DATA STORAGE AND MANAGEMENT FOR GLOBAL RESEARCH DATA INFRASTRUCTURES – STATUS AND PERSPECTIVES

Erwin Laure* and Dejan Vitlacil

PDC Center for High Performance Computing, Royal Institute of Technology (KTH), Stockholm, Sweden
Emails: erwinl@pdc.kth.se; dejan@pdc.kth.se

ABSTRACT

In the vision of Global Research Data Infrastructures (GRDIs), data storage and management plays a crucial role. A successful GRDI will require a common globally interoperable distributed data system, formed out of data centres, that incorporates emerging technologies and new scientific data activities. The main challenge is to define common certification and auditing frameworks that will allow storage providers and data communities to build a viable partnership based on trust. To achieve this, it is necessary to find a long-term commitment model that will give financial, legal, and organisational guarantees of digital information preservation. In this article we discuss the state of the art in data storage and management for GRDIs and point out future research directions that need to be tackled to implement GRDIs.

Keywords: Data storage, Data management, Data preservation, Standardization

1 INTRODUCTION

The Data Infrastructure challenge of moving from existing individual services to common interoperable distributed data systems, as defined by ISCU WDS (http://www.icsu-wds.org/), will allow data centres, services, and activities to become part of a new system that will be capable of ensuring long-term stewardship and providing quality-assessed data and data services to the international community.

As early as 1996, the task force on archiving digital information expressed the need for a trustworthy environment by declaring: "a critical component of digital archiving infrastructure is the existence of a sufficient number of trusted organizations capable of storing, migrating, and providing access to digital collections." (Task Force on Archiving of Digital Information, 1996)

In this article we review the state of the art in data storage and management for GRDIs in Section 2, develop a 10 years vision in Section 3, discuss current challenges in Section 4, and outline research directions leading towards this vision in Section 5, followed by a number of key recommendations in Section 6.

2 STATE OF THE ART

In order to enable continued access to data stored in digital electronic form, indefinitely into the future within acceptable limits, we need to protect it against both media deterioration and technological obsolescence. Storage media become obsolete as do devices capable of reading such media, and old formats and standards give way to newer formats and standards.

The first important issue is **time,** or more precisely, data lifetime. However, the meaning of data lifetime varies within multi-disciplinary environments. Long-term preservation for High Energy Physics is measured in decades, for life science in centuries, and for earth science in millennia. The preservation of digital data can be roughly divided into temporary, fixed-term, and perpetual preservation. For example, national or international systems of *digital archives*, as repositories of digital information that are collectively responsible for the long-term accessibility of intellectual heritage instantiated in digital form, are distinct from *digital libraries* in the sense that digital libraries are repositories that collect and provide access to digital information but might or might not provide for the long-term storage and access of that information.

The second important issue is **security.** This includes not only IT systems, such as servers, firewalls, and routers, and concepts, such as access to data, open content vs. defined access restrictions, etc., but also repository personnel and internal security, fire and flood detection systems together with suitable disaster preparedness and recovery plans. As with time, the meaning of security varies within different environments. For example, in life sciences patient data, copyrighted tools, and competing industries don't have the same meaning as in astronomy and astrophysics where there is an explicit policy on placing data in the public domain after one year or in high energy physics where security is often traded for performance and scalability. Similar differences can be found in **access** to data and **volumes** of stored data. Hence, different disciplines will need different storage architectures and different storage solutions, as pointed out by Girone (2011).

Current storage technology provides cheap, capacious, and reliable storage; however, demand for storage is constantly increasing, so do we really have storage technology that is ready for long-term stewardship?

The first problem we have to deal with is physical preservation of data – **bit preservation**. Storage vendors' claims about mean time to data loss (MTTDL), with examples of reliability estimates that span centuries or millennia, would make us think that the bit preservation problem is solved (Rosenthal, 2010). These claims are, however, projections based on how models of system components, such as disks and software, behave. The storage community needs to replace MTTDL with a metric that can be used to more accurately compare the reliability of systems in a way that reflects the impact of data loss in the real world. We need to define how well the bits need to be preserved, and we need to measure if actual systems achieve the required level of bit preservation.

Studies of disk replacement rates in large storage farms (undertaken by enterprises and institutions that collect petabytes of data with a long-term value) and storage failures show that manufacturers were very optimistic. Additionally, studies of silent data corruption in state-of-the-art storage systems show that we can try to control unintentional data change with mechanisms for change detection and correction, such as parity bits, checksums, cyclic redundancy checks, error-correcting code (ECC) memories, and multiple copies with quorum, but it is very unlikely that they will disappear completely (Kelemen, 2007). Therefore, apart from the expected bit error rate (BER) in network interface cards, low-level data corruptions exist in memories, disks, etc.; they are observed on a daily basis and have several origins.

Data can be corrupted in many ways: through hardware errors, data-transfer noise, firmware, software bugs, etc. Correction will cost time and money. We see, from time to time (e.g., with the introduction of new hardware, software, firmware, etc.), an increase in the corruption cases; a constant and careful monitoring of the situation is then required (Panzer-Steindel, 2007). The problem is often underestimated, as are cost and personnel involved. It is clear that more probes are required and increased monitoring needs to be done in order to increase bit-life. This, however, also means an increase of the originally required IO performance and also an increase in the computing capacity–in other words, an increase in the size and cost of data centres.

The second problem to solve is the selection of **storage media**, essentially disk vs. tape. Looking at the cost, typical disk-based solutions are about 15 times more expensive than typical tape-based solutions, and the cost of energy alone for typical disk-based solutions can easily exceed the entire total cost of ownership (TCO) for average tape-based solutions. Thus, tape typically provides better value in terms of TCO for the long-term preservation of irreducible images and binary data if one does not require fast access times (Reine, 2010). For access latency reasons much data reside on disks; however, not all of the data needs to be on disk (more than 95% of the data stored is rarely accessed beyond 90 days after it was created). To fit the different needs (access latency vs. cost), tiered storage architectures can be devised. More than acquisition cost, energy cost should be taken into consideration together with physical storage space, management, scalability, etc. Based on type of data, different technologies have to be taken into consideration, for example, primary disk mirroring, virtual tape libraries (VTLs), disk-to-disk (D2D), automated tape libraries, and disk-to-disk-to-tape (D2D2T) solutions (Beech, 2009).

Disk and tape markets have been shifting roles as disk encroaches on tape's traditional backup/recovery role, first by using faster disk arrays that appear as if they were tape libraries, i.e., VTLs, and then with disk-based deduplication to reduce the amount of storage needed. A combination of disk and tape storage may be the optimal strategy for many users to address these varied needs. A VTL or disk-based backup can provide the performance needed for recall of files for high-access applications. As data backed up to disk become infrequently or never accessed, they should be moved to tape for long-term retention. Tape technology can provide data security, compliance, and off-line protection (against viruses, hackers, system errors, operator errors, and so on) and a long-term, low-cost archive repository (Peters, 2011).

The tape world is developing in ways that make tape solutions more interesting; for example, the self-describing format of linear tape file system (LTFS) allows a tape to be treated like a disk or any removable storage media; also interesting is data integrity verification (Rector, 2011). Note that the National Center for Supercomputing Applications (NCSA) chooses a redundant array of independent tape drives (RAID) for data protection when costs of replication are prohibitive, as discussed by Butler (2011).

Several efforts aim at **federating** independent storage providers in a federated storage infrastructure, also known as Data Grids. For instance, the European Data Grid (Gagliardi, 2006) and EGEE (Laure, 2009) projects developed a pan-European computing and data infrastructure for eScience, driven primarily by the main pilot user community: the high-energy physicists who experiment on LHC and need to distribute, analyse, and store several petabytes of data every year. Access to storage and basic management of data exploits the storage resource manager (SRM - https://sdm.lbl.gov/srm/) interface. The European Grid Infrastructure (EGI - http://www.egi.eu) is now continuing these efforts, and a new European project, EUDAT (http://www.eudat.eu), aims at further consolidating the European storage infrastructure.

As one example from the US, the National Archives Center for Advanced Systems and Technologies (NCAST) developed the transcontinental persistent archives prototype (TPAP) (Moore, 2010), a research test bed for a distributed federation of different, and independently administered, computing platforms that interact as a single virtual repository. The original test bed was built on the concept of data grids and used the data-virtualization paradigm applied in the storage resource broker (SRB - http://www.sdsc.edu/srb/) software, which focused on strong consistency guarantees for the management of distributed data. A second-generation prototype was assembled based on the integrated rule-oriented data system (iRODS - https://www.irods.org), which enabled evolution of preservation policies through mapping of policies to computer actionable rules. With the present TPAP technology, it is possible to implement a preservation environment that is capable of enforcing preservation policies, automating administrative functions, and validating trustworthiness assessment criteria.

In addition to the technical aspects discussed above, frameworks for **auditing and certifying** trustworthy repositories are needed. These frameworks typically take into consideration several levels of certification: basic certification based on the data seal of approval (DSA - http://www.datasealofapproval.org/), extended certification based on self-audits using ISO16363 or DIN31644, and formal certification based on full external audits using ISO16363 or DIN31644. European efforts to build such frameworks (http://trusteddigitalrepository.eu) need to be put into international context to be truly applicable to Global Research Data Infrastructures. The end results of these efforts should be a set of criteria applicable to a range of digital repositories and archives, from academic institutional preservation repositories to large data archives and from national libraries to third-party digital archiving services.

For example, *Trusted Repositories Audit and Certification (TRAC): Criteria and Checklist* (http://www.dcc.ac.uk/resources/repository-audit-and-assessment/trustworthy-repositories) incorporates the sum of knowledge and experience, new ideas, techniques, and tools that result from cross-fertilization between US and European efforts. TRAC provides tools for the audit, assessment, and potential certification of digital repositories, establishes the documentation requirements for audits, delineates a process for certification, and establishes appropriate methodologies for determining the soundness and sustainability of digital repositories.

Finally, international **standards** play a crucial role in building solid and sustainable storage infrastructures. Industry organizations such as ARMA (http://www.arma.org/) and SNIA (http://www.snia.org/) are working on container and metadata standards for preservation and records management areas. For instance, SNIA is working on the extensible access method (XAM) initiative (http://www.snia.org/forums/xam) that standardizes the interfaces for storing, retrieving, managing, and searching data objects. XAM defines the general structure of XAM objects as a container for one or many data objects (documents, email, videos, images, audio, etc.) and metadata, i.e., information about the data objects (identifier, retention period, application that created the object, etc.). The XAM interface services include store and retrieve, search, manage, access control, and import/export possibilities. The difference between XAM and other storage interfaces, such as file systems, CIFS/NFS, or FTP, is that application-defined metadata fields are embedded with the actual data object inside the XAM object.

And a last word on **Cloud Storage**: thanks to their attractive “pay-per-use” model, clouds are increasingly considered viable alternatives to academic/community-driven data centres. Outsourcing data storage to commercial cloud providers can reduce IT and hosting costs and storage maintenance tasks, such as backup, data replication, and additional storage device purchases. These responsibilities are offloaded to the service provider, allowing organizations to focus on their own core business.

The key issue in using third-party cloud services is to ensure the conformance of the third-party provider with all the necessary functional, non-functional, and legislative requirements pertinent to certain communities. A key enabling factor for this will be the establishment of clear certification and auditing guidelines that cloud storage providers can be checked against in the same way as academic/community storage providers. Unfortunately, many of these requirements are currently implicit knowledge, and there is an implicit level of trust between the data communities and their storage providers. This tacit knowledge needs to be made explicit to give cloud-storage providers clear requirements they can respond to. In addition, reliability and security issues during wide-area data transfer need to be tackled, much as they have to be dealt with when using other federated storage infrastructures.

3 TEN-YEAR VISION

A common, globally interoperable, distributed data system, formed out of data centres, which may include academic, community owned, or commercial (cloud) data centres, that incorporates emerging technologies and new scientific data activities needs to be formed. These centres will have to be trustworthy under a common set of criteria. They will use common tools for the audit, assessment, and certification of digital repositories. They will have to have a common governing body that will support and coordinate this long-term mission.

What we need is global, multi-disciplinary, data-grid prototype activities that will—in tight collaboration with pilot user communities—give us a clear idea of all challenges. There is a need to establish this distributed data system as soon as possible in order to be able to analyse today’s projections and outcomes in 10 to 20 years.

In addition to technical infrastructure and capabilities, the long-term management of research data requires organizational sustainability to promote continuing stewardships. Providing a sustainable infrastructure for the preservation of research data requires organizational commitments, capacity, structures, and plans for data stewardship that are consistent with the missions of the organizations that accept the responsibility to serve in data stewardship roles (Downs, 2010).

4 CURRENT CHALLENGES

The more technological aspects of these recommendations can basically be delivered tomorrow. However, structural and financial questions, such as developing a bit-preservation service for trusted data retention, involve the concerted effort of data centres and funding institutions, especially if we are speaking about Global Research Data Infrastructure and distributed data system. They will also possibly involve several years of fine-tuning the structure and business model of such a service.

The key challenge is to define common certification and auditing frameworks that will allow storage providers and data communities to build a viable partnership based on trust. WDS's site certification (http://www.icsu-wds.org/wds-members/join-icsu-wds/criteria-membership-certification) is a first step in that direction.

5 RESEARCH DIRECTIONS PROPOSED

Beside the technical threats that long-term preservation of data is facing, we can't underestimate the importance of threats posed by organisational barriers caused by widespread uncertainty about legal and organisational requirements for managing the intellectual property that digital information represents. These latter threats might well prevent decisive preservation actions altogether.

The costs and the technical, legal, and organisational complexities of moving Global Research Data Infrastructures and digital information forward into the future raise great fear about the life of information in the digital future, namely, that owners or custodians who can no longer bear the expenses and difficulty will deliberately or inadvertently, through a simple failure to act, destroy the objects without regard for future use.

It is necessary to find a long-term commitment model that will give financial, legal, and organisational guarantees of digital information preservation.

It will be important to separate conceptually and physically data archiving and long-term stewardship of scientific data from data libraries. As with tiered storage architectures, we need long-term stewardship at tier 0, which shouldn't provide access to users but just to digital libraries at tier 1, which in their turn will then provide data to final users/consumers.

6 RECOMMENDATIONS

- Secure funding and sponsorship with respect to removing legal and economic barriers.
- Encourage the recognition of the global audit and certification approaches by all actors.
- Coordinate the appropriate organizations and individuals in the development of standards, criteria, and mechanisms for identifying and certifying repositories of digital information as archives.
- Engage actively in international policy efforts to design information infrastructure to ensure that longevity of information is an explicit goal.
- Monitor already existing technologies in use in long-term preservation centres and evaluate their weaknesses.
- Encourage fine turning of the structure and business model since this process will take many years.

7 REFERENCES

Beech, D. (2009) The evolving role of disk and tape in the data center, Best Practices for backup and long-term data retention. *Sylvatica Whitepaper*. Retrieved May 22, 2013 from the World Wide Web: http://www.datacenterscanada.com

Butler, M. (2011) Why RAIT for Blue Waters at NCSA*? Proc. 27th IEEE (MSST2011) Symposium on Massive Storage Systems and Technologies*, Denver, CO, USA.

Downs, R.R. & Chen, R.S (2010) Sustainable Governance for Long-Term Stewardship of Earth Science Data. *Proc. Earth and Space Science Informatics Workshop*, Fairfax, VA, USA.

Gagliardi, F., Jones, B., & Laure, E. (2006) The EU DataGrid Project: Building and Operating a large scale Grid Infrastructure. In Di Martino, B., Dongarra, J., Hoisie, A., Yang, L.Y., & Zima, H. (Eds.), *Engineering the Grid: Status and Perspective*. American Scientific Publishers.

Girone, M. (2011) EGI-InSPIRE Current Requirements and Outlook. *Proc. 27th IEEE (MSST 2011) Symposium on Massive Storage Systems and Technologies*, Denver, Colorado, USA.

Kelemen, P. (2007) Silent Corruptions. *Proc. HEPiX 2007 Spring Storage Day*, Hamburg, Germany.

Laure, E. & Jones, B. (2009) Enabling Grids for e-Science: The EGEE Project. In Wang, L., Jie, W., & Chen, J. (Eds.), *Grid Computing: Infrastructure, Service, and Application*. CRC Press.

Moore, R.W., Rajasekar, A., de Torcy, A., Conway, M., Ward, J., & Crabtree, J. (2010) NARA Transcontinental Persistent Archive Prototype, TR-10-04, RENCI Technical Report Series. Chapel Hill, NC, USA,

Panzer-Steindel, B. (2007) Data integrity (draft 1.3), April 2007. CERN/IT. Retrieved May 22, 2013 from the World Wide Web: http://indico.cern.ch

Peters, M. (2011) A Comparative TCO Study: VTLs and Physical Tape With a Focus on Deduplication and LTO-5 Technology. Enterprise Strategy Group White Paper. Retrieved May 22, 2013 from the World Wide Web: http://www.ithound.com

Rector, M. (2011) Tape Library Based Data Integrity Verification*, 27th IEEE (MSST2011) Symposium on Massive Storage Systems and Technologies*, Denver, CO, USA.

Reine, D. & Kahn, M. (2010) In Search of the Long-Term Archiving Solution Tape Delivers Significant TCO Advantage over Disk, Clipper Notes Report #TCG2010054RLH. Retrieved May 22, 2013 from the World Wide Web: http://www.oraclehplto.com

Rosenthal, D.S.H. (2010) Bit Preservation: A Solved Problem? *International Journal of Digital Curation* 5(1), pp 134-148.

Task Force on Archiving of Digital Information (1996) Preserving Digital Information. Mountain View, CA: Commission on Preservation and Access and Research Libraries Group. Retrieved May 22, 2013 from the World Wide Web: http://www.oclc.org

(Article history: Available online 1 July 2013)