

# Archival Services and Technologies for Scientific Data

**Jörg Meyer, Marcus Hardt, Achim Streit, Jos van Wezel**

Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

E-mail: joerg.meyer2@kit.edu, marcus.hardt@kit.edu, achim.streit@kit.edu,  
jos.vanwezel@kit.edu

**Abstract.** After analysis and publication, there is no need to keep experimental data online on spinning disks. For reliability and costs inactive data is moved to tape and put into a data archive. The data archive must provide reliable access for at least ten years following a recommendation of the German Science Foundation (DFG), but many scientific communities wish to keep data available much longer. Data archival is on the one hand purely a bit preservation activity in order to ensure the bits read are the same as those written years before. On the other hand enough information must be archived to be able to use and interpret the content of the data. The latter is depending on many also community specific factors and remains an areas of much debate among archival specialists. The paper describes the current practice of archival and bit preservation in use for different science communities at KIT for which a combination of organizational services and technical tools are required. The special monitoring to detect tape related errors, the software infrastructure in use as well as the service certification are discussed. Plans and developments at KIT also in the context of the Large Scale Data Management and Analysis (LSDMA) project are presented. The technical advantages of the T10 SCSI Stream Commands (SSC-4) and the Linear Tape File System (LTFS) will have a profound impact on future long term archival of large data sets.

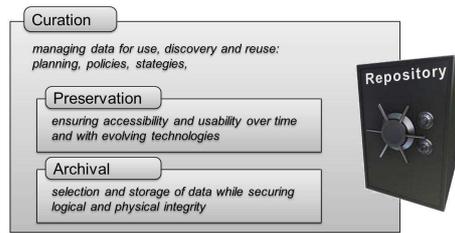
## 1. Introduction

Experiments, observations and simulations in various scientific fields produce more and more data. Data are a valuable good as often they cannot be reproduced at all or it takes a huge effort to generate them like in high-energy physics experiments [1]. Therefore, reliable archival services are indispensable. In addition, funding agencies like the German Science Foundation (DFG) recommend data archival periods of at least ten years. The following sections describe the chain from community support to data preservation and to modern tape technologies.

## 2. Large Scale Data Management and Analysis

Large Scale Data Management and Analysis (LSDMA) is a portfolio extension of the German Helmholtz Association with partners from four Helmholtz Centers including KIT, six Universities, and the German Climate Computing Center [2]. The aim of the project is to support communities in all respects of data life cycles. The close collaborations of community researchers and data experts from LSDMA takes place in so-called Data Life Cycle Laboratories (DLCLs). There are five DLCLs, called: Earth and Environment, Energy, Health, Key Technologies, and Structure of Matter. In addition to the DLCLs there is the Data Service Integration Team (DSIT) developing and providing generic services. DSIT comprises several work packages one of which is dedicated to the archiving topic. Its goal is to provide an archival service that allows





**Figure 1.** Relation between preservation, curation, and archival.

communities to easily archive their scientific data for example at the Large Scale Data Facility (LSDF) at KIT [3]. An archival system consist of a front-end that takes care of the ordered data input and a lower, bit preservation back-end which guarantees data integrity over time and also hides storage architecture changes from the front-end. The front-end can be tailored for different requirements or can be completely different for various communities. One of the first steps taken at KIT is the definition of a service and application programmers interface between the bit preservation level and the front-end. This computer center archive interface is a simple command list including put, get or delete, but will be further developed to include more sophisticated services.

### 3. Data Preservation

The term “curation” originally describes the work of curators in institutions like libraries, museums, or galleries. They are responsible for the selection and collection of documents and objects to preserve the cultural heritage. Digital curation is the analogy for digital objects, i.e. it involves data management for the (re-)usage and the discovery of data based on policies and strategies for the particular digital objects. Digital curation involves data archival and data preservation. Archival describes the selection of data to be stored and the actual long-time storage ensuring the logical and physical integrity of the data. The term “preservation” describes the aspects of ensuring the accessibility and usability over time and with evolving technologies. In Figure 1 the relationship of these terms is illustrated.

Magnetic tape systems provide storage capacities of many petabytes and are the least expensive technology to archive huge amounts of data. Other technologies like hard disks, RAM, or CPU caches and registers outperform tapes in access latency and bandwidth, however, they are much more expensive, provide smaller capacities, and consume more energy. Thus, tape libraries are the natural choice for long-term archives. The first step towards a sustainable archival service is to guarantee bit preservation, i.e. data will not get lost or corrupted. This is archived in two ways. Storage systems keep several replica of data. Corrupted data are identified by false checksums. The second way is to constantly monitor the readability of tapes. The Steinbuch Centre for Computing (SCC) at KIT operates several tape libraries used for the WLCG Tier-1 center GridKa, the LSDF, the University, and other projects. In Table 1, a list of libraries and their properties are given.

SCC applies two complementary tools to monitor the tape quality. The first one is a commercial tool from the Quotium company called StorSentry [4]. It runs in parallel to the backup or archival application and collects quality and performance related data during the production. The second tool was developed at SCC and reports recoverable and unrecoverable tape problems based on customized criteria. The source of information is a database with logging data from each unmount of a tape cartridge. Both tool can only identify problems with tapes being read

**Table 1.** Tape libraries at KIT.

library type	number of tapes	technology	capacity [PB]
GRAU	5800	LTO3, LTO4	3
Sun/Oracle SL8500	8000	LTO4, LTO5	8
IBM TS3500	6600	LTO4, LTO5	18

or written. In order to test all tapes a continuous reading of all data of all tapes independent of read/write requests of the production would be needed.

There is a trust relationship between data producers and data centers for the long-term data preservation. The Data Seal of Approvals [5] is an initiative to standardize the criteria for the quality of archives. The assessment is the result of a peer reviewed process. The approval guidelines are based on whether the data are discoverable in the internet, the data are accessible and have clear licenses and access rights, the data format is usable, the data are reliable, and the data can be referred to via some unique and persistent identifier. KIT is in the process of obtaining the Data Seal of Approval for part of its archives.

#### 4. Tape Technologies

The advance of recording density resulted in a doubling of the capacity of magnetic based storage every two to three years during the last two decades. This technological development is not limited to storage on rotating disk but happened also for storage on linearly written tape. In the shadow of numerous industry announcements of hard disk manufacturers about their latest products, storage capacity on tape has been increasing with similar pace albeit at a far lower absolute recording density than disk. Bit aerial density for tape is currently in the order of 35 Gbit/inch<sup>2</sup> whereas recent hard disks are in the order of 700 Gbit/inch<sup>2</sup>. However, density scaling for disks is slowing down whereas tape still has lot of headroom especially by increasing the number of tracks [6]. The latest tape drive technology is represented by the Oracle/StorageTek 10kD drive with a nominal capacity of 8.5 TB on a cartridge with a physical size far smaller than a state of the art 3.5 inch hard disk with a capacity of 4 TB. The listing in Table 2 shows that the increase in aerial density is for a large part contributed by the increase in track density, however the increase in linear density and improved compression algorithms are also an important factor. Conversely, an increase in capacity may be used for improved error correction coding while only moderately increasing the capacity of the standard media cartridge. The improved error correction will add to the number of benefits of tape for use in long time archival in which data integrity plays an increasingly important role next to its longevity and low power footprint.

A better protection against bit errors on tape must be supported by an improved and end to end error detection and error protection technology. The SCC-4 draft of the T10 Technical Committee of the International Committee on Information Technology Standards (INCITS) [7] includes a Logical Block Protection (LBP) for tape which allows easy verification of data. All drives in Table 2 can autonomously verify the contents of a cartridge if LBP has been used during write. The LBP safeguards the data path from computer to physical tape therefore eliminating errors introduced during data transfer. Backup and archive software such as IBM's Tivoli Storage Manager (TSM) can be instructed to write and verify using LBP.

Lastly, the new Linear Tape File System (LTFS) [8; 9] will prove to be significant for everyone storing data on tape for archival purposes. LTFS is an open source format for writing self-describing files on tape. The tape is written in two partitions, one storing an xml based record of the contents on tape and the second partition holding the actual data. The format not only

allows an exchange of media between unrelated sites and the easy reading and writing of media independent of the tape management system, but because of the unanimous adoption of the format by the tape industry reliable and vendor independent method for long time storage of digital data.

**Table 2.** Summary of properties of modern tape technologies.[10–16]

technology	capacity [TB]	I/O rate [MB/s]	bit-error rate	number of tracks	length [m]	brand
LTO5	1.5	140	$10^{-17}$	1280	850	consortium
LTO6	2.5	160	$10^{-17}$	2176	850	consortium
T10kC	5.0		$10^{-19}$	3584	1147	Oracle/STK
T10kD	8.5		$10^{-19}$	4608	1147	Oracle/STK
TS1130	1	160	$10^{-20}$	1152	825	IBM
TS1140	4	250	$10^{-20}$	2560	825	IBM

## 5. Summary

The number of data-intensive scientific communities is growing, leading to a demand for easy-to-use and sustainable archival services. Projects like LSDMA support communities the integration of the archival process as part of a typical data life cycle. Data preservation requires a sustainable and reliable quality of service. The Data Seal of Approvals intends to standardize the quality of archives at data centers. Bit preservation ensures the physical integrity of data. Besides storing distributed replica of valuable data, a constant monitoring of the storage media like tapes is required. The last section contains a review of modern tape technologies and a discussion of the advantages of the SCC-4 protocol and the linear tape filesystem.

## References

- [1] Blue Ribbon Task Force 2010 *Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access* (last visited on 2013-10-31) URL [http://brtf.sdsc.edu/biblio/BRTF\\_Final\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf)
- [2] van Wezel J *et al.* 2012 Data life cycle labs, a new concept to support data-intensive science (*Preprint arXiv:1212:5596*)
- [3] Garcia A O *et al.* 2011 *Proceedings of The 12th IEEE International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC-11/IPDPS-11)* 1467
- [4] Quotium webpage on storsentry (last visited on 2013-10-31) URL <http://www.quotiumdatasentry.com>
- [5] Sesink L, van Horik R and Harmsen H 2008 Data seal of approval, data archiving and networked services (dans) (last visited on 2013-10-31) URL <http://datasealofapproval.org/en/>
- [6] Argumedo A J, Berman D, Biskeborn R G, Cherubini G, Cideciyan R D, Eleftheriou E, Häberle W, Hellman D J, Hutchins R A, Imaino W, Jelitto J, Judd K, Jubert P O, Lantz M A, McClelland G M, Mittelholzer T, Narayanaswami C, Ölçer S and Seger P J 2008 *IBM Journal of Research and Development* **52** 513–528
- [7] Webpage of the technical committee t10 (last visited on 2013-10-31) URL <http://www.t10.org>

- [8] Pease D, Amir A, Real L V, Biskeborn B, Richmond M and Abe A 2010 in Khatib *et al.* [9] pp 1–8
- [9] Khatib M G, He X and Factor M (eds) 2010 *IEEE 26th Symposium on Mass Storage Systems and Technologies, MSST 2012, Lake Tahoe, Nevada, USA, May 3-7, 2010* (IEEE Computer Society) ISBN 978-1-4244-7153-9
- [10] IBM webpage on lto5 (last visited on 2013-10-31) URL <http://www-03.ibm.com/systems/storage/tape/oem/lto5/full-high/specifications.html>
- [11] IBM webpage on lto5 (last visited on 2013-10-31) URL [http://www-03.ibm.com/systems/storage/media/lto\\_5/specifications.html](http://www-03.ibm.com/systems/storage/media/lto_5/specifications.html)
- [12] IBM document on lto6 (last visited on 2013-10-31) URL [http://www-01.ibm.com/common/ssi/rep\\_ca/2/877/ENUSZG12-0272/ENUSZG12-0272.PDF](http://www-01.ibm.com/common/ssi/rep_ca/2/877/ENUSZG12-0272/ENUSZG12-0272.PDF)
- [13] IBM system storage lto ultrium 6 tape drive performance white pape URL <http://public.dhe.ibm.com/common/ssi/ecm/en/tsw03182usen/TSW03182USEN.PDF>
- [14] Oracle webpage on t10kc (last visited on 2013-10-31) URL <http://www.oracle.com/technetwork/articles/systems-hardware-architecture/st-t1000-ficon-perf-176630.pdf>
- [15] Oracle webpage on t10kd (last visited on 2013-10-31) URL <http://www.oracle.com/us/products/servers-storage/storage/tape-storage/t10000d-ds-1991052.pdf>
- [16] Oracle webpage on t10kd (last visited on 2013-10-31) URL [http://docs.oracle.com/cd/E48725\\_01/en/E20715/html/appendix\\_a.htm](http://docs.oracle.com/cd/E48725_01/en/E20715/html/appendix_a.htm)