EOSDIS Evolution Plan
November 17, 2005

Mary Ann Esfandiari
Deputy Program Director for EOS Operations
ESDIS Project Manager
(301) 614-5048
Mary.A.Esfandiari@nasa.gov

Contents

- Key Points
- EOSDIS Evolution's 1st Step
  - MODAPS
  - GES DAAC
  - LaRC DAAC
  - ECS/SDPS
- Benefits
- Next steps

Key Points (1 of 3)

- ESDIS embarking on an EOSDIS Evolution process
  - Began in Feb 05 with goals set by HQ appointed Study Team
  - Plan based on informal responses received from EOSDIS Elements (DAACs, SIPS, ECS, Middleware)
- Changes designed to
  - Improve efficiency and introduce more autonomy, agility and scalability through infusion of newer commodity based hardware
  - Move control over processing, archive and distribution for specific science data to science teams
  - Reduce complexity of existing system through development of several DAAC-unique systems
  - Reduce annual operational costs by 15-25% within 3 years
- Changes will happen gradually over next 2-3 years and are planned to
  - Reduce risk associated with operational changes
  - Have "proof of value" periods before taking next steps
  - Continue full operations while evolving

Key Points (2 of 3)

- Lots of work to do
  - Some new interfaces
  - Some new development, integration and testing
  - Some new data flow paths
  - Some new tools

- New changes and revitalization will pave the way to the future
  - Reductions in operational costs provide available funds for new Earth Science applications, new missions, potential new Earth Science Data Records for decadal studies and applications to support measurement-based data
  - Improvements in data processing, archive and distribution can result in faster science findings and available data to serve global change studies

Key Points (3 of 3)

- **Evolution process will require more teaming, partnerships, and positive, pro-active relationships among all of us**
  - DAACs
  - SIPSs
  - ECS/EMD
  - ECHO and GCMD
  - HQ Science Mission Directorate
  - EOS Program Office

- **My office door is always open**
  - Comments, suggestions, questions, and ideas

MODAPS

Transfer responsibility for MODIS processing, archive and distribution from GES DAAC to MODAPS

Initial Step: Atmosphere data

Features:

- Processing on demand for Level 1
- All commodity-based for processing, on-line data storage and access
- Based on extension to existing MODAPS science team production systems

Benefits:

- Reduction in archive growth through on-demand processing
- Faster access to products, reduced reprocessing time from all on-line storage
- Reduced costs due to use of commodity disks, reduction in operations at DAAC
  - Resulting changes to GES DAAC include: 10% fewer products; 90% reduction in total volume and similar reductions in archive growth/day
- Closer involvement and control by science community

GES DAAC

Consolidate GES DAAC data holdings into one DAAC-unique system (S4PA)

Features:

– Transition of AIRS, HIRDLS, OMI, MLS and V0/V1 managed heritage data sets
– Phase out of ECS in FY08 timeframe
– Based on existing on-going effort (development already in progress)
– Reduction of archive volume (due to transfer of MODIS data)

Benefits:

– Reduction in operations costs due to elimination of multiple systems
– Reduction in sustaining engineering costs due to use of simpler, scalable software and reduction in dependency on COTS products
– Increased system automation due to single system, simpler operational scenarios
– Improved data access due to planned use of increased on-line storage and commodity disks/platforms
– Phased elimination provides risk mitigation for MODAPS effort

LaRC DAAC

Consolidate LaRC DAAC data holdings into one DAAC-unique system (ANGe)

Features:

- Initial Step: Transition CERES data from LaTIS heritage system to ANGe
- Based on existing on-going effort (development already in progress)

Benefits:

- Reduction in operations costs due to elimination of multiple systems
- Reduction in sustaining engineering costs due to reduction in dependency on COTS products
- Increased system automation due to single system, simpler operational scenarios
- Improved data access due to planned use of increased on-line storage and commodity disks/platforms (from ECS simplification plan)

Future Steps (contingent on success of Initial Step) include transition of MISR, MOPITT, TES and Phase out of ECS in FY08 timeframe

ECS/SDPS

Rearchitect ECS to simplify sustaining engineering and automate operations

Features:

- Simplify software architecture (eliminate 15 components & 750K SLOC)
- Move towards disk-based archive
- Leverage new hardware technology (e.g., commodity-based systems; shared storage) to reduce hardware maintenance costs

Benefits:

- Low risk approach based on proven data pool technology
- Increased system automation; simplified hardware/software configuration
- Reduction in operational costs at ECS DAACs
- Improved data access due to increased on-line storage and commodity disks/platforms
- Provides risk mitigation for GES DAAC and LaRC DAAC ECS phase out efforts

Plan Summary

- In EOSDIS Evolution 1st Step, four EOSDIS elements are affected:
  - GES DAAC
  - LaRC DAAC
  - MODAPS
  - ECS/SDPS
- Minimal or no impact to other elements at this time
  - Some interface changes
- Evolution will a continuing process
  - Continue to work with HQ Study team to plan future steps
- Today’s presentation is focused on Step 1
  - Charts available at : http://eosdis-evolution.gsfc.nasa.gov/

Key Benefits of Evolution Plan

- **Maximize Science Value**
  - Data access easier and data products quickly available to science community
  - MODIS data more closely integrated with science community
  - Pathfinder for migration of other data into science communities
- **Manageable Risk**
  - Minimizes software development efforts
  - Builds upon existing systems – no “blank paper” new development
  - Utilizes steps within plan as proof of value before proceeding
  - GES DAAC provides risk mitigation for MODAPS development effort
  - Reduced footprint for ECS provides fallback for GES and LaRC development efforts, support for NSIDC and LP DAAC, commodity hardware for DAAC-unique solutions
- **Substantial Cost Savings**
  - Expected savings of 10-25% annually beginning in 3rd year

Next Steps

- **Begin detailed work planning with affected elements**
  - **establish baseline, develop implementation plans, detailed integrated schedule**
  - **Reporting mechanism to track progress**

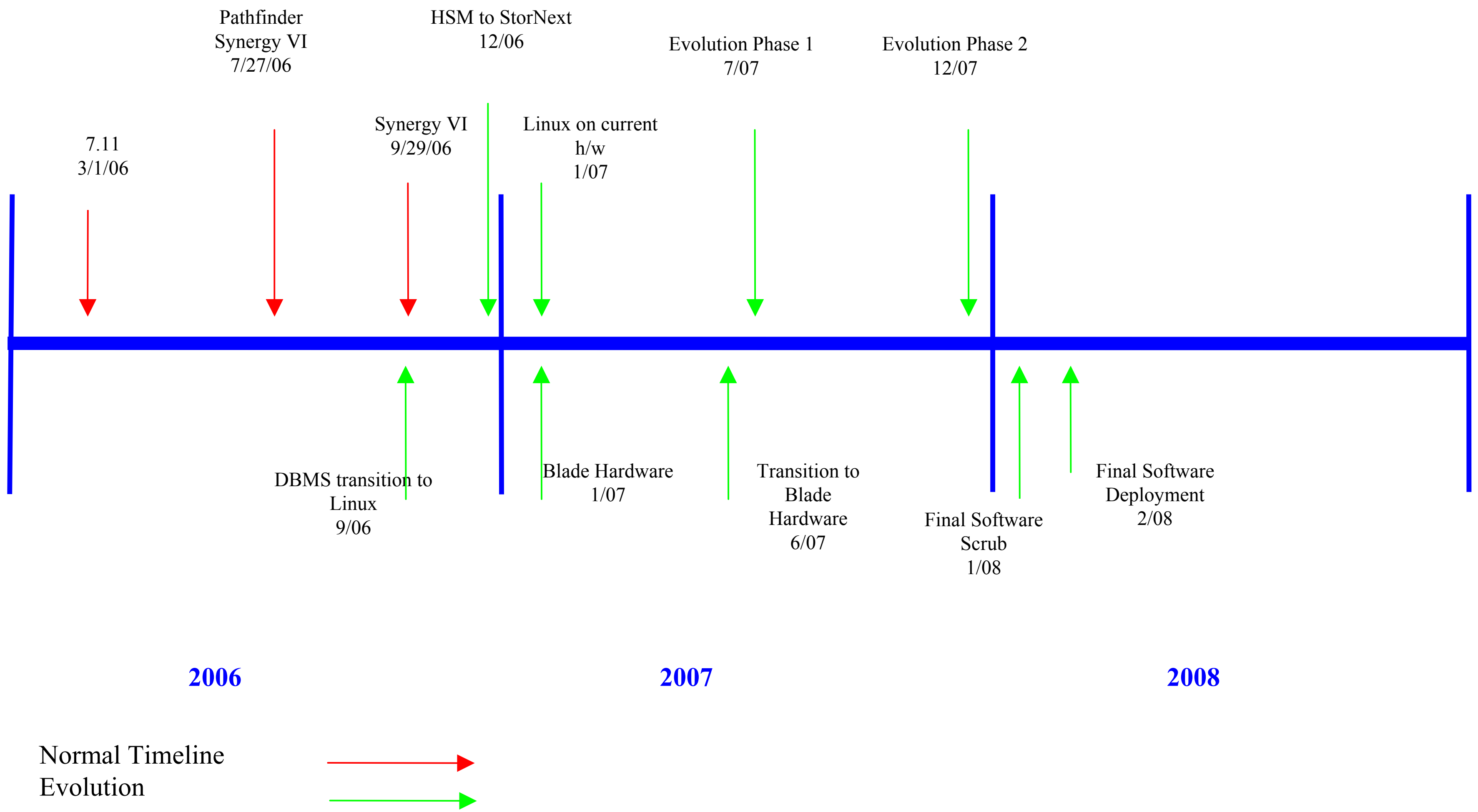
- **Address ESDIS Project level concerns:**
  - **Ensure smooth transition for users**
  - **Minimize disruption at EOSDIS elements**
  - **Coordinate interfaces and other inter-element issues**
  - **Ensure sustaining engineering for areas that are not covered in new changes (i.e. networks, configuration management, ICDs, etc.)**

- **Team work and the best efforts from all of us will maximize our ability to achieve success**

LP DAAC Impacts

- Phase 1 impacts are solely related to the ECS systems changes
- ESDIS delivered Task Proposal Request to EMD on January 27th, 2006
- EMD has 25 days to respond to the request
- When the task is approved, LP DAAC will work with NSIDC, ESDIS, and EMD on the new system
  - Support automation study
  - Evaluate DAAC Unique Extensions (DUEs)
- Final delivery is expected during the 2nd quarter of FY08

Deliveries Timeline



BACKUP

EOSDIS Evolution 2015 Vision Tenets

Vision Tenet	Vision 2015 Goals
Archive Management	▪ NASA will ensure safe stewardship of the data through its lifetime. ▪ The EOS archive holdings are regularly peer reviewed for scientific merit.
EOS Data Interoperability	▪ Multiple data and metadata streams can be seamlessly combined. ▪ Research and value added communities use EOS data interoperably with other relevant data and systems. ▪ Processing and data are mobile.
Future Data Access and Processing	▪ Data access latency is no longer an impediment. ▪ Physical location of data storage is irrelevant. ▪ Finding data is based on common search engines. ▪ Services invoked by machine-machine interfaces. ▪ Custom processing provides only the data needed, the way needed. ▪ Open interfaces and best practice standard protocols universally employed.
Data Pedigree	▪ Mechanisms to collect and preserve the pedigree of derived data products are readily available.
Cost Control	▪ Data systems evolve into components that allow a fine-grained control over cost drivers.
User Community Support	▪ Expert knowledge is readily accessible to enable researchers to understand and use the data. ▪ Community feedback directly to those responsible for a given system element.
IT Currency	▪ Access to all EOS data through services at least as rich as any contemporary science information system.

EEE Study Team Members

Name	Organization	Title
Moshe Pniel, Chair	NASA JPL	Manager, Astrophysics Formulation Program Office
Martha Maiden*	NASA HQ	Program Executive for Data Systems, Science Mission Directorate
Mary A. Esfandiari*	NASA GSFC	Deputy Program Director for EOS Operations & ESDIS Project Manager
Dr. Walt Brooks	NASA ARC	Chief, Advanced Supercomputing Division of the Information Sciences and Technology Dir.
Dr. Peter Cornillon	University of Rhode Island	Professor of Oceanography
Dr. Scott Denning	Colo. St. Univ	Assistant Professor, Atmospheric Science
Dr. Jim Frew	UC - Santa Barbara	Associate Professor, School of Environmental Science and Management
William Green	Retired (Cal Tech)	Manager of Infrared Processing and Analysis Center and Spitzer Science Center
Dr. Bernard Minster	UC - San Diego	Professor of Geophysics, Scripps Institution of Oceanography

* *Ex Officio*

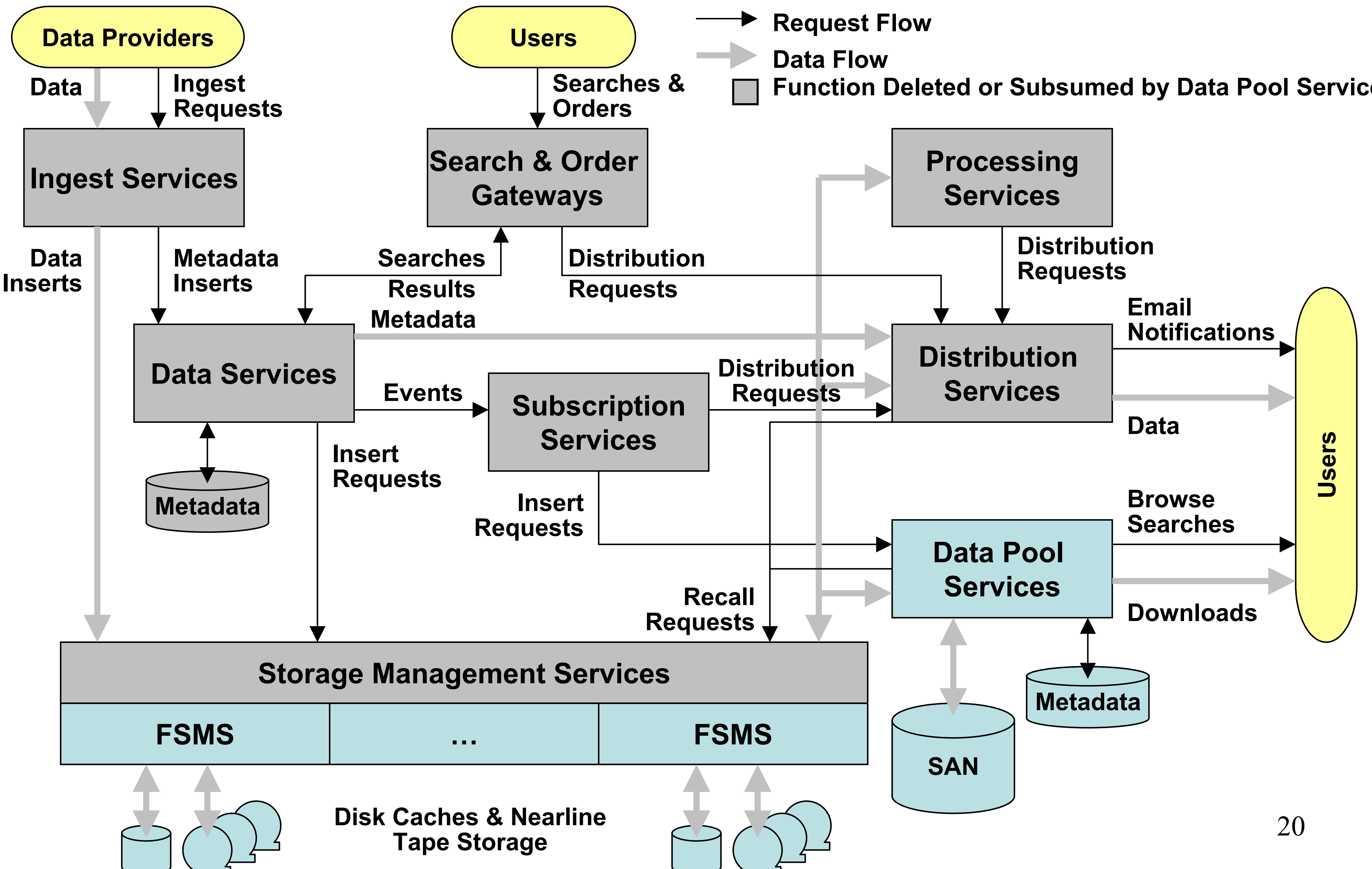
EEE Technical Team Members

Name	Organization	Title
Mary A. Esfandiari, Lead	NASA GSFC	ESDIS Project Manager & Deputy Program Director for EOS Operations
Dawn Lowe, Implementation Lead	NASA GSFC	ESDIS Deputy Project Manager
Jeanne Behnke	NASA GSFC	Acting, ESDIS Science Operations Office Manager
Chris Bock	NASA GSFC	Hardware Systems Manager, ESDIS SSDO
Michelle Ferebee	NASA LaRC	Langley Assistant DAAC Manager
Kathy Fontaine	NASA GSFC	Global Change Data Center
Michael Goodman	NASA MSFC	Earth and Planetary Science Branch
Pat Liggett	NASA JPL	Manager, Physical Oceanography DAAC
Ken McDonald	NASA GSFC	External Development & Interfaces, Long Term Archive
Ed Masuoka	NASA GSFC	MODAPS SIPS Lead
Dan Marinelli	NASA GSFC	Deputy Manager, ESDIS SSDO
Karen Moe	NASA GSFC	Earth-Sun System Technology Office
Robin Pfister	NASA GSFC	ECHO, Middleware Lead, ESDIS Project
Dr. H. Ramapriyan	NASA GSFC	Assistant Project Manager, ESDIS
Dr. Skip Reber	NASA GSFC	ESDIS Project Scientist
Curt Schroeder	NASA GSFC	EMD COTR
Ed Sofinowski	SGT, Inc.	Process & Technical Consultant
Dr. Scott Turner	Aerospace Corp.	Lead for Cost Assessment, Alternatives Analysis
Bruce Vollmer	NASA GSFC	GES DAAC Engineer

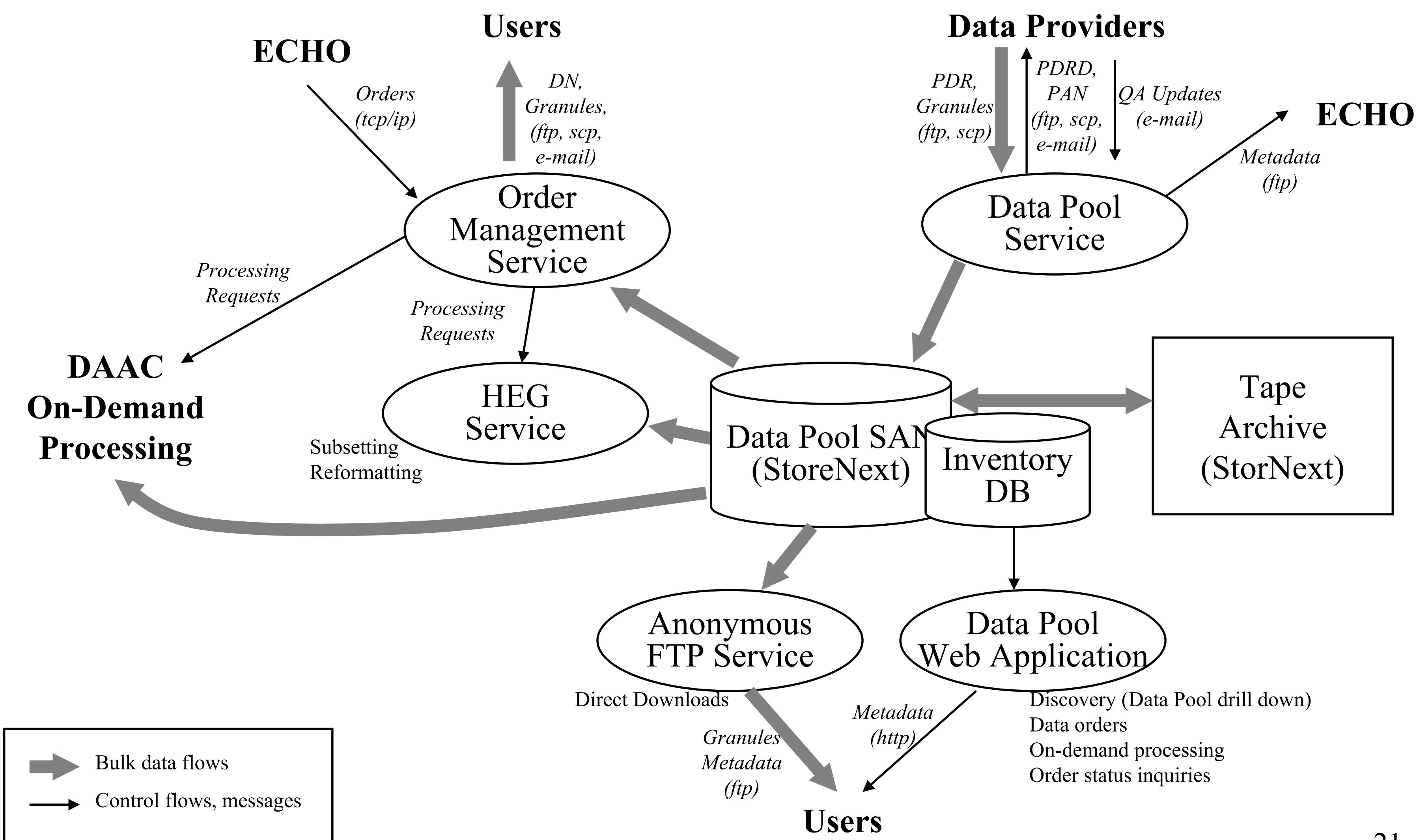
EEE Technical Team Consultants

Name	Organization	Title
Dr. Gene Feldman	NASA GSFC	Manager, OCDPS
Dr. Chris Lynnes	NASA GSFC	GES DAAC Lead, System Engineering
Dr. Erich Stocker	NASA GSFC	Manager, TSDIS/Precipitation Processing System
Dr. Victor Zlotnicki	NASA JPL	Program Manager for Climate Variability

SDPS Software Evolution Roadmap



Software Changes
Target Software Architecture



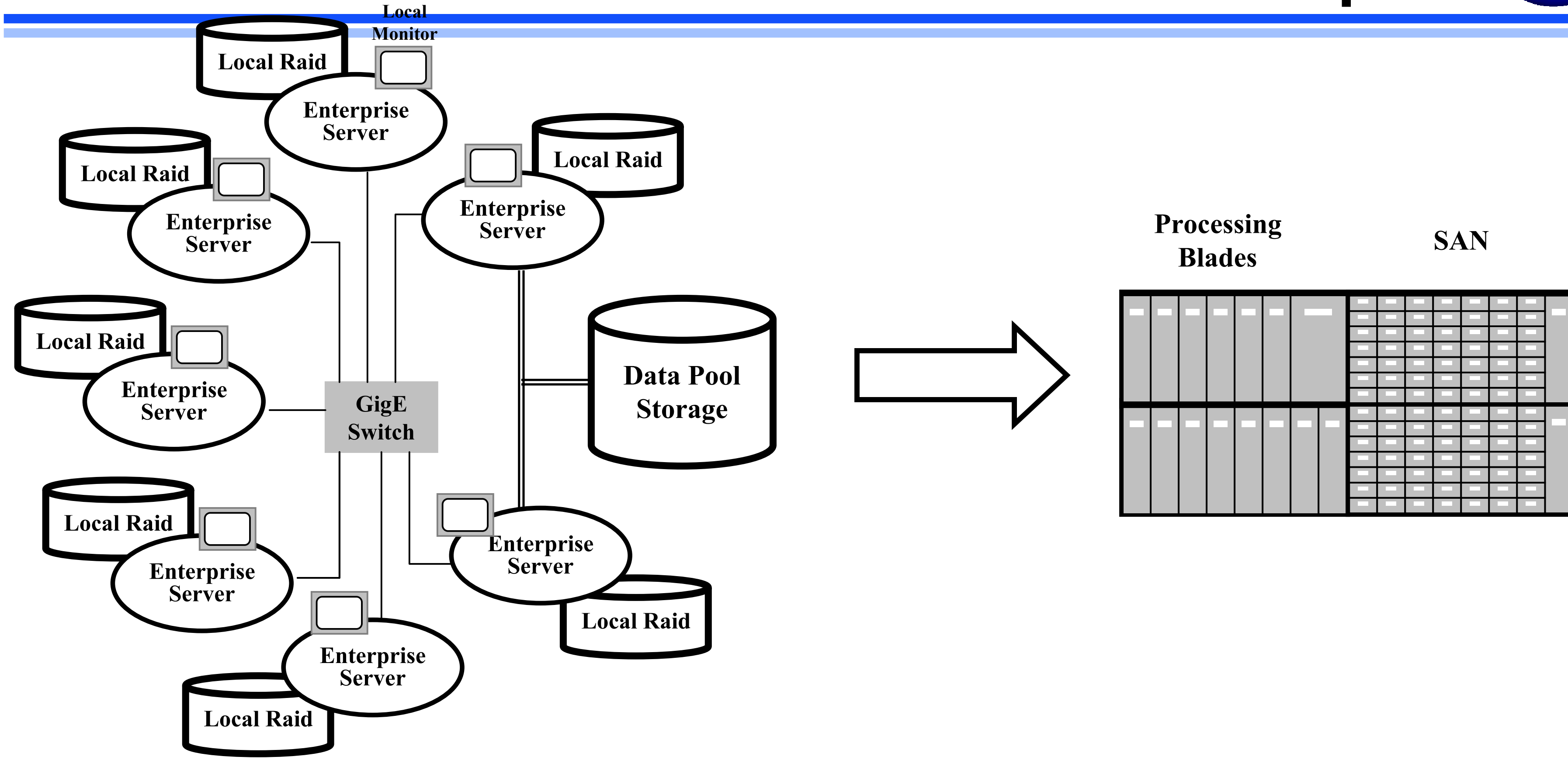
Software Changes Technical Approach

- S4P processing system is used for all processing at GES, LP, and ASDC DAACs, replacing the SDPS PDPS processing system
  - SCLI or external subsetter interface used to request processing inputs
  - SIPS interface used to return processing outputs for archiving or distribution
- ECHO provides user search, browse, and order submission functions enabling these functions to be removed from SDPS
  - Data Pool drill down search interface would be retained
- Machine-to-Machine Gateway phased out (provided through ECHO)
- ECHO clients fully replace EDG
  - External users search ECHO for collections and granules
  - External users perform browse via ECHO, including browsing MISR and AIRS collections
  - External users submit data orders via ECHO
  - DORRAN order approval at LP DAAC will occur outside of SDPS
  - V0 Gateway is not longer supported. ECHO will submit orders via a TBD interface (e.g., XML/SOAP)
  - SDPS will no longer support
    - Returning packaging and processing options
    - Returning cost estimates
    - Exporting lists of valid values for science metadata in ODL (export via XML will continue)
    - Performing access control checks for user orders

Software Changes
Technical Approach

- Media ingest and ingest via TCP/IP (only used by Landsat) no longer supported
- Distribution, ingest and archiving moved to the Data Pool service
- Support for media ingest is dropped (not currently used)
- Additional automation is added (based on joint DAAC-EMD analysis)
- Results
  - Current custom code baseline is 1,260,000 lines of C++/Java code
  - Resulting custom code baseline is 400,000 lines of C++/Java code
  - About 750,000 lines of existing (complex) C++ code is removed
  - 63,500 lines of new code must be developed
  - 32,500 lines of existing code must be reworked
  - Elapsed time to implement ~24 months

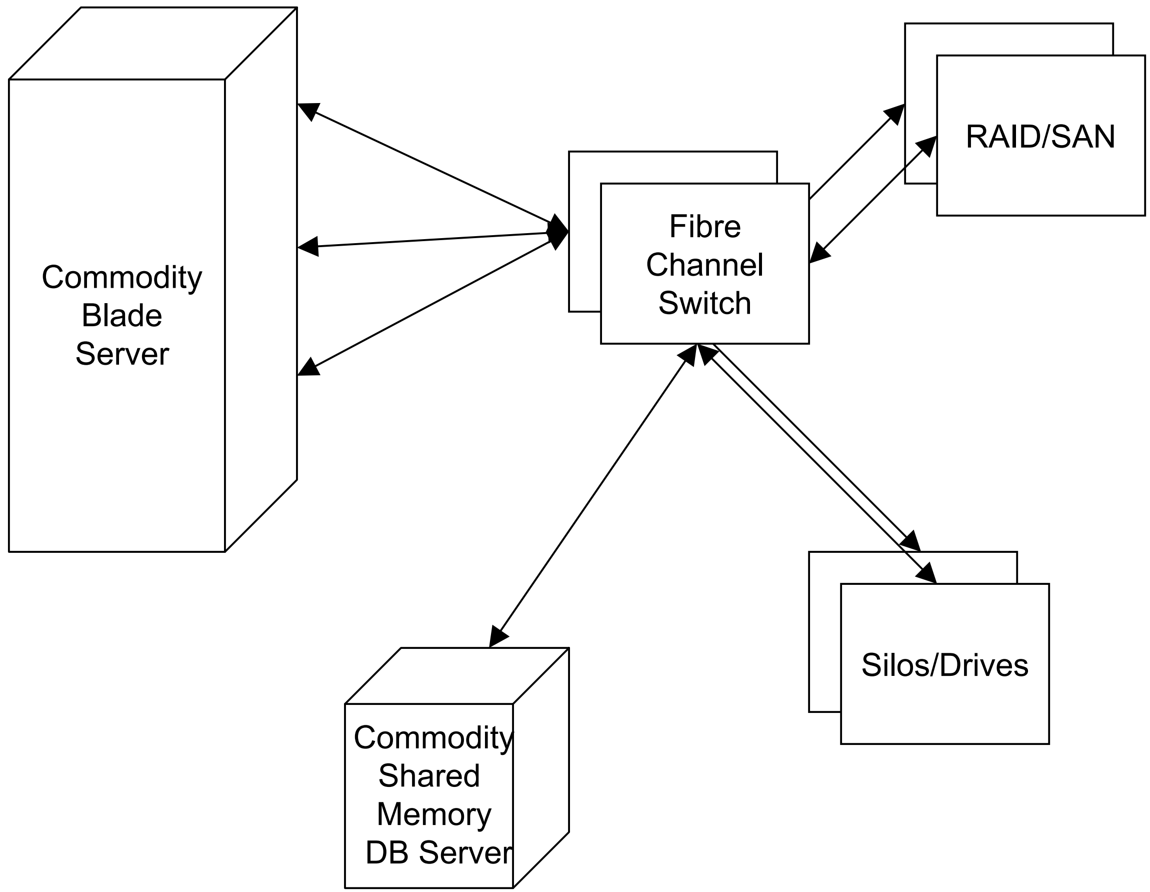
SDPS Hardware Evolution Roadmap



TODAY:
- Server platforms with local RAID
- Solaris or Irix OS
- Switched GigE Network
- Some platforms connected to SAN

FUTURE:
- Rack mounted blades
- Linux OS
- Rack mounted SAN
- Optical fiber network infrastructure
- All platforms connected to SAN

Hardware Changes
Target Hardware Architecture



Commodity Blade Server

Commodity-Based
Shared Memory
DB Server

Shared SAN with
Shared Filesystem

Hardware Changes
Technical Approach

- Replace existing SGI and Sun servers with Linux-based commodity blade server
- Consolidate DBMS servers. Replace existing SGI and Sun servers with single large-shared memory commodity server
- Transition from AMASS to StorNext Storage Manager (SNSM)
  - SNSM will run on expanded SNFS metadata server
  - SNSM can read AMASS written tapes so no data migration is needed