Short outline for short article

RBLandau 20170914

Executive Summary:

- (assertion) Digital document collections cannot be safeguarded simply by making a few copies; too many copies are required for adequate protection in real-world situations. The curator needs to audit the integrity of the several copies at intervals.

- (technique) If a copy of a document is corrupted or lost, it can be repaired from other extant copies. A document is permanently lost only when all of its copies have been damaged or lost.

- (assertion) It is important that the several copies of documents be stored independently, so that no single incident is likely to affect multiple copies. Such incidents include natural disaster, regional conflict, local terrorist attack, government censorship, local economic downturn, business failure, business realignment, etc.

- (scope assertion) This paper does not deal explicitly with threats posed by inimical or incompetent human agencies. Deliberate hacking and poor management practices are beyond our current scope.

- (motivation) The quality of digital storage media is highly variable, particularly over long periods of time. Stored data may deteriorate slowly, in small pieces such as disk sectors, or may fail in large blocks, such as disks, disk arrays, or entire storage services. Redundant storage techniques such as RAID protect against only some such failures.

- (assertion) A curator cannot always assess the reliability of a storage vendor of a particular type or location. A strategy to preserve digital documents must be robust over a wide range of reliability and physical conditions.

- (assertion) This paper presents simulation results that suggest that a reasonably small number of copies, audited and repaired regularly, suffice to preserve digital document collections over a very wide range of error conditions.

- (assertion) Document auditing methods must examine every document at some interval. Methods that examine only random subsets of documents sampled with replacement leave some documents unaudited and vulnerable to loss.

- (assertion) Moderately frequent auditing of documents, e.g., annual auditing, suffices to preserve document integrity.

- (assertion) Document size, type, and fragility due to, e.g., encoding, compression, or encryption, need not be significant factors in the choice of storage methodology.

- (assertion) It is possible to divide collections into subsets by, e.g., value, and choose different storage strategies for the subsets. High value documents may be stored more widely and audited more frequently than low value documents.

- (conclusion) To preserve collections of digital documents, store a modest number of copies on independent storage services and audit their contents at regular intervals. This strategy can protect collections over long periods and against a wide variety of storage quality levels, short term quality variations, and large scale failures.

Outline

The Problem

- Most digital data is stored on disks. Disks fail, a little at a time or all at once. Disk storage services can also fail, slightly or totally.
- Large collections of digital documents need to be preserved perfectly, or almost perfectly, for long periods of time.
- There is little empirical guidance available to help curators plan how to preserve their collections at reasonable cost.

The Approach

- Library clients make several copies of all documents in the collection.
- Copies are stored on commercial cloud storage services or private datacenters.
- Copies may suffer errors that corrupt a single document, and storage services may suffer errors that destroy all documents stored there. In either case, such losses are silent, latent errors that are not signaled to the client. The client must actively discover the error.
- Clients audit documents on a regular schedule. The auditing process actively patrols for document errors by retrieving the full content of document copies, or other convincing fixity information, from the servers.
- Auditing will be done in repeated cycles. Each cycle may contain all documents, or several systematically chosen subsets, or several random subsets sampled with or without replacement. For example, an annual audit cycle may be broken into quarterly segments, with a different quarter of the collection being visited each calendar quarter.
- Auditing in several segments per cycle can even out the bandwidth requirements for retrieving documents.
- Any documents found to be corrupted or missing during an auditing cycle will be repaired by refreshing the copy from other extant copies.
- Note that it is also possible that copies could be stored on offline media, such as tape or optical disk. Such media are not fundamentally different from the online media that we are concerned with here; they have their own rates of bit rot and total loss, but auditing and repairing documents stored on such media is operationally very different. This investigation does not attempt to cover offline media.

The Simulations

- The discrete event simulations operate on a collection of fixed size for a fixed duration.
- Clients place document copies on multiple servers. For any single simulation, all servers have the same statistical characteristics: sector lifetime, server lifetime, glitch rates, shock rates.
- Documents suffer sector errors that occur at random times and locations. These errors corrupt the document content. The base rate of arrival of errors is usually constant during the simulation, though it

may be increased for short periods by glitches (q.v.). The rate may be varied over a wide range to represent a variety of physical disks and operating conditions.

- Servers may suffer "glitches" that increase the local sector error rate for a short time in a single server. Glitches are intended to represent temporary operational problems such as HVAC failures, noisy AC power and such. Glitches arrive randomly at a tunable rate.
- Servers may also suffer from "shocks" that reduce their expected lifetime. Shocks are intended to represent economic downturns, floods, wars and such. Shocks arrive randomly at a tunable rate.
- Random events in the simulations all independent and identically distributed Poisson arrivals. The arrival rates are all expressed in terms of half-lives rather than mean exponential lifetime or bit error rate.
- Document losses are assessed as the fraction of the collection that is permanently lost during the simulation period.

Results and Conclusions

- Auditing is essential to reduce the number of copies needed to limit permanent document losses.
- Auditing should be regularly scheduled and should examine every document during every auditing cycle.
- The recommended baseline for preserving collections over a very wide range of operational conditions: five copies, annual auditing, total auditing (i.e., all copies examined each cycle). The auditing cycles may be segmented to reduce variations in bandwidth required for auditing, so long as the sampling for segments is done without replacement.
- Excessively frequent auditing, e.g., weekly or monthly, is unnecessary overkill.
- Glitches of moderate frequency and moderate impact appear similar to increased local sector error rate.
- Baseline auditing can protect even against moderately frequent and severe shocks. For major, severe shocks -- e.g., non-regional wars, the sinking of Atlantis -- other steps may be necessary.

Details: Simulation Parameters

- Document size can vary, but is not a factor in the results (document losses as a percentage of the collection size).
- Storage shelf size can vary but, again, is not a factor in the results.
- The number of independent copies can vary from 1 to 20.
- Sector error rates vary over a wide range to account for disk technologies, and manufacturing and batch variations. Reliable, or even plausible, numbers are very difficult to obtain from industry.
- The strategies for auditing cycles can vary: frequency of auditing cycles, segmentation of a cycle, and sampling without or with replacement.
- Minor glitches are intended to model short term physical problems that increase bit error rates on disks, e.g., HVAC failures. These vary by frequency, impact level, and duration.
- Major shocks are intended to model large scale problems that affect entire services or groups of services. These vary in frequency, impact level, duration, and span of impact.
- In an attempt to simplify understanding of many parameters, all the time-related parameters -- sector error rates for disks, arrival rates for glitches and shocks, expected server lifetime -- are expressed as half-lives of the relevant objects rather than error rates or mean exponential lifetimes.

Details: Variations Tested for Robustness

- A number of simplifying assumptions have been used to reduce the sample space to manageable size. Most are simply choices of scale and discrete values for the many tunable parameters.
- Vary the number of copies allocated to independent storage services.
- Vary sector error rates for small disk errors.
- Vary auditing strategy, from none to frequent, number of segments per cycle, and whether segment sampling is done with or without replacement.
- Vary glitches, frequent or rare, minor or major, short or long, affecting the document error rates on single servers.
- Vary shocks, frequent or rare, minor or major, short or long, affecting the survival rate(s) of one or more servers at a time.
- Vary document size and error rate together: larger docs should predictably represent bigger targets for random errors.
- Each simulation cycle is repeated a number of times with different seeds for the pseudorandom number generator. It is initially performed with a modest sample size, approximately twenty; some cases have been sampled with much larger sample sizes for validation.
- Simulations are performed with repeatable seeds for the pseudorandom number generator, to enable accurate replication of the experiments.
- Document size and sector error rate scale linearly and predictably against each other: larger documents are larger targets for randomly placed errors.
- The size of a storage shelf scales linearly and predictably with the hit/miss rate of randomly placed errors: errors on a full shelf will rarely miss an occupied area; errors on a partially occupied shelf may land in unoccupied areas.

What Is Not Modeled

- Deliberate attacks on collections or parts of collections.
- Attacks on the auditing mechanism.
- Media or format obsolescence.
- Operational mistakes and bad practices.
- Partial document loss and manual repair from minor damage.
- Characterization of individual disk drives, models, or error correction or storage technologies.
- Operational considerations of RAID, erasure coding, or other redundancy by storage services.
- Hardware or software failures of storage controllers.

Future Work

- Develop standards for auditing protocols between clients and servers.

Software Available Soon (RSN)