The International Journal of Digital Curation
Issue 3, Volume 4 | 2009

The Significance of Storage in the "Cost of Risk" of Digital Preservation

Richard Wright, Ant Miller,
BBC Research and Development, London
Matthew Addis,
IT Innovation Centre, University of Southampton

Abstract

As storage costs drop, storage is becoming the lowest cost in a digital repository – and the biggest risk. We examine current modelling of costs and risks in digital preservation, concentrating on the Total Cost of Risk when using digital storage systems for preserving audiovisual material. We review the vital role of storage and show how planning for long-term preservation of data should consider the risks involved in using digital storage technology. Gaps in information necessary for accurate modelling – and planning – are presented. We call for new functionality to support recovery of files with errors, to eliminate the all-or-nothing approach of current IT systems, which in turn reduces the impact of failures of digital storage technology and mitigates against loss of digital data[1].

[1] This article is based on the paper given by the authors at iPRES 2008; received May 2009, published December 2009.
The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.

Significance of Storage

As storage costs continue to drop by roughly 50% every 18 months[2], there are two effects:

- **Storage looks free (but isn't):** the cost of storage devices becomes negligible, but power, space, cooling and management costs remain significant, especially in large data centres (Barroso & Holze, 2009).
- **Storage is abundant:** much more storage is used.

The following figure shows how hard drive storage has increased over the last 25 years[3].

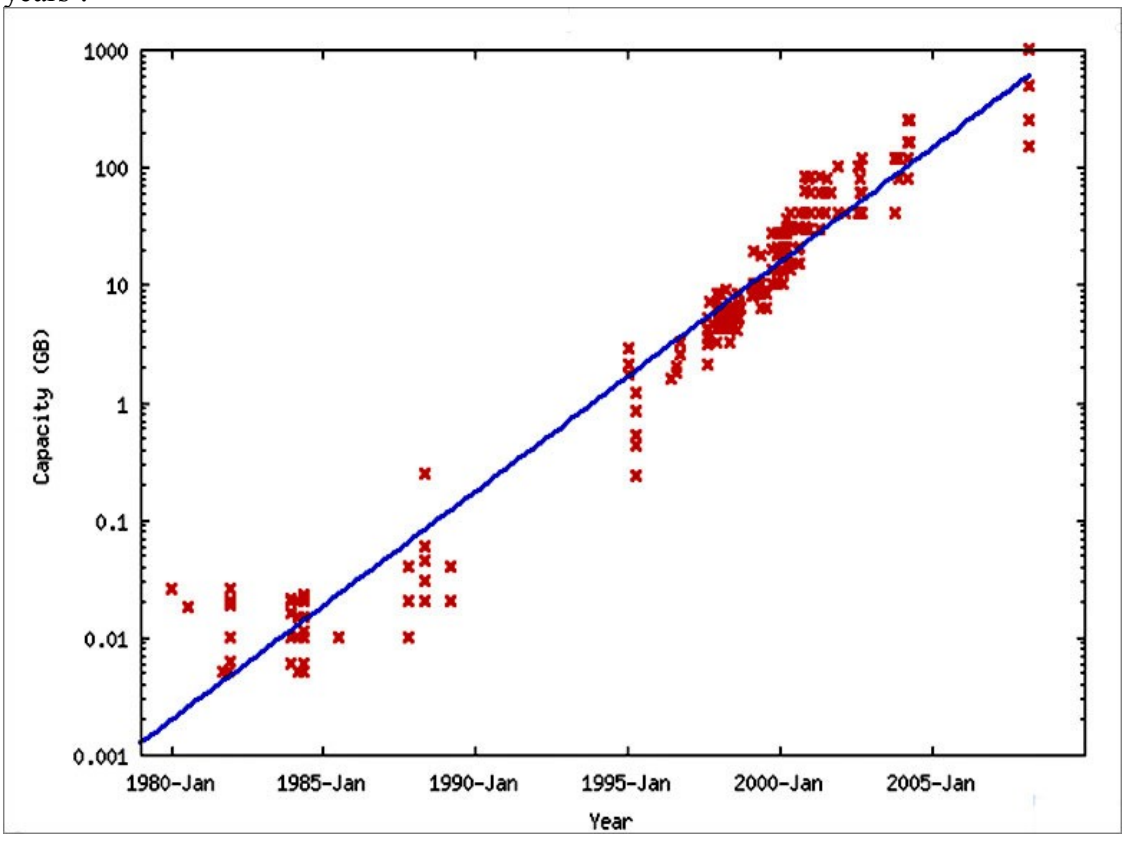

Figure 1. Increase in capacity of hard drives over the last 20 years.

The largest available disk size (for a desktop computer) has increased from 5 MB to one terabyte – a factor of 200,000 (which is about 18 doublings in about 25 years, so very close to doubling every 18 months).

The "growth of risk" is of course much larger: a factor of 200 000 in disk size, times the increase in the usage of disks (about 10,000 over the same period (Lawson, 2008)).

[2] Wikipedia definition of Moore's Law http://en.wikipedia.org/wiki/Moore's_law
[3] Wikipedia (image source) http://en.wikipedia.org/wiki/Image:Hard_drive_capacity_over_time.png

This "growth of storage" also divides into two effects:

- the number of storage units increases(globally, and number used by any given institution)
- the amount of data stored on each unit also increases

The increase in storage units results in an ever increasing number of users being responsible for, or dependent upon, storage systems that have thousands of individual storage devices (hard drives, optical disks, data tapes). The increase in the amount of data stored on each device makes the failure of each device more significant in terms of the volume of data potentially lost. A 3.5" floppy disk with 1.4 megabytes (MB) of data represented a few dozen files. A 650 MB CD could hold 500 times more data: thousands of files, or one hour of audio. A USB-attached terabyte hard drive is 700,000 times bigger than a floppy, and 1,400 times bigger than a CD. It could, for example, hold the entire contents of an institution's audio collection (such as several years' work by many people, collecting oral history recordings).

The increase in storage units (devices) means that statistics on failure rates that were once seen as "safe" are now appreciable risks. An advertised Mean Time Between Failure of 1,000 years looks very safe to a person buying a new hard drive (though it will be obsolete in five years). Schroeder and Gibson (2007) give results on a survey of major datacentres holding 100,000 disks, and found annual failure rates ranging from one to 13 %, averaging around 3% - far higher than an MTBF of 1,000 years. Similar results have been seen in other large scale applications of hard drive storage, for example, Pinheiro, Weber & Barosso (2007). This failure rate means that owners of a thousand of those same hard drives will need systems (e.g., big RAID arrays) and processes (e.g., continual hot-swapping and rebuilding) to ensure these failures are managed.

But it does not stop here. There is a widespread assumption that mass storage technology, e.g., RAID disk and offsite tape backup, solves the problem of failures in storage units. The reality is that data corruption or loss can be caused by failures in hardware, bugs in software, and human errors at all levels. Worse still, this corruption can happen without detection or correction. These are the so called 'latent' or 'silent' errors as described by Baker et al. (2006) and Rosenthal (2008).

Field studies of large disk-based systems, for example, Keleman (2007) and Jiang, Hu and Zhou (2008), reveal endemic silent data corruption in hard drive storage, including in 'enterprise class' systems that are explicitly designed to prevent data loss. In the CERN study reported by Keleman, as much as 1 bit in every 10^9 was on average irreversibly corrupted. Errors occured in the very systems, e.g., RAID controllers, that are designed to mitigate against failures lower down in the stack, and to protect against bit or sector level errors on hard drives.

The implication is clear. Any organisation using mass storage systems with a requirement for long-term data integrity should employ an ongoing and proactive programme of data-integrity checking and repair at an end-to-end systems level. It should not be assumed that any component of the system (networks, storage, memory, processing) is somehow 'safe', that is, immune from failures and data corruption problems.

Cost Modelling

We present an approach to risk that combines the dimensions of cost, risk (uncertainty) and value (benefits). This model builds upon and extends work on cost modelling by both the digital library and audiovisual communities. Early on in the development of digital libraries there was the fundamental work on preservation strategies by Beagrie and Greenstein (1998), Hendley (1998), Granger, Russell and Weinberger (2000) – and eventually something about the audiovisual sector from EU PRESTO project (Wright, 2002). The state of the art was brought together, and specifically labelled "life cycle", in the important paper of Shenton (2003).

Since then, there have been entire projects and conferences devoted to *life-cycle models and costs*. At a conference organised by the Digital Preservation Coalition and the Digital Curation Centre (DPC/DCC, 2005) there were reports from the LIFE[4] and eSPIDA[5] projects, both specifically about costs, though the eSPIDA work was more generally concerned with a formal method for including intangible benefits (value) in business cases. More pertinent to the present paper, it also specifically introduced the issue of uncertainty into the modelling process. Cost and value models are only one part of the wider activity of economic modelling, which in turn is just part of achieving sustainable digital preservation and access. Details and examples can be found in the Blue Ribbon Task Force (2008) interim report in this area.

Specific digital library and digital preservation cost models reported at the 2005 DPC/DCC conference included work from Cornell University, The National Archive in the UK, and the Koninklijke Bibliotheek in the Netherlands as well as two papers arising from PrestoSpace[6]. In all these models and studies, and for digital library technology in general, little is said about storage (except in the PrestoSpace work). Digital libraries assume that storage will be there (somewhere), and will work and continue to work. In estimating Total Cost of Ownership (TCO), the complexity of the models just mentioned is devoted to digital library processes, not storage devices (or their management). In digital library/repository TCO models, storage cost is generally modelled as a single number per year, and the model simply "adds up" those numbers.

We hope that current work in preservation theory and methodology, with use of file-description metadata, will support and encourage the ability of storage systems to return less-than-perfect files in a usable fashion. Examples of work with relevance to file description include Planets (file characterization) and Shaman:

- MPEG-21 DIDL[7] = Digital Item Declaration Language
- Planets XCEL, XCDL = eXtensible Characterisation Languages (Becker, Rauber, Heydegger, Schnasse & Thaller, (2008); Thaller, (2008))
- Shaman = multivalent approach (Watry, 2007)

[4] LIFE http://www.life.ac.uk/
[5] espida http://www.gla.ac.uk/espida/
[6] PrestoSpace http://www.prestospace.eu; http://digitalpreservation.ssl.co.uk/index.html
[7] Cover Pages http://xml.coverpages.org/mpeg21-didl.html

Cost-of-Risk Modelling

Estimation of cost involves uncertainties. Some uncertainties can be represented as variances in cost estimates (uncertainty about how much costs may vary from the predicted value), but a whole range of uncertainties are related to things that may or may not happen, and should be formally identified as **risks**.

A risk is the likelihood of an incident along with the business consequences (positive or negative) (Addis, 2008).

Examples of possible incidents that put content at risk include:

- Technical obsolescence, e.g., formats and players
- Hardware failures, e.g., digital storage systems
- Loss of staff, e.g., skilled transfer operators
- Insufficient budget, e.g., digitisation too expensive
- Accidental loss, e.g., human error during QC
- Stakeholder changes, e.g., preservation no longer a priority
- Underestimation of resources or effort
- Fire, flood, meteors

Traditional risk modelling[8] (and its use in project management) looks at lists of such incidents, and their attendant likelihoods (assessing likelihood may have the largest uncertainty of the whole process!) as contained in a risk register, and then proceeds to predict the consequences – the impact – of each item.

Possible consequences for preservation from the above list of incidents would include:

- Corruption or loss of audiovisual content
- Interruption to services
- Inefficiencies and increased costs
- Corner cutting and increased risks
- Failure to meet legal obligations
- Loss of reputation or loss of customers

A more comprehensive approach to the whole issue of uncertainty in preservation is to include the concept of value (benefit). The work of eSPIDA has already been mentioned.

[8] Risk definition and risk management: JISC: http://www.jiscinfonet.ac.uk/InfoKits/risk-management/ PRINCE2: http://www.ogc.gov.uk/methods_prince_2.asp

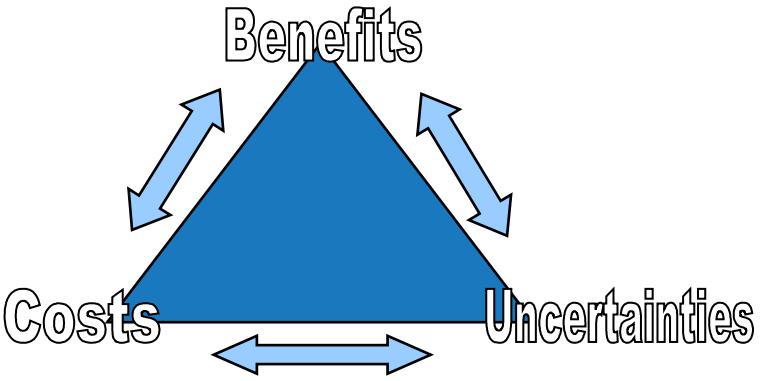

Figure 2. Interplay of costs, uncertainties (risks) and benefits (value) when planning preservation.

The combination of uncertainty, cost and benefits forms a three-way interaction, as shown in Figure 2. The key point about this approach is that applies to the whole issue of business-case planning, not just to the more narrow issues of risk analysis and cost modelling.

A typical preservation scenario, which can be optimized by use of the cost-of-risk approach, is shown in Figure 3.

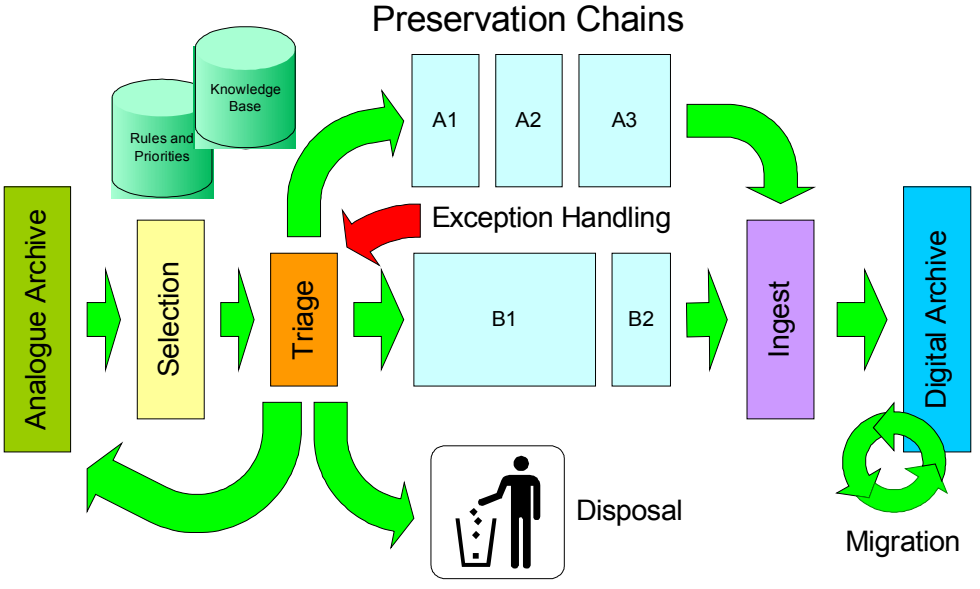

Figure 3. Workflow model for digitisation of analogue auidovisual material.

This shows a *triage* applied to items entering a workflow, so that items with a low risk of problems go through a low-cost workflow, while high-risk and high-value items go through a more comprehensive – and therefore expensive – workflow. The red arrow marked *exception handling* shows a response to an error: a low risk item fails in the simple workflow path *A1, A2, A3*, and so is shuttled back to the alternate workflow *B1, B2*. The proportion of items going through the various branches determines the overall cost – which rises if there is an unforeseen high level of exceptions (Addis & Veres, 2007). All the knowledge about risks, probabilities, costs and values – and what to do about them – can be combined in such a workflow. This form of process modelling allows overall costs to be estimated, and detailed simulations also allow time and effort to be estimated (to within the accuracy of the parameters of the model).

This integrated approach to cost, risk and value allows all the factors affecting preservation planning, funding and management to be considered in one set of interactions, rather than being taken separately. For quantitative modelling, all three factors need to be converted to a common unit of measurement. As cost and benefits are already commonly thought of in financial terms, the task is then to also express the uncertainties in monetary units: the cost-of-risk.

Full details require a much longer presentation. The essential issue for a quantitative approach to costs, benefits and risks is to have a common unit of measurement that allows direct comparison and hence proper evaluation of trade-offs. Standard practice in project management divides risk probabilities and impacts into two (high vs. low) or at most three categories (high, mid, low). However, an actual number is needed for the probability of occurrence of a risk, and a cost (if money is to be the metric unit) of the potential impact. How such costs can be assigned is entirely dependent on the situation. From a risk perspective, there has already been a great deal of detailed work specifically relevant to preservation in the DRAMBORA[9] project. This includes example risks that surround the long-term storage of digital content, with some examples shown below of how these might be interpreted for audiovisual preservation.

DRAMBORA Risk ID	Title[10]	Example
R30	Hardware Failure	A storage system corrupts files (bit rot) or loses data due to component failures (e.g., hard drives).
R31	Software Failure	A software upgrade to the system looses or corrupts the index used to locate files.
R32	Systems fail to meet archive needs	The system can't cope with the data volumes and the backups fail.
R33	Obsolescence of hardware or software	A manufacturer stops support for a tape drive and there is insufficient head life left in existing drives owned by the archive to allow migration
R34	Media degradation or obsolescence	The BluRay optical discs used to store XDCAM files develop data loss.
R35-R38	Security	Insufficient security measures allow unauthorised access that results undetected modification of files.
R39	Disasters	All content is in a small space through use of high density storage systems (e.g., tape robot) which makes the archive vulnerable to large-scale loss in a fire or flood.
R40	Accidental System Disruption	An operator accidentally deletes one or more files.
R55, 56, 59	Loss of integrity or authenticity	There is no audit trail for the changes made to content, which mean preservation actions are not taken or are inappropriate.
R60	Unsuitable backups	The backup tapes can't be read.
R61	Inconsistent copies	There are two copies of the content but they are different due to corruption of one of them, but which one is correct can't be identified.
R64, R69	Content Identifiers	The identifier used to locate a particular file in the system is lost or corrupted.

Table 1. Example DRAMBORA risks.

[9] DRAMBORA http://www.repositoryaudit.eu/
[10] In some cases the title has been shortened or paraphrased to make it easier to understand.

These risks can then be quantified using an estimate of the frequency of their occurrence, by using, for example, the DRAMBORA likelihood rating from 1 to 6[11] based on statistics from the literature or experience, e.g., on data corruption rates seen in storage systems as described earlier in this paper. In this way, methodologies such as DRAMBORA can be applied to relatively specific areas such as media failures or loss of individual files.

From a preservation cost perspective, there are some examples in the Blue Ribbon Task Force interim report and further specific costs for audiovisual preservation in Addis (2008). Various case studies exist with comprehensive cost breakdowns, for example the National Archives (Riksarkivet, RA) in Stockholm as described in Palm (2006). The recent and extensive Google report on Web Scale Computing also provides details of cost-modelling for large-scale storage and processing infrastructures (Barroso & Holze, 2009). This including the cost of mitigating against failures of all the components of these infrastructures.

Figure 4 shows consideration of risk as the central metaphor in strategic planning.

Figure 4. Risk Management Process (reproduced from 'A Risk Management Standard' by the Institute of Risk Management[12]).

[11] 1=Once in a hundred years or less often, 2=Once in ten years, 3=Once in 5 years, 4=Once a year, 5= Once a Month, 6=More than once a month

[12] The Institute of Risk Management http://www.theirm.org/publications/PUstandard.html

Cost of Risk and Mitigation of Loss

The effort within the digital library community to define and construct trusted digital repositories pays little attention to storage. The *trust* issue is typically defined and examined mainly at the institutional level, with less emphasis placed at the level of IT systems or individual device or file failures. DRAMBORA and TRAC correctly take an holistic approach to trust, but to be used effectively it is essential to go down to the detail of storage. The only physical reality of the content of a trusted digital repository consists of files which of necessity exist on some form of storage.The "atomic level" of success or failure of a repository is the success or failure of an attempt to read individual files. Such success or failure is clearly fundamental to the concept of trust for the whole repository.

Effort by those working in the storage area of the IT industry is focused on reducing the likelihood of read errors (device failure or file-read error). There is no concept, within standard IT systems, of a built-in and simple-to-use approach that provides access to partially-recoverable files[13]. If the inevitable low-level errors cannot be corrected by the built-in error detection and correction technology, the read fails and the file fails to open. Worse still are cases where the storage system then interprets the error as a fault with the media, e.g., a data tape labels it as such, and then stubbornly refuses any further attempts to read that media or any of the files on it. This happens when the system safeguards itself, for example, protecting tape heads from being damaged by faulty tape, but it can be at the expense of not maximising retrieval of the data in that system.

There is nothing that the ordinary user can do at this point, and even the all-powerful system manager can only look at backups to see if there is another copy of exactly the same file. This is not to say that there is no way to access what is left of the file. There is technology[14] to attempt to read corrupted files or failed hard drives, but such technology falls in the category of *heroic measures*: e.g., requiring special skills, a lot of time or sending the file or drive to an external company that will attempt a recovery using proprietary technology, at a substantial price[15].

Physically, a file with a read error is not an all-or-nothing situation. There will still be a stream of data (somewhere in the *stack* of operations between the user and the hardware) which is likely to be mainly correct, and is also likely to even have indications of which bytes are incorrect (because of lateral parity errors). For simple error detection and correction schemes, a common situation underlying an inability to read a file is a single block of data that has two or more such errors, so that the longitudinal parity check is ambiguous. At that point, a whole file of many blocks of data is called unreadable, because two bytes – at known locations – fail their parity check and so are known to be erroneous.

---

[13] This statement is based on the long term computing experience of the authors, communications with colleagues and general knowledge of computer operating systems and storage management. For example, a search of the 48 tutorials about storage on the SNIA (Storage Networking Industry Association). website found no mention of partial recovery or indeed any form of recovery from file read-errors. http://www.snia.org/education/tutorials/2009/fall/

[14] For example, File Recoverer from PCTools http://www.pctools.com/file-recover/ or ddrescue from GNU http://www.gnu.org/software/ddrescue/ddrescue.html . Operating systems also include repair tools, e.g. to fix corruption of files or filesystems, such as fsck, chkdsk. The problem is that none of these are easy to use at the scale of data often seen in digital preservation projects.

[15] Recovery tool box http://www.recoverytoolbox.com/ This company is just one of many offering tools that may be able to repair a corrupted file.

Returning to the definition of risk as having two factors: *probability* and *impact*: the ability to read *most* of the data in a corrupted file would, in certain cases, greatly reduce the *impact* of the error. This is the area of risk reduction that is being examined by the UK project AVATAR[16] (Addis et al, 2008). AVATAR is also looking at the whole issue of optimization and management of storage, from the perspective of archiving and long-term preservation.

Reducing the impact of a storage failure is a method for *mitigation of loss* (Knight, 2007). The issue of loss and recovery from loss has been identified as a neglected area in digital-preservation thinking, but its importance has been highlighted by the growing awareness of the phenomenon of bit rot (Panzer-Steindel, 2007).

Despite the best efforts of the IT industry, despite mean time between failure of hard drives exceeding one million hours, and despite tests of storage functionality yielding read-error estimations of one failure in 10^{17} read attempts – errors do occur. The first author was, in 2008, personally experiencing one file-read failure per month – and in each case these are total failures, with no possibility of mitigation (beyond the commercial route of heroic measures). Large data centres have now begun to publish data on errors, including the report from CERN mentioned earlier that showed bit-rot levels affecting up to 1 byte in 3×10^7 (Panzer-Steindel, 2007), meaning one error in 30 megabytes of data. At the file level, one file in 1,500 was affected.

Redundancy and Risk.

Standard practice for reducing risk of loss is to have another copy. The use of second (or higher) copies is a method of reducing impact: a file-read error or a device failure has much less impact if recourse can be made to a backup copy or system.

At a more sophisticated level, arrays of hard drives are used to gain the benefits of redundancy at lower cost. RAID[17] technology achieves protection for the loss of one of N drives in a set of N+1 – so the net cost is N+1 drives, rather than the 2N that would be required by simple redundancy. RAID has now advanced (e.g., RAID6) to the point where multiple disks can fail without data loss, which means data can still be accessed safely whilst individual disks are being replaced and live rebuilding takes place. This allows disk systems to be built that are resilient to hardware failures and data read errors. For large data centres, the problem is shifted, to some extent, from risk of loss from device failure to having the right support processes to "feed" large systems with a constant supply of new drives and have the people in place to do so.

Whilst RAID can mitigate against detected failures, it does not solve the problem of 'silent' errors. Indeed, if there are additional errors in the software or firmware used to implement RAID, then this can actually make the problem worse. Neither does the use of RAID avoid the need for multiple copies in more than one geographic location to mitigate against catastrophic loss scenarios, e.g., fire or flood in a data centre. However, many copies of the content in many places can result in prohibitive costs. For example, in the audiovisual domain, Standard Definition digital video has an uncompressed data rate of about 200 MBit/s and even when stored with compression, e.g., 50MBit/s DV, this means that multiple Petabytes of storage are required for a

[16] AVATAR-m: http://www.it-innovation.soton.ac.uk/projects/avatar-m/
[17] RAID: Redundant Array of Inexpensive Disks – an efficient method of achieving device-level redundancy. http://en.wikipedia.org/wiki/Redundant_array_of_independent_disks

typical broadcast archive. HD requires five times as much space. In digital cinema, scanning or creating images at 4K resolution requires up to 30 times the data rate of SD. For 3D cinema with twin data streams at up to 144 fps the volumes are truly vast. This size presents a real problem. Estimates are that it costs $1.5M per annum for 1PB of storage using online disks, with the cost of tape (in robots) being approximately a third of this (Moore, D'Aoust, McDonald & Minor, 2007). Multiple copies (more than two) can mean unacceptable costs at this scale. When considering the cost of storage it is important to consider how the cost is broken down, e.g., between equipment, maintenance, space, power and cooling etc., and how these factors are dependent on each other. Both Moore et al. (2007) and Barroso and Holze (2009) provide example breakdowns for large-scale storage operations. Over half the cost can be made up of staff, utilities and space. Lowering these costs, either by relocating storage to where these elements are cheaper or by choosing a different type of storage technology can alter the probability of loss. As an extreme example, data tapes on shelves can achieve high densities with a naturally low power requirement. However, unless they are checked and migrated regularly, which introduces a further labour and equipment cost, then the increased probability of data loss can negate the initial savings compared to using tape robots or other automated mass storage systems.

At the same time as redundancy is added to storage systems to reduce risk, redundancy is being taken out of the files stored on those systems, as a way to save space. Compression, lossless or lossy, is based on the innate redundancy (entropy) of the original data. When the redundancy is removed from a file, a complex transformation has to be applied to the resulting data in order to transform it back to the original (or close to the original, in the case of lossy compression).

To Encode Or Not To Encode.

Not encoding, in particular not using compression, typically results in files that have minimal sensitivity to corruption. In this way, the choice not to use compression is a way to mitigate against loss.

As an example, consider uncompressed audio. A .WAV file is simply a header followed by a sequence of numbers – one number per sample of the desired audio waveform. If the audio is sampled at 44.1 kHz (the rate used on CDs), each sample represents about 23 micro-seconds of data. Losing one byte of data results in one bad sample, but there is no spread to any of the rest of the data.

Hence an uncompressed audio file can be perfectly usable despite loss of one byte. Indeed, experiments have shown[18] that a .WAV file with 0.4% errors is almost undistinguishable from the original, whereas an MP3 file with the same level of errors either will not open at all, or will have errors affecting most of the audio, and rendering it unusable. The same logic applies to video, images – and even to text if represented as a sequence of characters (with embedded mark-up, as in the old days of "printer control characters" as escape sequences within a text "stream").

Using compression, be it lossless or lossy, can save on storage space, and in turn allow more copies to be held for the same cost. However, compression can make the files much more sensitive to data corruption. The CERN study (Panzer-Steindel, 2007 p.3) found that a single bit error would make a compressed file unreadable, with a

[18] First author's own experiments, unpublished

probability of 99.8%. The same applies to the use of content-specific compression, e.g., image or video compression for media files. For example, Heydegger (2008) developed a 'robustness indicator' on the sensitivity of image formats to bit level corruption and then investigated how compression affects robustness. This work is notable as it includes JPEG2000, which is emerging as a strong candidate for preservation in the AV community (Pearson & Gill, 2005) including digital cinema[19]. Tests by Heydegger showed that corrupting only 0.01% of the bytes in a compressed JPEG2000 file, including lossless compression, could result in at least 50% of the original information encoded in the file being affected. In some cases, corrupting just a single byte in a JPEG2000 image would cause highly visible artefacts throughout the whole of that image.

This sensitivity to corruption is traded for a saving in storage space, although the trade-off is not always simply one for the other. For example, Heydegger found that one byte of corruption had the following effect:

- a 10 MB uncompressed TIFF file had just .00001% errors (meaning just that one byte was affected)
- a lossless JPEG2000 file had 17% errors for a saving of 27% in storage
- a lossy JPEG2000 file had 2.1% errors for a saving of 62% in storage

As an example of the affect of data loss on image files, here are two examples: a BMP file (uncompressed) and a GIF file (compressed). The BMP file has 1,400 errors, one in every 256 bytes. The GIF file has a single error

Figure 5. BMP file with one error per 256 bytes (1,400 errors).

[19] Enhanced Digital Cinema project (EDCINE) http://www.edcine.org/intro/

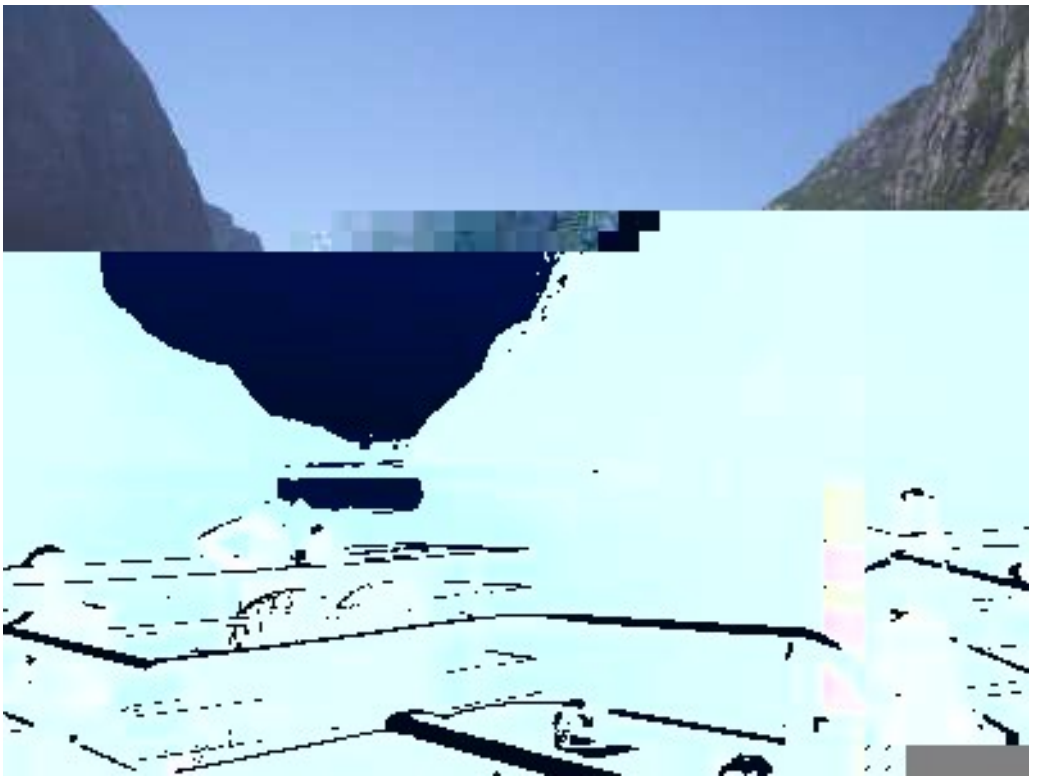

Figure 6. GIF file with a single error (in 14 KB).

From the above results, it is evident that removing redundancy through compression increases impact of corruption, i.e., the "cost of error". The compression increases the proportional damage caused by an unrecoverable read error. However, if there is no mechanism for using files despite read errors, then it is of no practical significance whether a one-byte error causes major damage, or only very local and minor damage. If the file cannot be read in either case, the error-magnification factor caused by compression is hidden.

If less-than-perfect files can be passed back to the user, or to a file-restoration application, then the increase in "cost of error" caused by compression can be legitimately compared with the decrease in cost of storage. As the cost of storage devices reduces, and as storage management improves in efficiency, preservation strategies based solely or largely on storage costs are less and less satisfactory.

An unsolved issue in preservation strategy is whether it is better (lower "cost of risk" for the same or less total risk) to use lossless compression and then make multiple copies (externalized redundancy) as a way to reduce the impact of storage errors – or to avoid compression and exploit the internal redundancy of the files. The problem at present is that there is little or no technology (within conventional storage systems, or conventional digital repositories) to support the second option.

It is also possible to encode files in a way that deliberately increases their robustness to corruption, JPEG2000 wireless (JPWL) being an example[20]. Redundancy and error checking are built in to improve robustness to errors introduced during transmission over wireless channels. However, whilst this approach, and

[20] JPEG 2000 http://www.jpeg.org/jpeg2000/j2kpart11.html

source/channel coding more generally[21], is used for robust transmission through space, i.e., from one geographical location to another, it has yet to be applied to long-term transmission through time where the channel is the storage system and noise is introduced by that channel, e.g., silent corruptions.

Either way, the results of Heydegger combined with the 'bit rot' headline findings of NetApp or CERN would imply that maintaining integrity of very large files is nearly impossible. For example, if the bit corruption rates of 10^{-9} reported in the CERN study occurred in the audiovisual domain, then for data files that are 10^{13} bits in size (approx 1TB, which is an hour of uncompressed HD), it would seem inevitable that these files will become corrupted quite rapidly when stored on disk. Yet this is not the case. The headline figures reported, whilst attention grabbing, neglect the distribution of the corruption. Studies show that corruption is typically at the block level rather than bit level, it tends to be spatially correlated, e.g., successive blocks on a disk are more likely to be corrupted than blocks at random, and it affects media in batches (e.g., a bad batch of hard drives from a particular manufacturer) (Barroso & Holze, 2009; Krioukov et al, 2008;). This is why corruption of large files exists, but is not endemic. Further studies are needed on how this pattern of corruption files translates to loss of content.

The question of which strategy to take depends upon more than just the ability of file systems to return files with partial errors. An holistic approach to risk management means dealing with disaster recovery (fire, flood, theft, etc.), human error (accidental corruption, deletion, miscataloguing, etc.), and technology obsolescence (formats, software, devices, etc.). All these risks provide a strong motivation for having multiple copies in multiple places using multiple technical solutions. If an offsite copy of uncompressed video is created to address disaster recovery, then lossless compression may allow two offsite copies for the same cost. Three copies in three places may well be enough to reduce the risk of loss due to individual storage failures to a level where no further measures are needed beyond those of conventional storage systems, e.g., RAID. However, this has yet to be proven and is unlikely to apply in all circumstances.

Given so many options (to compress or not to compress; how many copies to make; if, how and when to detect and repair corruption etc.), there is a need for new frameworks that allow the multitude of strategies to be compared and implemented. Both AVATAR-m (Addis 2008, Addis, Lowe & Middleton, 2009) and PrestoPrime[22] have started work in this area. Whilst only tackling subsets of the overall cost of risk of loss problem, they do recognise that a 'one size fits all' approach is unlikely to succeed and hence they focus on how to provide a way for various strategies to be evaluated and combined.

The right approach is not just one of finding the best technical solution. It is also a question of finding an approach that is in itself easy to understand and is amenable to risk assessment and management. As stated earlier, no part of the system should be considered infallible, including those parts responsible for error correction. For example, consider the use of erasure-coding[23] techniques in wide-area storage

[21] Wikipedia http://en.wikipedia.org/wiki/Information_theory
[22] PrestoPRIME http://www.prestoprime.org/
[23] Wikipedia http://en.wikipedia.org/wiki/Erasure_coding

networks, for example, clouds, for example, as used by Permabit[24]. The erasure code introduces redundancy that mitigates failures in the individual storage nodes. If there are many such nodes then an erasure code can be much more efficient than brute force replication (Weatherspoon & Kubiatowicz, 2002). On the other hand, encoding and distributing a file in this way means that there is no way to recover any of that file if the 'index' is lost that describes how the encoding and distribution has been done. The software/system/vendor becomes the risk rather than the storage, i.e., one danger has been, to some extent, traded for another.

Despite various techniques being available, it is not until file-reading systems are willing and able to return files despite errors, and include media-specific reconstruction techniques to "fill in" where errors are known to exist, that there will be an effective way to exploit file-error recovery as a method to mitigate against loss. This gap currently prevents a whole class of "cost of risk" strategies from being used to complement conventional techniques.

The frustration for audiovisual archivists is that digital technology has taken us one step forward, and now is taking us two steps back. The ability of analogue videotape recorders to cope with loss of data (dropout) was limited, and black lines would appear in the resultant images. Digital tape recorders had much better built-in compensation: the *concealment* option would allow a missing line to be replaced by a neighbouring line, and expensive machines could even replace entire frames with an adjacent (in time) frame. Now file-based digital technology has *no* ability to cope with loss (corruption; uncorrectable errors), beyond the "external redundancy" option of multiple copies.

One could accept that files will remain "all-or-nothing" entities – you either get everything in them or you lose the lot. The strategy then becomes one of splitting assets, e.g., a video sequence, into multiple files and implementing data integrity measures at the application level, currently being investigated in the AVATAR project (Addis et al, 2009). For example, an audiovisual program could be split into separate files for shots, scenes, frames, regions of interest, audio, video or many other ways. The most important parts would then be assigned to one or more storage systems with appropriate levels of reliability – avoiding the "all eggs in one basket" problem. An asset can be separated into pieces based on knowledge of the importance, to the user, of the various parts of the content – knowledge that a file system or storage device will never have. The disadvantage is increased technology and management costs, which introduces complexity and new risks – both are a violation of the "simplest is best" principle.

Conclusions

Comprehensive and integrated planning for preservation can be accomplished through the use of a three-factor model, based on costs, benefits and uncertainties. The cost-of-risk concept allows all three factors to be quantified on a common, monetary scale.

[24] Permabit Technology Corporation http://www.permabit.com/

When considering how best to store files with some assurance of long-term integrity, the choices are complex and include how many copies to use, how to encode them, where to store them, how frequently to check them, how to repair them efficiently, and how to do all this in a cost-effective way.

There is no single answer, rather a range of options where the "best" one in a given context will change over time. This calls for new frameworks that allows these choices to be quantified, compared, assessed and implemented to suit the differing needs of archives, including their budgets and appetite for risk.

Starting with a simple strategy is frequently the best approach. One way to reduce the cost of risk, and hence the best chance for mitigation of loss, is simply not to compress the data. Storing only uncompressed data would appear to add cost rather than reduce it – but storage costs are typically a small part of a preservation project or strategy (labour is always the dominant cost), and storage media cost is dropping by 50% every 18 months.

But the full benefit of uncompressed files (in terms of mitigation of loss and consequent reduction of impact) will remain irrelevant unless and until the storage industry and digital repository architects produce systems that allow access to less than perfect files.

References

Addis, M. (2008). Cost models and business cases (for audiovisual curation and preservation). *HATII-TAPE Audiovisual Preservation Course*. University of Edinburgh. Retrieved November 24, 2009, from http://www.hatii.arts.gla.ac.uk/news/tape.html

Addis, M., Beales, R., Lowe, R., Middleton, L., Norlund, C., & Zlatev, Z. (2008). Sustainable archiving and storage management of audiovisual digital assets. *IBC 2008*, Amsterdam.

Addis, M., Lowe, R., & Middleton, L. (2009). A new approach to audiovisual archiving. In *63rd Broadcast Engineering Conference*, Las Vegas, April 18-23, 2009.

Addis, M., & Veres, G. (2007). *Knowledge database and report on tape condition.* PrestoSpace Public Report, Deliverable D6.2. Retrieved November 24, 2009, from http://www.prestospace.org/project/deliverables/D6.2.pdf

Baker, M., et al. (2006).A fresh look at the reliability of longterm digital storage, *EuroSys'06*, April 18-21, 2006, Leuven, Belgium.

Barroso, L. A., & Holze, U. (2009). *The datacenter as a computer: An introduction to the design of warehouse-scale machines.* Google Inc. Synthesis Lectures on Computer Architecture no. 6. Published by Morgan and Claypool.

Beagrie, N., & Greenstein, D. (1998). *A strategic policy framework for creating and preserving digital collections*. British Library Research and Innovation Report 107. London: British Library. Retrieved November 24, 2009, from http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/framework.pdf

Becker, C., Rauber, A., Heydegger, V., Schnasse, J., & Thaller, M.. (2008). A generic XML language for characterising objects to support digital preservation. *Proceedings of the 2008 ACM symposium on Applied computing,* Fortaleaza, Brazil. pp 402-406. Retrieved November 24, 2009, from http://portal.acm.org/citation.cfm?id=1363786&jmp=indexterms&coll=GUIDE&dl=GUIDE

Blue Ribbon Task Force. (2008). *Sustaining the digital investment: Issues and challenges of economically sustainable digital preservation*. Interim Report of the Blue Ribbon Task Force on Sustainable Digital Preservation And Access (BRTF-SPDA). December 2008. Retrieved November 24, 2009, from http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf

Digital Preservation Coalition/Digital Curation Centre. (2005). Report for the DCC/DPC workshop on cost models for preserving digital assets. In *Cost Models for Preserving Digital Assets*. British Library Conference Centre, July 26, 2005. Retrieved November 24, 2009, from http://www.dpconline.org/graphics/events/050726workshop.html

Granger S., Russell, K., & Weinberger, E. (2000). *Cost elements for digital preservation.* CEDARS Project. Retrieved November 24, 2009, from http://www.webarchive.org.uk/wayback/archive/20050409230000/http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc

Hendley, T. (1998). *Comparison of methods and costs of digital preservation*. Research and Innovation Report, 106. British Library. Retrieved November 24, 2009, from http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html

Heydegger, V. (2008). Analysing the impact of file formats on data integrity. *Proceedings of Archiving 2008*, Bern, Switzerland, June 24-27.

Jiang, W., Hu, C., & Zhou, Y. (2008). Are disks the dominant contributor for storage failures? A comprehensive study of storage subsystem failure characteristics. In *6th USENIX conference on File and Storage Technologies (FAST'08)*, San Jose, California. February 2008.

Kelemen, P. (2007). Silent corruptions. In *CERN IT*, LCSC 2007, Linköping, Sweden.

Knight, S. (2007). Manager innovation centre and programme architect national digital heritage archive, NLNZ. Remarks 'from the floor' on the significance of efforts to mitigate against loss, at the SUN PASIG meeting, November 2007, Paris. Retrieved November 24, 2009, from http://sun-pasig.org/nov07_presentations.html

Krioukov, A., Bairavasundaram, L.N., Goodson, G.R., Srinivasan, K, Thelen, R., Arpaci-Dusseau, et al. (2008). Parity lost and parity regained. In *6th USENIX conference on File and Storage Technologies (FAST'08),* San Jose, California. February 2008, pp. 127-141. Retrieved November 24, 2009, from http://www.usenix.org/events/fast08/tech/krioukov.html

Lawson, S. (2008, April 22). Seagate ships one-billionth hard drive. *Computerworld.* Retrieved November 24, 2009, from http://www.computerworld.com/action/article.do?command=viewArticleBasic&taxonomyName=storage&articleId=9079718&taxonomyId=19&intsrc=kc_top

Moore, R. L., D'Aoust, J., McDonald, R. H., & Minor, D. (2007). Disk and tape storage cost models. In *Archiving 2007*.

Palm, J. (2006). *The digital black hole*. Retrieved November 24, 2009, from http://www.tape-online.net/docs/Palm_Black_Hole.pdf

Panzer-Steindel, B. (2007). *Data integrity*. April 8, 2007. Retrieved November 24, 2009, from http://indico.cern.ch/getFile.py/access?contribId=3&sessionId=0&resId=1&materialId=paper&confId=13797

Pearson, G., & Gill, M. (2005). An evaluation of motion JPEG 2000 for video archiving, *Archiving 2005* (April 26- 29, Washington, D.C.), IS & T. Retrieved November 24, 2009, from http://www.imaging.org/ScriptContent/store/epub.cfm?abstrid=32265

Pinheiro, E., Weber, W.-D., & Barosso, L.A. (2007). Failure trends in a large disk drive population. *5th USENIX Conference on File and Storage Technologies (FAST'07)*.

Rosenthal, D. (2008). Bit preservation: A solved problem?. In *Proceedings of iPRES 2008*: *The Fifth International Conference on Preservation of Digital Objects*, British Library, London. Retrieved November 24, 2009, from http://www.bl.uk/ipres2008/index.html

Shenton, H. (2003). Life cycle collection management. *LIBER Quarterly*, *13*.

Schroeder, B., & Gibson, G. A. (2007). Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you? *5th USENIX Conference on File and Storage Technologies*. February 2007, pp. 1-16. Retrieved November 24, 2009, from http://www.usenix.org/event/fast07/tech/schroeder/schroeder_html/

Thaller, M. (2008). *Characterisation.* Planets presentation at Digital Preservation Planning: Principles, Examples and the Future with Planets. Retrieved November 24, 2009, from http://www.planets-project.eu/docs/presentations/manfred_thaller.pdf

Watry, P. (2007). Digital preservation theory and application: Transcontinental persistent archives testbed activity. *International Journal of Digital Curation*, *2*(2), pp. 41-68. Retrieved November 24, 2009, from http://www.ijdc.net/ijdc/article/viewArticle/43/0

Weatherspoon H., & Kubiatowicz, J. D. (2002). Erasure coding vs. replication: A quantitative comparison. In *1st International Workshop on Peer-to-Peer Systems (IPTPS '02)*. Retrieved November 24, 2009, from http://oceanstore.cs.berkeley.edu/publications/papers/pdf/erasure_iptps.pdf

Wright, R. (2002). *Broadcast archives: Preserving the future.* PRESTO Project. Retrieved November 24, 2009, from http://presto.joanneum.ac.at/Public/ICHIM%20PRESTO%2028_05_01.pdf