Building Cost Effective Remote Data Storage Capabilities for NASA's EOSDIS

Stephen Marley

Mike Moore & Bruce Clark

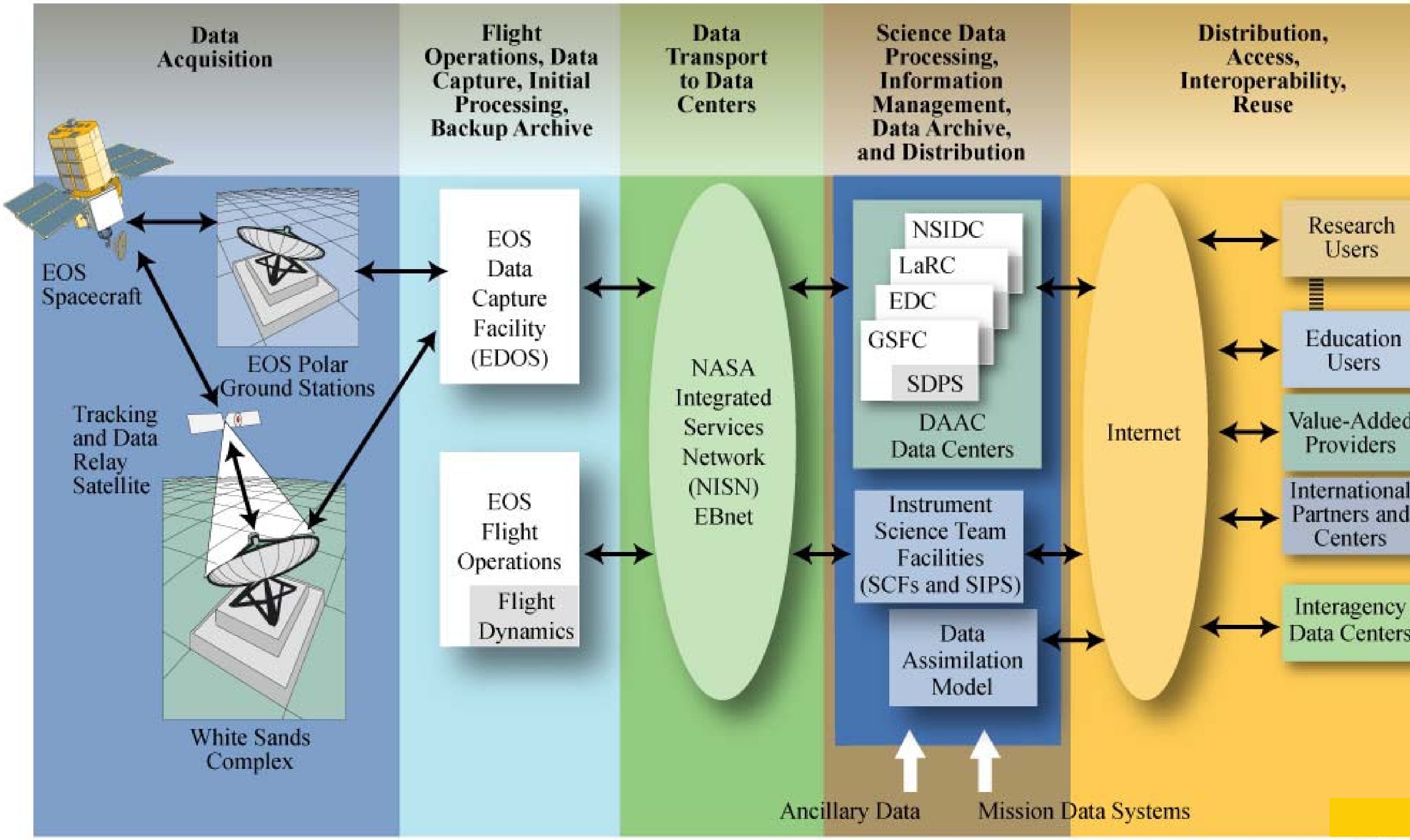
8-Apr-2003

IEEE/NASA MSST2003

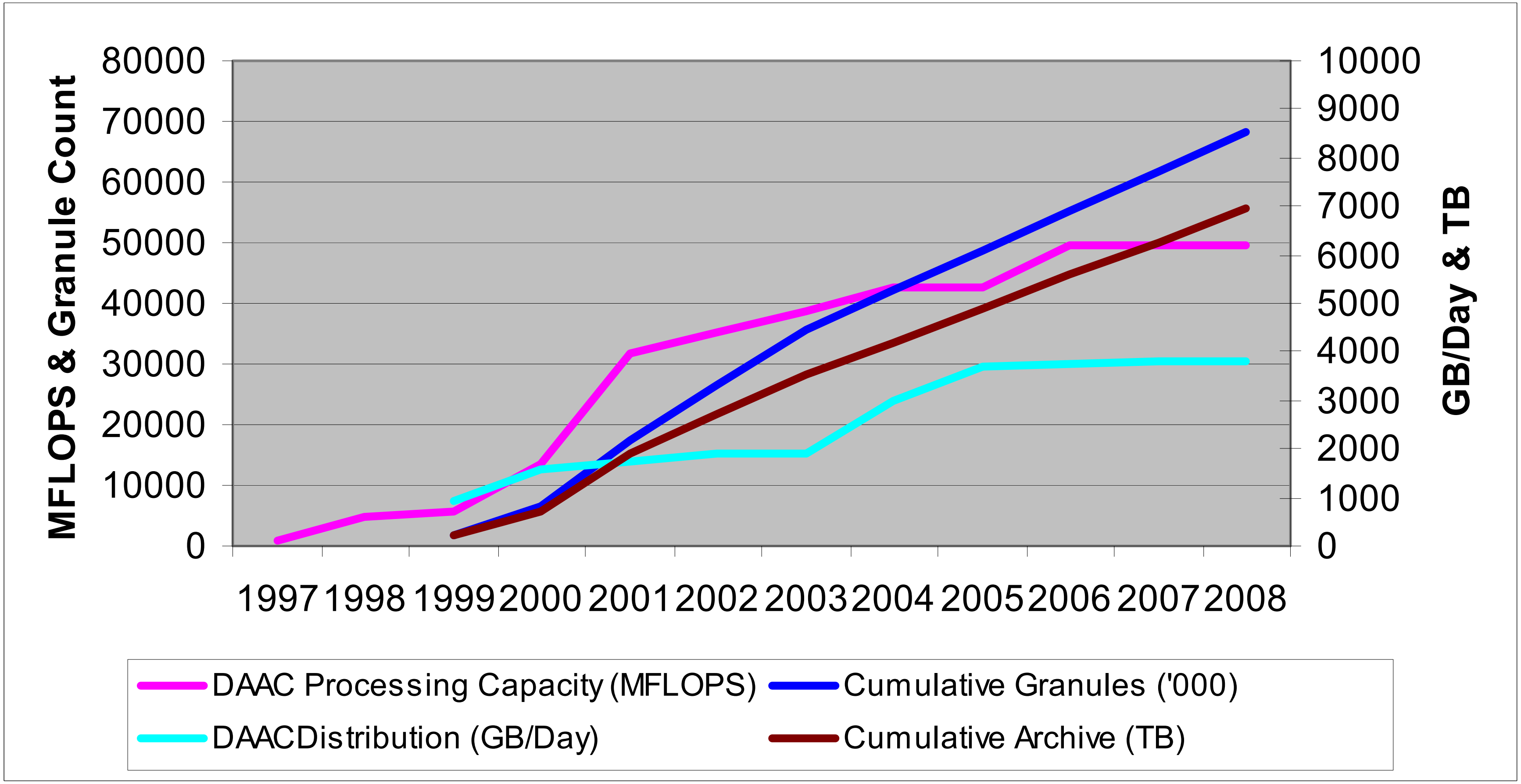
Agenda

- Overview of the EOS Mission
- Remote Data Store Program
- Modeling Petabyte Data Stores

EOSDIS Mission



Mission Capacities



Agenda

- Overview of the EOS Mission
- Remote Data Store Program
- Modeling Petabyte Data Stores

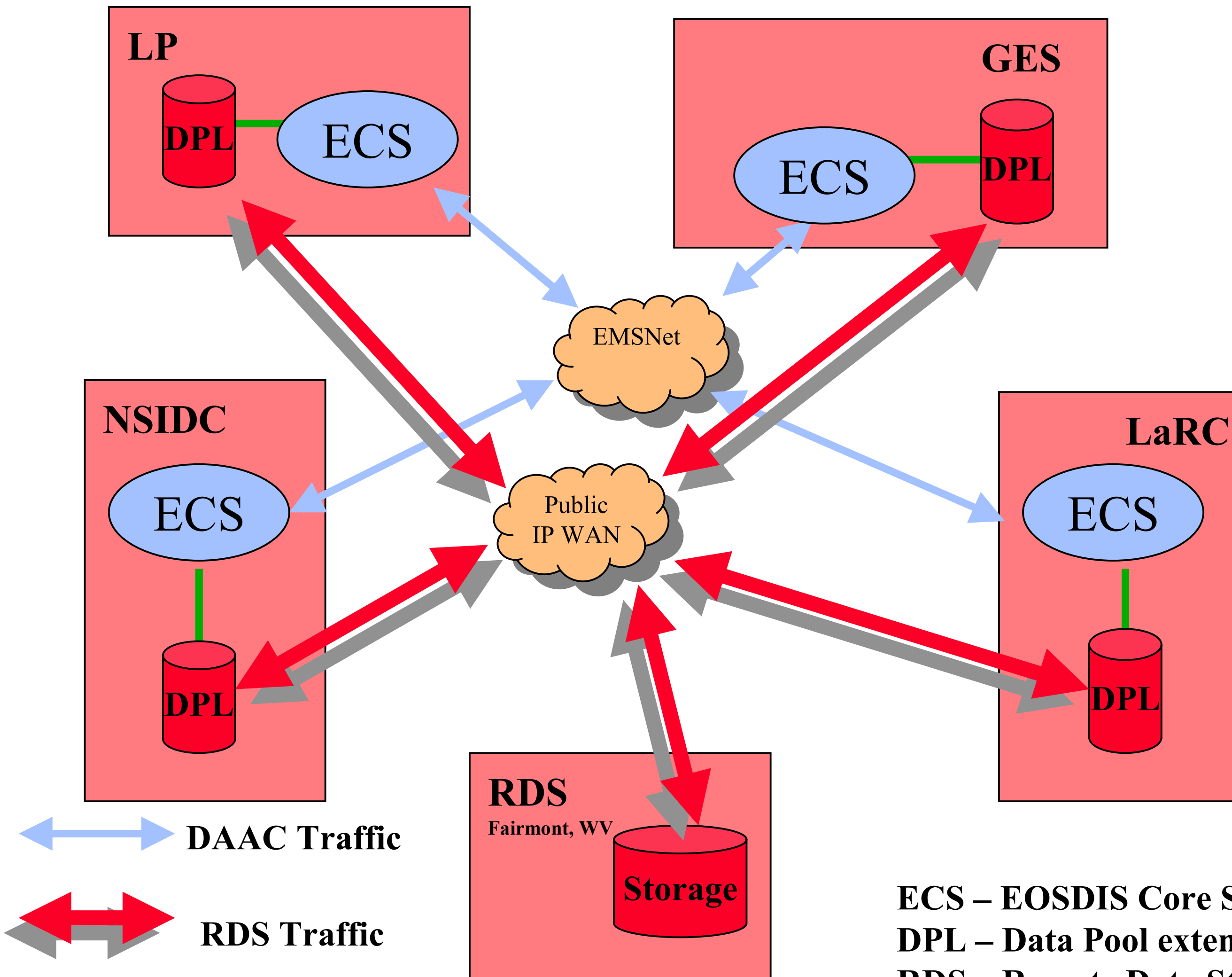
Remote Data Store

- **Remote Data Store (RDS) Goal**

  "To provide automated, secure, seamless and efficient remote on-line redundant storage, recovery and access of operational EOS mission data products."

- **NASA Goal**
  - **Establish a partnership with industry to develop and evolve distributed scientific data services for a petabyte scale data set**
- **Relationship to Existing EOS Data Services**
  - **EOSDIS DAACs remain the primary data production; archive and distribution service provider for the EOS Missions**
  - **RDS provides additional support to the EOSDIS mission**
    1. RDS will provide an additional capability for the remote storage of part of a national data resource
    2. RDS will begin to research emerging data storage & access technologies for large distributed scientific data sets
    3. Provide additional data service capacity for data managed by the EOSDIS DAACs

RDS Concept



ECS – EOSDIS Core System
DPL – Data Pool extension to ECS
RDS – Remote Data Store

RDS Expectations

- ***Automated*** – Operations are a key cost driver for long-term data archives. To the maximum extent practicable, the RDS needs to leverage storage management tools to minimize operational maintenance activities
- ***Secure*** – This data is considered a National Data Asset, and as such needs to be secure from both physical and electronic attack whether intentional or otherwise
- ***Seamless*** – This means storage location independent access
- ***Efficient*** – Not all data will need to be accessible with same level of service. Efficient access is defined as providing the most cost effective level of service consistent with data use
- ***Remote*** – Implying an IP WAN scale of distribution
- ***On-line*** – As technology evolves the definition of what on-line means can get blurred. For our purposes, we define on-line to imply access latencies not to exceed a few seconds, but potentially as fast as milliseconds
- ***Redundant Storage, Recovery & Access*** – Moving beyond mere backup or mirroring. Redundant storage implies an intelligence to the data duplication that reflects varying levels of data criticality and access loads. This intelligence allows flexibility in the way the system can respond to individual storage, recovery and access requests

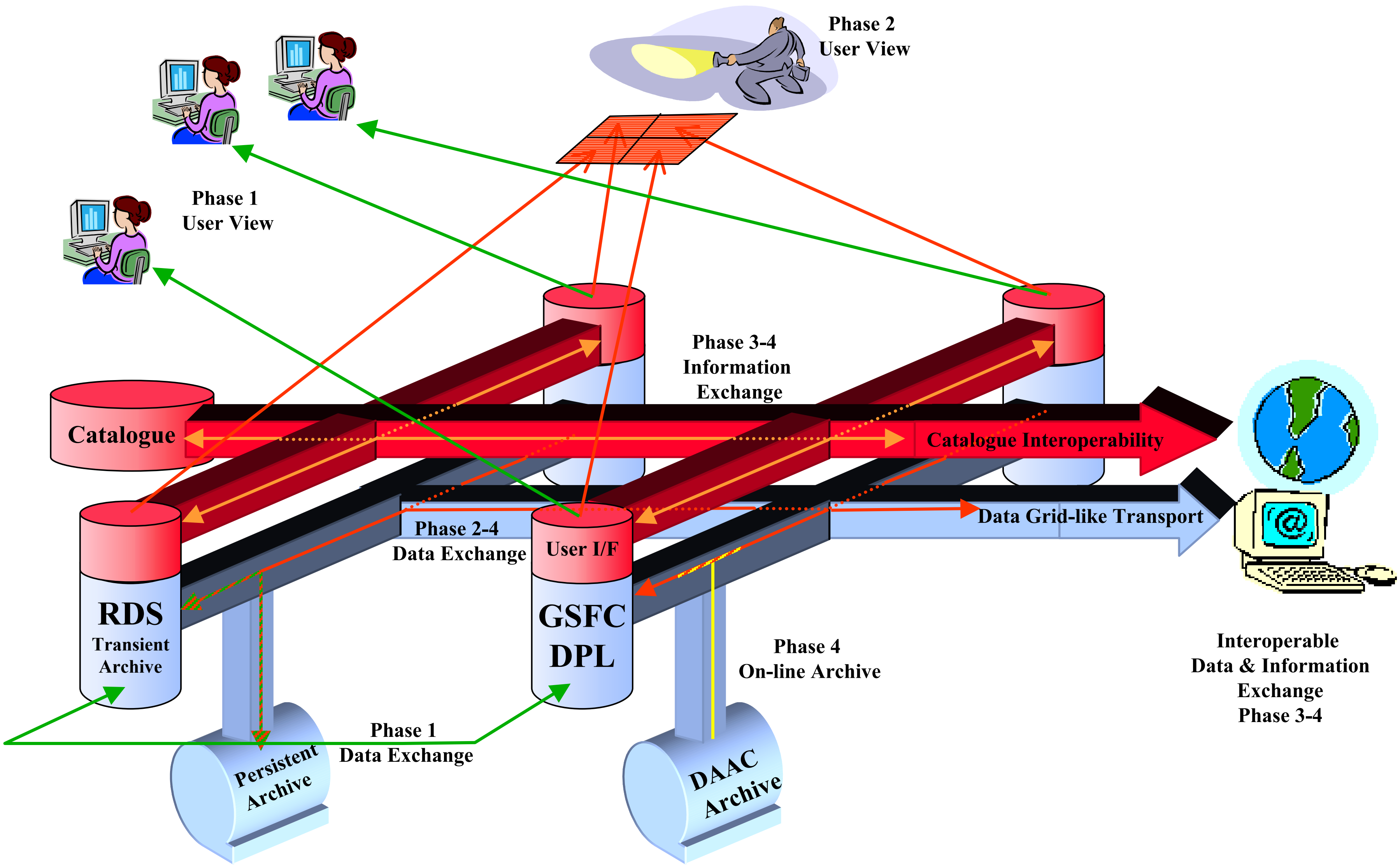
Key Operational Goals for RDS

- **The RDS will provide two types of service in support of the EOSDIS Mission:**
  - a. Data Integrity Services
    - – The RDS will coordinate Remote Data Backup and Recovery Services as well as Redundant Shared Data Pool Data Storage across the EOSDIS Data Center Enterprise.
  - b. Data Access Services
    - – The RDS will present a consolidated user view of the Data Center Enterprise holdings that are both unified and seamless:

    Unified – meaning that only a single reference for each product is presented to the end-user or application independent of the number of copies of the product that are stored in the enterprise for redundancy and load sharing purposes;

    Seamless – meaning that the physical location (both geographic and storage technology) of products is transparent to the end-user or application.

RDS Vision



Phase 2
User View
Phase 1
User View
Phase 3-4
Information
Exchange
Catalogue
Catalogue Interoperability
Data Grid-like Transport
Phase 2-4
Data Exchange
User I/F
RDS
Transient
Archive
GSFC
DPL
Phase 4
On-line Archive
Interoperable
Data & Information
Exchange
Phase 3-4
Phase 1
Data Exchange
Persistent
Archive
DAAC
Archive

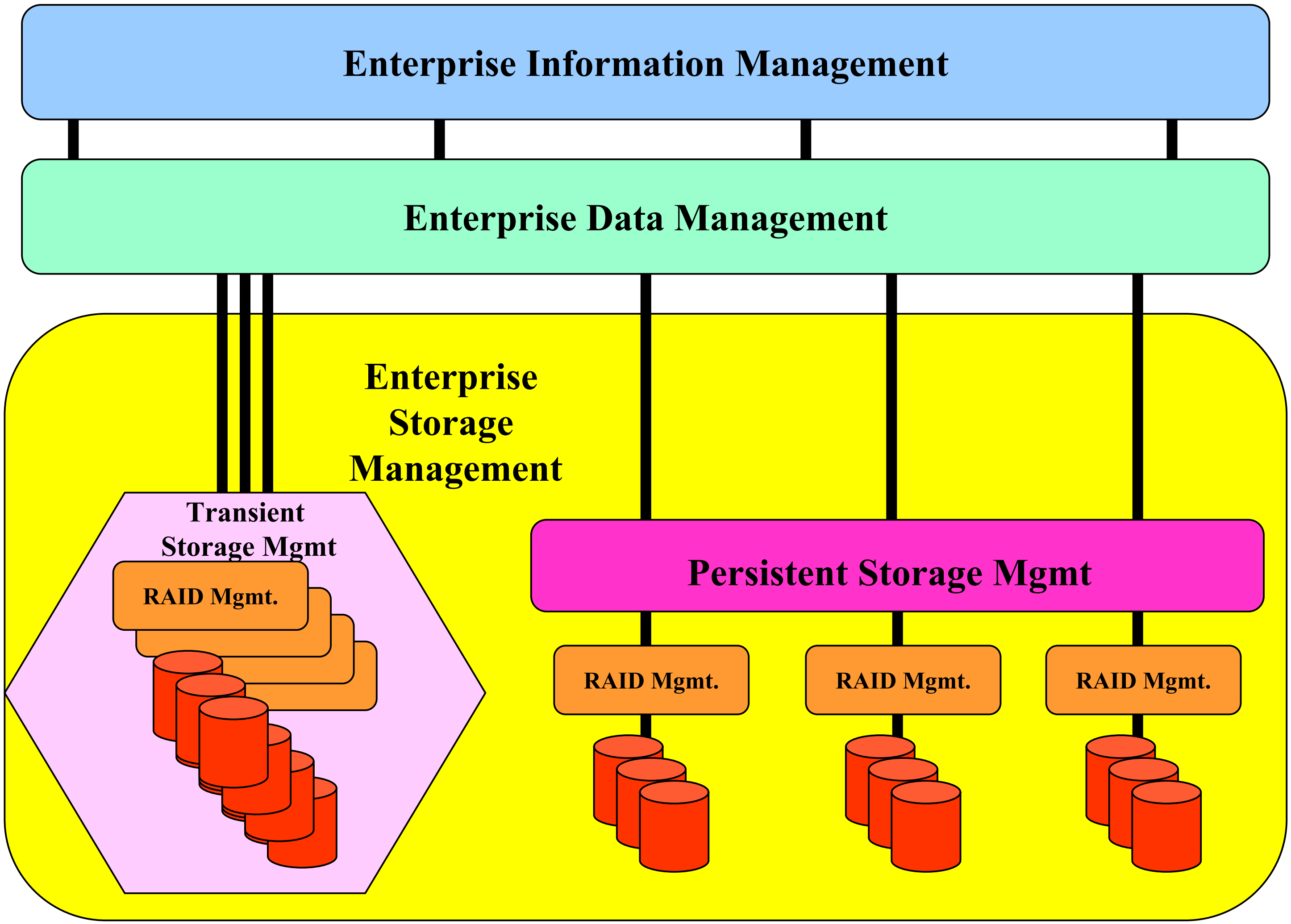
Key Research Goals for RDS

- **Vendor Independence**
  - Science Data Archives are evolving from short-term "Mission" focused storage to long-term "Data Asset Management"
  - Single vendor "fixed" solutions become unviable because technology will evolve beyond any given storage technology or vendor solution
  - The RDS needs to be built on open industry standards that will permit the integration of new technology.
  - This will permit NASA to use the RDS solution as a research test-bed for new technologies. Examples for research include:
    1. Heterogeneous Storage over IP; Object-based Storage Devices (OSD) and Services
- **Enterprise Information Management**
  - Volumes of data available for research are increasing.
  - Need to reference data in a way meaningful to the end user and not the storage system
  - The RDS needs to provide an environment which enables the implementation of Enterprise Information Management services on top of the storage services. Examples for research in this area include:
    1. Data Grid; Data Pool

Architecture Considerations

- When defining the solution architecture, the following key drivers need to be integrated together.
  - *Persistent, secure physical storage:* Both long term and short term storage solutions must provide a high data integrity environment, that is tolerant of hardware and application software failure. In addition, although the data itself is not sensitive, it is part of a national data set and considered a national resource, and so needs to be securely stored and accessed.
  - *Cost effective persistent storage:* The scale of the data being archived is tremendous. The total cost of ownership (TCO), therefore, becomes a critical issue. Not just the costs of physical disk, but also the server architecture and management environment are all important drivers.
  - *Higher performance transient storage:* EOS data access patterns vary considerably as the data ages. Data is accessed by user and applications at a higher frequency within the first 30 days of acquisition. This was the one of the drivers for the generation of Data Pools at the DAACs
  - *Enterprise Data Management:* Although the individual DAACs define the contents of the Data Pools (i.e. what data is promoted from the ECS archive), the management of the data across the enterprise for data and service integrity needs to be managed through software at an enterprise level, and not just through inter-DAAC agreement or management mandate.
  - *Enterprise Information Management:* Sitting above the Enterprise Data Management, this layer provides the domain or science context that justifies having the data available to a broader community. This layer presents the information space that the data represents to the end-user, in terms that the end user recognizes.

Candidate Architecture



Key Capabilities

Layer	Capabilities
Enterprise Information Mgmt	Application View; Storage Repository Abstraction Thematic Abstraction; Knowledge Based Access; Data Authorship

Layer	Capabilities
Enterprise Data Mgmt	Homogeneous Storage View; Policy Mgmt

Enterprise Storage Management

Persistent Storage Systems		Transient Storage Systems	
Layer	**Capabilities**	**Layer**	**Capabilities**
Object-based Storage Devices	Large Block "Fixed Content"; Abstracted File System	**SAN FS Management**	Shared Server
		Enterprise Device Mgmt	Storage Network Management
		SAN Fabric Management	SAN Connectivity

Layer	Capabilities
RAID Management	Logical Volume Management Duplexing; mirroring; flash store; Striping

Layer	Capabilities
RAID Hardware	Storage Buffering

Standards

- During a multiphase development such as RDS it is critical for the long-term success of the program to adhere to appropriate standards throughout all phases.
- Telecommunications
  — By adhering to defined telecommunication and network standards in the early stages of implementation, RDS will be better positioned to evolve to inter-SAN communications technologies as they mature.
- Storage Interconnectivity
  — All storage technologies implemented under RDS will be expected to conform to current IEEE and SNIA standards, with a strategy that takes into account the evolving nature of existing (or draft) iSCSI, FCIP and iFCP storage interconnectivity standards as well as other standard initiatives as they emerge.
- Storage Management
  — RDS intends to leverage emergent initiatives like those supported by SNIA in support of its enterprise storage management goals (e.g. Bluefin)
- Applications/GIS
  — It is anticipated that data services based on FGDC metadata standards and OGC applications standards will be implemented on the Data Pools within the next 2 years. It is expected that RDS will utilize these services.

Technology/Standards Maturity

- The level of maturity of distributed, heterogeneous storage technology currently falls short of the ultimate needs for RDS.
- The need for RDS is immediate, and so the deployment of the RDS systems and functionality will use a phased approach that follows the on-going evolution and maturity of storage technologies that support or compliment the RDS mission.

Phasing Roadmap

Layer	Phase 1	Phase 2	Phase 3	Phase 4
Enterprise Information Management	No Prototype solution	Yes Preliminary Implementation	Yes Leverage Integrated Storage & Data Management	Yes Leverage shared file system & integrated DAAC Archives
Enterprise Data Management	No Prototype solution Management by Operator Procedure	Optional If cost and interoperability technology plan permit	Yes Integrate across RDS/DAAC Enterprise	Yes Integrate existing DAAC Archives
Transient Storage	Optional If cost and interoperability technology plan permit	Yes Prototype Interoperability	Yes Implement Interoperability Prototype shared File System	Yes Implement Shared File System
Persistent Storage	Yes Initial Archive Capacity	Yes Supplemental Capacity	Yes Supplemental Capacity	Yes Supplemental Capacity
RAID Management	Yes	Yes	Yes	Yes
RAID Hardware	Yes	Yes	Yes	Yes

Next Steps

- RDS project will begin to test some of the underlying assumptions and conclusions of this paper.
- Initial SAN-based transient archive solution is in place
- Grid-enabled Content Addressable Storage (CAS) based persistent archive prototype is currently being prepared for deployment.
- Will provide a significant test-bed capability against which to evaluate emerging storage and storage management solutions.

Agenda

- Overview of the EOS Mission
- Remote Data Store Program
- Modeling Petabyte Data Stores

Modeling Petabyte Archives

- The cost of RDS needs to take into account not only the cost of the acquisition but also the cost of operations and the cost of maintenance.
- Total Cost of Ownership Approach
  - Technology refresh is a way of life
  - Take advantage of technology refresh
  - Identify the real costs of operations into the future

Total Cost of Ownership

- **TCO Elements Include:**

  a. Hardware: includes racks, servers, disks, cabling, acquisition costs, and maintenance costs.

  b. Software: includes acquisition costs, maintenance costs, license management, upgrades, and monitoring.

  c. Personnel: includes operations staff, maintenance support staff, vendor staff, and consulting staff.

  d. Availability: includes component MTBF, system MTBF, and enterprise MTBF

  e. Performance: includes time to first byte, total volumes served, and time to complete.

  f. Facilities: includes equipment floor space, power, cooling, and staff floor space.

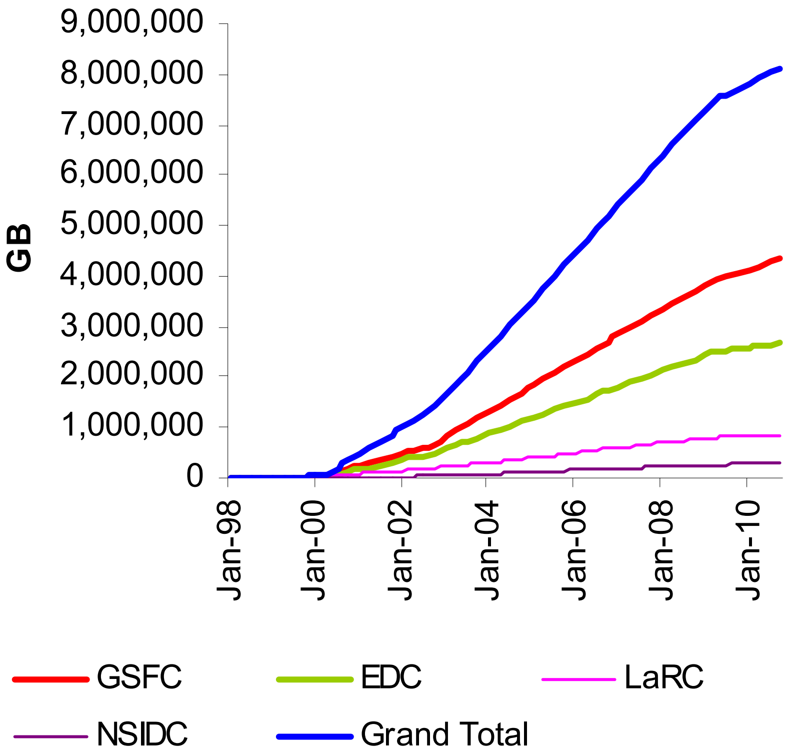
TCO Contribution by Phase

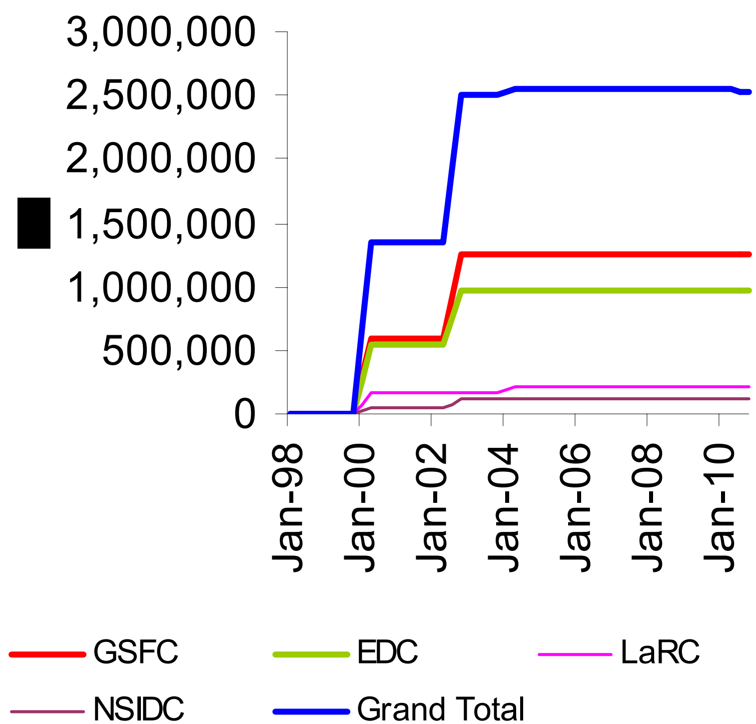
Contributing Element	Phase 1	Phase 2	Phase 3	Phase 4
Hardware	High Initial Acquisitions	Neutral Improved Cost Performance of Storage Environment	Lower Improved Cost Performance of Storage Environment	Low Improved Cost Performance of Storage Environment
Software	High Initial Acquisitions	High Functionality Evolution	Lower Functionality Enhancements	Low Functionality Complete
Personnel	High Establishing Operations	Neutral Improved cost performance of operations	Neutral Distributed data management costs incurred	Low Operations routine
Availability	Neutral Distributed Data Storage. Backup Only	Lower Site Interoperability Established	Low Site Interoperability Routine	Low Site Interoperability Routine
Performance	Neutral Sites Unconnected	Lower Site Interoperability Established	Low Network cost performance improvements	Low Network cost performance improvements
Facility	High Establishing Operations	Neutral Improved cost performance as build-out costs depreciate & storage footprint improves	Lower Improved cost performance as build-out costs depreciate, automation increases, & storage footprint improves	Low Improved cost performance as build-out costs depreciate, automation increases & storage footprint improves

Mission Model Description

- **Mission Modeling Variables –factors that affect the volume of data being generated at the Data Centers**
  - **Mission Assumptions:**
    1. Satellites: Terra; Aqua; Aura; ICEsat (significant volume products only)
    2. Products: L0 – L1A (archive only); L1B-L4 (Archive & Production)
    3. Volumes: Mission daily production baselines
    4. Duration: Design Life of Mission x 1.5
    5. Reprocessing: 7x production (retained for 6 months)

Predicted Data Volumes

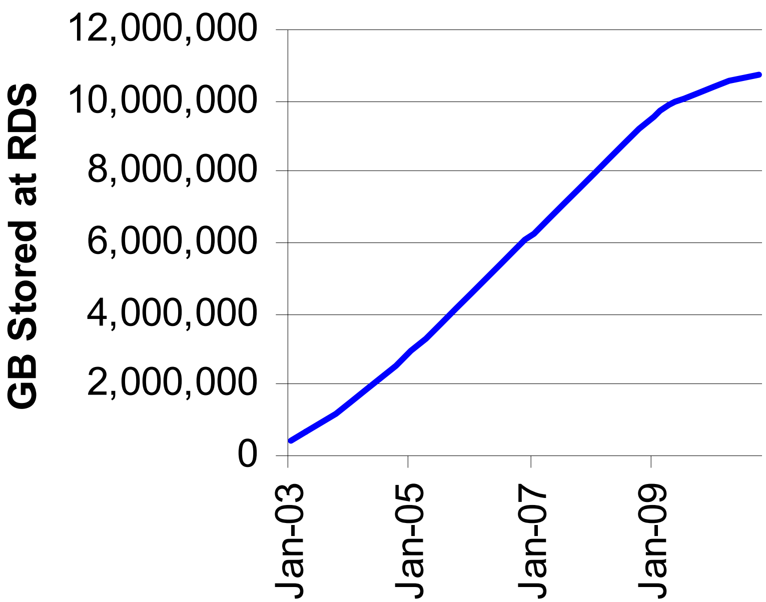


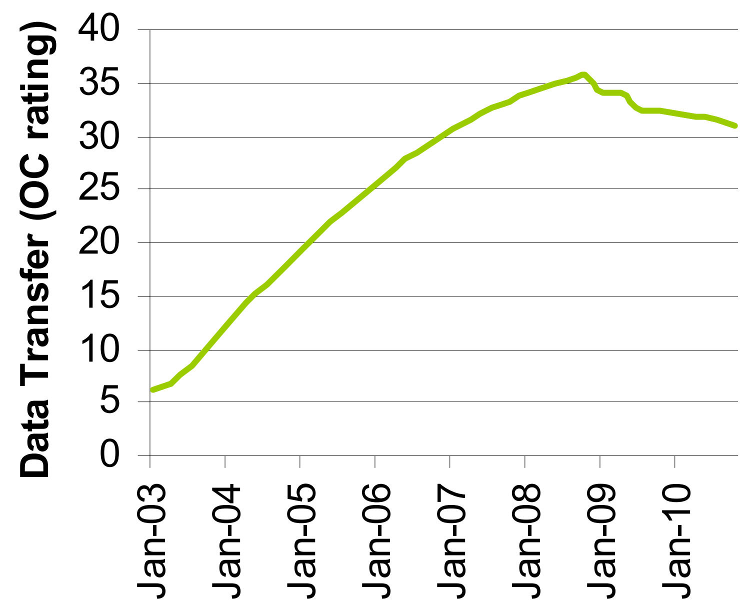


Persistent Volume – Growth projected into the future.

Transient Volume – Capacity for Reprocessing

Capacity Ramp Up





RDS Total Storage – Notional plan to achieve 100% mid-2010

Communication Capacity – User access modeled at 1x production

Staffing Model

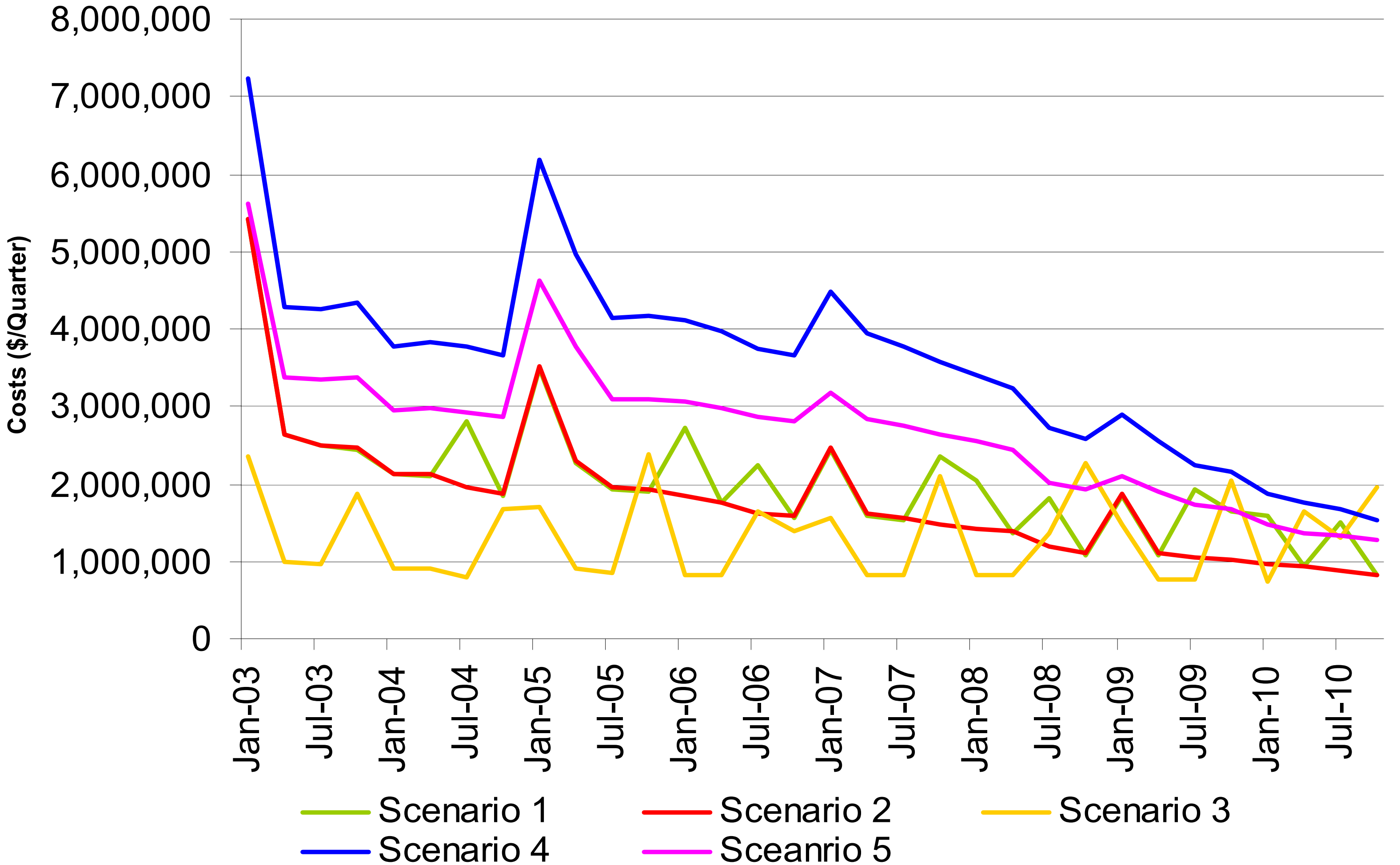
- **Staffing levels associated with hardware NOT volumes**
  - Assumes a significant improvement in the enterprise management of disk systems
  - Assumes a technology refresh rate that permits "genocidal sparing"

Activity	Count	Shift Work
Disk Rack Mgmt.	5 Racks/Staff	Yes
Tape Mgmt.	1 Staff	Yes
Communications Mgmt	1 Staff	No
Storage Policy Mgmt.	1 Staff	No
System/Install Engineer	1 Staff	No
Center Manager	1 Staff	No
System / Storage Administrators	2 Staff	Yes

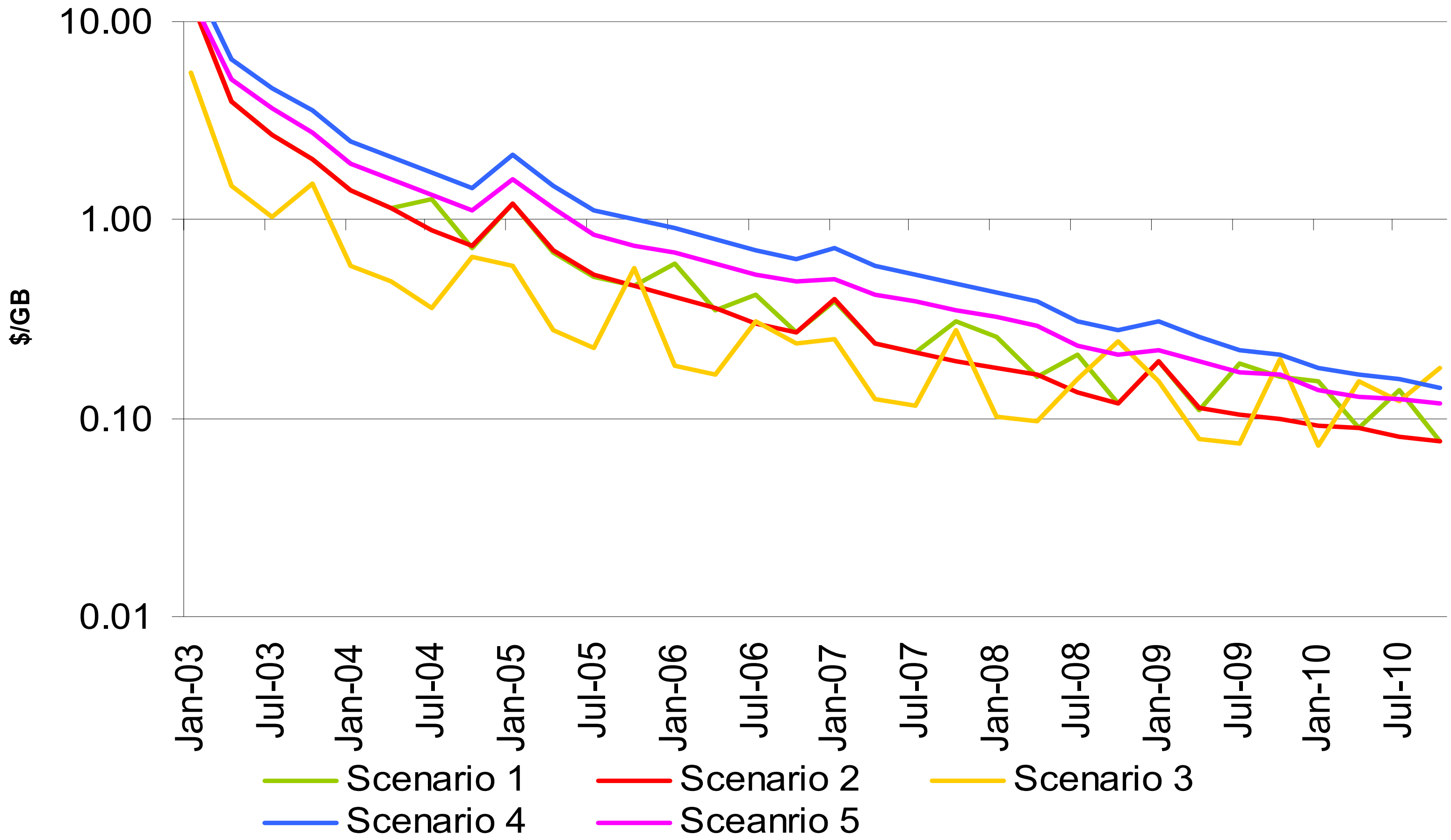
Modeling Scenarios

Scenario	Longer-Term Archive	Short-Term Archive	Comment
1	Near-Line	On-line (SCSI)	Traditional static data architecture. Adequate support for “re-processing campaign” access, poor support for extensive random archive access applications.
2	Off-Line	On-line (SCSI)	Used when the archive is accessed in large “campaign” sized chunks.
3	Near-line	20% On-line (SCSI) 80% Near-Line	Same concept as Scenario 1, but for data access patterns that drop off even more steeply.
4	On-Line (SCSI)	On-line (SCSI)	The high performance solution for the fastest possible access at all times to all data.
5	On-Line (ATA)	On-Line (SCSI)	Compromise between Scenario 1 & Scenario 4. Provides higher performance than near-line solutions at a better cost performance than all SCSI.

Hardware Costs



Total Cost Performance



Modeling Preliminary Conclusions

- **The results of taking a Total Cost of Ownership (TCO) approach to model large-scale (multi-petabyte) scientific data archives has led to the following conclusions:**
  - On-line storage staffing needs is the most critical cost factor in future data center TCO.
  - Facilities costs are and will remain an insignificant cost consideration at 5-10% of the TCO.
  - ATA vs. SCSI can offer considerable savings in the near-term for on-line storage, and will be competitive with near-line tape storage in the longer-term.
  - In the longer-term, if staffing is manageable, the TCO for data storage will become relatively insensitive to the storage media choices.

Summary

- **The development of long-term data storage solutions for EOS data that supports the changing needs of the science communities is an on-going activity**
- **Constantly changing technology is a way of life, and should not hinder the start of this work**
- **RDS represents a solution path that provides both a practical solution path using today's technology, and a platform from which to evolve solutions as technology changes**