Backups

CSH5 Chapter 57

Backups

M. E. Kabay & Don Holden

1

Topics

- Backup Bloopers
- Definitions and Needs
- Strategies
- Archives, Maintenance and Retention
- Storage
- Disposal
- Costs
- Optimizing Frequency

2

Backup Bloopers

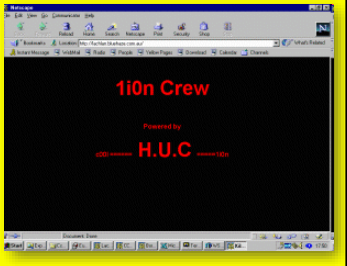
- BerkshireNet
- Margot Kidder
- Digital Technologies Group
- Amtrak Reservations
- Newcastle University Prof
- Sun Valley Ski Resort
- Stanford University Grad School of Business
- McAfee's QuickBackup 2.04
- US Customs
- US National Archives
- MIT "Cheating Scandal"
- Office of the Vice President of the USA
- Norwegian Banks
- Swisscom Mobile GSM Network Failure
- Macomb County (Michigan) Sheriff's Dept
- Electronic Voting Machines

3

BerkshireNet

Feb 96 -- RISKS 17.83

- BerkshireNet, Pittsfield, MA
  - Swastikas and racist messages
  - Masqueraded as provider's administrator
  - Erased data on two computers
  - Shut down system
- Down 12 hours
  - no current backup
  - older deleted files restored
  - but several days of data lost



4

Margot Kidder

Sept 96 -- *People Online* via RISKS 18.46

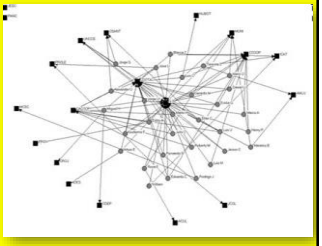
- Computer virus was last straw
- Led to nervous breakdown
- Unidentified virus apparently destroyed only copy of her book
- No backup

5

Digital Technologies Group

Oct 96 -- AP

- Digital Technologies Group lost all computer files and backups
  - U$17,000 direct costs
  - loss of months of work
  - 1-week shutdown
  - seriously damaged credibility as ISP
- Disgruntled ex-employee
- Alleged perpetrator arrested



6

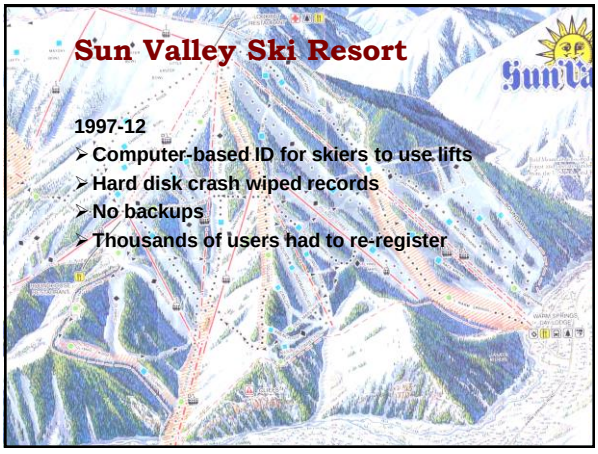
Newcastle University Prof

1997-08

- Prof. David Newell
- School of Clinical & Laboratory Research
- Expert in cancer research and anti-cancer drugs
- Computer + 5 floppy disks stolen
- Contained only copy of his research data
- Thieves kindly returned his disks

7

Sun Valley Ski Resort



1997-12

- Computer-based ID for skiers to use lifts
- Hard disk crash wiped records
- No backups
- Thousands of users had to re-register

Stanford University Grad School of Business

1998-03

- Installed additional disk capacity on servers
- Reloaded files from a corrupt backup tape
- All faculty and student files destroyed
- Many faculty and graduate students lost research data – because they didn't keep their OWN backups!

9

US Customs

1998-10

- Hard disk failure
- Chain of system failures
- 6 hour downtime
  - 80,000 requests backlogged
- Sys admins shut down modem lines
  - Importers had auto-redial
  - Each redial counted as new request
- Customs Service deleted 45,000 import requests
- Required resubmission from importers

10

US National Archives

1999

- System problem wiped 43,000 messages from e-mail system
- But contractor responsible for backups had made no backups
- Someone had also turned off system logging
- Assistant Archivist said, "Safest way to save important messages is to print them out."

11

MIT "Cheating Scandal"

massachusetts institute of technology

2000-03

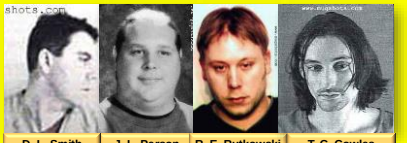
- Boston Globe reported "cheating scandal"
  - 22 students in cell bio class had wrong grades
- Turned out a grad student had sorted the names but not included the grades. . . .
- *No backups,* so no records of correct grades
- No logs, no tracking, no *nothing*

Office of the Vice President of the USA

- 1998-04
  - Migrated e-mail server to Windows NT4
  - Moved e-mail files to a new partition
  - Forgot to include that partition in backup schedule
- 1999-05
  - Staffers discovered zero backups for e-mail for over a year

Macomb County (Michigan) Sheriff's Dept

- 2002-05: RISKS 22.08
- Lost 50,000 photos of criminals
  - Hard drive crashed
- Had some pictures printed out as backups
- But had made no electronic backups at all on any medium

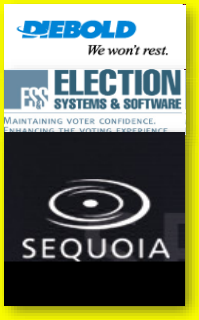


D. L. Smith MELISSA | J. L. Parson W32.Blaster.B | R. E. Rutkowski Ur-spammer | T. C. Cowles Ur-spammer

14

Electronic Voting Machines

- 2004 election period saw growing concern about electronic voting machines
  - Entirely proprietary (secret) code
- Most had no audit trail of any kind
  - Trust the electronic counts without means of verification
  - Documented cases of votes being switched to opposite party
- Demands growing for paper trail
  - Print ticket to allow user to see vote supposedly registered by machine
  - Place tickets in ballot box as backup measure in case of recount



15

Definitions

- Backups – copies of data files or records
  - *Should be* on different device
    - ✓Tape, hard disk, CD, DVD…
  - *May* be on same disk (*.wbk, *.bak etc.)
- Units & abbreviations
  - KB = kilobyte = 1024 bytes (characters)
  - MB = megabyte = 1024 KB = 1,048,576 bytes
  - GB = gigabyte = 1024 MB = 1,073,741,824 bytes
  - TB = terabyte = 1024 GB = 1,099,511,627,776 bytes
  - PB = petabytes = 1024 TB = 1.126 E15 bytes
  - EB = exabytes = 1.153 E18 bytes

16

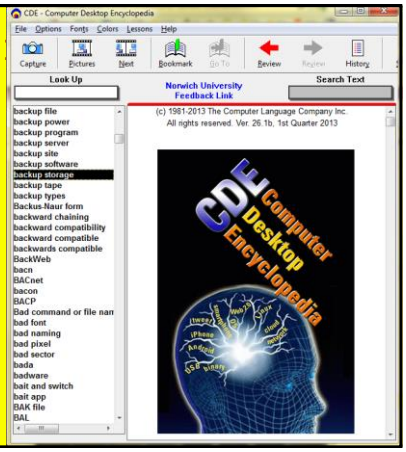
Needs

- Provide valid information in case of data corruption or loss
- Satisfy legal requirements for access to storable data, as for audit purposes;
- In forensic examination of data to recognize and characterize a crime, and to identify suspects;
- For statistical purposes in research;
- To satisfy requirements of due care and diligence in safeguarding corporate assets;
- To meet unforeseen requirements.

18

Backup

- Following slides are built using information and images from the *Computer Desktop Encyclopedia v 26.1b.*
- Copyright © 2013 Computer Language Corp. All rights reserved.
- Used with permission.



19

Mag Tape Cartridges

Capacities range from a few GB to 100 GB per unit (2013)

A Lotta Cartridges

These are most of the tape and disk cartridges (as well as cassettes) that have come on the scene since the mid-1990s.

20

Tape Libraries (PB total)

From CDE; courtesy Iomega Corporation and Storage Technology Corporation

Optical Disks

- CD (compact disk)
  - 650 MB capacity
  - ~3 Mbps transfer rate
- DVD ("digital videodisk" → "digital versatile disk" → "DVD"
  - 4.7 GB single-side single-layer
  - 8.5 GB single-side double-layer
  - 9.4 GB double-sided single-layer
  - 17 GB double-sided double-layer
  - Transfer rates ~200 Mbps

22

DVD Types

Standards from the DVD Forum

- DVD-R: DVD-Recordable
  - Write-once, read-many
  - "DVD-dash-R" or "DVD-minus-R"
  - 4.7 GB standard
- DVD-RW: DVD-Read-Write
  - Write many, read many
  - "DVD-dash-RW" or "DVD-minus-RW"
  - 4.7 GB standard
  - Rewrite 1,000 times

23

DVD Types (cont'd)

Standards from the DVD+RW Alliance

- DVD+R: DVD+Recordable
  - Write-once, read many
  - 4.7 GB
- DVD+RW: RVD+ReadWrite
  - Write many, read many
  - 4.7 GB
  - Supports CAV recording format especially good for data random access
  - DVD+RW drives typically read DVD-R media as well

24

Blu-ray Disks for Backups

- Developed in 2002 by Sony, Hitachi, Philips…
  - Same size as CDs and DVDs
  - Single-layer: 25 GB
  - Dual-layer: 50 GB
- Blu-ray Burners
  - Internal
  - External
  - In 2013, about $60
- Blu-ray software
  - E.g., *WinDVD, PowerDVD, Handy Backup*
  - 25 - 128 GB/disk depending on layers (1, 2, 3, 4)
  - Speeds up to ~288 Mbps (may increase in future)

25

Hard Drives for Backups

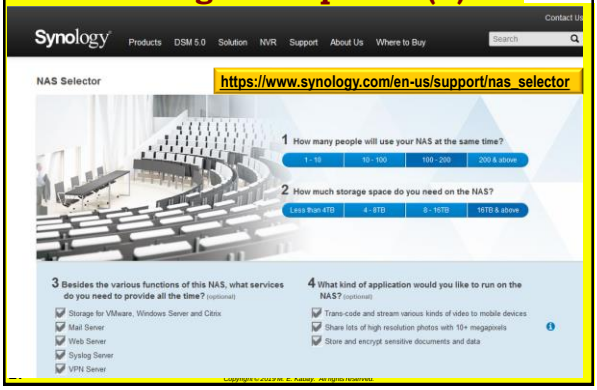
- For consumers: small external drives
- USB 2.0/3.0, Firewire
  - Maxtor, Seagate, Western Digital, ….
  - In Mar 2014, ~$30/TB for 3 TB drive
- For corporate servers
  - Network Attached Storage (NAS)
  - More expensive (2-10x)
  - Fail-safe engineering
  - RAID technology
  - Specialized processors
  - Rack-mounted or standalone
  - Install disk drives to suit
- Transfer rates ~3Gbps

WD MyBook 3 TB $125

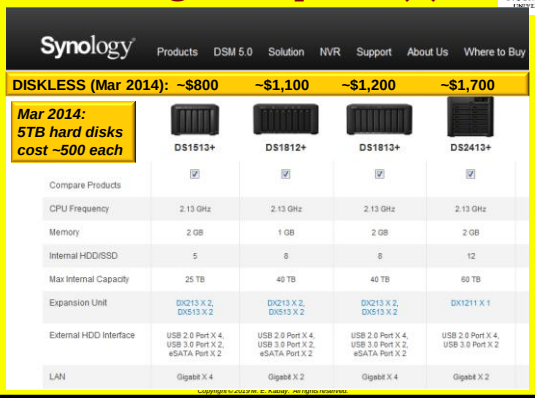
Synology DS412+ $1159 for 16 TB

26

Browsing NAS Options (1)



Browsing NAS Options (2)

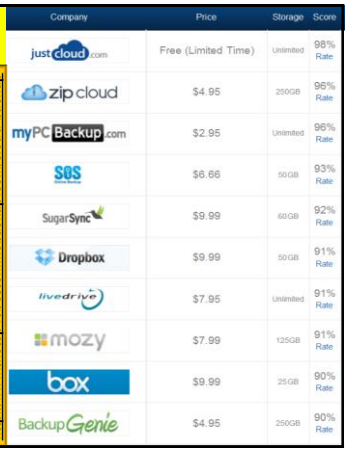


28

Cloud Backup (1)

- Once connected to 'Net
  - Backup from anywhere
  - Retrieve from anywhere
- Warnings:
  - Confidentiality
  - Shared folders may result in accidental deletion or sabotage
  - Xfr rate depends on 'Net bandwidth

http://www.thetop10bestonlinebackup.com/cloud-backup



29

Cloud Backup (2)

- Consider for organizations too
- See "Why Cloud Backup: Top 10 Reasons"*
  1. Achieve disaster recovery with secure offsite cloud backup
  2. Freedom from manual and complex tape backup tasks
  3. Predictable costs for simpler budgeting
  4. Reliable and guaranteed data recovery
  5. Minimized risks and costs of downtime
  6. Fast data restores
  7. Take advantage of service provider's expertise and resources
  8. Offload regulatory compliance requirements to service provider
  9. Well-managed cloud more secure than your own
  10. Ease of setup and use, set-and-forget, no training required

30

* http://www.ironmountain.com/arma/docs/top_ten_reasons.pdf

Strategies

Everything

- Exclusive access
- Types by coverage
  - Full
  - Differential
  - Incremental
  - Delta
- Types by target
  - System
  - Application
- Testing

Everything changed since the last *full* BU

Everything changed since the *last BU of any type*

All the *records* that have changed since the last BU

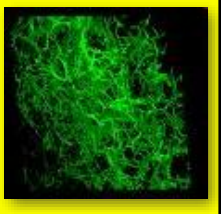
WARNING:
Always ask for a *definition* of *"Daily* "or "*Partial"* BU

31

Archives, Maintenance and Retention

What to keep? How long?

- Retention policies
- Rotation of media
- Media longevity and technology changes
  - Media degradation (e.g., fungus*)
  - Software and hardware changes
    - ✓Application programs
    - ✓Operating systems
    - ✓Transport mechanisms
    - ✓Data encoding schemes



* See Mendham, S. (2005). A little care goes a long way. *PC World.* http://www.pcworld.idg.com.au/index.php/secid;3;id;461315701

32

Storage

Humidity 40-60%
Temp 50-70F

- Environmental protection
- Onsite
- Offsite
  - Care during transport
  - Homes
  - Safes
  - Banks
  - Data vaults
  - Online backups

Location Storage Type

Bad idea

Access?

33

Offsite Storage

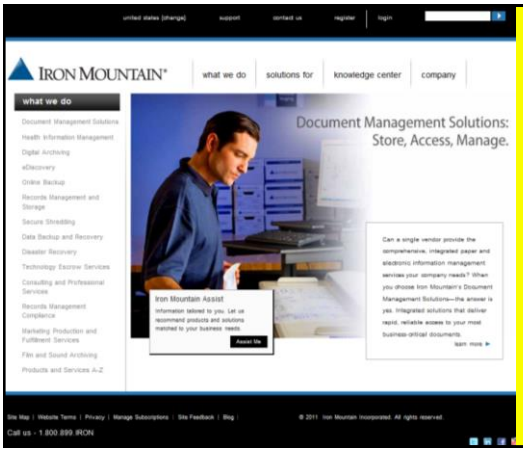
IRON MOUNTAIN®

- Document Management Solutions
- Health Information Management
- Digital Archiving
- eDiscovery
- Online Backup
- Records Management and Storage
- Secure Shredding
- Data Backup and Recovery
- Disaster Recovery
- Technology Escrow Services
- Consulting and Professional Services
- Records Management Compliance
- Marketing Production and Fulfillment Services
- Film and Sound Archiving

These functions are copied verbatim from Iron Mountain's list of services.

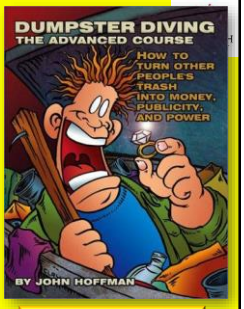
http://www.ironmountain.com/services/

34



Disposal

- Scavenging
  - Dumpster® Diving
  - Scratch tapes
  - Diskettes, cartridges
- Data and media destruction
  - Erase ≠ destroy
  - Reformatting inadequate
  - Degaussers inadequate
  - Random overwrites
    - ✓Military-grade = 7 passes



http://tinyurl.com/6dhnu2y

Stop data remanence!

"Wipe"

36

Disposal (cont'd)

- Physical destruction works
  - Recommended: specialized shredders
  - Incineration (special types)
- Secure collection points
  - Lockbox magnetic garbage
- Mobile data destruction services
  - Paper shredders
  - Special facilities for tapes

Royal Media Shredder

http://tinyurl.com/24mwtb

37

Secure Data Destruction

- Equipment for destroying documents:
  - Paper shredders (crosscut ONLY – see oval)
  - See "Cross Cut Shredders – 5 Reasons Why They Are Better" < http://tinyurl.com/65cuw52 >
- Considerations when evaluating services
  - Costs
  - Types of media processed
  - Onsite vs offsite
  - Chain of custody assurance
  - Regulatory compliance
  - Environmental protection

38

Costs

- Tape costs
  - Tapes / BU
  - Cost/tape
- Time costs
  - Hours/BU
  - Operator cost/hour
- Fixed costs
  - Storage space, racks, insurance, transport, hardware rental/maintenance, software. . .

39

Optimizing Frequency

E(x) = P(u)*C(u) - P(n) *C(n)

Where

- x is particular strategy
- E(x) is the expected value or cost of the strategy
- P(u) is the probability of having to use the backup within a single day
- C(u) is money saved by not having to redo all the work
- P(n) is probability of not having to use the backups at all
- C(n) is cost of making and storing BU

40

Personal Notes on Kabay BUs

- Two ~identical computers: MAIN & SPARE
- 2 1-TB portable USB drives for alternate-day use
- ViceVersa synchronization software (MS)
- AM:
  - Synch MAIN → USB (including Backups)
  - Synch USB → SPARE
- PM
  - Synch SPARE → USB
  - Synch USB → MAIN
- Daily incremental BU 23:00 → MAIN HD & MAIN Flashdrive
- Monthly FULL BU MAIN → TB drives in fire safe
- Since 2014: more extensive use of Dropbox for active files

41

Now go and study

42