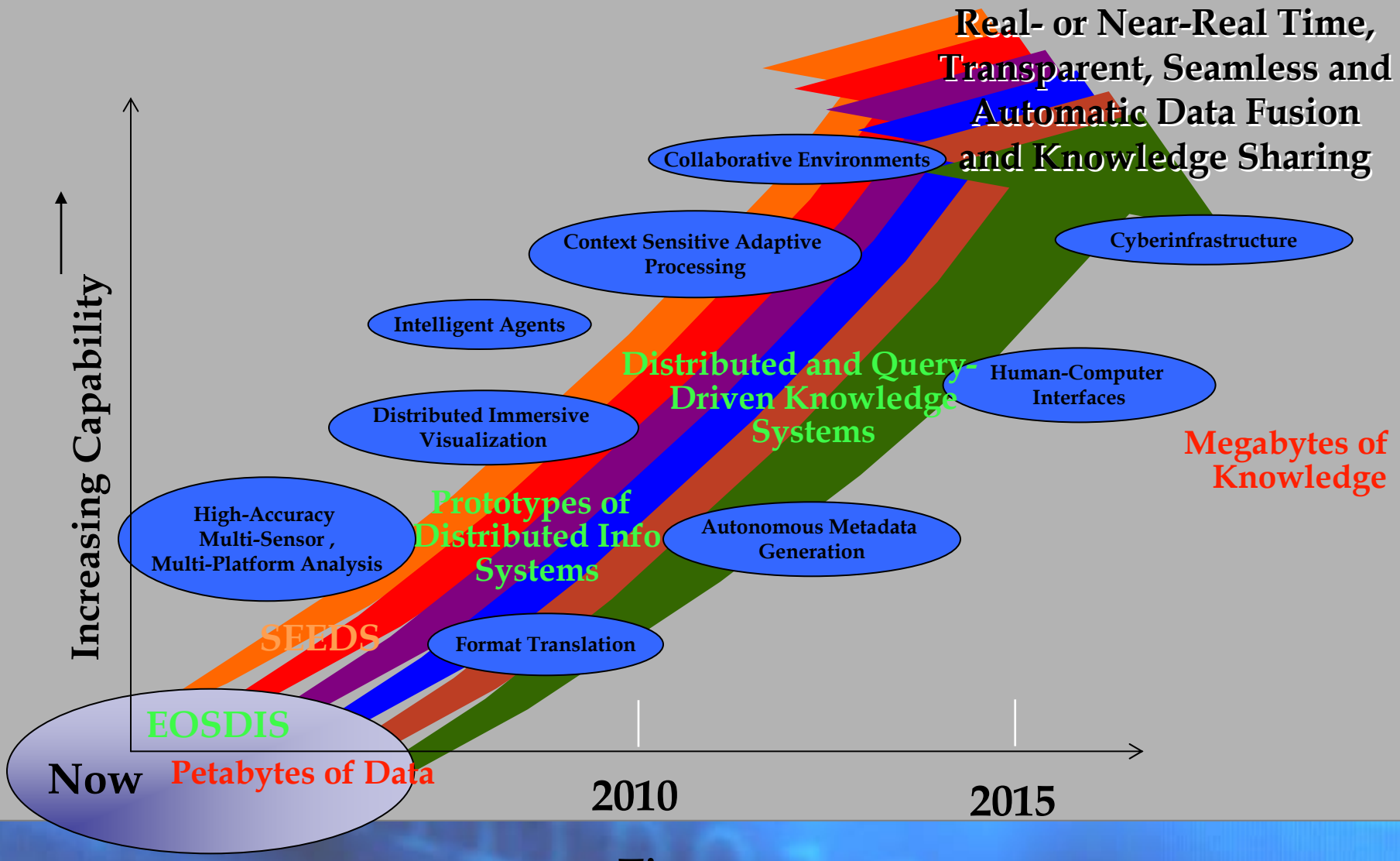
- **DAAC Minimum Levels of Services**
  - **e.g., Long-Term Data and Products**
- **Peer-to-Peer or Peer-to-Institutions**
  - **Prototypes: GCMD; ECHO; Mercury;**
  - **Nepster/Direct Readout; UNITE /ESML**



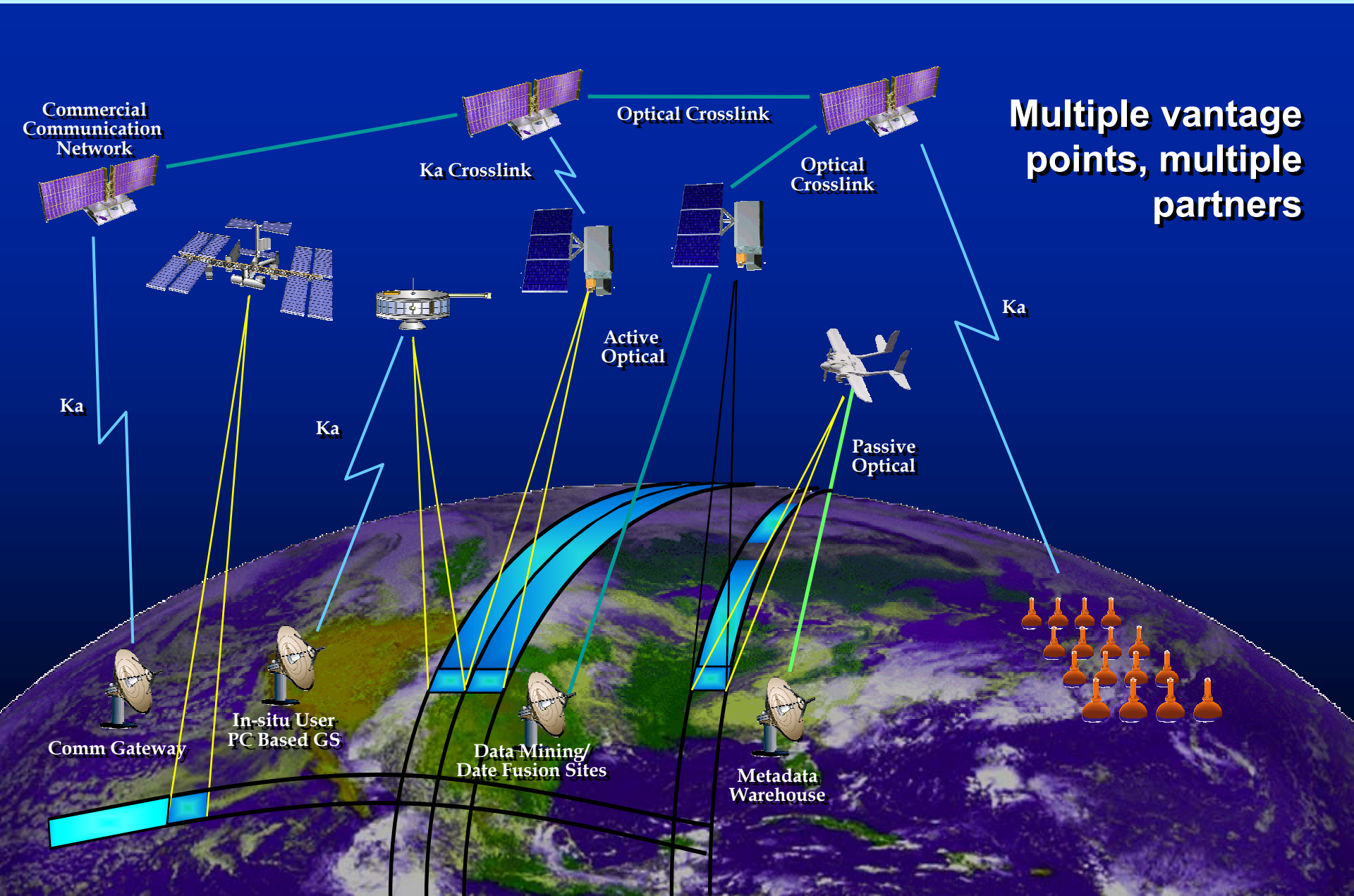
Earth System Science Today & Tomorrow:
Linking Measurements-Missions-Models To Improve Scientific Impact



Towards an ESE Knowledge Environments Vision



Our Vision -- An Integrated Global Observing System



Long-term Archive: NASA's Policies and Plans

LTA Policy

- NASA's agency mandate is research
  - ESE mission data products are managed by NASA's Distributed Active Archive Centers for the life of the missions
- NASA's partner agencies (NOAA and USGS) have a mandate for operations
  - These agencies will provide LTA services for ESE mission data products
- NASA-NOAA MOU on EOS - 1989
- NOAA will use its best efforts to…Assume responsibility at a time to be agreed upon for active long-term archiving and appropriate science support activities for atmospheric and oceans data for the EOS program
- NASA-USGS MOU on EOS - 1993
- USGS will fund long-term archive functions. USGS will fund archive and distribution functions, including operations and maintenance costs for EOS and related data more than 3 years old.

Call to Action from the Science Community

- It is essential that the LTA program preserve key long term data and information including the definitive version of the EOS Level 1 data and derived products...and any other data sets/products needed to interpret them...and any other data sets needed for calibration, and validation... (USGCRP, LTA Workshop Report, 1998).
- A long-term archive should be established and operated in the simplest way possible to meet user needs and program goals. A long-term archive is not only for today's generation of users but also for the next generation of scientists and citizens whose needs have yet to be expressed but must be provided for (NRC CES, Issues in the Integration of Research and Operational Satellite Systems for Climate Research, 2000).
- NOAA should begin now to develop and implement the capabilities to preserve in perpetuity the basic satellite measurements, i.e., radiances and brightness temperatures (NRC CES, Ensuring the Climate Record, 2000).

USGS Perspective

- USGS is responsible for LTA of ESE land data
- USGS's NSLRSDA
  - National Satellite Land Remote Sensing Data Archive at EDC
  - Established by legislation in 1992
  - Provides archive for MSS, TM, AVHRR (LAC to 1989), Corona, and ETM+
  - $5M annual budget
  - Archive Advisory Committee with representation from academia, government, industry, NGOs, and international partners

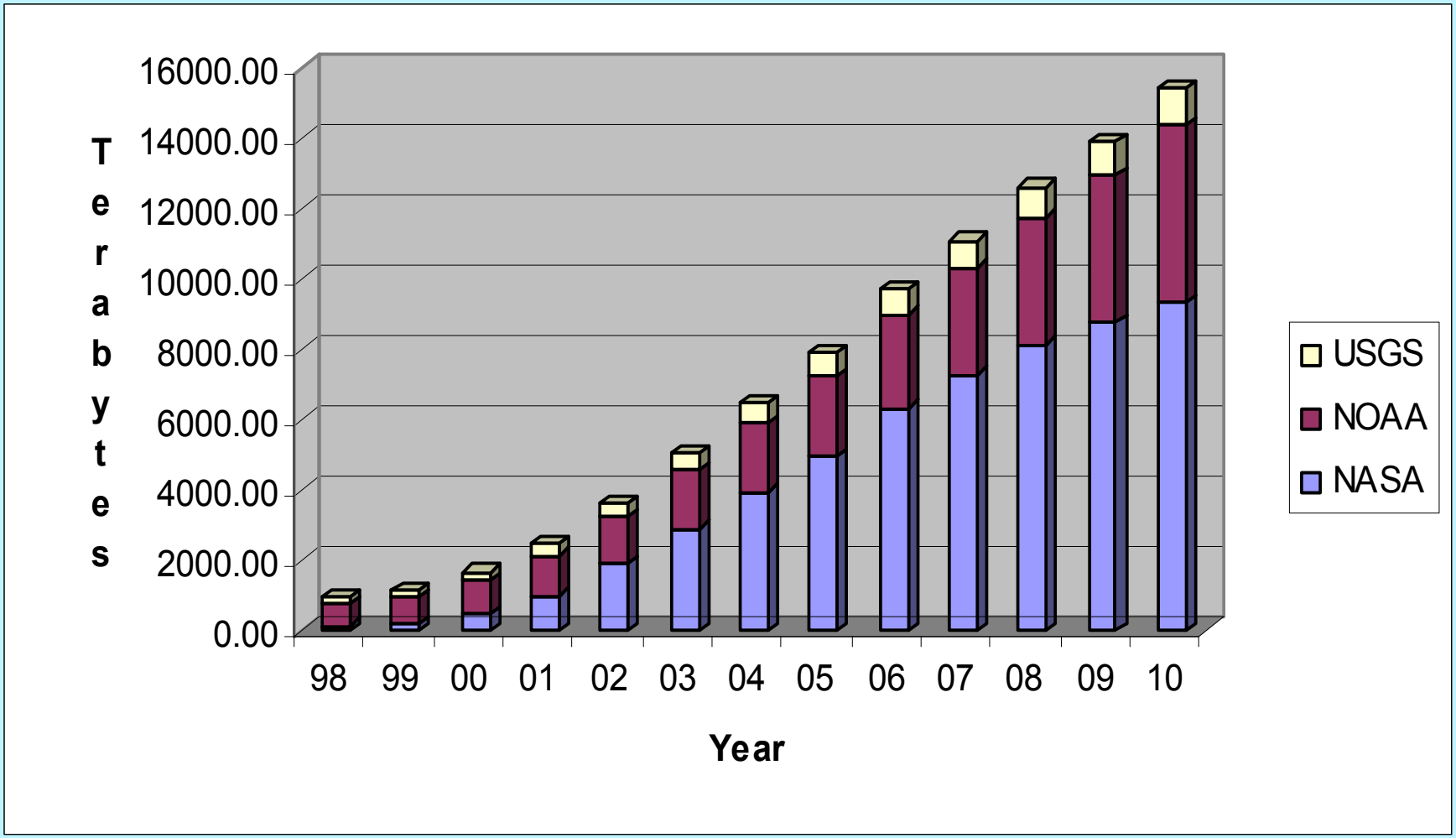
USGS LTA Status

- Global Land Cover Characteristics Completed
- Global Land 1-KM AVHRR FY04
- N. American Landscape Char. FY04
- Remaining V0 Data Sets* FY04
- EOS Data Sets* TBD
- Significant challenges exist...
  - in defining the levels of service applied to LTA of NASA data
  - in defining the time-table to make the transition to LTA
  - and in securing the resources needed to do the job

A form of archive-in-place is under consideration as a mechanism for USGS LTA of NASA data. An annex to the NASA/USGS MOU for the long-term archive of NASA's land science data that specifically addresses the transfer of "EOSDIS Version 0" data sets has been generated and is signature cycle.

Backup Slides

National Earth Science Data Holdings



Graph updated May 31, 2003

Assumptions

- Chart in page 2 shows data generated by NASA, NOAA and USGS in the respective bars, without regard to where they are archived. (e.g., Landsat-7 and ASTER data are included under NASA even though they are archived at USGS's EROS Data Center)
- Estimates for NPP and GPM volumes are included in NASA bars
- NASA data volumes are shown based on design lifetimes (typically 5 years) for the various satellite missions; generally satellites last significantly longer, and as data are accumulated beyond design lifetimes, the volumes rise faster than shown in later years
- Allocations of data volumes to agencies will change as long-term archives are implemented with NOAA and USGS

NOAA Strategy and Plans

- NOAA is responsible for ESE climate data
- Discussion on LTA of NASA Earth Science Data in NOAA on-going since about 1993
- NOAA's CLASS (Comprehensive Large-Array data Stewardship System)
  - NOAA initiative for archive, access and distribution to NESDIS data products
  - Will provide data services for NPP and NPOESS
  - "Statement of Intent" signed by G. Asrar and G. Withee for inclusion of NASA Earth Science LTA within CLASS (see back-ups)
  - A joint NASA-NOAA team developed a draft Implementation Plan as called for in the Statement of Intent
  - A joint NASA/NOAA team is working to define the interface and levels of service requirements that will enable a test of transfer of MODIS low level (Level 0/1)data from NASA's Goddard DAAC to CLASS to occur in FY03

NASA-NOAA COLLABORATION ON LONG-TERM ARCHIVE
STATEMENT OF INTENT

In a 1989 Memorandum of Understanding (MOU) NOAA and NASA agreed that NOAA would assume responsibilities for the NASA Earth Observing System (EOS) data and other of NASA's related atmospheric and oceanographic data. The MOU it specifically calls for NASA and NOAA to

> "... [Generate] a joint plan for coordinated development of the short- and long-term archives...and associated science support activities...in accordance with a schedule to be agreed" and to
>
> "Prepare by an agreed date an initial Program Definition and Implementation Plan. The plan will identify and describe the scope of major elements covered by this agreement, including estimate funding requirements by each agency and implementation schedules."

Under the direction provided by the MOU cited above, this Statement of Intent calls for NASA's Earth Science Enterprise and NOAA's National Environmental Satellite, Data and Information Service to take the following immediate steps.

1. Agree that the NOAA NESDIS Comprehensive Large Array-data Stewardship System (CLASS) shall serve as the national atmospheric and oceanic long-term data archive.
2. Agree that appropriate atmospheric and oceanic data records from NASA's Earth Science Enterprise program will be included in this national archive.
3. Agree to merge the activities associated with the currently established Long-term Archive (LTA) and CLASS study groups into a unified Integrated Product Development Team (IPDT) to focus on the above goals.
4. Agree to charge the IPDT to use the LTA adaptive approach outlined at the June 15, 2001 meeting as the starting point for future activities.
5. Appoint Rob Mairs of NOAA and Martha Maiden of NASA as Co-leads of the IPDT and charge them to report back to the Assistant Administrator of NESDIS and the Associate Administrator for Earth Science by September 30, 2001 with an overall strategy and a joint Program Definition and Implementation Plan for the development of this national archive for atmospheric and oceanic data products.

For the National Aeronautics and Space Administration

_____

Dr. Ghassem Asrar
Associate Administrator,
Office of Earth Science

For the National Oceanic and Atmospheric Administration

_____

Mr. Gregory W. Withee
Assistant Administrator
for Satellite and Information Services

"Statement of Intent" signed July 3, 2001