Süsteemihaldus
MTAT.08.021
System Administration
Storage and Backup
Outline

- Storage
- RAID - Redundant Array of Inexpensive Disks
- LVM – Logical Volume Management
- Backup
Storage Hierarchy

• “Classical” view, as seen from the processor:
  - primary
  - secondary
  - tertiary
  - off-line
Storage Media

• magnetic storage
  - hard disk drive
    • 4,200 rpm to 15,000 rpm
    • Up to 3.0 Gbit/s SATA2
  - floppy disk, tape

• optical storage
  • CD (780nm) (150 KiB/s at 1X speed)
  • DVD (640nm) (11.08Mbps)
  • HD-DVD, Blue-Ray (405nm) (36.55Mbps)

• semiconductor storage
  - “Flash” storage, Solid State Drive (SSD)
Magnetic storage basics

"Ring" writing element

Longitudinal Recording (standard)

Recording layer

"Monopole" writing element

Perpendicular Recording

Recording Layer

Additional Layer
Optical storage basics

CD
- \( l = 800 \text{ nm} \)
- \( p = 0.6 \mu\text{m} \)
- \( \varnothing = 1.6 \mu\text{m} \)
- \( \lambda = 780 \text{ nm} \)

DVD
- \( l = 400 \text{ nm} \)
- \( w = 600 \text{ nm} \)
- \( p = 740 \text{ nm} \)
- \( \varnothing = 1.1 \mu\text{m} \)
- \( \lambda = 650 \text{ nm} \)

HD DVD
- \( l = 200 \text{ nm} \)
- \( w = 320 \text{ nm} \)
- \( p = 400 \text{ nm} \)
- \( \varnothing = 620 \text{ nm} \)
- \( \lambda = 405 \text{ nm} \)

Blu-ray
- \( l = 150 \text{ nm} \)
- \( w = 130 \text{ nm} \)
- \( p = 320 \text{ nm} \)
- \( \varnothing = 480 \text{ nm} \)
- \( \lambda = 405 \text{ nm} \)

0.1 mm
0.6 mm
1.1 mm
1.1 mm
0.1 mm
What are these?
Hard Disk Drive Interfaces

- **SAS (Serial Attached SCSI)**
  - 3.0 Gbit/s at introduction
  - 6.0 Gbit/s available February 2009
  - 12.0 Gbit/s in development

- **SATA (Serial ATA)**
  - 1.5 Gbit/s SATA1
  - 3.0 Gbit/s SATA2
  - 6.0 Gbit/s SATA3

Artjom Lind (artjom.lind@ut.ee)
28.02.2013
https://courses.cs.ut.ee/2013/syshald
Hard Disk Drive Interfaces

• FC (Fibre Channel)
  – 1GFC  200 Mbps  1997
  – 2GFC  400 Mbps  2001
  – 4GFC  800 Mbps  2004
  – 8GFC  1600 Mbps  2005
  – 10GFC 2550 Mbps  2008
  – 16GFC 3200 Mbps  2011
  – 32GFC 6400 Mbps  2014 (announced)
Hard Disk Drive Interfaces

- IDE, ATA, Also: Parallel ATA
  - (Integrated Device Electronics,
  - (Advanced Technology Attachment)
  - 16 MB/s originally
  - later 33, 66, 100 and 133 MB/s
  - Dead in 2003
- SCSI (Small Computer System Interface)
  - SCSI-15MB/s (1986)
Other possible Interfaces

• FireWire (IEEE 1394)
  - 50–400 MB/s (1995)
• USB
  - USB 1 12 Mbps 1996
  - USB 2 480 Mbps (35Ms/s) 2000
  - USB 3 5,000 Mbps (625 Mb/s) 2008
• Ethernet ?
HDD: characteristics

- interface, r/w transfer rate, cache
- transfer rate of the interface ≠ “real” read/write
- transfer rate
- capacity
- physical dimensions
  - 3,5”, 2,5”
- hot-swap-ness
- rotation speed (rpm)
  - 4200, 5400, 7200, 10k, 15k
Storage Virtualization: RAID

- no virtualization – JBOD
- DAS – Direct Attached Storage
- RAID (Redundant Array of Independent Disks)
  - dividing and replicating data among multiple disks
  - the result is presented as one virtual disk
- implemented in hardware and/or software
Storage Virtualization: RAID

RAID 0 (striped set) – not really “redundant”
Storage Virtualization: RAID

RAID 1 (mirror)
Storage Virtualization: RAID

- combined RAID “levels”:
  - RAID 0+1 (01)
  - RAID 1+0 (10)
Storage Virtualization: RAID

RAID 5 (striping with parity)
Storage Virtualization: RAID

RAID 6 (striping with dual parity)
Storage Virtualization: RAID

- RAID 0 – fast r/w, poor reliability
- RAID 1 – fast r, good reliability, expensive
- RAID 1+0 – fast r/w, good reliability, very expensive
- RAID 5, 6 – fast r, slow w, better economy
  - large caches are used to compensate slow writes
- write caching needs battery-backed cache
- RAID is not a backup strategy
Storage Virtualization: LVM

• LVM – Logical Volume Management
• volume – an integral chunk of storage
  – physical volume – disk or RAID array
  – logical volume – built from physical volumes by concatenating, striping, mirroring, slicing
  – functionality (all on-line): extending-reducing, moving data between physical volumes
  – also uses some of the RAID terminology
Storage Virtualization: NAS

- NAS (Network Attached Storage)
- different level of abstraction compared to DAS
  - still needs a “physical” storage layer beneath
- operating on the file system level
  - NFS (Network File System)
  - CIFS (Common Internet File System)
Storage Virtualization: SAN

• SAN – Storage Area Network
  – many-to-many connections
  – dedicated storage network
  – operating on the block level

• block devices over “regular” networks
  – (S)ATA over Ethernet (AoE)
  – iSCSI (Internet SCSI)
Data Backup

• data protection process
  – protects against:
    • hardware failures
    • software bugs
    • user errors
    • administrator errors

• archiving process
  – ...as sometimes required by law
Data Backup

• the main (sole?) purpose of backup is:
  • enabling a successful restore
Backup Costs

• reliable backup systems are expensive, but to lose your data could cost even more!

• backup system must grow with storage

• files are backed up regardless of the importance of their contents
  – use technology and policies to make storage more effective
Backup Strategy

• before the backup we must know (from the risk analysis process):
  – what kinds of critical data do we store and process?
  – where are these data sets located?
  – are some of them more critical?
  – how long restore times can we afford?
Backup Strategy

• the level of the Data Backup service:
  – what data sets to backup (coverage)
  – how often to backup (schedule)
  – how many older backups to retain (history)
  – restore requirements (speed, granularity)
  – archiving requirements
Backup Levels

- full backup (level 0)
  - incremental level 1
  - incremental level 2
- In the backup schedule, the levels are combined
  - level 0 + level 1
  - level 0 + level 2
Backup Schedule

- daily full backups are rare
- most schedules combine full + incremental
  - how often to make full backup?
  - what is the data “change rate”?
- media recycling, archiving
Restore

• restore / recovery
  – file recovery
  – file system recovery
• you can only recover what you have backed up!
• test the recovery
  – test the “full recovery”
Data Backup Security

- physical security
  - large-scale disasters may affect all of your equipment
  - off-site backups are recommended

- network security
  - sensitive data does not stop being sensitive on the backup tapes
  - security measures must be equal or better compared to the system being backed up
Backup Software

• which client platforms are supported?
• do your databases and repositories need special backup interfaces?
• are you being charged by the quantity of data or by the number of clients?
Backup Hardware

• tape drives
  – (S)DLT - ((Super) Digital Linear Tape) – SDLT-320, SDLT-600, DLT-S4, DLT-V*
  – LTO (Linear Tape-Open), LTO-2, LTO-3, LTO-4...

• tape library, jukebox

• disk based backup
  – Disk-to-Disk backup (D2D)
  – Disk-to-Disk-to-Tape backup (D2D2T)
Storage problems

- hard drives tend to be the most unreliable components in the system
- user needs tend to grow faster than storage capacities
- storage capacities tend to grow faster than storage data transfer rates
- storage capacities tend to grow faster than backup system capacities
Some Unix utilities:

- Software NAS (expose block device over network)
  - nbd-server
  - nbd-client
- dd
  - Byte-wise read/write utility
- Rsync
  - File-wise synchronization utility
- Tar
  - File-wise archive utility