Evolution of EOSDIS, Science Data Systems and the DAACs

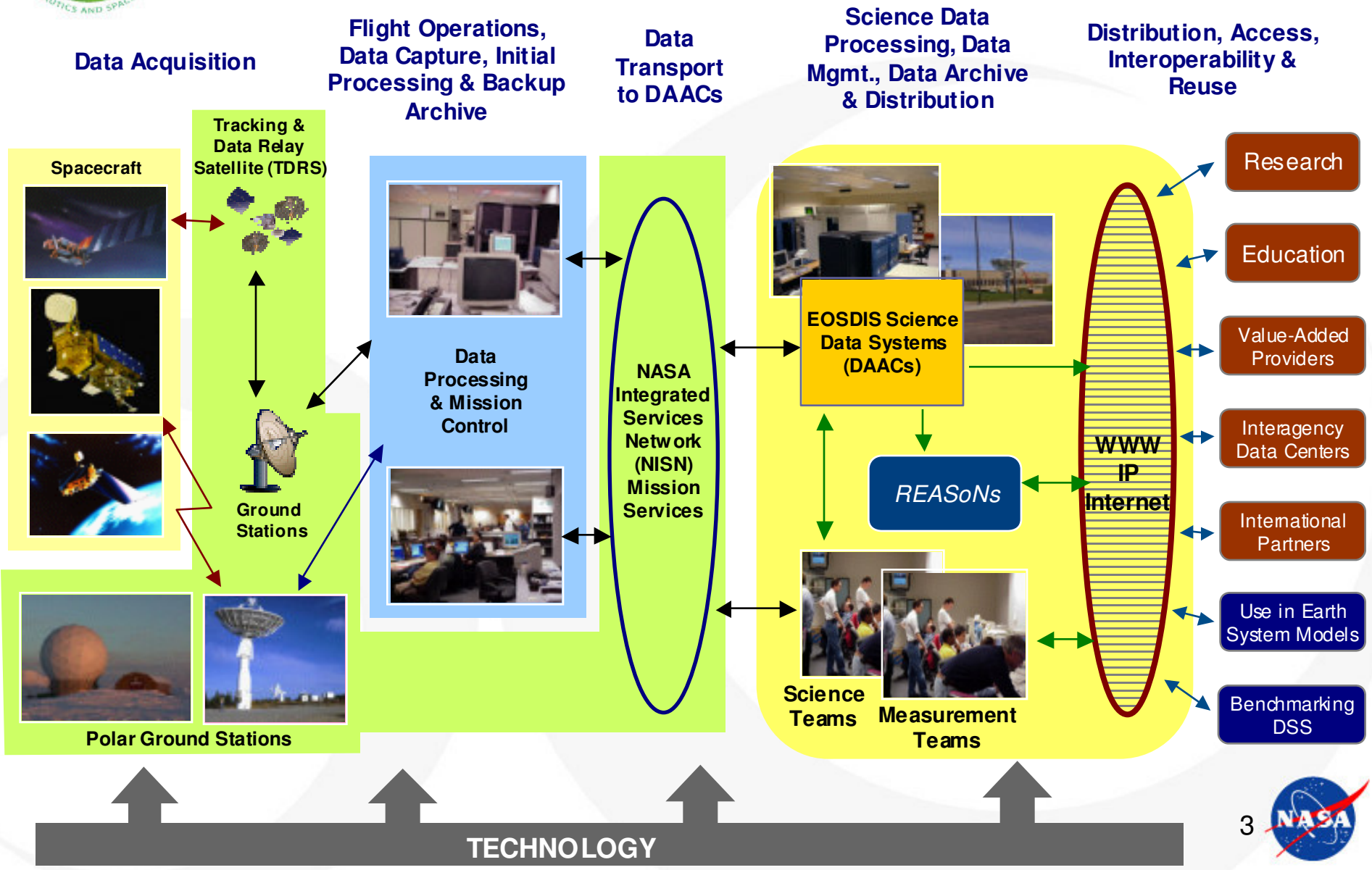
MODIS Science Team Meeting
July 13-15, 2004
Baltimore

Martha Maiden
Program Executive for Data Systems
NASA Headquarters

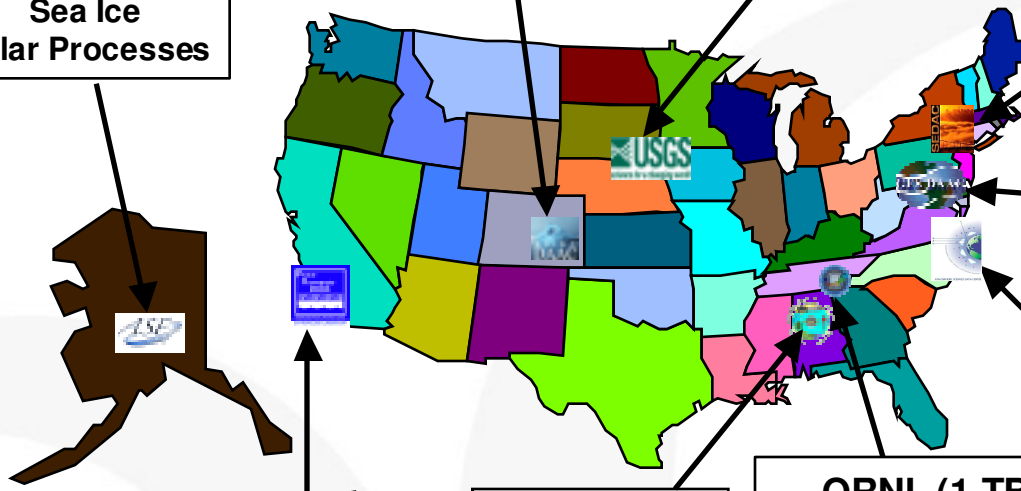
Overview of Presentation

- Current state of NASA's ESE Data and Information Systems.
- EOSDIS elements and plans for evolution.
- The Ocean Discipline Processing System and the Precipitation Processing System - prototypes for understanding the elements of measurement-based data systems, variation amongst such systems ("one size does not fit all"), and how such systems will drive evolution of ESE Data and Information System.

ESE Data System Architecture (current)



ESE discipline focused DAACs: Serving and Exploring Earth Science Data, Information, and Information Technology…as only NASA can



ASF (256 TB)
SAR Products
Sea Ice
Polar Processes

NSIDC (67 TB)
Cryosphere
Polar Processes

LPDAAC-EDC (1143 TB)
Land Processes
& Features

SEDAC (0.1 TB)
Human Interactions in
Global Change

GES DAAC-GSFC (1334 TB)
Upper Atmosphere
Atmospheric Dynamics, Ocean Color, Global Biosphere, Hydrology, Radiance Data

ASDC-LaRC (340 TB)
Radiation Budget,Clouds
Aerosols, Tropospheric
Chemistry

ORNL (1 TB)
Biogeochemical
Dynamics
EOS Land Validation

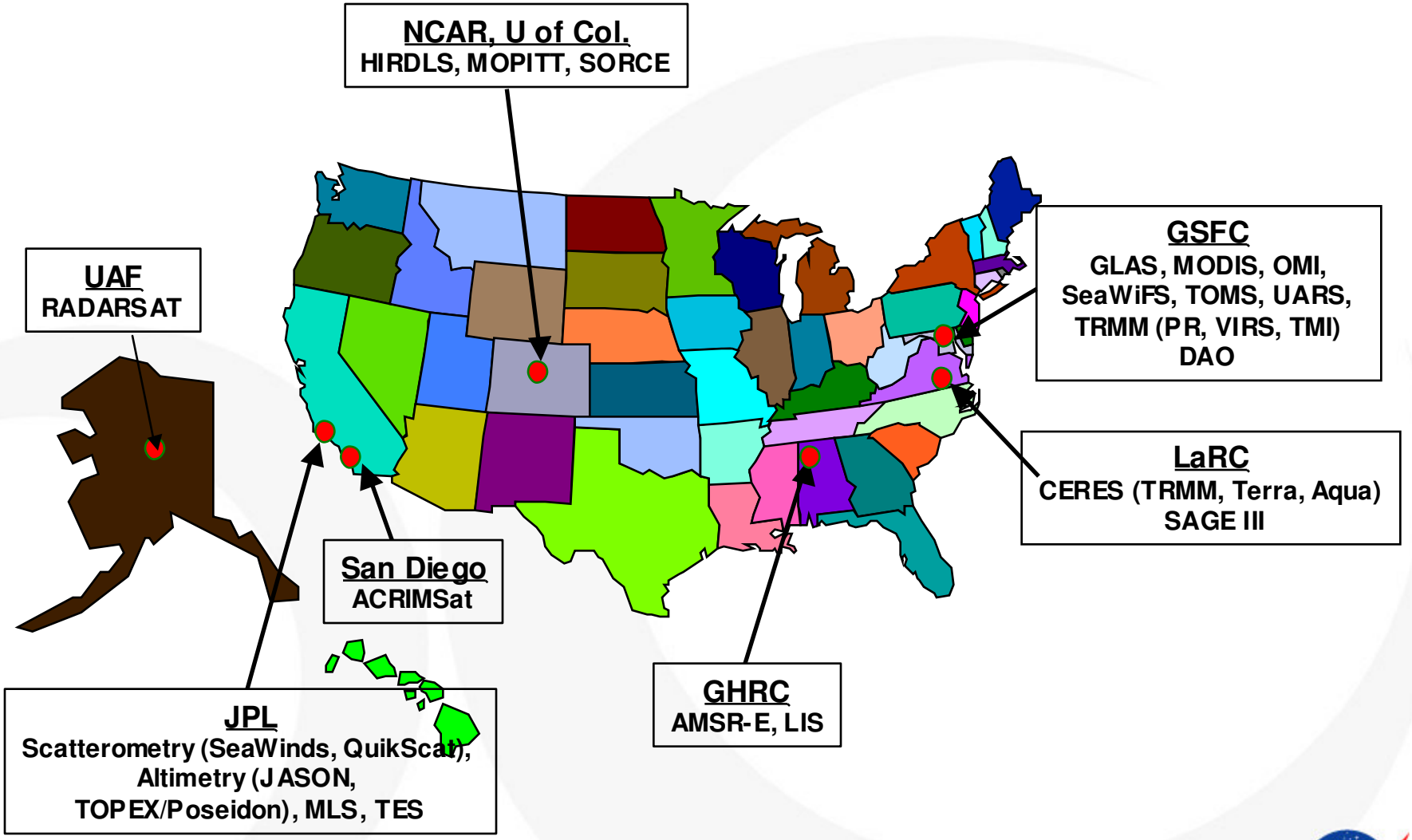
GHRC (4TB)
Global
Hydrology

PODAAC-JPL (6 TB)
Ocean Circulation
Air-Sea Interactions

The DAACS:

- Focus on intelligent use of NASA's information through sound data management
- Ensure unencumbered distribution of NASA Earth science data to ALL users
- Provide complete user services and data expertise services
- Exploit advanced technologies to further facilitate the ESE mission

ESE Funded Science Investigator-led Processing Systems (SIPSs)



Cross-Cutting Science Observations to Support ESE Science Roadmap

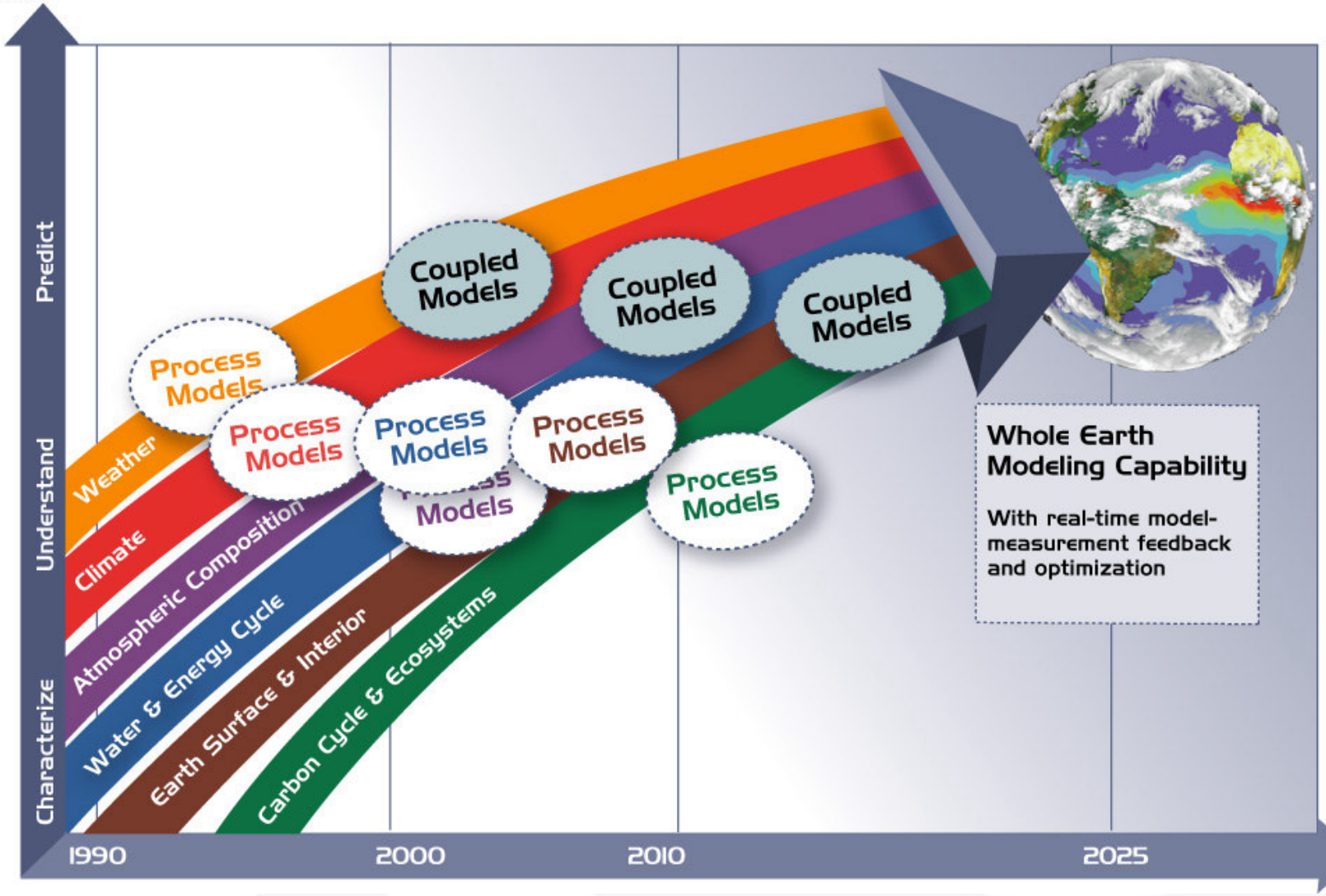
		Missions in Operation																									Missions in Implementation									Missions in Formulation									
		ERBS	UARS	TOPEX Poseidon	ERS-2	RadarSat I	IRS-P3 SIMBIOS	TOMS-EP	TRMM	QuikScat	Terra	ACRIMSAT	CHAMP	SAC-C	Jason-1	SAGE-III Meteor 3M-N1	Aqua	ADEOS-II (SeaWinds)	GOES I - M	POES	Landsat	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
Science Roadmaps	Climate Variability and Change								S		S	S				S	S							P			S					P	S	P					S			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

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



Drivers of Evolving NASA's Earth Science Data & Info Systems

- Missions to Measurements
  - ESE is moving from mission-based data systems to those that focus on Earth science measurements.
  - ESE's DIS will be a resource for Earth science-focused communities enabling research, and will be flexible, scalable and suited for the particular community needs.
  - Continue on the pathways for acquiring observations to understand processes and develop Earth system models.
- Maturity of ESE sharpens focus on environmental view of data: SIPS moving to "CoMPS", Community-based Measurement Processing Systems
  - Pathfinder Datasets were early prototype in NASA
  - Need intercalibrated time series of geophysical records, NRC has termed "Climate Data Records"
  - 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

Drivers of Evolving NASA's Earth Science Data & Info Systems, continued

- Integration for Earth System Science
  - Facilitate information synthesis. (Initial EOSDIS organization was zeroth-order attempt and can be improved.)
  - Increase data usability by science research, application, and modeling communities
- The Advance of Information Technologies
  - NASA will remain at the forefront of IT development and will partner with other agencies to ensure the strategic use of IT resources to avoid obsolescence and enable enhanced performance.
  - The lowering cost of IT infrastructure enables ESE data systems to take advantage of improving computation, storage and network capabilities.
- Facilitate the Transition from Research to Operations
  - Work with Federal partners to transition operational elements of data systems to other agencies while maintaining core data system functions necessary for conducting NASA ESE mission and goals.

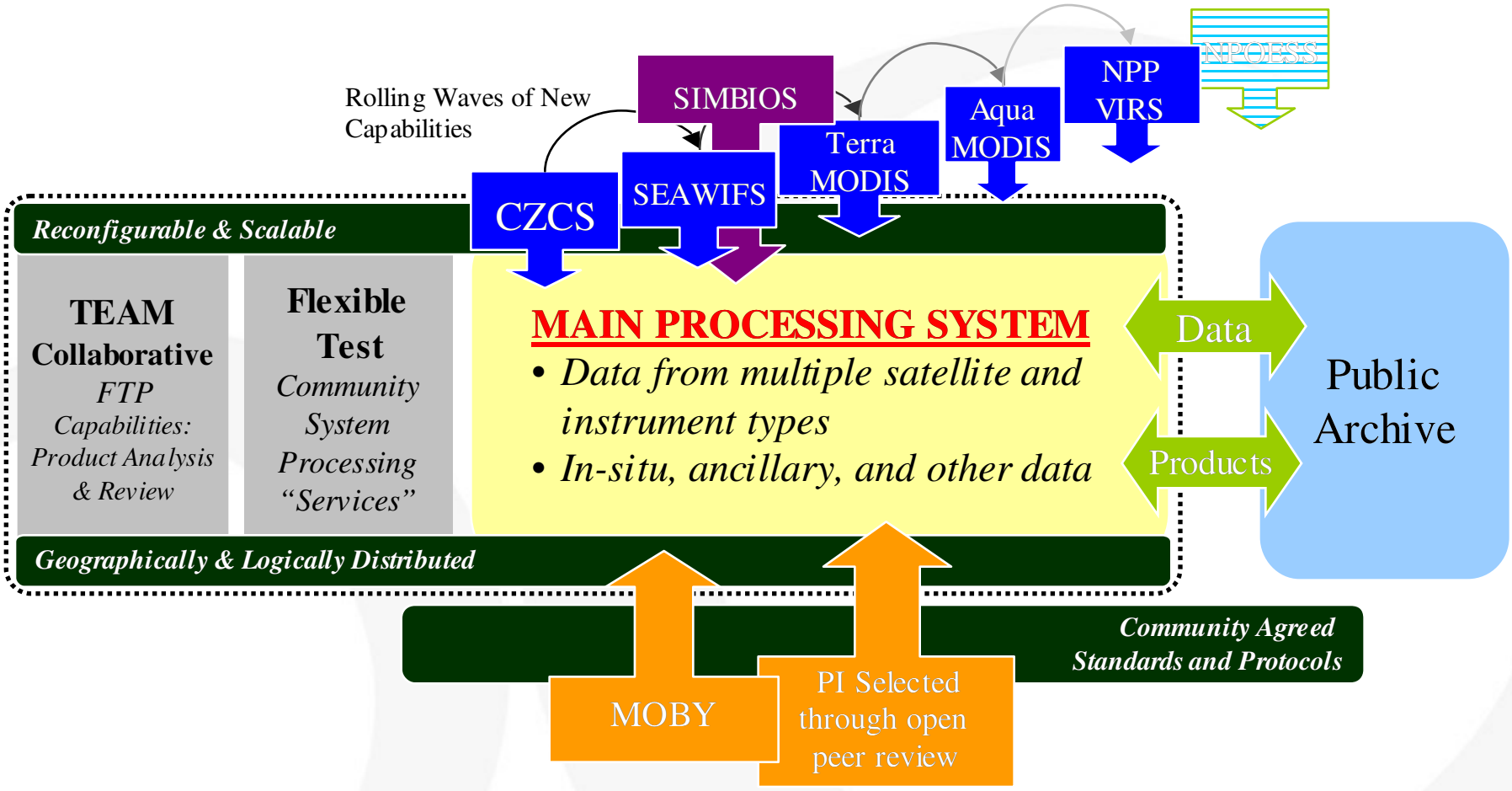
Evolution of EOSDIS Elements Study

- Charter signed by Ghassem Asrar June 04, 2004
  - ESE is moving from mission-based data systems to those that focus on Earth science measurements.
  - ESE's DIS will be a resource for Earth science-focused communities enabling research, and will be flexible, scalable and suited for the particular community needs.
  - Continue on the pathways for acquiring observations to understand processes and develop Earth system models.
- External Study Team
  - Chair, Moshe Pniel/Cal Tech
- Internal Technical Working Group
  - Lead, Glenn Iona/GSFC
- Study should by done in FY05

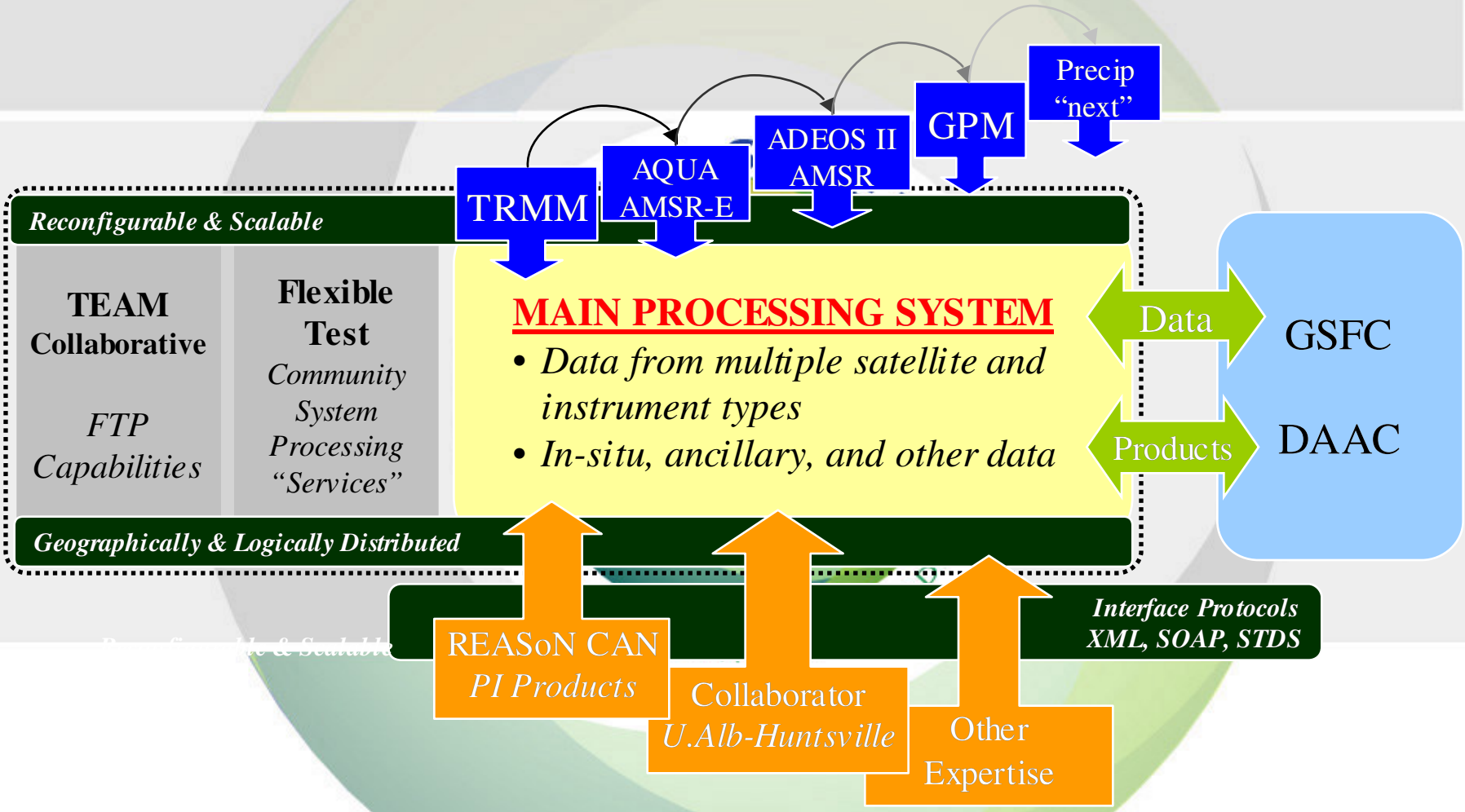
Evolution of EOSDIS

- Primary goal is evolution to meet the future ESE objectives and priorities
  - *Decompose* into functional elements
  - *Consider alternatives* to move towards a more distributed, heterogeneous data and information environment with a fully interoperable architecture
  - *Develop element options* and *action plan*
  - *Emphasize* science value and cost control
- Additional Goals
  - Increase life-cycle cost effectiveness
  - Increase end-to-end data and data system efficiency
  - Improve support for data utilization by end users
- Consider: ESE Prototype Measurement Systems, REASoN, Earth Science Working groups, ESTO, IT activities

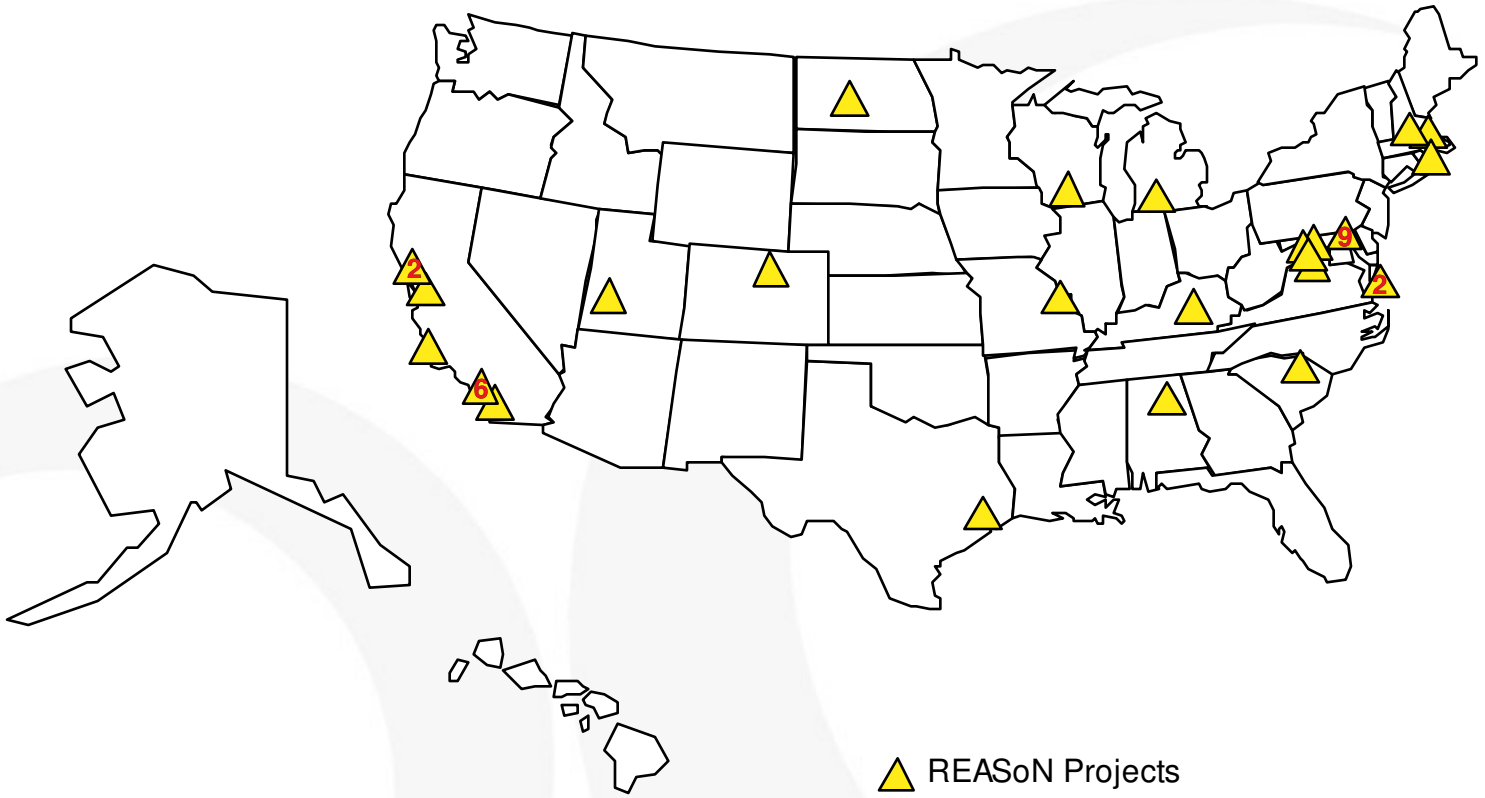
The Ocean Color Processing System



The Precipitation Processing System



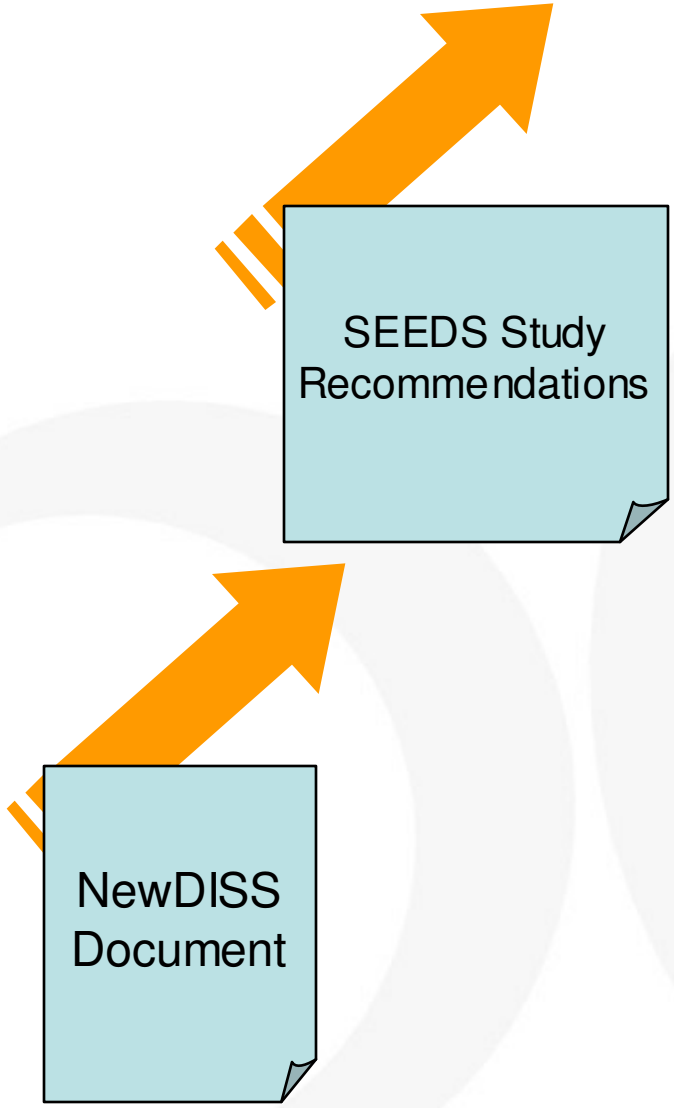
REASoNs - Distributed and Heterogeneous



42 projects projects producing data and information and/or services competitively selected through the Research, Education and Applications, Solutions Network Cooperative Agreement Notice (REASoN CAN) for development of next-generation architectures.

Processes to Manage Next Generation Distributed Data Systems

Earth Science Data Systems Working Groups (DSWGs)



- **Technology Infusion**
  - Co chairs: Karen Moe, ESTO, Rob Raskin, JPL
- **Metrics Planning and Reporting**
  - Co chairs: H. K. Ramapriyan, ESDIS, Paul Davis, University of Maryland
- **Standards and Interfaces**
  - Co chairs: Richard Ullman, ESDIS, Ming Tsou, San Diego State University
- **Reuse**
  - Co Chairs: Robert Wolfe, Raytheon, Steve Ackerman, University of Wisconsin
- **Data Life Cycle - * New in FY05**

- Kick-off meeting held January 8, 2004 in Orlando, FL
- Second Joint Working Group meeting will be held Oct 18-19 at Greenbelt Marriot
- Meeting open to all (as is done in IETF). REASoNs represented.

Ocean Color (OC) Data Stewardship at the GSFC Earth Sciences (GES) Data and Information Services Center (DISC)

featuring existing on-line access, tape backup, data analysis tools, and full user services

OCDPS: Selected MODIS Oceans Color

→ = data flow

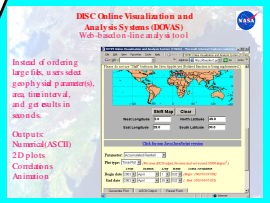
OC REASoN CAN (Gregg): CZCS, OCTS, SeaWiFS seamless time series (Produced at OCDPS; Archive/Distribution at GES DISC)

GES DISC Ocean Color Panorama

Providing Institutional Data Stewardship

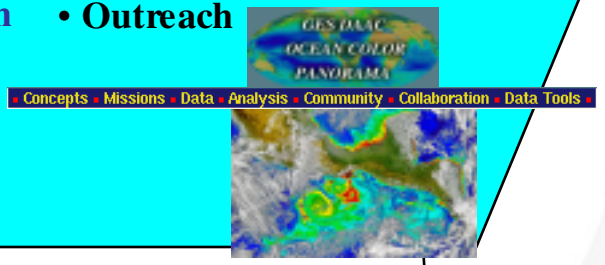
On-line: OC REASoN CAN data, V0 OC data, MODIS OC data

Architecture:
- Low Cost on-line archive
- On-line visualization and analysis tools
- Direct data access



Data Services:
- Value added data services & tools
- Insertion of new technology to facilitate data management
- Continued data services evolution based on community needs
- Advances services that promote interoperability and interdisciplanary studies

User Services:
- Supports data access and usage by scientists, modelers, decision support systems, applications, students, international users
- Outreach



~FY04 | ~FY05 | ~FY06 | ~FY07

GES DAAC V0 System: CZCS, OCTS, SeaWiFS

GES DAAC: Tape Backup; No User Services (Uses existing reliable archive; Minimal overhead)

CZCS, OCTS, SeaWiFS seamless time series, heritage CZCS, OCTS, SeaWiFS, MODIS Oceans (from OCDPS), all MODIS Oceans data

Ocean Color & Precipitation Processing System: Measurement System Findings

- Priorities set by ESE Research Science Program Managers in consultation with the funded measurement team & science community
- Requirements are science-driven based on the specific measurement needs and resource constraints
- Measurement team has science and system expertise to perform periodic assessments and trade-offs
- Measurement focus allows science research with seamless data sets across missions
- Measurement sets evolve via community collaboration
- Flexible and dynamic framework
  - To support a dynamic suite of standard and research products
  - To allow measurement system optimization for functionality and services provided
  - To support distributed functions with science collaborators
  - Negotiated support for common services to support measurement team

Key Drivers (Preliminary)

EOSDIS

- Level 1 requirements
  - Requirements set by HQ PEs and science PMs
  - Formal requirements control process
  - Fixed standard products
- Mission focus
- Broad user community with diverse needs
- Infrastructure framework
  - Archives & stewardship
  - Networks for EOS and other missions
  - Media, electronic distribution
  - Data interoperability & access

Measurement Systems*

- Science driven requirements
  - Single science PM sets priorities
  - Local authority to make trades within resources
  - Dynamic product suite
- Measurement focus
- More cohesive community focus
  - Strong community partnerships
- Distributed functions
  - Tailored to serve primary community
  - Can leverage community inputs
  - Can rely on EOSDIS for common services

* Based on Ocean Color & PPS

Preliminary Findings: Common Services

Support for Measurements Teams

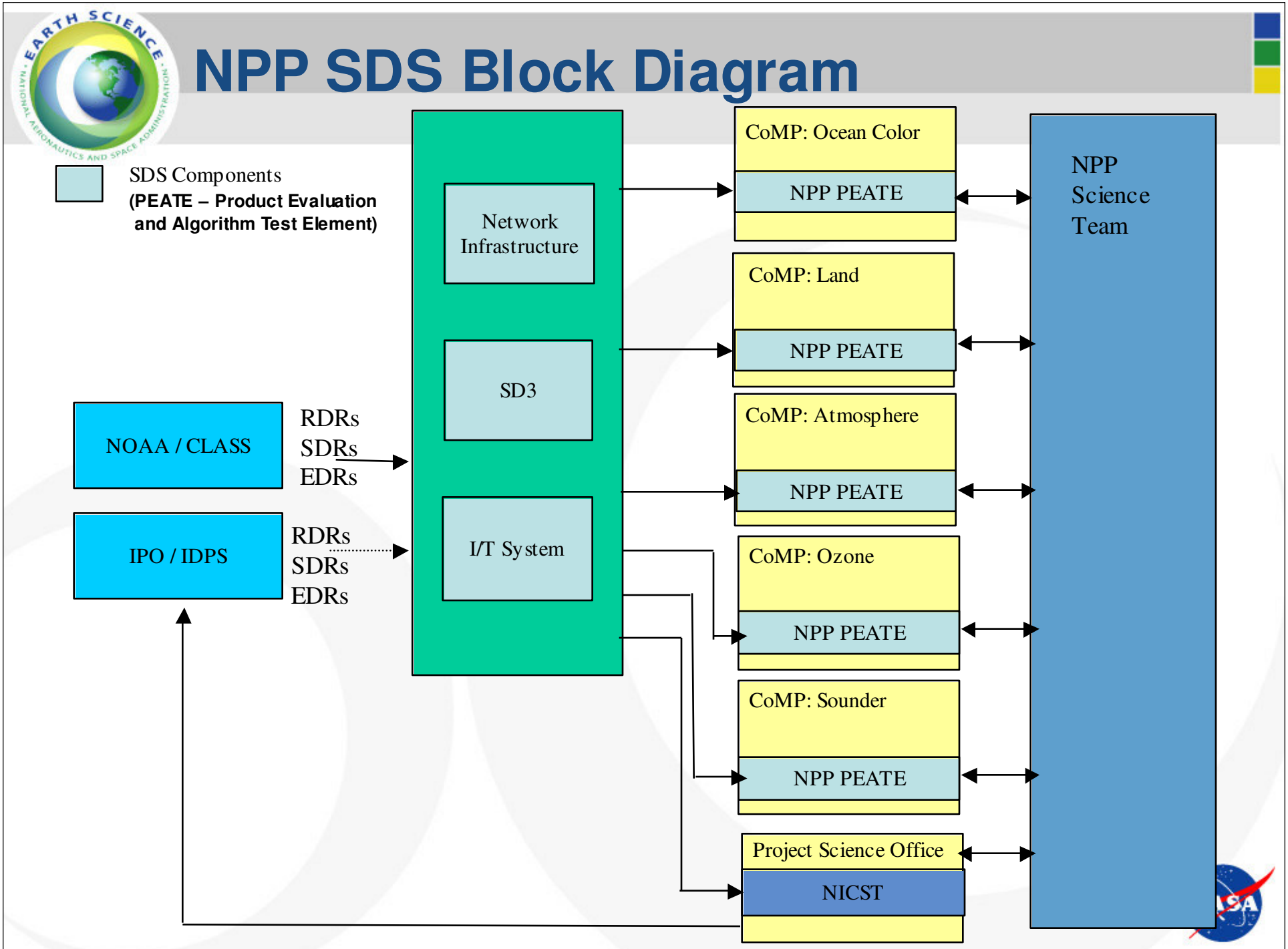
- Data Archive & Stewardship
  - Preservation of climate data records & source data
  - Metrics & data accounting *
  - Data life cycle engineering
  - Long Term Archive Transition
- Distribution
  - Media (phase out in 3-5 yrs?)
  - More products on line
  - Repackage (reduce granule size, preferred formats)
  - User Services / help desk
- Infrastructure
  - Data capture and transport
  - Communication networks
  - Enable reuse *
  - Infuse technology *

Additional support for User Community

- Search & Access
  - Cross-discipline access to climate data records
  - Enable custom web portals
  - Broker data & services (e.g. ECHO)
- Standards*
  - Facilitate definition of common spatial/temporal grids
  - Facilitate open interfaces

* Earth Science Data Systems WGs

NPP SDS Block Diagram



Data System Functions: Getting the Right Mix

- Measurement data system prototypes have established that scientific data processing and community distribution of products can be performed by the measurement teams. “One size does not fit all” for system functions performance.
- Making data available to broad user communities and interdisciplinary science may best be performed by NASA established data centers. ESE Science Focus Areas such as Carbon Cycle cross traditional disciplines.
- Continuous optimization of the distributed system so that activities can be relocated will allow for the most efficient use of scarce resources, maturation of science data, etc.
- To ensure the ‘flow’ of data through a distributed, heterogeneous data system, the interfaces between the pieces become even more critical.
- Community-based standards and protocols can inform and be informed by inter-community working groups (ES DSWG domain-specific functions).

Backup

Features of the GES DISC Ocean Color Community Services

- Data and metadata integrity ensured by consistency checking between metadata, database inventory, and on-line inventory
- Consistent version and filename control maintained
- Researcher data request and order tracking tools
- User query response tracking (User Assistance System).
- Convenient and customized (responsive to ocean color community requests) data access and ordering interfaces.
- Comprehensive and accurate mission and data documentation.
- Distribution on media (tape or CD-ROM) as a viable option to support ocean color-researchers from third-world and other technically less-advanced countries
- DAAC distribution methods, developed in collaboration and consultation with the SeaWiFS Project and NASA HQ, insures maximum, multi-tiered, distribution capability.
- Reprocessings conducted in a manner that minimized user confusion and reduces potential "mixture" of data versions, insuring data integrity and valid scientific results.
- For all archived data sets, the GES DISC Data Support Teams have a tradition of surmounting existing limitations to get data to users who need it.

Evolution of EOSDIS Elements Study Team Members

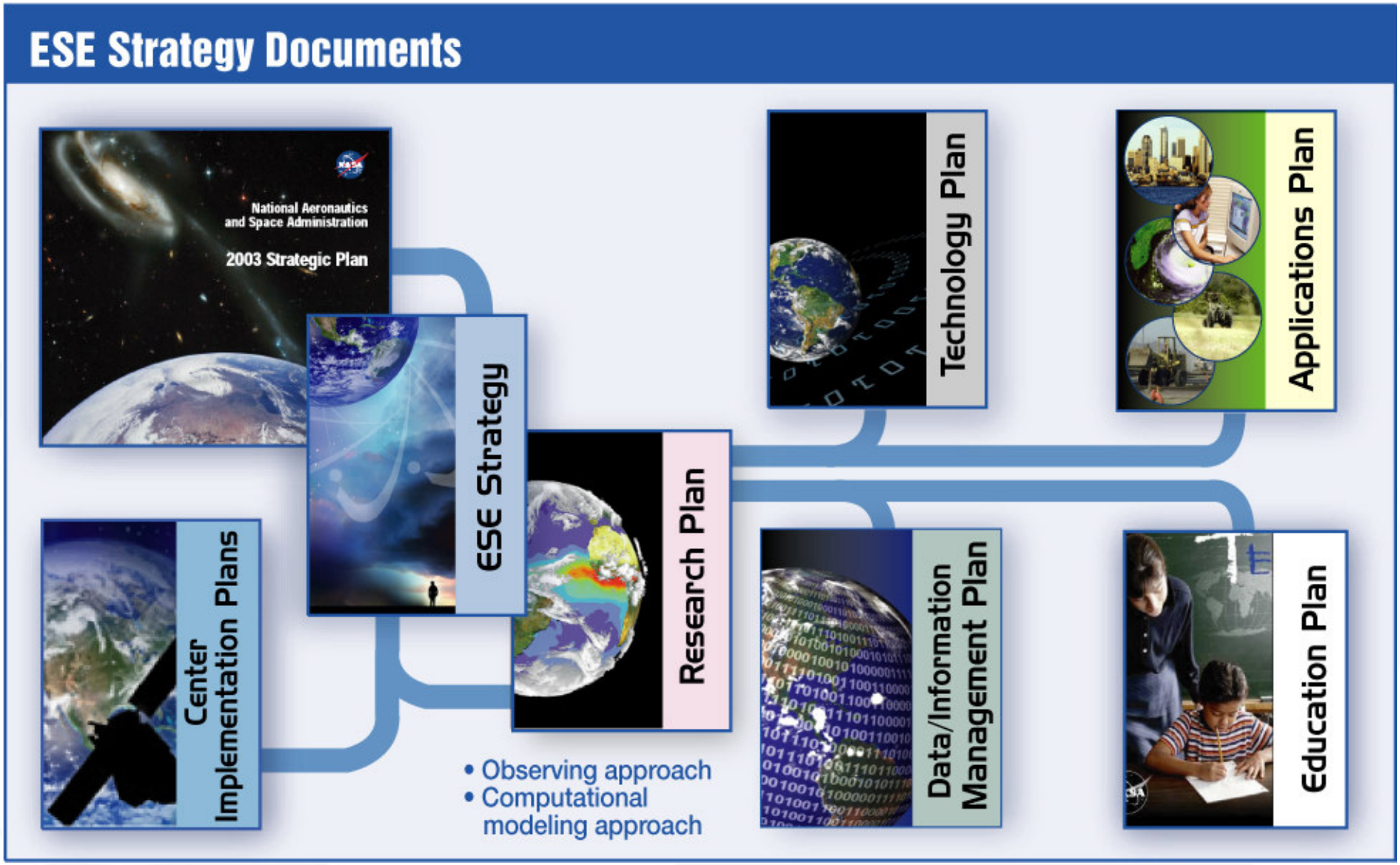
Technical Working Group Members

Glenn Iona/GSFC – Lead
Karen Moe/GSFC ESTO – Co-Lead
Michelle Ferebee/LaRC
Tom Fouser/JPL
Michael Goodman/MSFC
Vanessa Griffin/GSFC
Dawn Lowe/GSFC
Ed Masuoka/GSFC
Robin Pfister/GSFC
Rama Ramapriyan/GSFC
Bruce Vollmer/GSFC

Study Team Members

Moshe Pniel, Chair - Cal Tech
Walt Brooks/ARC
Peter Cornillon/URI
Scott Denning/U Colo
Jim Frew/UCSB
William Green/retired, JPL
Bernard Minster/UCSD
Ex Officio:
Mary Ann Esfandiari/GSFC EOS Program, ESDIS Project
Martha Maiden/Headquarters PE for Data Systems

ESE Strategy Documents



EOSDIS Science Operations

Science Operations components include:

- *Eight Distributed Active Archive Centers*
- *Global Change Master Directory*
- *EOS Data Gateway and EOS Clearing House (ECHO) Operations*
- *Interfaces with Science Investigator-led Processing Systems*
- *ESE User and Usage Metrics Collection and Analysis*
- *EOSDIS Contributions to the ESE Outreach Program*
- *Integration of cross-element collaborative activities:*
- *User Services Working Group, DAAC Outreach*

The ESDIS Science Operations Office provides integration needed to accomplish the ESE data and information goals and objectives

The Ocean Color Processing System

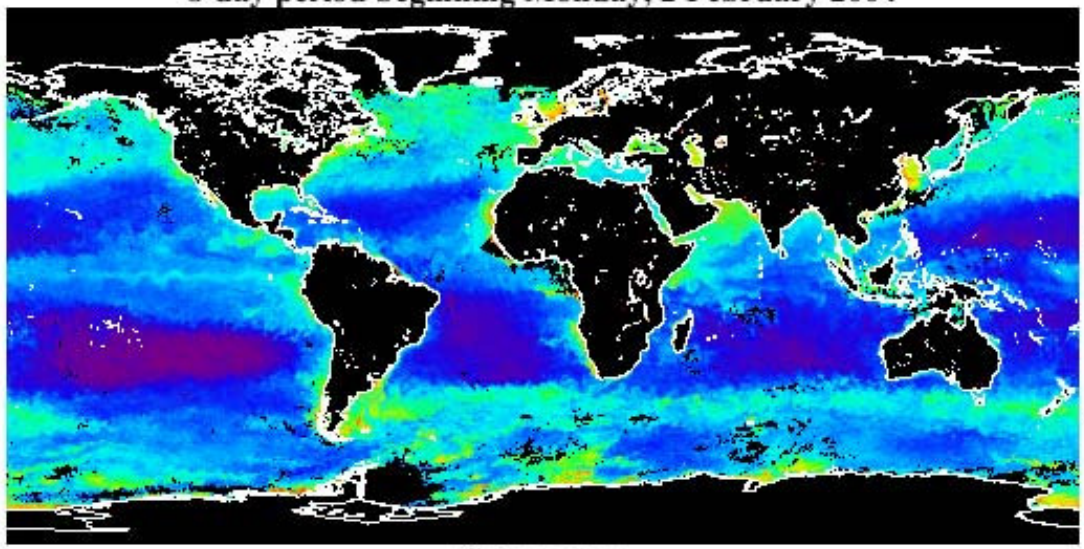
SeaWiFS
GAC LAC HRPT MLAC
MODIS (Aqua)

Radius (*km*) about map click or typed-in location:
36 400 800 1200 1500

Select swaths containing *(at least)*:
any part 25 % 50 % 75 % all of the area of interest.

Display results 10 at a time.

8-day period beginning Monday, 2 February 2004



Chlorophyll

Comment Help

Select one or more regions:
AdriaticSea
AegeanSea
Antarctica
ArabianSea
AralSea
Australia
Azores

or specify boundary coordinates or a single location:
N:
W: :E
S:

or check this box to select all of this time period's scenes.
Find swaths

Reconfigure page

Mission													
	2002	J	F	M	A	M	J	J	A	S	O	N	D
	2003	J	F	M	A	M	J	J	A	S	O	N	D
	2004	J	F	M	A	M	J	J	A	S	O	N	D

December 2003

S	M	T	W	T	F	S
	1 ***	2 ***	3 XXX	4 XXX	5 XXX	6 XXX
7 XXX	8 XXX	9 XXX	10 XXX	11 ^^^	12 ^^^	13 ^^^
14 ^^^	15 ^^^	16 ^^^	17 ^^^	18 ^^^	19 000	20 000
21 000	22 000	23 000	24 000	25 000	26 000	27 ***
28 ***	29 ***	30 ***	31 ***			

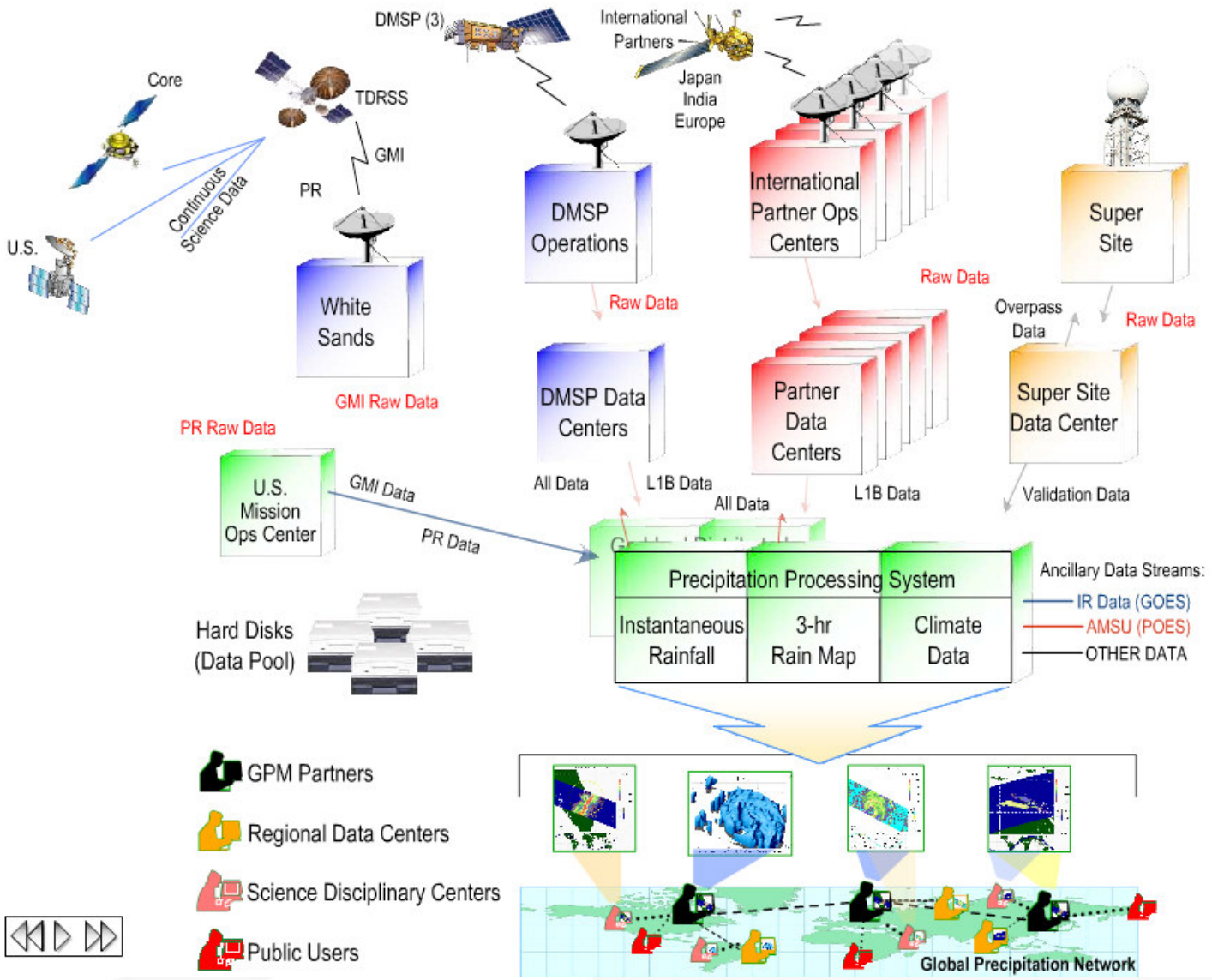
January 2004

S	M	T	W	T	F	S
				1 000	2 000	3 000
4 000	5 000	6 000	7 000	8 000	9 ***	10 ***
11 ***	12 ***	13 ***	14 ***	15 ***	16 ***	17 XXX
18 XXX	19 XXX	20 XXX	21 XXX	22 XXX	23 XXX	24 XXX
25 ^^^	26 ^^^	27 ^^^	28 ^^^	29 ^^^	30 ^^^	31 ^^^

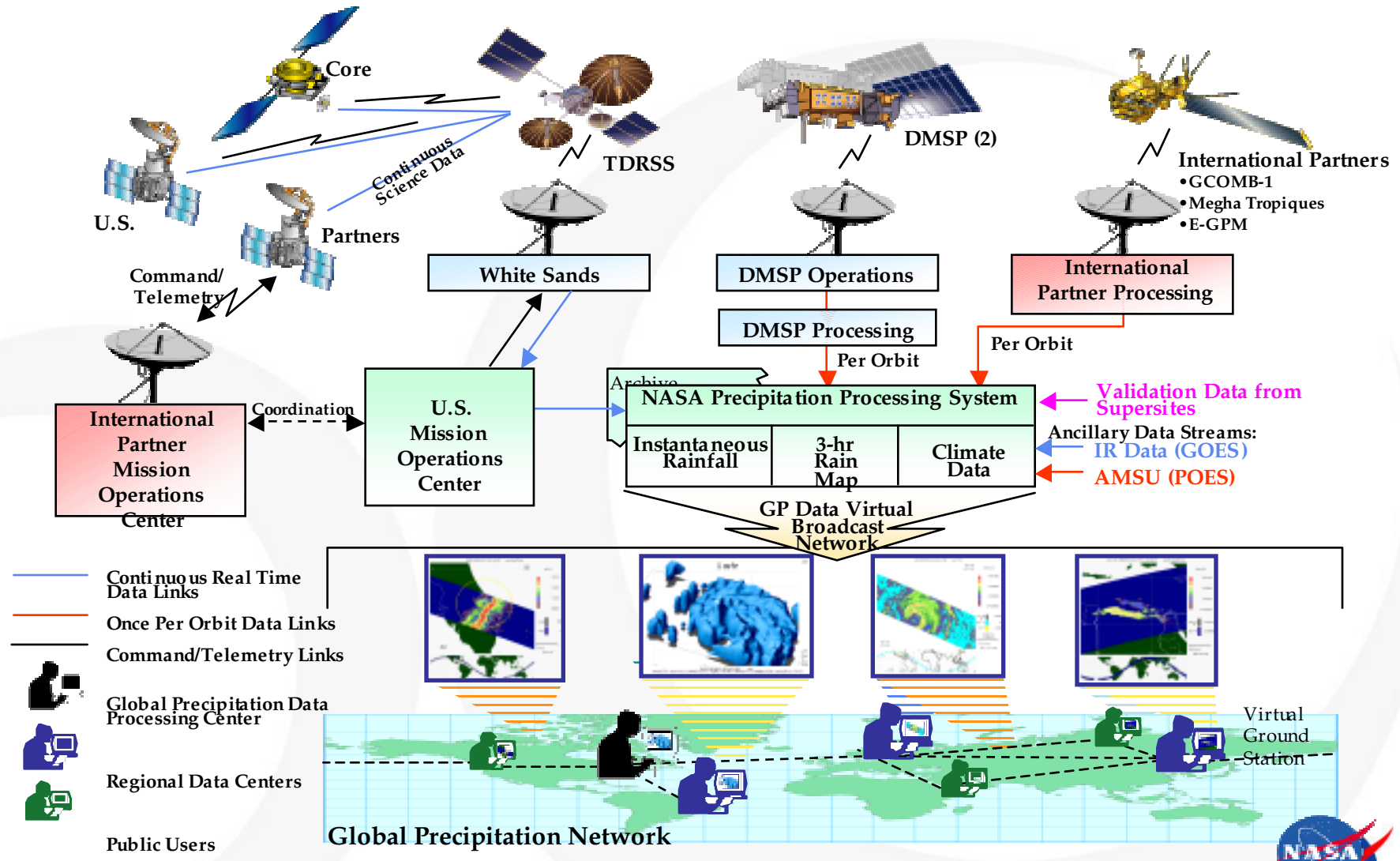
February 2004

S	M	T	W	T	F	S
1 ^^^	2 000	3 000	4 000	5 000	6 000	7 000
8 000	9 000	10 ***	11 ***	12 ***	13 ***	14 ***
15 ***	16 ***	17 ***	18 XXX	19 XXX	20 XXX	21 XXX
22 XXX	23 XXX	24 XXX	25 XXX	26 ^^^	27 ^^^	28 ^^^
29 ^^^						

Precipitation Processing System



Precipitation Processing System Depicted for Global Precipitation Mission era



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	For the National Oceanic and Atmospheric Administration
_____	_____
Dr. Ghassem Asrar Associate Administrator, Office of Earth Science	Mr. Gregory W. Withee Assistant Administrator for Satellite and Information Services

"Statement of Intent"
signed July 3, 2001