Why Is Off-site Storage Difficult?

Robert C. Bell | Partner Services Manager

3 December 2013

CSIRO ASC Data Store: History

- 1990: JSF: Cray Y-MP2/216 – big, expensive, fast disc
  - Turned a compute problem into a storage problem!
  - DMF from November 1991

Multi-site Storage

- Why?
- What are we trying to achieve?
- Why is it difficult?
- Examples from history
- Suggestions for future

Multi-site Storage – Why?

- Risk assessment
- Loss of data and service

Excerpts from previous talk – "Ways to lose data"

- Threats to data
  - red highlighting of cases where offsite storage may help
- Examples of threats which may lead to the loss of data.
- External
  - Terrorism
  - Targeted terrorism (a terrorism attack that targets all copies of some data)
  - War
  - Natural disaster (storm, hail, lightning, flood, cyclone, tornado, earthquake, tsunami, fire, landslide, liquefaction, drought, meteorite, cosmic rays). Though natural disasters, many of these can be exacerbated by human activity.

UQ Flood (courtesy ABC)

Threats to data

- External
  - Nuclear accident, arson, fire, water damage, electrical fire.
  - Subsidence
  - Financial or organisational failure, insolvency, bankruptcy, breach of contract
  - Ram raid
  - Robbery
  - Political campaign, sabotage, riot, insurrection, civil war
  - Climate change

Threats to data

- Facilities
  - Building failure
  - Infrastructure failure, e.g. power supply, water supply, telecommunications, plant, floor collapse, fire suppression damage to equipment
- Hardware
  - CPU, Memory, Channels, Network, Switches, Controllers, Disc (bit rot), Tapes, Hardware/media obsolescence
- Software
  - In OS
  - In filesystem
  - In database
  - In firmware, e.g. RAID, storage arrays, discs, controllers
  - In middleware, e.g. HSM, iRods, rsync, netCDF, tar, zip.

Threats to data

- Software
  - In shells and scripting languages
  - In applications
  - In software rot

  (Where although the data may be intact, the means to access and manipulate the data may be lost. For example, no current versions of Microsoft PowerPoint can read PowerPoint presentations from a decade or so ago.)
  - The Victorian Electronic Records Strategy (VERS) developed by CSIRO with the Public Records Office of Victoria was one of the first good examples of a successful implementation of a strategy to avoid this problem. The strategy involves storing information in two formats, both with open definitions: PDF, to preserve the look of a document, and XML to preserve the meaning through metadata.

Threats to data

- Human
  - Disgruntled employee, particularly administrator
  - Revenge
  - Human failure (wrong command, wrong machine, wrong window, wrong directory, wrong instructions; lack of procedures or failure to follow, faulty scripts)
  - Finger trouble (rm * .o instead of rm *.o)
  - Network intrusion, particularly if root is compromised
  - (In a paper from the 1990s, possibly from IDC, it was reported that 68% of incidences of data loss were due to human error. This is clearly the most common cause, and yet we are beguiled at times into strenuous efforts to protect our data from hardware errors.)

Identifying risk

- Risk assessment is commonly done by considering the chances of a threat, and the impact of a threat.
  - Events with a low chance and low impact are usually ignored.
  - Events with high chance and high impact must have mitigation.
- A matrix for risk assessment is usually developed.

Risk mitigation

- Risk mitigation is the identification and implementation of controls that will lower the risk.
- Multi-site storage is one mitigation strategy.

Off-site Storage – What are we trying to achieve?

1. Protection of data from loss
   - Replica, or backup, or replica with backup
   - Access to offsite data? How, where, who by, etc?
2. Continuation of service: all about timescales
   - Continuous dual-site running with failover (banks)
   - Restoration within minutes, hours, days, weeks, etc
     - The shorter the target, the more prepared you need to be …
   - Problem of managing and reconciliation after loss of a site or communication between sites (how does each site know whether it is in charge if it can't reach the other site? STONITH)

Off-site Storage – Why is it difficult?

1. So many possible threats to data
   – some simple strategies won't work, e.g. mirroring!
2. So difficult to prepare strategies for continuation
3. So rarely funded to make it happen
4. Hard to do with disc-only implementations, since the second site is often directly tied to the first, and backups next to impossible with modern capacities
5. DMF is an advantage and a complication – added layer of protection (gets around backup problem by trickling data out continually), but have to protect databases.

Off-site Storage – Examples from history

1. Threat to data in early 1992 – failing partner

   Started to use the dual tape copies in DMF to provide one copy of the data off-site, along with the dumps of the file systems and DMF databases – done monthly.

   No attempts at providing access to the off-site data.

   Within a year, provided the way to transfer data from old to new Cray system.

Off-site Storage – Examples from history

- 1 (continued).
- Developed into continuing strategy, and in use for about a decade.
- Take the second copy tapes (and dumps) off-site at intervals.
- Pros
  - at least something is done!
- Cons
  - Delays when second copies are needed on site because the first copy is bad
  - Off-site tapes become sparse:
    - can't easily merge tapes without destroying separation of copies!
  - Hard to find remote tapes if stored in boxes, or even on racks
  - Labour intensive!
  - Tapes deteriorate in transit and in remote storage (undetectable!)
- Was the inspiration for dmsilo – available from SGI

Off-site Storage – Examples from history

- 2. As above, but have a tape library remotely, to act as an automated rack.
- Pros
- can easily find tapes
- Cons
- as above, plus
- increased cost
- floor-space requirements

Off-site Storage – Examples from history

- 3. House tapes off-site as above, but have tape library and drives off-site.
- Pros
- can read tapes remotely, over the network
- Cons
- extra cost of drives
- Need fast network to drive tapes remotely.
  Note 1: The Monash LaRDS system does this, with tape drives on the other side of Blackburn Road using FC over dark fibre).
- Note 2: UQ and NCI NF as well
- Note 3: CSIRO SC currently testing this.

Off-site Storage – Examples from history

- 4. Take tapes off-site as above, but have tape library and drives and server off-site.
- Pros
  - Can read tapes remotely
  - Can do merges remotely
- Cons
  - Extra cost and complication of another server to drive tapes.
- CSIRO SC looking at acquiring a server to act as disaster recovery system remotely, and able to provide limited access to the data.
  - Simple re-build of primary site from dumps and DMF databases
  - Plan for read-only
  - Not doing remote DMF mover

Off-site Storage – Examples from history

- 5. Make three copies of tapes, and take one copy off-site.
- Pros
  - don't compromise copy separation
  - always have two copies on-site to cover bad tapes
- Cons
  - Extra costs - three copies.
  - Still don't know how to do merges ‘safely’,
    except by swapping entire second and third copy sets.

Off-site Storage – Examples from history

- 6. Use DMF ftp MSP to send copies off-site
  (buddy site with CSIRO Aspendale Vic, JCU to UQ).
- Pros
  - At least something is done!
- Cons
  - Slow: holds up clearing of disc
  - File names are obfuscated remotely
  - No user visibility of data at remote site

Off-site Storage – Examples from history

- 7. Arrange with another large-scale storage site to do cross-copying (filesystem to filesystem), so the second copy of each sites' files are kept at the remote site.
- Pros
  - no extra tape costs
  - second copy can be easily recovered over the network
  - full file visibility at both sites
- Cons
  - managing the synchronisation could be hard
  - not scalable (inotify?, rsync, lsyncd, etc)
  - all files, or selected?
  - no ‘backup’ protection unless done with DMF (or rsync options)
  - need to decide the time delay
  - Hard to do deletes (except for Gareth's reverse rsyncs)

Off-site Storage – Examples from history

- 8. Use external services, e.g. cloud
- Pros
  - Seems easy
- Cons
  - Maybe expensive
  - Cloud data does get lost
  - Security and privacy issues
  - May not give coverage back in time
  - Provider may not survive

Off-site Storage – Examples from history

- 9. Use Yotta-Yotta – intercepts FC actions to disc at first site, and replicates actions at second site.
- Pro
  - Full concurrent read-write access to both sites!
- Con
  - Cost
  - Complexity
  - Need extremely good network
  - Difficulty of reconciling after interruption
  - No longer in business!

- Lots more!

Off-site Storage – Examples from history

- 10. Current CSIRO SC practices
- Copy full dumps on Sundays to remote site, along with DMF databases, etc.
  - User filesystem: /datastore – small files are not migrated (< 28 kbyte)
  - Full dump contains these files: 42% of files (about 11 million), 35% of on-line disc space, 0.2% of data.
- Copy incremental dumps to remote site daily
  - Any new/changed small files
  - No coverage for large files – assumed to be re-generatable, or protection in users' hands
- Commenced making third copies at times to remote tape drives

•Off-site Storage – Suggestions for future

- SGI bundled solution!
- Attempt in RFQ in 2003-2004!
  - **Replication**
  - There will be situations where CSIRO will require parts of the file system to be automatically replicated to remote sites. It is ***highly desirable*** for the proposed system to accommodate this aspect together with consideration of this file as a 'copy' which can be retrieved [15].
  - *[15] Complies.*
  - *Remote copy to a remote site can be achieved in numerous ways under DMF. SGI will need to understand the data volumes, network bandwidths and business objectives of this process in order to propose an appropriate configuration.*

•Off-site Storage – Suggestions for future

- Keep working on the problem!
- Cooperation between sites
  - Just writing second copies at remote site does not give quick recovery if the primary site is out of action
- Split holdings into categories
  - identify critical data needing to be highly protected
  - use DMF tags to manage different classes of data
- SGI developments may help

Thank you

CSIRO IM&T Scientific Computing Services
Robert C. Bell

t +61 3 9669 8102
e Robert.Bell@csiro.au
w www.hpsc.csiro.au