Science Operations Concepts for EOSDIS
Part I - Data Products Resource Allocation

Draft Plan
Version 1.0

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Science Issues

- Apparent lack of consensus on how to maximize support for data products and capabilities while reducing cost of developing and operating EOSDIS.
- A coherent science operations concept is essential to provide the science community with the resource baseline, cost linkage, and leverage to be able to prioritize the allocation of resources to science data products.
- 3x volume and 30x processing increase (from ECS contract baseline) imply that science requirements growth may be uncontrolled.
- If the science community does not manage and prioritize science requirements (to cost) there is a risk that critical decisions could be made on a relatively uninformed bottom-line basis.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Science Issues cont.

- Insufficient information for the science community to validate the derivation/interpolation of system sizing estimates from investigator provided volume and processing requirements
- The costs of the EOS data products have received a disproportional share of attention as compared to the benefits (implication that benefits are poorly understood/ articulated in comparison to costs).
- Apparent erosion in linkage between data products and specific questions about global change such that all data are apparently equally valuable and hence equally vulnerable to deletion in response to budget cuts.

Goals

- Maximize the science benefit for a given data product implementation budget by examining the balance between data production, data archive, data distribution, architecture, and other infrastructure of EOSDIS
- Manage risk through the control of growth and allocation of computational resources.
- Promote the capability to support resource allocation at the project [macro] and instrument team [micro] levels.
- Recognize where policy decisions or clarifications are needed and provide through scientific consensus.
- Simplify the Project's efforts to coordinate and deliver support to the science community.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Proposed Strategy

- **The elements of the operations concept include:**
  - establish a baseline science data capacity that defines a resource envelope into which the science data product requirements are prioritized;
  - use a phased algorithm implementation to provide a flexible framework for scaling the implementation of science data processing requirements;
  - implement an increasingly rigorous set of product evaluation criteria in order to reduce complexity, promote automation and reduce operational costs of the standard products;
  - introduce a configuration management of science data requirements with a progressively more detailed starting with overall capacity (planning phase) to specific product placement at DAACs
  - establish a science advisory process to: review data products and prioritize the resource allocation to products based upon science readiness, operational cost and demand

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Processing Capacity Baseline

- **Goals:**
  - Provide quantitative cost-contrained basis for resource allocation.
  - Define baseline such that science community can realize benefits from technology improvement (e.g., to expand capacity).
  - Ensure that science products do not absorb brunt of cost-cutting.
- **Approach:**
  - Provide a more reliable cost estimate of the data product requirements.
  - Use the cost estimate to establish overall product capacity baseline defined in terms of at-launch capability and post-launch growth.
  - Reassess overall baseline capacity annually with respect to technology advances, costs of other EOSDIS elements, and budget changes.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Estimate Data Product Costs

- **Goals:**
  - support a more reliable estimate of data products' for PDR
  - provide a more descriptive basis for describing algorithm requirements
- **Approach:**
  - compile data from instrument teams to meet the information requirements of the system engineering and cost modellers
  - assess IDS investigators requirements for data products
  - provide information including algorithm process, data flows, and product availability schedules to modellers by October 3
  - determine fraction of estimated cost that is covered by anticipated budget

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Prioritize Resource Allocation

- **Goals:**
  - prioritize product capacity resource allocation (include post-launch products and reprocessing) as necessary to conform to budget profile
  - establish capacity allocations by DAAC and instrument team
- **Approach:**
  - develop nominal capacity allocation by scaling DAACs and instruments equally according to ratio of estimated cost to budget
  - iteratively:
    - » assess science objective and system engineering feasibility and breakage associated with current working allocation
    - » examine product rephasing, scrubbing, and deletion alternatives
    - » reallocate DAAC and instrument team capacity
  - converge on resource allocation in FY95 in time for initial H/W buy and review as necessary

Phased Product Implementation

- **Goals:**
  - reduce start-up costs of EOSDIS by scaling investment commensurate with realistic schedule for algorithm maturity and acceptance.
  - avoid making irreversible changes due to premature product deletion.
  - defer decisions to ensure that resource allocation can be made on the basis of observed rather than predicted performance.
- **Approach:**
  - promotion of products is linked to product capacity baseline and subject to science advisory review.
  - Experimental Products - "latest/greatest" algorithms run at SCFs or TLCF with DAAC support for archival and distribution.
  - Provisional Products - fully integrated algorithms run at DAACs on a sub-sampled basis, e.g., 5% of nominal volume (sub-sampling is product specific).
  - Standard Products - full scale production and archival at DAACs with reprocessing.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Science Configuration Management

- **Goals:**
  - Manage capacity allocations more so than product requirements.
  - Localize responsibility for maximizing science return from DAAC and instrument team capacity allocation.
- **Approach:**
  - Establish science data processing requirements review starting in FY95.
  - Facilitate interchange of information among instrument teams of how to maximize processing efficiency.
  - Overall 1-3 year advance planning of data product resource allocation documented in Science Data Plan.
  - DAAC month to month operational planning documented in DAAC Science Operations Plan.
  - Develop plan for coordination and scheduling of algorithm update and reprocessing coordination.
  - Convene appropriate advisory group(s) to deal with exceptions.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Science Advisory Process

Scenario One - Long Term Planning

Participants: UWGs, SOFT, IWG, EAP, SEC, IDS, Project Scientists, Instrument Representatives

Agenda: Allocation of resources for system-wide maintenance and growth:

- DAAC, instrument team, product allocations
- processing, reprocessing, new product allocations
- document results in Science Data Plan

Lead: Cross-organizational science representatives

Review: Same plus ESDIS Project

Implement: ESDIS Project

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Science Advisory Process

Scenario Two - Operational Planning

Participants: DAAC Scientist and Manager, UWG, Instrument Representatives

Agenda: Allocation of DAAC resources:

- production, reprocessing, distribution
- H/W and staff
- growth
- DAAC unique services and extensions
- document results in DAAC Science Operations Plan

Lead: DAAC Scientist / DAAC Manager

Review: EOSDIS Project Scientist, ESDIS Project

Implement: DAAC

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Discussion

- **Publish EOS Data Product Reference Guide:**
  - product abstracts (w/ description of benefits)
  - availability schedule
  - selected tabular information
  - hardcopy and Mosiac (w/ keyword search)
- **SPSO:**
  - substantially revise DB structure to incorporate process info
  - ingest information from AHWGP
  - issue updated data product report (requirements & allocations)
  - provide information access/exchange through on-line server
  - terminate on-line SPDB
- **Cost modelling:**
  - need to understand what a product/parameter/file costs in order to evaluate various alternatives for capacity resource allocations

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Even More Discussion

- **Begin holding data production workshops in FY95:**
  - scrub data processing requirements
  - coordinate data product implementation and scheduling
  - science/system engineering issues and resolutions joint papers could be made available electronically
  - product QA and validation
  - I&T, production planning, and reprocessing scenarios
  - establish minimal functional and performance requirements necessary to sustain production of and access to data products
- **Educational Modules:**
  - how supported -- investigator, contract out
  - subscription-based access saves on-line loading
  - begin working prototype

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov

Benefits

- Science community has methodology for managing science requirements to match resource envelope.
- Phased implementation preserves product set and SDR architecture within cost constraints.
- Prioritization of products based upon observed, rather than predicted benefits, costs, and demand.
- System sizing based upon realistic algorithm implementation schedule with deferred hardware acquisition as opposed to worst-case system sizing at launch.
- Consideration of cost in algorithm review will maximize the number of products that can be implemented for a given budget.

S. Wharton, EOSDIS Project Scientist
wharton@eosdata.gsfc.nasa.gov