

NATIONAL AERONAUTICS
AND SPACE ADMINISTRATION

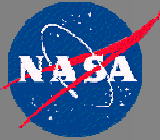
Reusing Software to Build Data Processing Systems: NPP Science Data Segment Case Study

Shahin Samadi, Ryan Gerard, Mary Hunter, James J. Marshall
(Innovim / NASA GSFC)

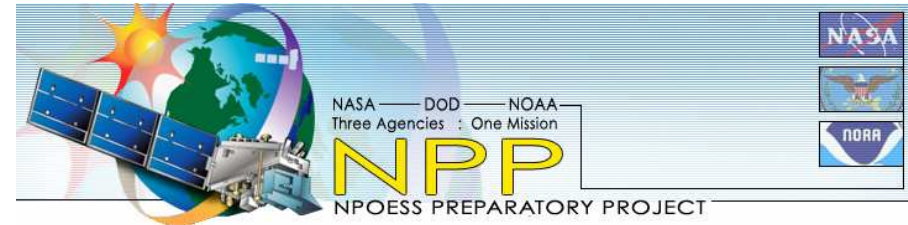
Robert J. Schweiss, Robert E. Wolfe, Edward J. Masuoka
(NASA GSFC)

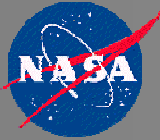
Presented by Robert E. Wolfe

2007 IEEE Aerospace Conference
March 4, 2007



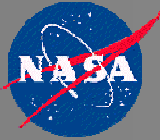
- NASA is saving resources and reducing risk in the development of the Science Data Segment (SDS) for the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP)
 - SDS reuses many components of data systems developed for NASA's Earth Observing System (EOS)
 - Part of NASA's movement from Missions to Measurements – where measurements from multiple missions (instruments) are processed to build long-term climate quality data records
 - Reuse allows scientists and developers to devote their efforts to new requirements and innovation
- Two examples of software reuse are presented
 - Includes expect cost savings



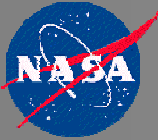


- The National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP)
 - Remotely-sensed atmosphere, land, ocean, ozone, and sounder data
 - Serves the meteorological and global climate change scientific communities
 - Risk reduction for NPOESS
- NPP will be one of the missions that will contribute to and participate in the Global Earth Observation System of Systems (GEOSS)
 - Collaboration
 - Lessons learned about reuse
 - Methods for implementing software reuse

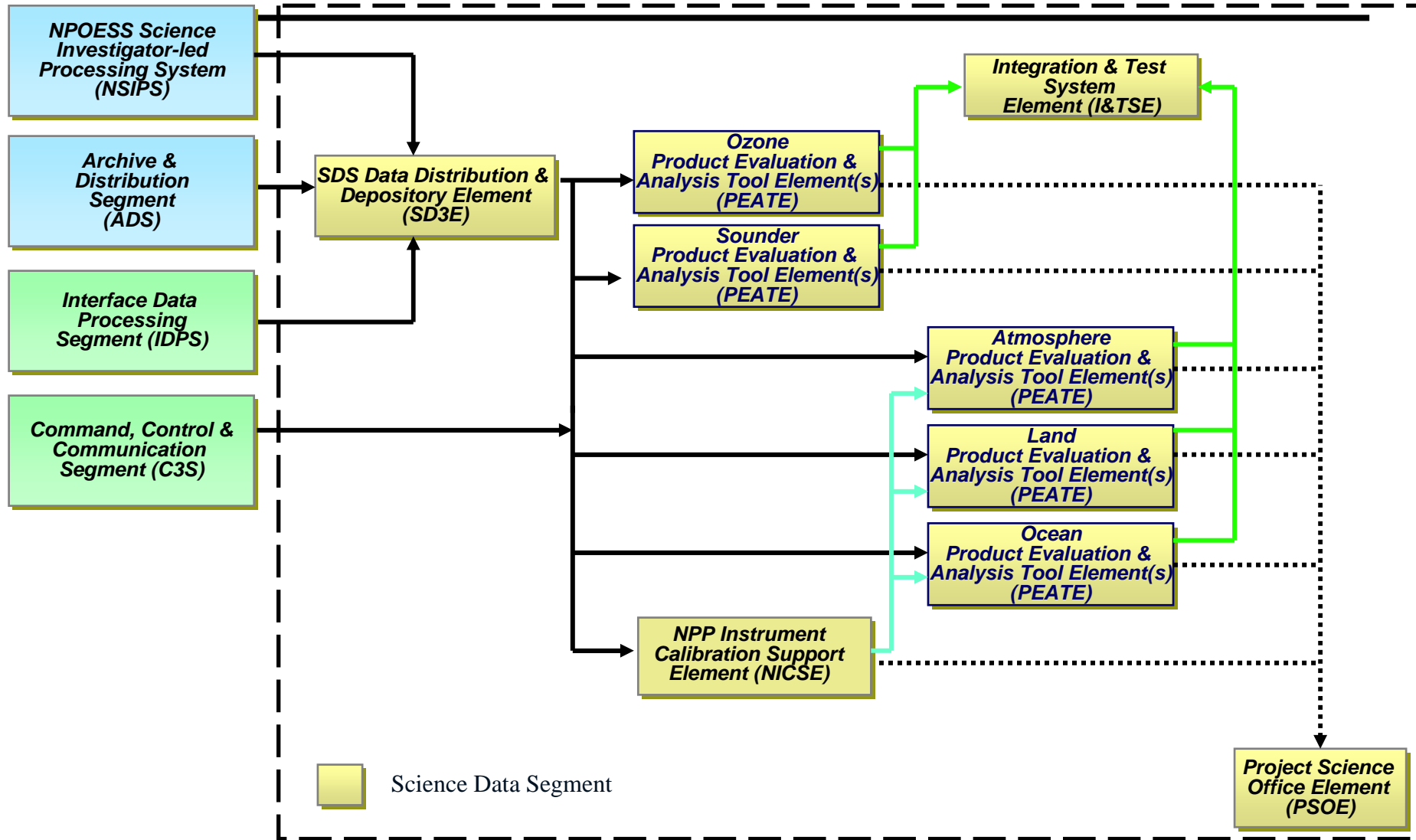


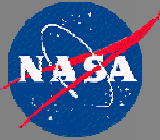


- Science Data Segment (SDS)
 - Purpose is to assess and verify the quality of NPP products to accomplish climate research
 - Receive and evaluate Raw Data Records (RDRs), Sensor Data Records (SDRs), and Environmental Data Records (EDRs)
- The SDS is composed of nine elements:
 - SDS Data Distribution and Depository Element (SD3E)
 - Integration and Test System Element (I&TSE)
 - Project Science Office Element (PSOE)
 - NPP Instrument Calibration Support Element (NICSE)
 - 5 Product Evaluation and Analysis Tools Elements (PEATEs)



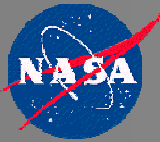
Simplified SDS diagram





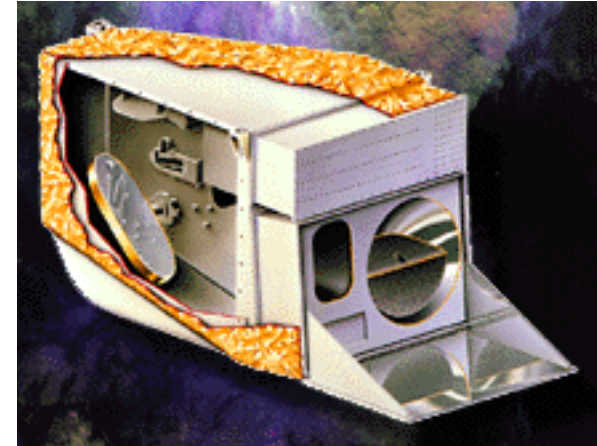
Overview of SDS Elements

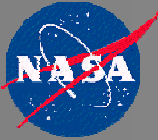
- SDS Data Distribution and Depository Element (SD3E)
 - Ingests, stages, and distributes the NPP data products
- Integration and Test System Element (I&TSE)
 - Demonstrates algorithm enhancements and/or calibration improvements
 - Generates Intermediate Products
- Project Science Office Element (PSOE)
 - Provides management direction and science guidance to other elements
 - Manages algorithm and calibration recommendations for submission to the NPP/NPOESS Algorithm Change/Configuration Control Board (CCB)
- NPP Instrument Calibration Support Element (NICSE)
 - Supports assessing and validating pre-launch and post-launch radiometric and geometric characterization and calibration assessment of the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument data
- Product Evaluation and Analysis Tools Elements (PEATEs)
 - Performs EDR Evaluation/Characterization and algorithm improvements coordinated with External Science Investigators, Science Team, and NICSE (for Land, Ocean, and Atmosphere PEATEs)
 - *The five PEATEs are: Atmosphere, Land, Ocean, Ozone, Sounder*



Background of MODAPS

- The Moderate Resolution Imaging Spectroradiometer (MODIS) Adaptive Processing System (MODAPS) was developed to produce global products for the MODIS instrument on the Terra and Aqua spacecraft.
- MODAPS has a history of reuse
 - Used to produce atmospherically corrected and Earth-located radiances for various instruments onboard the Land Remote-Sensing satellites (Landsats) and the NOAA 7 through NOAA 14 spacecraft
- Most of the development was spent in refining software to produce improved science products.
 - Only minor effort was expended on modifying the overall processing and distribution framework of MODAPS





Energy Balance Product Suite

- Surface Reflectance
- Land Surface Temperature, Emmissivity
- BRDF/Albedo
- Snow/Sea-ice Cover

Vegetation Parameters Suite

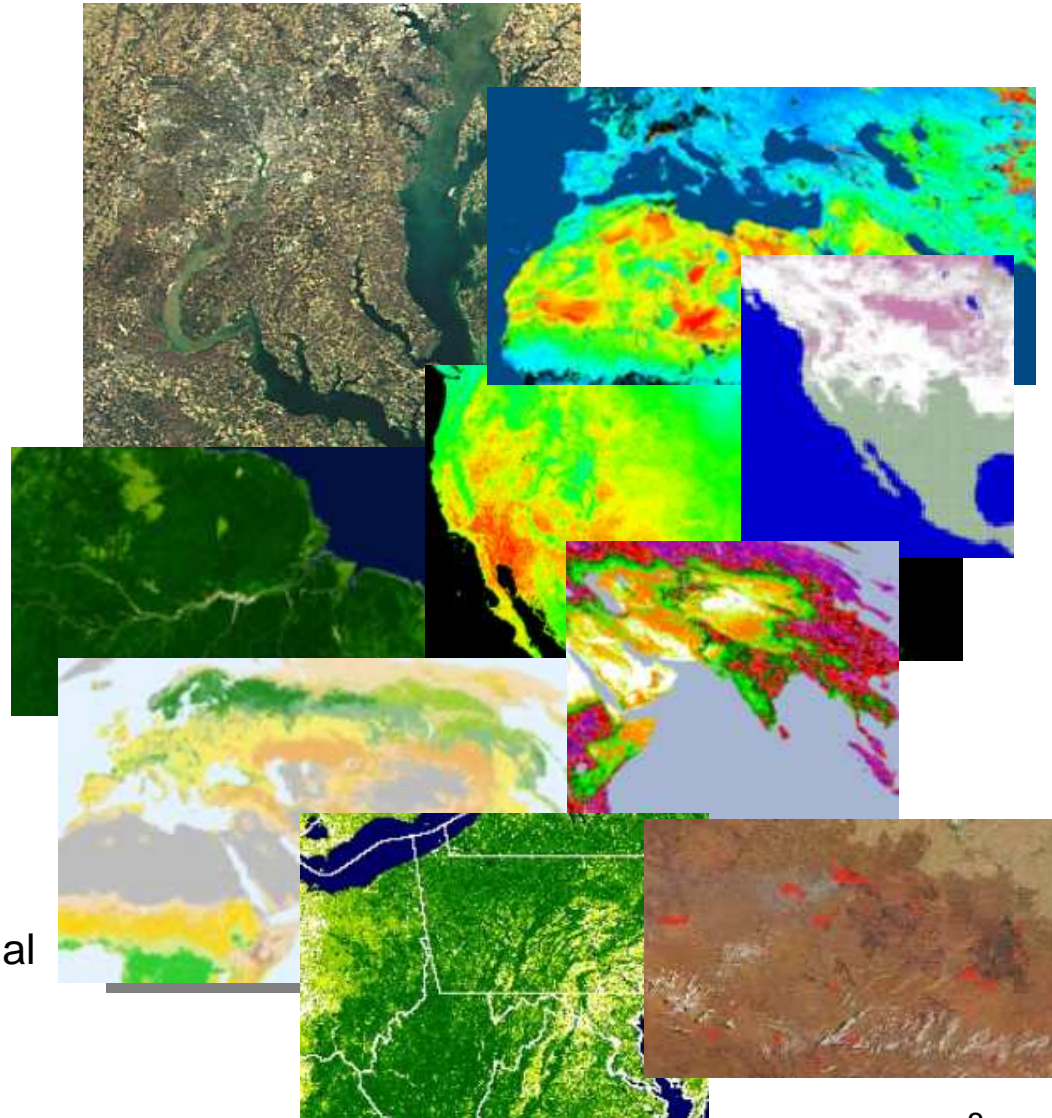
- Vegetation Indices
- LAI/FPAR
- GPP/NPP

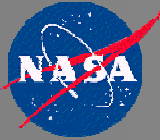
Land Cover/Land Use Suite

- Land Cover/Vegetation Dynamics
- Vegetation Continuous Fields
- Vegetation Cover Change
- Fire and Burned Area

Global Products

- Daily, 8-day, 16-day, monthly, annual
- 250m, 500m, 1km, 0.05 deg.





Description of MODAPS

- The main subsystems of MODAPS are:
 - *Ingest* and *Metadata*, used to ingest data products
 - *Operator GUI* and *Database*, used to archive and access data products
 - *Scheduler*, *Loader*, and *Data Production*, used to control the storage, processing, and distribution of data products
 - *Land Data Operational Product Evaluation* (LDOPE), used for quality assurance and assessment
 - *Land and Atmosphere Archive and Distribution System* (LAADS), *Archiver*, and *Export*, used to search, order, and deliver products from the archive

- MODAPS Web Sites:

- MODAPS services:

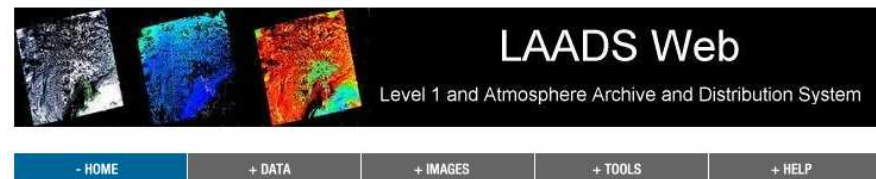
<http://modaps.nascom.nasa.gov/services>

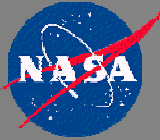
- LAADS distribution

<http://ladsweb.nascom.nasa.gov/>

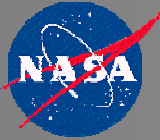
- Land Quality Assurance

http://landweb.nascom.nasa.gov/QA_WWW

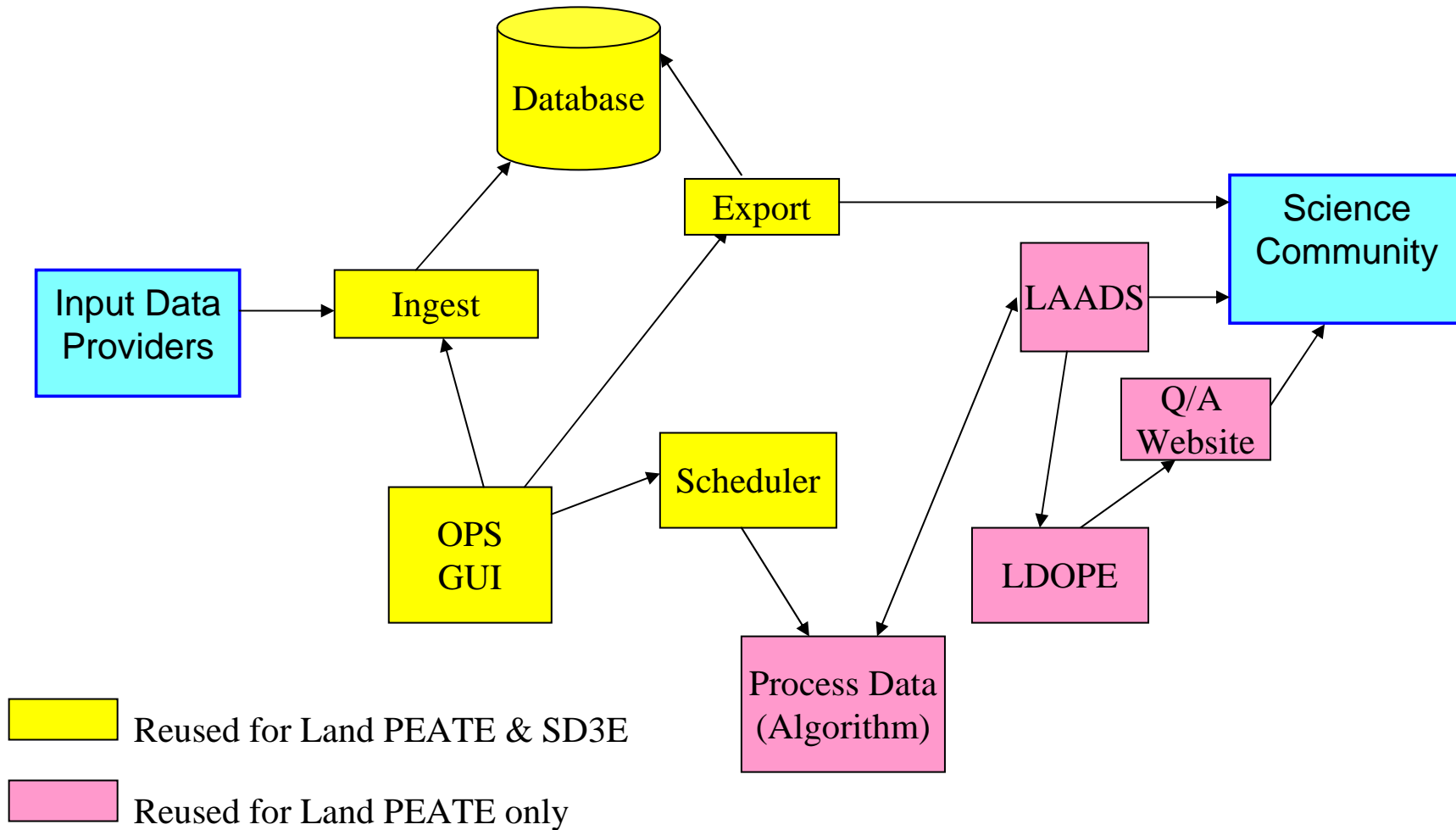


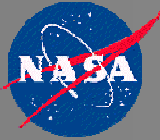


- Reusing MODAPS to create the NPP Data Adaptive Processing System (NPPDAPS) for the Land PEATE
 - VIIRS instrument and associated algorithms has heritage from the MODIS instrument
 - Similar processing required to produce VIIRS EDRs
- MODAPS subsystems reused by the Land PEATE:
 - Ingest
 - Operator (OPS) GUI
 - Database
 - Export
 - Scheduler
 - Land and Atmosphere Archive and Distribution System (LAADS)
 - Land Data Operational Product Evaluation (LDOPE)
 - Process Data
 - Quality Assessment (Q/A) web site



MODAPS Block Diagram



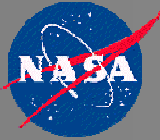


Reuse of MODAPS for NPPDAPS

- Approximately 97% of the core software components were reused without modification.
- Primary reasons modifications were required:

Subsystems	Reason for Modification
Ingest, Metadata	Differences in file naming conventions
Loader	Differences between MODIS and VIIRS production rules
LDOPE	Need to assess NPP SDRs and EDRs and support new data products
Q/A Web Site	Need to accommodate new HDF formats and VIIRS Image bands
Archiver	Storage of VIIRS data

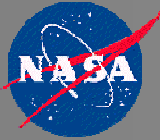
- *The OPS GUI, Database, Scheduler, Product Generation, LAADS, and Export subsystems were reused without modification.*



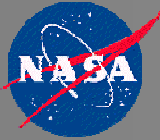
NPPDAPS Reuse Estimates

<i>Estimated Reuse for NPPDAPS</i>			
Subsystem	Percent Reuse	With Reuse [staff-months]	Without Reuse [staff-months]
Ingest	90.7%	3	25
Metadata	88.4%	4	37
OPS GUI	100%	0	161
Database	100%	0	56
Scheduler	100%	0	18
Loader	99.3%	1	24
Product Generation	100%	12	30
LDOPE	90.9%	24	281
LAADS	100%	0	1
Archiver	95.4%	2	37
Export	84.4%	9	60
Total	88.6%	55	730

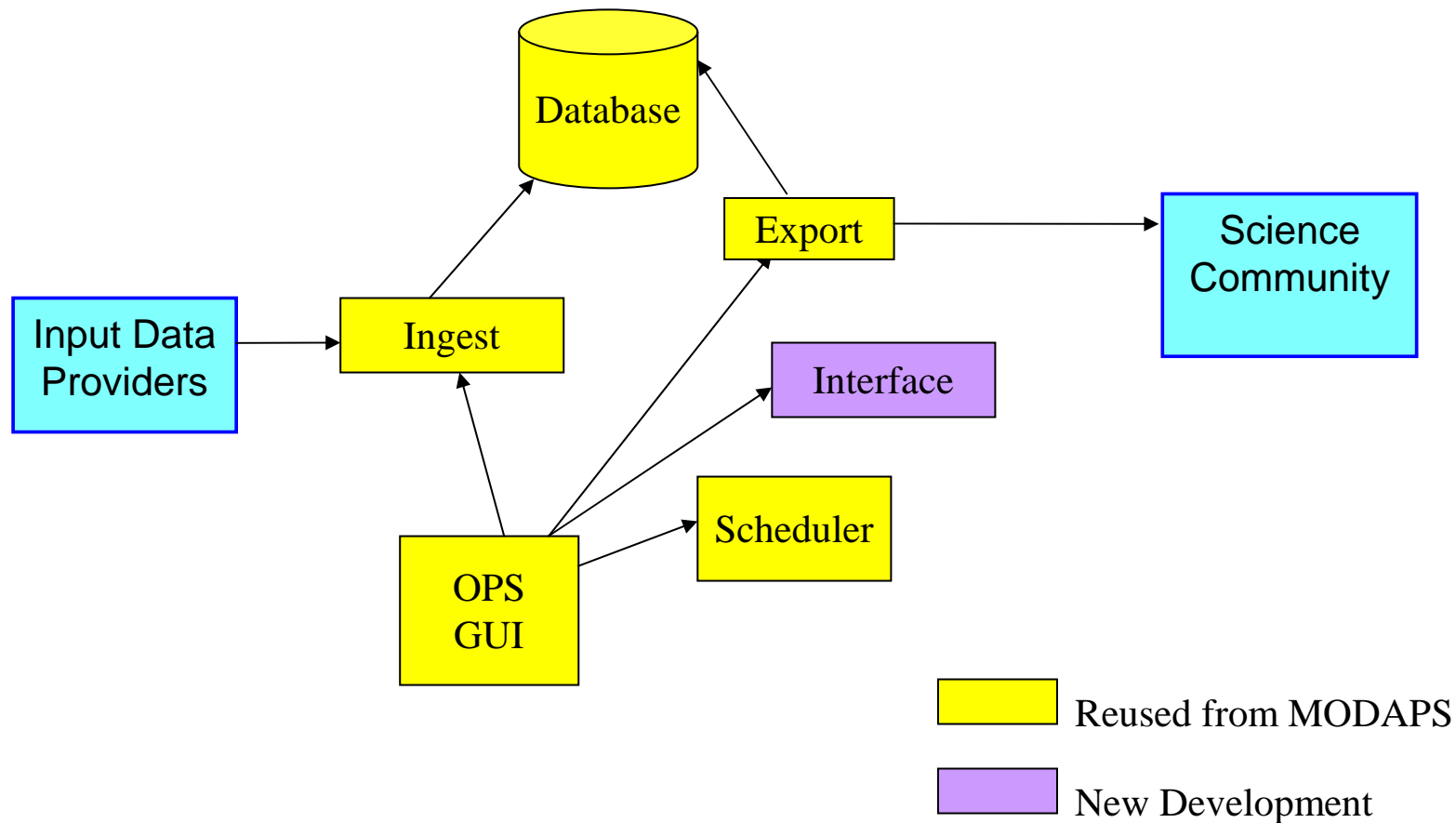
Note that the effort without reuse is more than 10x the effort with reuse.

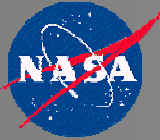


- The main subsystems of the SD3E are:
 - *Scheduler*, used to manage all tasks and processes
 - *Ingest Controller*, used to verify received data products and make them available to PEATEs via FTP server
 - *Interface Controller*, used to handle subscriptions and ad-hoc requests for data products
 - *Utilities*, used to monitor the health of and maintain the SD3E
- MODAPS subsystems reused by the SD3E:
 - Ingest
 - Operator (OPS) GUI
 - Scheduler
 - Export
 - Database
- MODAPS architecture was also reused
 - Local inbound and outbound FTP directory structure reused
 - Choice of hardware (e.g., disk controller, RAID configuration, server vendor) was made based on MODAPS experiences

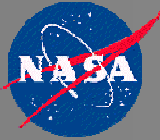


SD3E Subsystem Reuse



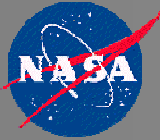


- Approximately 53% of the code in SD3E was reused with very little or no modification.
- Primary reasons modifications were required:
 - Interfacing with new external interfaces
 - Ingesting NPP data products
 - Supporting mission-specific verification methods (e.g., integrity files)
 - Parsing mission-specific metadata and storing it in the database
 - Storing NPP data products locally for 32 days
- Some key modifications:
 - Use of PostgreSQL database instead of Sybase database
 - Addition of new routines/functionality to SD3E Utilities subsystem to read metadata from HDF files
 - New database tables to support new external interfaces
- The SD3E Interface Controller is new development.



Advantages of Reuse

- Reuse of MODAPS provides a large savings in development, reducing the time and effort required to develop new systems that are similar to MODAPS.
 - Land PEATE development is more than 10x faster
 - SD3E development is more than 2x faster
- The resident knowledge of current MODAPS developers and scientists can be reused through collaboration.
- Reuse of existing code:
 - Reduces the amount of new code that must be written
 - Reduces the amount of testing that must be done
 - Helps ensure the reliability of the new system
- Lessons learned from MODAPS are reused to help new system developers make better choices.

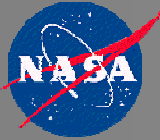


SD3E Reuse Estimates

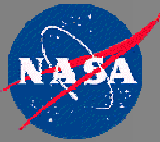
<i>Estimated Reuse for SD3E</i>			
Language	Percent Reuse	With Reuse [staff-months]	Without Reuse [staff-months]
Perl	52.8%	6	13
SQL	53.6%	1.5	3
Total	53.0%	7.5	16

Note that the effort without reuse is more than 2x the effort with reuse.

<i>Significantly Reused Subsystems</i>		
Subsystem	Language	Percent Reuse
Scheduler	Perl	90%
Utilities	Perl	75%
Table Definitions	SQL	66.6%
Stored Procedures	SQL	65%



-
- Reuse of existing systems with some new development and some modification greatly reduces the level of effort, schedule, and risk to create the NPP SDS data system
 - The general functions and software of NPP SDS will be applicable to systems that are included in GEOSS.
 - The lessons learned from this example of reusing existing software will be valuable to the future missions that are included in the GEOSS plan.
 - Realizing the benefits of existing expertise will allow future developers to devote the time and effort necessary to accommodating additional GEOSS requirements (e.g., near real-time data distribution).



- NASA Earth Science Data Systems (ESDS) Working Groups
 - Coordinator, Kathy Fontaine (<http://esdswg.gsfc.nasa.gov/>)
- Software Reuse Working Group
 - Co-Chair, Robert E. Wolfe (robert.e.wolfe@nasa.gov)
 - General Information:
Ryan Gerard (ryan.gerard@gsfc.nasa.gov)
James J. Marshall (james.marshall@gsfc.nasa.gov)

<http://softwarereuse.nasa.gov/>

The screenshot shows the website for the Earth Science Data System Software Reuse Working Group. The header includes the NASA logo and 'GODDARD SPACE FLIGHT CENTER'. There are links to 'Visit NASA.gov' and 'Goddard Home Page'. The main banner features the text 'Earth Science Data System Software Reuse Working Group' with a graphic of a globe and a recycling symbol. Below the banner is a navigation menu with links for '+ HOME', '+ REUSABLE ASSETS', '+ RESOURCES', '+ OPEN SOURCE', and '+ FUNDING OPPORTUNITIES'. The main content area has a 'news' sidebar with articles like 'From Local to Global Coordination' and 'Accepted: Fall 2006 AGU Meeting Reuse Session'. The main section is titled 'Earth Science Software Reuse' and contains a welcome message and information about the portal's goals and survey findings.