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

Steve Marley Ph.D.
Raytheon Company
Stephen_R_Marley@raytheon.com

Mike Moore
NASA GSFC
mike.moore@gsfc.nasa.gov

Bruce Clark
EDS
bclark@eos.east.hitc.com

Abstract

NASA's Earth Observing System (EOS) Program is now collecting unprecedented volumes of data (nearly one petabyte/year) to aid the nation in its understanding of the Earth's near and long-term climate processes. These data are a national resource that must be carefully preserved to maximize the nation's return on the EOS Program. To address this need NASA has initiated the development of a Remote Data Store (RDS) backup and recovery capability that will operate independent of, but closely allied to, the current Distributed Active Archive Centers (DAACs). This paper outlines the following: the evolutionary technology path that will ultimately provide automated, secure, seamless and efficient remote on-line redundant storage; and the recovery and access of operational EOS mission data products. Significant factors that affect the total cost of operations are discussed, as well as emerging technologies and standards. Preliminary modeling points to the operational staffing levels as a dominant cost component. If on-line storage management techniques cannot improve to the point where a small staff can manage petabytes of data, the viability of disk-based storage solutions for large scientific data repositories is unlikely.

1. Introduction

NASA's EOS Data and Information System (EOSDIS) has been collecting data operationally since the launch of Landsat 7 in April 1999. Since then, the Terra and Aqua satellites have been successfully launched. Currently, the four primary DAACs (EDC, GSFC, LaRC, & NSIDC) are archiving raw and processed products in excess of three terabytes each day. See [1] for more details concerning DAACs.

The EOSDIS Core System (ECS) is the heart of the data processing and archive at each of the DAACs. Engineered by a team of aerospace partners and led by Raytheon, ECS was developed over a 10-year period to be the world's largest civil satellite data and information system. Details of the system design are available from the EOSDIS Core System Project [2]. At a high-level, the user services for each DAAC can be divided into two components:

- A traditional "search-and-order" style interface to data stored in the deep near-line archives; and
- A newer "navigate-and-get" style interface to a SAN based on-line data store of recently generated data products.

This second access mechanism, known as Data Pool (DPL), was developed in recent years in response to an increased demand from an increasingly broader user community for a simpler but more capable method for accessing EOS data [3]. RDS will use Data Pool as its interface to EOS data.

The purpose of the EOSDIS Remote Data Store (RDS) project is to demonstrate and ultimately implement an integrated scientific data resource solution that uses generally available Commercial-Off-the-Shelf (COTS) hardware and software storage technologies to provide enterprise-class data transport, access, security and scalability. The overall context is shown in Figure 1-1. The four existing DAACs currently exchange mission critical data through private government networks. RDS will permit the exchange of on-line data through public networks to provide a remote backup of this same data. It will be implemented across four phases, using an approach that tracks the ongoing evolution and maturity of storage technologies that support or complement the RDS mission.

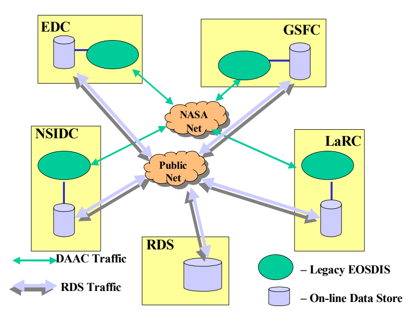

Figure 1-1 RDS Context

In Phase 1, the RDS project will demonstrate data storage interoperability (i.e. data exchange) between the Earth Science (GES) DAAC on-line data store at NASA's

GSFC facility in Greenbelt, MD and an RDS solution implemented at NASA's IV&V facility in Fairmont, WV.

In Phase 2, data storage functionality will be extended to provide site-to-site data recovery in conjunction with a shared data storage environment where local and remote users and applications can view and access data located anywhere within the implemented RDS storage environment. In addition, in Phase 2, the remaining ECS DAAC on-line data stores will be integrated into the RDS storage environment.

In Phase 3, the shared RDS storage environment will be enhanced to provide a managed unified storage capability, where physical location dependencies have been abstracted from the user view. In addition, the access to data in the RDS storage environment will be improved such that DAAC, user, and third party application data needs are serviced most efficiently, regardless of where the data is stored in the RDS / DAAC enterprise.

Finally, in Phase 4, the RDS storage enterprise will be further enhanced to provide access to existing near-line data archive systems through the implementation of a unified view of the RDS enterprise data allowing for more seamless data access.

Phase 1 is currently being deployed. The timing of Phases 2 & 3 are technology-driven, whereas Phase 4 is mission driven.

Phase 2 deployment will depend on the successful development of a domain access client built on top of the solution for Enterprise Data Management (see Figure 3-1). For this reason it is not expected that Phase 2 will be fully implemented until mid-2004.

Phase 3 is driven by the ability of distributed network storage devices to interoperate over IP WAN. Assuming that protocols are definitized during 2003 and that 24 to 36 months will be required to develop reliable commercial products based on those standards, we would expect to move to Phase 3 sometime during 2006.

Phase 4 can occur anytime after Phase 2. Indeed, the ability to make tape libraries available on-line is as much a commitment to do so as it is a technology question. That said, the tape archives are fulfilling their current mission, and we do not see a driving need to have the archives directly accessible until after Phase 3. Other mission drivers will require archives to be directly accessible sometime between 2006 & 2008.

2. RDS Vision

The ultimate goal of the RDS is:

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

Specifically we have interpreted this goal as follows:

Automated – Operations are a key cost driver for longer-term data archives. To the maximum extent practicable, the RDS needs to leverage storage management tools to minimize operational and 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 transparent access to data independent of storage location.

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.

3. Architectural Components

When defining the solution architecture, the following key elements need to be integrated (see Figure 3-1):

- **Persistent, secure physical storage:** Both longer-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 stored securely.
- **Cost effective static longer-term 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 software environment and operational labor are all-important drivers. The impact of longer-term archive architecture on TCO is discussed in detail later in this paper.

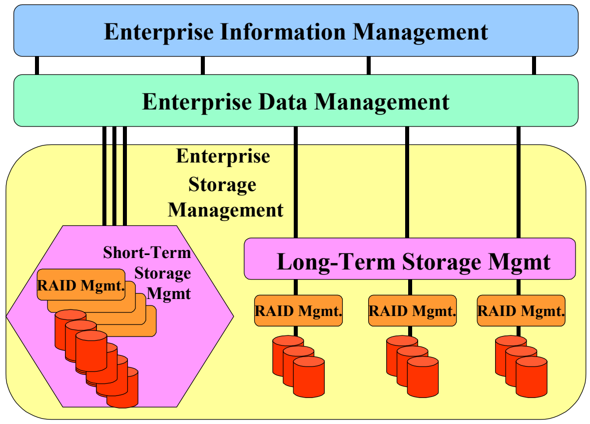

Figure 3-1 RDS Architectural Components

- **Higher performance dynamic short-term 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 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 short-term storage (i.e. what data is promoted from the ECS near-line archives), the efficient tactical management of data and service integrity across the enterprise needs to be administered through software at the enterprise level, in support of the strategic inter-DAAC agreements and management mandates.
- **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.

4. Technologies & Standards Drivers

RDS will have to interface with heterogeneous data storage systems as a matter of practical reality:

- There is a significant legacy data store in ECS;
- Legacy storage solutions are currently implemented as SANs and HSM tape solutions with a mixture of COTS products as well as custom code. There is no assumption that this architecture will remain valid into the future;
- Storage systems will evolve (both in technology and management) and so all solutions will become obsolete and need replacing;
- The same can be said for server technology;
- Lastly, individual data centers may choose different evolutionary paths and schedules in the future based on their perceived priorities. So, not only will technologies evolve across the enterprise, they will evolve at different rates, and in different directions.

The most practical approach (from a customer's perspective) to dealing with heterogeneity is to encourage the adoption of a common standards framework. This will ensure that not only will components interoperate within a particular vendor's solution set, but they will also interoperate with other vendor components. However, the definition and adoption of standards is a complex and drawn-out process where strategic business priorities often dominate over ideal technical solutions. Any single customer will have little practical impact on either the direction or speed of the standards process, as larger market forces are dominant.

RDS cannot afford to be beholden to standards processes, nor can it define and adopt its own standards, and expect to be compatible with standards as they emerge. The RDS needs to adopt a more evolutionary approach.

In the early phases, when storage system interoperability is at a minimum, the Enterprise Information Management layer services will have to provide interoperability through domain-based applications. As the storage industry begins to develop standards and frameworks, then the task of heterogeneous storage management can be pushed down the storage stack. There are several trends in industry standardization that will help:

- Storage Management – initiatives such as the Storage Management Initiative (formerly Bluefin [4]) are working towards a common management framework for storage solutions.
- Storage Architectures – Object-based Storage Devices (OSD) [5] and object-based file systems such as Lustre [6] are being investigated in SNIA and various academic institutions.
- Protocol Initiatives – iSCSI, iFCP, FCIP [7], for example, will permit standardization of block-level protocols over IP
- Enterprise Data Management – a number of companies have initiatives to provide EDM across their storage solutions. One of the more interesting Open Source initiatives is the Data Grid (for example see [8,9,10]).

Some of these initiatives will gain market-place acceptance, and others will not. Although RDS will benefit from the successful adoption of standards, it will not attempt to anticipate which initiatives will be adopted by the industry, but will wait until those standards are adopted, and products built to those standards are generally available. In this way, RDS will be somewhat protected from being drawn into a technological blind alley which is either expensive to operate, or too expensive (or disruptive) from which to extract itself.

The RDS solution will evolve via a clear approach that demonstrates how new standards-based technologies can be inserted into the solution to provide cost effective evolution from initial capabilities through to Phase 4 functionality.

5. Technology Roadmap

RDS intends to implement generally available and mature storage and data access technologies that represent "best of breed" solutions in its multi-phased approach. Because of the evolutionary nature of shared storage and distributed storage technologies, RDS intends to implement "leading edge" rather than "bleeding edge" technologies. It is the intent of RDS to maintain a heterogeneous storage hardware architecture that takes advantage of advances in storage and data management products and enables RDS to take advantage of improvements in price and performance of storage components. All technologies implemented under RDS will be expected to meet applicable industry standards. Roadmap details by phase are (see Figure 5-1):

Phase 1 - Remote data storage across IP WAN Prototype. In this phase, the basic capability of data storage interoperability across Data Center installations (DAACs and RDS) at distributed locations will be demonstrated. Phase 1 is not expected to provide sufficient operational storage or bandwidth to provide a comprehensive backup service for all DAAC Data Pools. The Phase 1 Prototype will integrate with the operational GES DAAC Data Pool such that the connectivity between the Data Pool and the Prototype is considered operational, but the data flows are for demonstration purposes only, and not considered operational.

The architecture of the storage will depend on the available vendor solutions. However, it is our expectation that Phase 1 will consist of the basic architecture for both the short and longer-term archive capability. This architecture will need to be viable for the next 10 years, with a technology roadmap that could permit evolution to a viable longer-term archive.

Phase 1 may also consist of a shared storage installation to provide redundancy with the existing DAAC Data Pools. However, as heterogeneous storage component interoperability is somewhat immature and vendor dependent, this option will only be exercised if a viable interoperability roadmap is available.

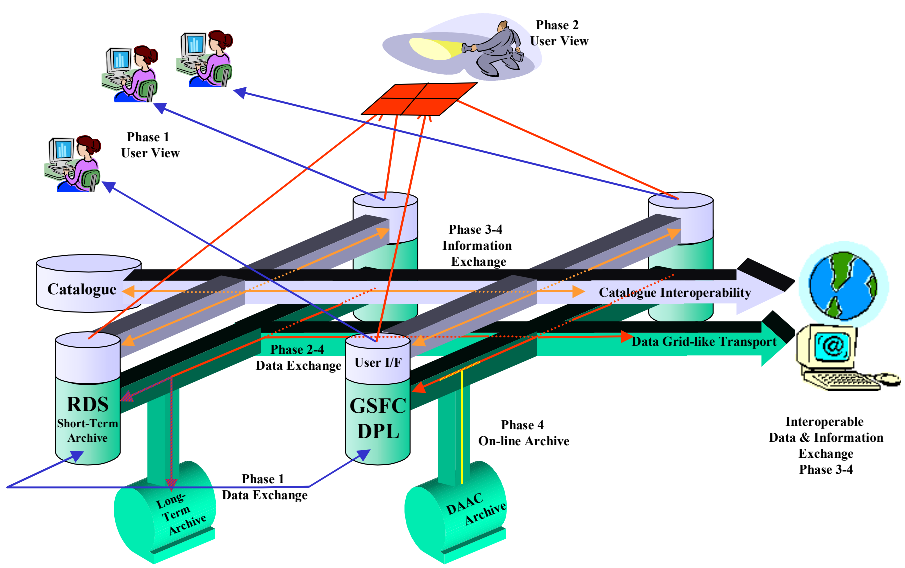

Figure 5-1 RDS Phased Implementation – The users view on the data holdings becomes increasingly coherent as the system evolves

RDS intends to implement appropriate storage management tools, using both software and hardware, to allow for the local administration, management and monitoring of the RDS data storage resources.

Higher-level enterprise services are not expected to be deployed in Phase 1. These technologies are considered formative at this stage, and a future prototyping effort will be used to validate any proposed solution against the longer-term vision for the RDS.

Phase 2 - Shared Data Storage and Recovery. In this phase, the basic storage capability will be enhanced to provide site-to-site data recovery and a seamless shared data storage environment where users and applications can view and access data located on any storage in the RDS network.

In this Phase, heterogeneous shared storage interoperability at the block-level will be prototyped (if generally available) for implementation in Phase 3.

RDS intends to implement enterprise data management tools if mature, using both software and hardware, to allow for the local administration, management and monitoring of the RDS data storage resources.

Phase 2 will also see the preliminary implementation of the enterprise information management tools that will provide a unified application view and services against the distributed data holdings.

In Phase 2, the remaining ECS DAAC Data Pools will be integrated as data sources into the RDS storage environment and the RDS data storage capacity and communications services will scale to meet the anticipated operational volumes.

Phase 3 - Redundant data storage with load sharing. In this phase, the shared data storage environment will be enhanced to provide a managed, unified storage capability, where physical location dependencies have been eliminated from the user view. Also in this phase, data is redundantly stored across the RDS / DAAC enterprise such that DAAC, user, and third party application data needs are serviced most efficiently.

Phase 3 will see the implementation of block-level interoperability over IP WAN for all the shared storage in the RDS / DAAC enterprise. Shared file systems and services across the shared storage will be prototyped, if available, for implementation in Phase 4.

RDS will continue to utilize enterprise data management tools with increasing levels of functionality to allow for the local and remote (in support of service continuity) administration, management and monitoring of the RDS and DAAC data storage resources.

This Phase will implement a closer integration of the enterprise information management tools and services with the enterprise data management tools as those tools become more capable and automated.

Phase 4 - On-line remote archives. In this phase, the RDS storage network will be further enhanced to provide access to existing near-line data archive systems, through the same unified, seamless access methodology.

Phase 4 will see the implementation of common shared file systems across the RDS/DAAC enterprise. This will facilitate both efficient enterprise data management and truly seamless enterprise information management.

RDS will continue to utilize enterprise data management and enterprise information management tools with increasing levels of functionality to allow for the local and remote (in support of business continuance) administration, management and monitoring of the RDS and DAAC data storage and archive resources.

To complete the RDS picture, we need to link the architectural concepts and roadmaps that we have described with practical details related to the RDS mission. In the following sections we take a cost modeling approach to modeling RDS storage architecture. We first describe some of the unique aspects of EOS data, and then develop the basis of an RDS data center model. Lastly we will discuss some of the preliminary results of that modeling.

6. Cost Model Drivers

A significant issue with balancing the potential solutions for an RDS-like system is the Total Cost of Ownership. Unlike commercial transaction-based solutions, the data stored in RDS has characteristics that impact the long term cost model. Specifically:

- **Replaceable Persistence** – The derivation of meaningful physical results from raw satellite observations is an intricate process that matures as the scientific understanding of the data and the physical process matures. The net effect is that although physical concepts such as "average sea surface temperature fluctuations in the mid-Atlantic" are persistent, the data products that contribute to that physical concept are continually being reprocessed and improved. Thus, the high level concepts have to be re-mappable to changing source data.
- **Bulky** – Individual data products may not be that large (rarely exceeding a gigabyte). However, when considered as a data set some logical products can exceed a petabyte.
- **Static** – Unlike transaction-based data applications, scientific data is static. By this, we mean that the individual data records are not in a continuous state of flux, being updated, deleted or overwritten. As described above, there is a constant background of reprocessing that will replace whole data sets, but within a given data set products are written once.
- **Evolving Usage** – As data sets mature, the usage patterns for them change considerably. Usage

shortly after acquisition is predominantly for time-sensitive applications such as higher-level product generation and emergency response. As the data becomes dated then the focus shifts towards long-loop analytical applications, where geospatial association is increasingly relevant.

Total Cost of Ownership (TCO) is usually defined as "the comprehensive cost / benefit analysis of a particular technology delivery." The contributing elements of TCO are:

- **Hardware**: includes racks, servers, disks, cabling, acquisition costs, and maintenance costs.
- **Personnel**: includes operations staff, maintenance support staff, vendor staff, and consulting staff.
- **Facilities**: includes equipment floor space, power, cooling, and staff floor space.
- **Communications**: includes the provisioning of external capacity, security and redundancy.
- **Software**: includes acquisition, development and integration costs, maintenance costs, license management, upgrades, and monitoring.
- **Availability**: includes component MTBF, system MTBF, and enterprise MTBF.
- **Performance**: includes time to first byte, total volumes served, and time to complete.

Each of these factors will evolve slightly differently throughout the lifetime of RDS:

Hardware, Personnel and Facility costs are high in the beginning as the operational environment is established. Although costs will come down as storage device density and maintainability continue to improve, the recurring costs for these elements will continue to provide the most significant cost model contributions.

Communications costs are similarly high in the beginning as service is established. However, unlike the costs above, communications costs are directly related to volume of "business" being transacted as opposed to data volume being stored. In addition, the cost of service is expected to come down as providers' capital costs are depreciated, and the broadband market develops.

Software costs are also high in the early phases, and will not tend to show significant reductions until after all the necessary functionality is developed, at which point industry standard software maintenance models will apply.

Availability and Performance costs are highly dependent on the mission requirement. RDS is pre-operational in the early phases. So, although the cost of provisioning a particular level of service is expected to decrease over time, the actual required level of service will increase as RDS moves towards operations.

7. Cost Model Description

The modeling of RDS cost elements requires a detailed understanding of the many contributing variables. These variables can be classified as follows:

- **Mission Modeling Variables** – these relate to the factors that affect the volume of data being generated at the Data Centers and include the baselined production volumes at each product level, the duration of the satellites and reprocessing policy.
- **RDS Data Center Variables** – these relate to the systems deployed at the data center and include storage models; data center capacity ramp up; staffing models; and storage policy (on-line vs. near-line vs. off-line). Communications are modeled from capacity perspective, but we have not attempted to cost the provisioning of the service. The remaining cost contributors (software, availability, performance) have not been modeled at this stage.

These two classes are discussed in detail below.

Mission Modeling Variables

The DAACs serve a number of purposes within NASA's EOS mission. Their first priority is to provide mission data production and archive services. However, the secondary priority of providing timely access to those same data products is becoming of increasing importance as the science community becomes more familiar with those products. In addition, as described below, the access requirements of data reprocessing are not insignificant.

The EOS philosophy has, therefore, always been one of a heterogeneous archive environment where both the higher performance short-term and higher capacity longer-term storage needs are addressed individually.

Longer-Term Archive

For the purposes of sizing the RDS storage, we have made the following assumptions to simplify the scope:

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

Mission life is not readily determinable. However, it is noted that both of the first two missions (Terra & Aqua) were inserted into orbit with minimum fuel cost, and NASA has traditionally engineered its satellites to last well beyond design life. As can be seen in Figure 7-1, these assumptions lead to a mission-life archive requirement of in excess of eight petabytes by 2010.

Short-Term Archive

The longer-term archive requirement is not the only demand for storage in the EOSDIS mission. The data systems also need to support user access and a considerable amount of reprocessing. Reprocessing is a side effect of the scientific process. As processing algorithms and instrument calibration improves, the data already acquired continually undergoes regular reprocessing. Reprocessing rates are expressed in terms of a multiplier of the production rate required to keep up with the satellite data acquisition rate. Although originally designed for more, the ECS system was originally implemented for a 2x reprocessing capacity. The processing capacity has been steadily improved over the years as more capable technology has become available, and as the scientific understanding of the data has improved. Currently ECS can support a sustained 3x capacity with a 7x peak capacity.

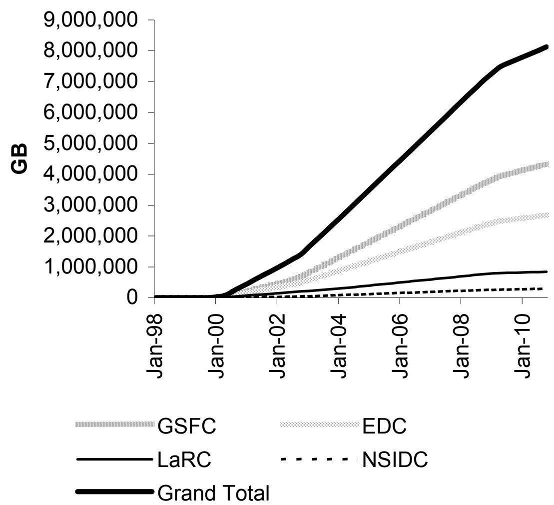

Figure 7-1 Longer-Term Archive – Growth over the last 2 years and projected into the future.

Reprocessing has a significant impact on the short-term storage requirements because there is a delay between the generation and archival of the reprocessed data of up to 6 months, during which time the data is held awaiting validation. At the end of 6 months, the new products are either deleted if proven faulty, or archived in place of earlier versions. The impact of reprocessing at 7x is shown in Figure 7-2.

Although the short-term archive capacity remains reasonably constant throughout the mission life, at 2.5 petabytes it represents a significant fraction of the total storage requirement and actually dominates the storage in the early years.

The mission characteristics are summarized in Figure 7-3 by each site. Reprocessing is the major contributing factor to all these characteristics except the End-of-Life Capacity.

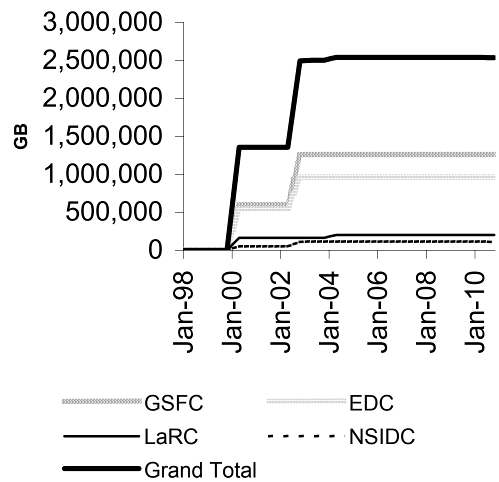

Figure 7-2 Short Term Archive – Capacity for Reprocessing at 7x

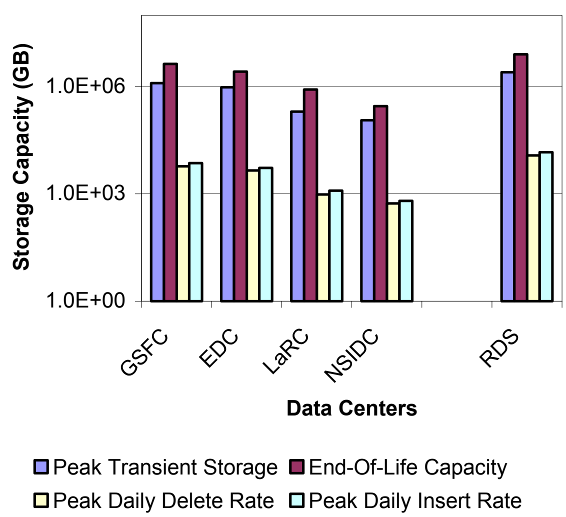

Figure 7-3 Mission Characteristics – Key peak capacities

RDS Data Center Variables

Next, we discuss the key modeling variables that affect the RDS data center, including storage density trends, capacity ramp-up, and staffing and facilities costs.

Storage Density Trends

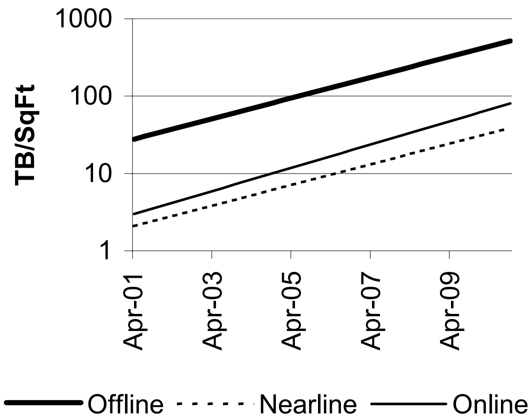

Figure 7-4 Storage Density Trends – Fully burdened machine room density.

All current data indicates that storage densities will continue to double about every 2 years into the future. Taking this into account and interpolating from current densities yields the curves in Figure 7-4.

Here we plot the curves for on-line, near-line and off-line. We have taken full account of the associated hardware (racks, robots, shelving etc.) when calculating the footprint.

Capacity Ramp-Up

The phasing of the RDS capacity ramp-up to full End-of-Mission capacity is somewhat arbitrary, as it is not driven by any specific mission requirements. We assumed that full capacity in the RDS is not achieved until 2010, and so we applied a capacity ramp up as shown in Figure 7-5.

Folding this capacity ramp-up with the data insert and reprocessing discussed above and adding an additional 1x production requirement for data distribution out of the archive (i.e. every product generated is distributed once) enables us to model the communications requirement (see Figure 7-6).

Staff Modeling

Staff modeling for on-line storage is not trivial. There are no good metrics available for how staffing needs would scale into the future. If we take the current industry recommendation (1 Full Time Equivalent or FTE per 200TB), then as the on-line storage scales to 10PB or more, this would imply a staffing level of 50 staff per shift (assuming 24x7 operations). This is clearly impractical for data center operations of the scale we are contemplating.

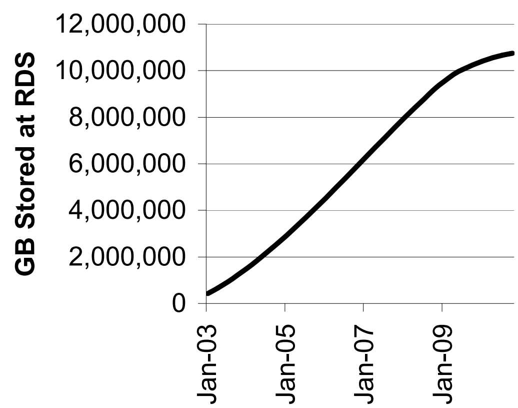

Figure 7-5 RDS Capacity Ramp-Up – Notional plan to achieve 100% by end of EOS mission life.

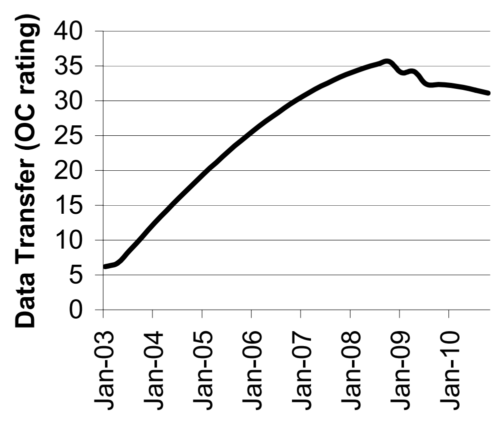

Figure 7-6 Communication Capacity – User access modeled at 1x production

Looking closely at the activities of operations staff for on-line systems reveals that staffing scales more with LUN management and physical hardware than it does with absolute volume. This is especially true for large-scale science data archives where data is static in nature, and not subject to the dynamic management that transaction-based systems require. In addition, because of the ever-increasing demand for on-line storage, we assumed a policy of "genocidal sparing" integrated with a constant data migration policy. Specifically, when we lose a device we migrate the data to another LUN before catastrophic failure, and simply leave the failed device in place, to be removed when the whole rack is retired (usually within 2 years at current industry rates). In this approach operations staff are not constantly rebuilding LUNs, but are instead spending their efforts bringing newer (i.e. higher capacity) drives on-line. Staffing can then be associated with the hardware being directly managed. The staffing model used for data center operations is shown in Table 1. In addition, office space will have to reserved for vendors who will be on-site almost constantly assisting with equipment commissioning and decommissioning.

Table 1 Staffing Model – Staffing requirements for standard data center operations exclusive of user services

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

Power & BTU Model

We assume for power and BTU that the power per rack of disk remains constant as densities increase. This is because as drive form factors change to pack more but smaller disks into the same rack the increase in the number of spindles in a rack is offset by a decrease in the power consumption per spindle.

Technology Insertion

Another variable in the modeling is the rate at which new technology is inserted. For the purposes of our modeling, we considered increasing storage density as the technology driver. It was recognized that media replacement was an important cost driver, and so we included migration to higher density media as a continual activity initiated two years after archive for data stored on disk, and four years after archive for data stored on tape media. We chose not to include optical media in our Longer-Term Archive model for this iteration as our opinion the cost curves are not sufficiently reliable for inclusion in a credible cost model without further study.

8. Model Results

We defined several storage scenarios (Table 2) and modeled the impact of each scenario on the data center costs.

Table 2 Storage Scenarios – Alternative approaches to data storage at RDS.

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.

For each of these scenarios we combined the RDS data center model described above with an internally developed hardware cost model based on current component costs and industry trending.

For each scenario we looked at three components:

- **Facilities** – including floor space costs and total power costs (including cooling and hardware needs)
- **Hardware** – including storage, racks, robots etc.
- **Personnel** – including shift staff and management.

Figure 8-1 shows the predicted facilities costs throughout the duration of RDS. The principle cost contributor, not surprisingly, is the power needs driven by the on-line disk storage. However, the floor space costs are not insignificant.

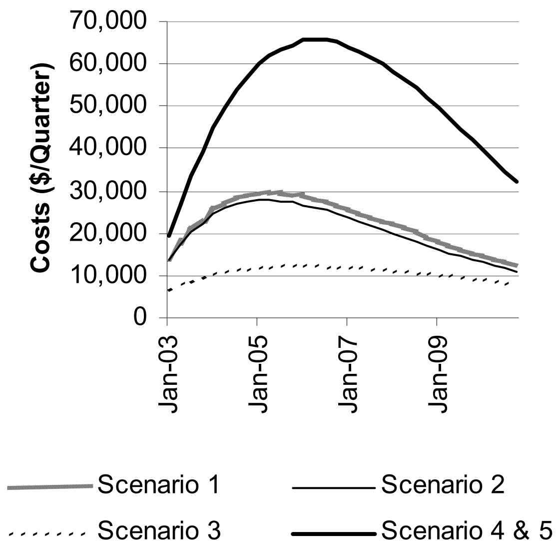

Figure 8-1 Cost of Facilities – Quarterly operational costs exclusive of build out, communications, personnel and hardware.

For scenarios 4 and 5 where only on-line storage is modeled, the total cost of floor space is dominated by the office space requirements of shift operations, and is comparable to the total facilities cost of the cheapest scenario (Scenario 3) which has the minimum on-line capacity.

Despite the broad variation in facilities costs displayed in Figure 8-1, facility costs are not the significant contributor to data center TCO. Both personnel and hardware costs are over an order of magnitude larger than the facilities costs.

Personnel costs (Figure 8-2) are dominated by the shift staff caring for the on-line storage. Although, initially this cost is small when compared to the hardware start-up costs of the data center (Figure 8-3), as the RDS matures and storage prices drop, personnel costs for all scenarios trend towards being a dominant cost even if we can assume that the improvement in operator staff productivity that was described earlier is achievable.

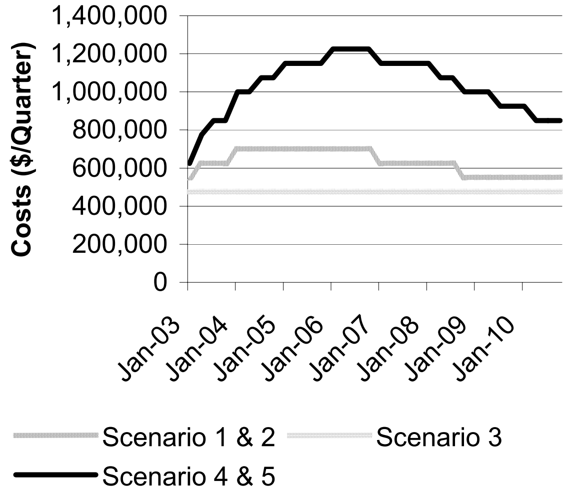

Figure 8-2 Cost of Personnel – Quarterly personnel costs exclusive of user data services.

Figure 8-3 shows the predicted hardware costs as RDS scales to several petabytes. These curves include the cost of technology insertion. Disk drives are swapped out after two years. Tapes and tape drives are modeled for replacement every 4 years. All scenarios show a predictable decrease in operational costs driven by the increase in storage density per device. However, the operational costs tend to converge in the out years despite the fact that the density curves for disk and tape media are in step (Figure 7-4). This comes about because the technology replacement and capacity enhancement costs for the peripheral hardware for tapes (drives, libraries, robots etc.) becomes the more significant contributor to the TCO of near-line storage. This is particularly noticeable in Scenario 3, where near-line storage services a significant fraction of the short-term storage requirement. As shown in Figure 8-3, for RDS beyond 2009 the hardware costs for on-line disk solutions (Scenarios 4 & 5) are comparable or lower than those for near-line tape dominated solutions (Scenarios 1 & 3). In addition, they are more predictable from quarter to quarter because the peripheral hardware costs can be bought in smaller increments.

However, the implication of this for operations is nevertheless critically sensitive to the operational staffing requirements of on-line disk. Figure 8-4 shows the operational costs in terms of $/GB for RDS into 2010. Using the staffing and facility models described earlier in this paper, Figure 8-4 shows that the TCO for a SCSI based on-line archive (Scenario 4) remains marginally more costly than a mixed tape/SCSI solution (Scenario 1 & 3). However, an ATA based on-line longer-term archive solution (Scenario 5) becomes very cost competitive.

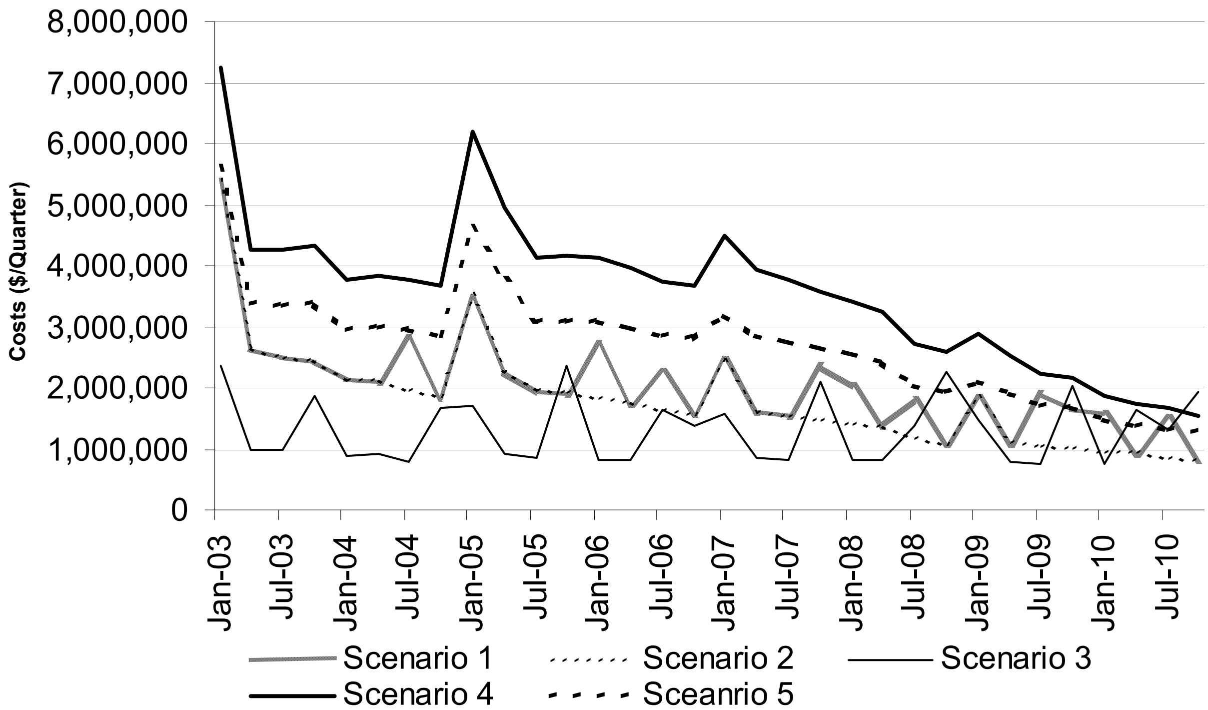

Figure 8-3 Cost of Hardware – Quarterly hardware costs inclusive of media replacement costs.

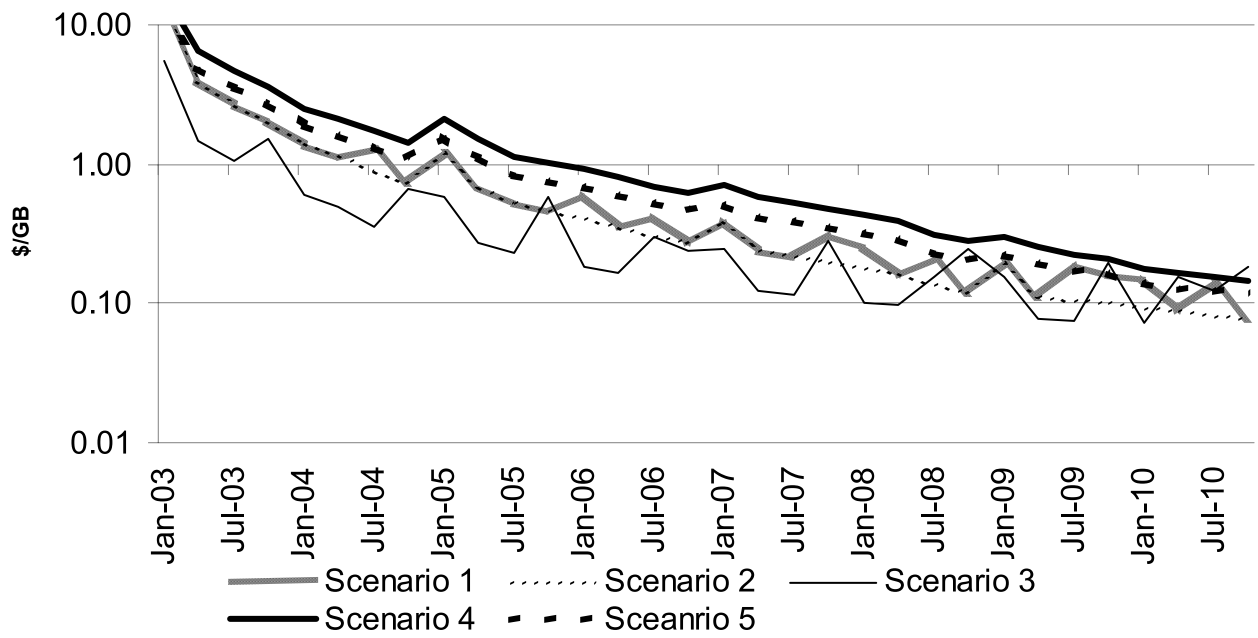

Figure 8-4 Consolidated Data Center Costs – Predicted operational cost trends per unit storage exclusive of communications and user services costs.

9. 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 percent 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.

In the future, the RDS project will begin to test some of the underlying assumptions and conclusions of this paper. The initial SAN-based short-term archive solution is in place and a Grid-enabled Content Addressable Storage (CAS) based longer-term archive prototype is currently being prepared for deployment. This will provide a significant test-bed capability against which to evaluate emerging storage and storage management solutions.

10. Acknowledgements

This work has been funded under NASA Contract NAS5-60000. The authors gratefully acknowledge the effort of many people in the review of this paper, and would like to recognize Thomas M Ruwart in particular for his contribution of the hardware cost model.

11. References

[1] *NASA DAAC Alliance* at http://nasadaacs.eos.nasa.gov/index.html

[2] Release 6B Segment/Design Specification for the ECS Project, *EOSDIS Core System Project* [305-CD-610-002], August 2002

[3] J.M. Moore & D. Lowe. Providing rapid access to EOS data via Data Pools, *Proceedings of the 47th Annual Meeting of* SPIE [4814-56], July 2002

[4] *Storage Network Industry Association: Bluefin* at http://www.snia.org/tech_activities/bluefin/

[5] *Storage Network Industry Association: OSD* at http://www.snia.org/tech_activities/workgroups/osd/

[6] *Lustre* at http://www.lustre.org

[7] *Storage Network Industry Association: IP Storage Forum* at http://www.snia.org/tech_activities/ip_storage

[8] *The Data Grid Project* at http://eu-datagrid.web.cern.ch/eu-datagrid/

[9] *The Globus Project* at http://www.globus.org/

[10] *The Information Power Grid* at http://www.nas.nasa.gov/About/IPG/ipg.html