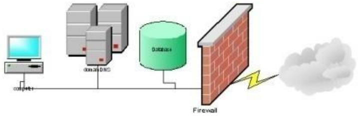
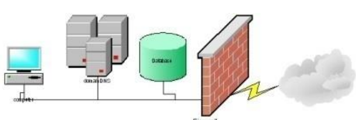
Management of Information Systems

Data Management



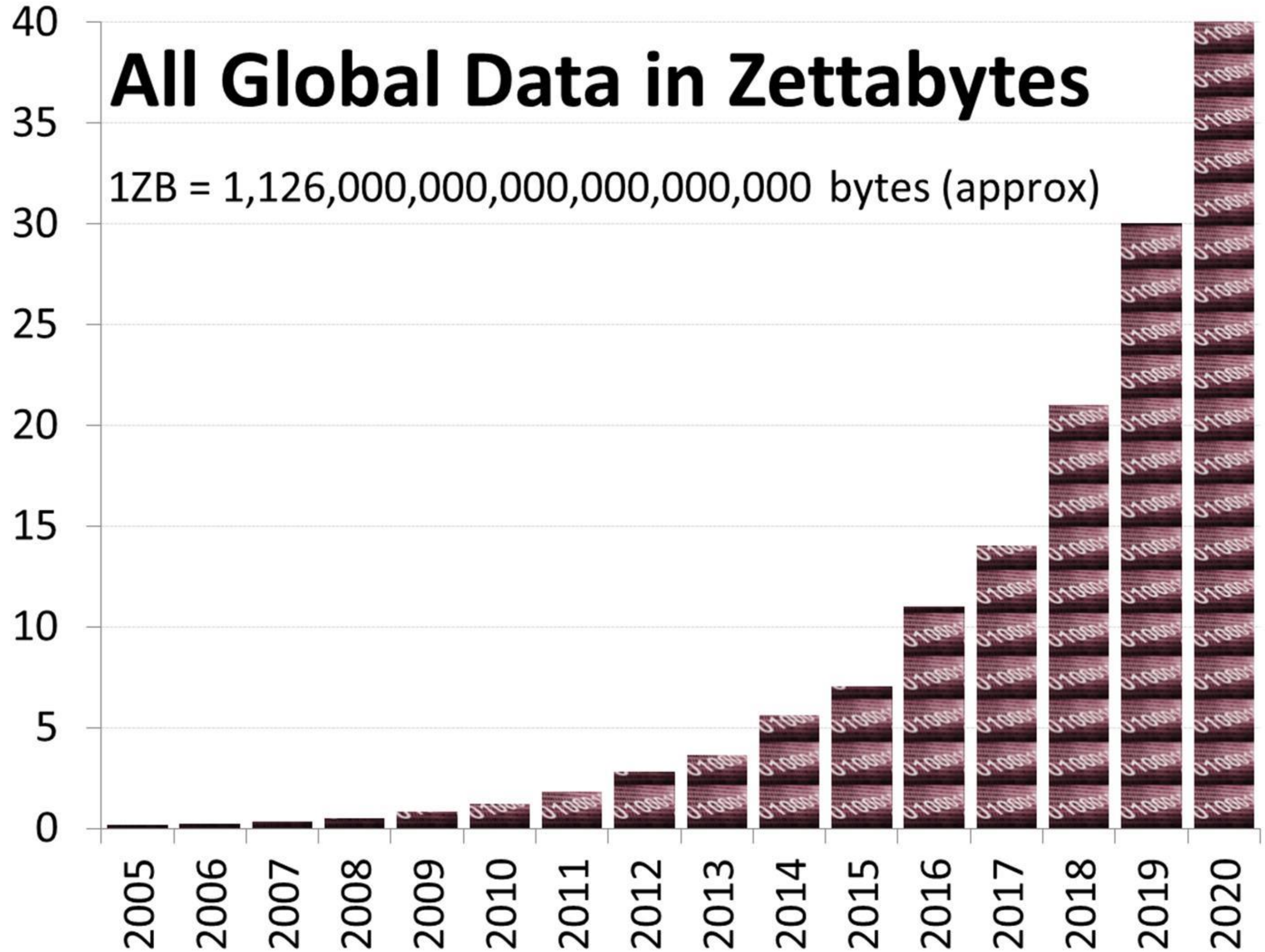
Data Management

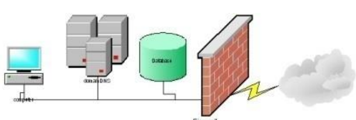
- Types of data
- Data storage types
- Reliability of disks (RAID)
- Data storage systems (DAS, SAN, NAS, IP SAN) and transmission technologies
- Data copy methods (Volume/Flash copy)
- Virtualisation



Data Generated

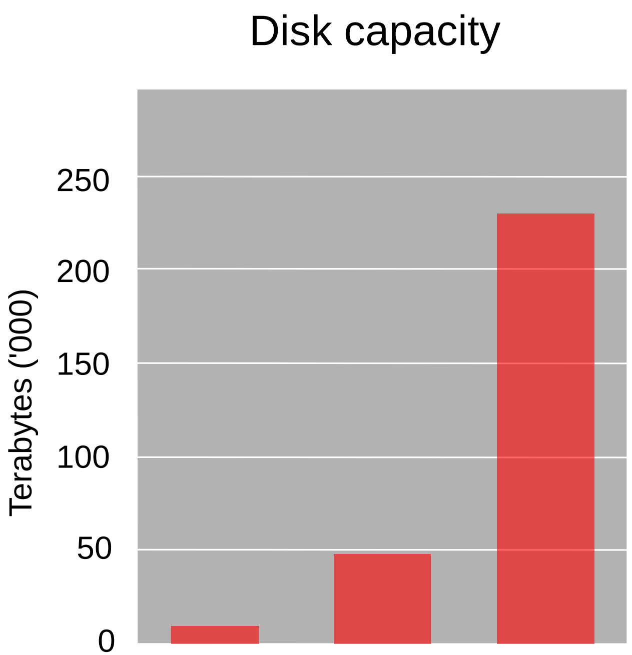
- Zetta = 10^{21}
- In the past two years as much data genereated as much till then

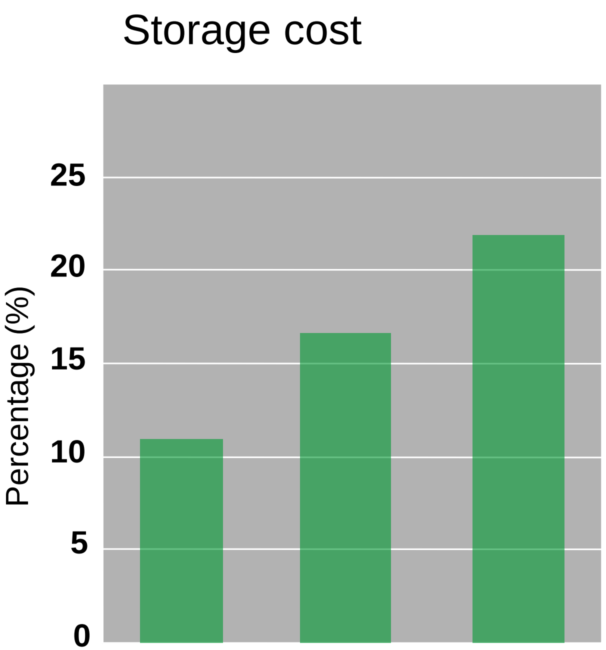


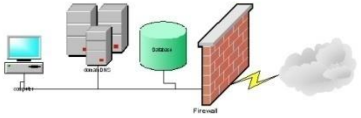


Storage Capacity, Cost

More data doesn't mean more information – but means more cost!



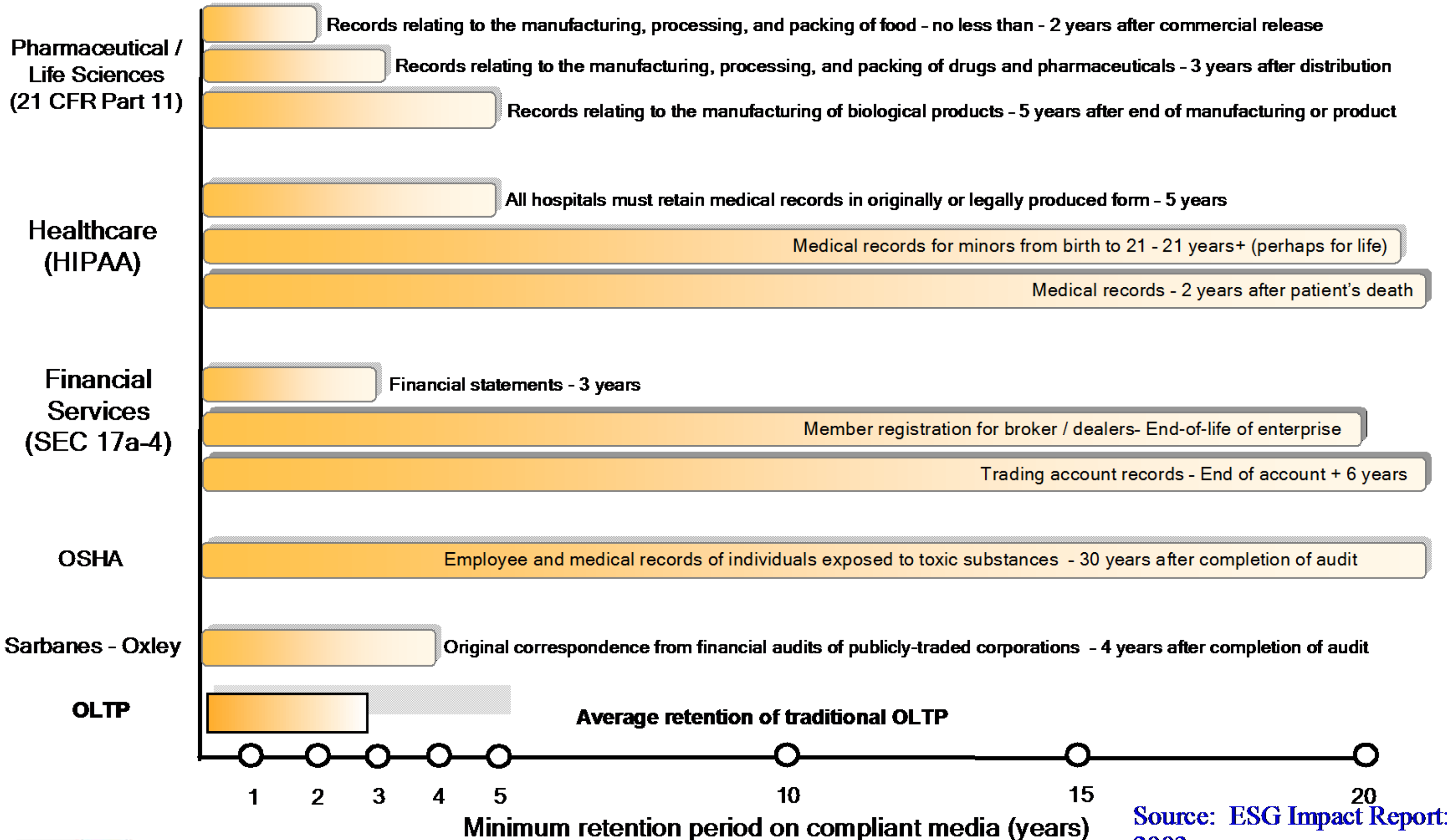




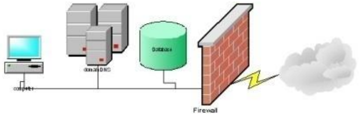
Data Types

Data type	Action
Empty storage capacity	Requires more
Non business files	Delete
Duplicated data	Delete / Share
Redundant application data, log files, dump files, temporary files	Clear – often
Stuck, orphan data	To archive
Valid data	Protect, reach and strore this data

Minimum retention periods

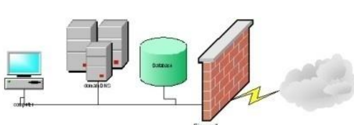


Source: ESG Impact Report: Compliance, May 2003



Data Stream Types

Structured	**Unstructured**
Known source and goal	Unknown
Server to Server	Client to client, Client to Server
Business devices	Personal devices
Special applications	E-mail, Web
Hard security	Soft security
Private networks	Public networks
Planable	Hard to plan, statistic
Database	File servers
It can be monitored, and planned	**It can only be regulated by policies**

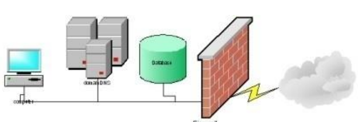


Data Management

1. Simplification and automation of IT infrastructure and its management
Goal: lower TCO (TCO, *Total Cost of Ownership*), higher ROI (ROI, *Return on Investment*)

2. Business continuity, security and safety
Goal: business continuity

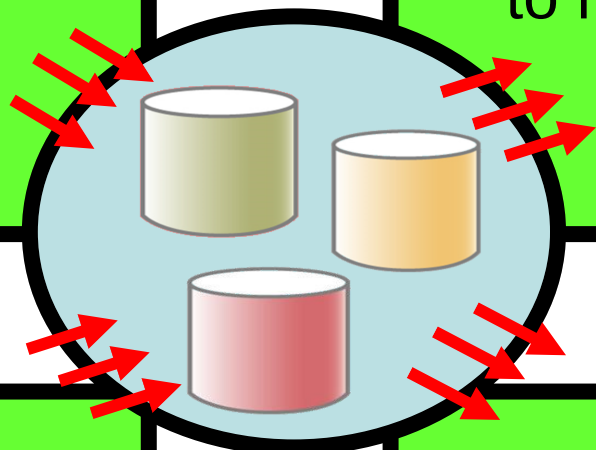
3. Efficient information lifecycle-management
Goal: store the information proper of its value, take legal regulations into consideration



Information and Data Life Cycle, Data Types

Record and code information

Analisation of data form to retrieve the original information



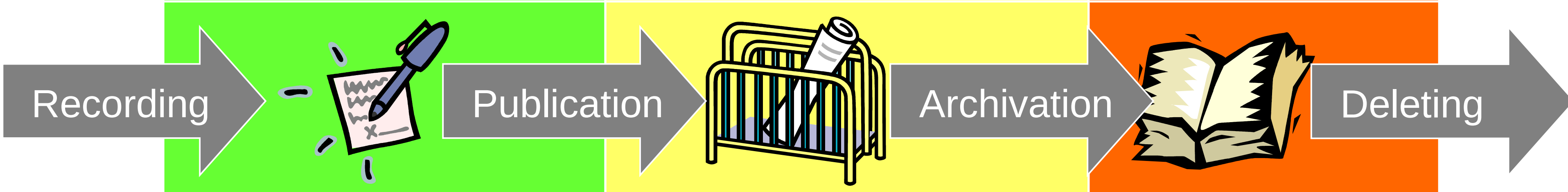
Making index and meta-data to search and to catalogue

Datamining –

logical, statistical processing of bulk data

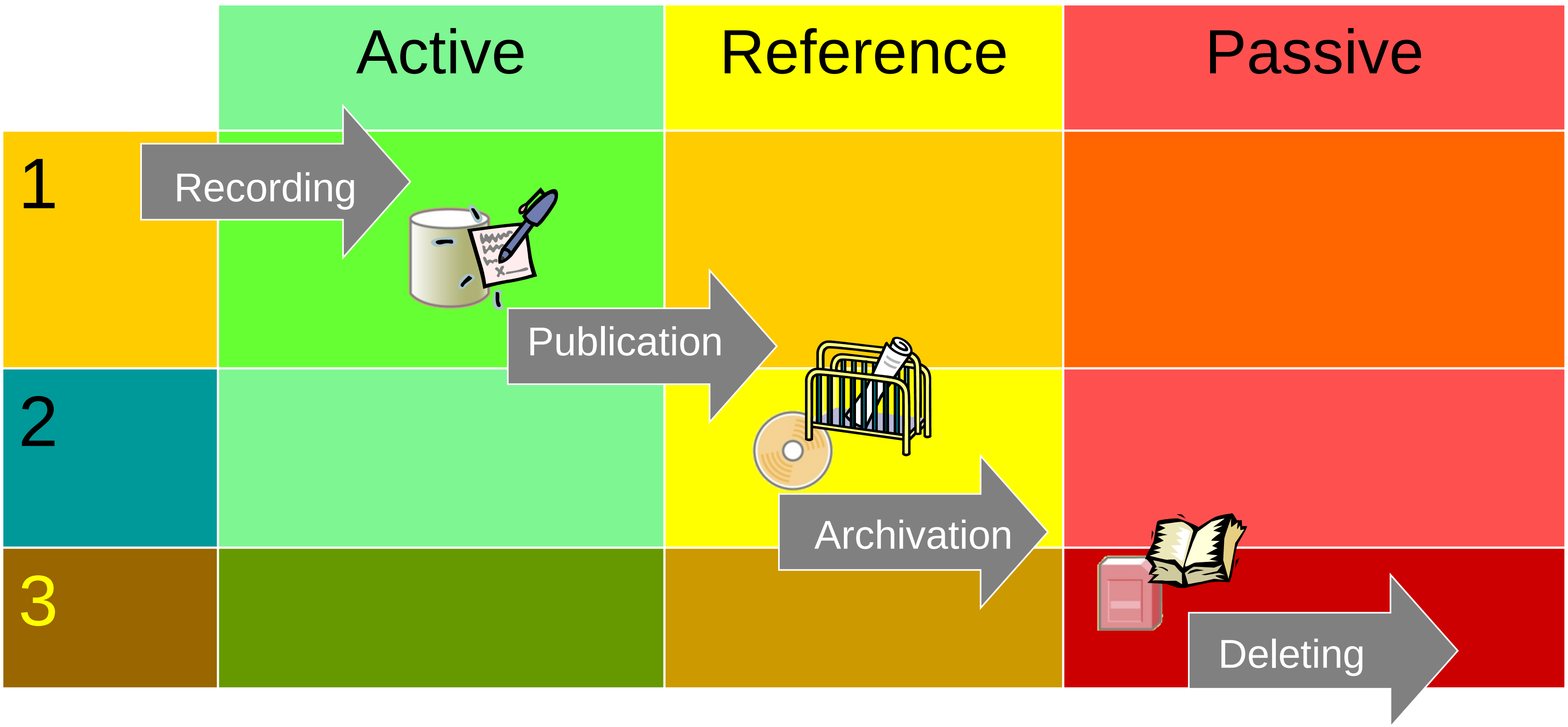
Need of Data - Changes

Active	Reference	Passive

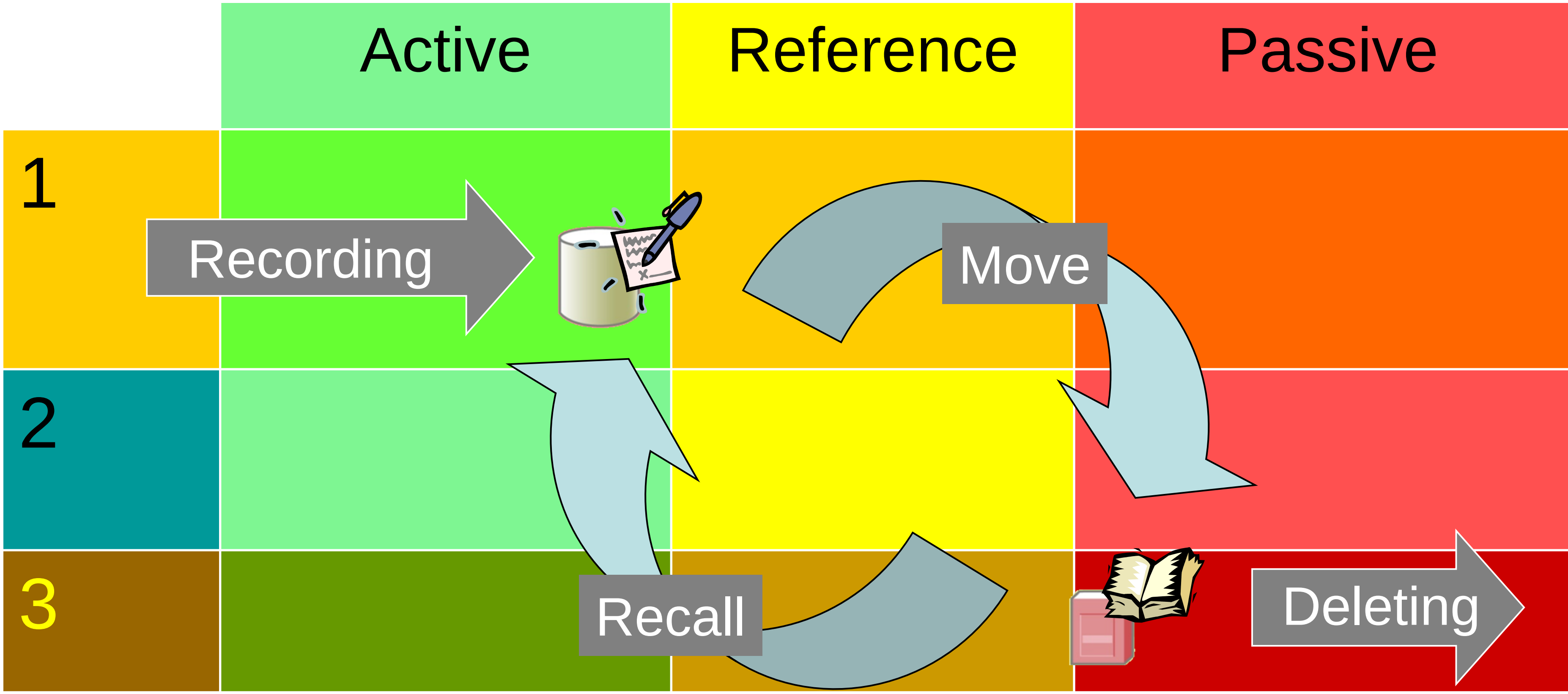


Active	Reference	Passive
Application DB	Riports	Archives
Online Transaction Processing (OLTP)	Transaction riports	Back-ups
Document handling	Public documents, Web sites, e-mail messages	Scanned-pictures Security video records
Software development	Audio/Video records	

Optimal Storage Technology: Select Accoring to Data Value

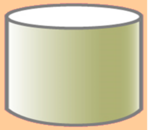
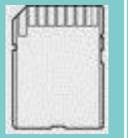


Hierarchycal Storage Management (HSM)



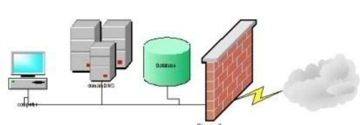
HSM parameters: size, type of data and the time of the last access

Storage Devices

	Advantages	Problems
Disk 	• „immediate” data access • „random” read/write (R/W)	• Disk replace • Power supply, cooling • Life time 3-4 years!
Flash memo-ry 	• No moving parts • Fast	• (Yet) expensive
Optical 	• Secondary storage –WORM (*Write Once Read Many*)	• Drop behind the development of disk and tape equipments, SOHO device (*Small Office Home Office*)
Tape 	• 10-20x cheaper than the disk storage • Storage time: 30 years	• Non-immediate data access • Serial read/write (R/W)

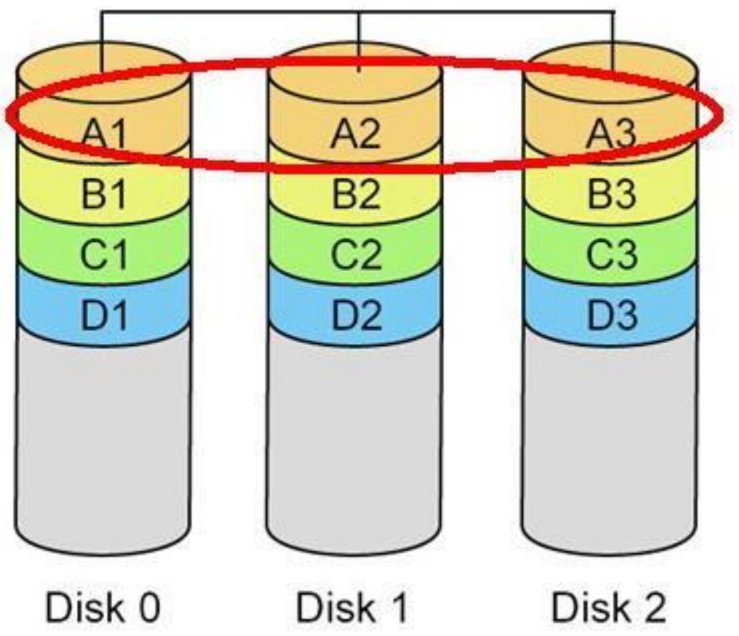
Reliability of disks

- Data stored on disks – protect from demage
- Originally one high-reliability disk
  - SLED: Single Large Expensive Drive
    - Expensive…
  - MTBF (Mean Time Between Failure)
    - Appr. 750 000 hours (appr. 85 years)
    - But a large(??) disk array with 1000 disk
    - MTBF: 750 000 hours / 1000 = appr. 1 month



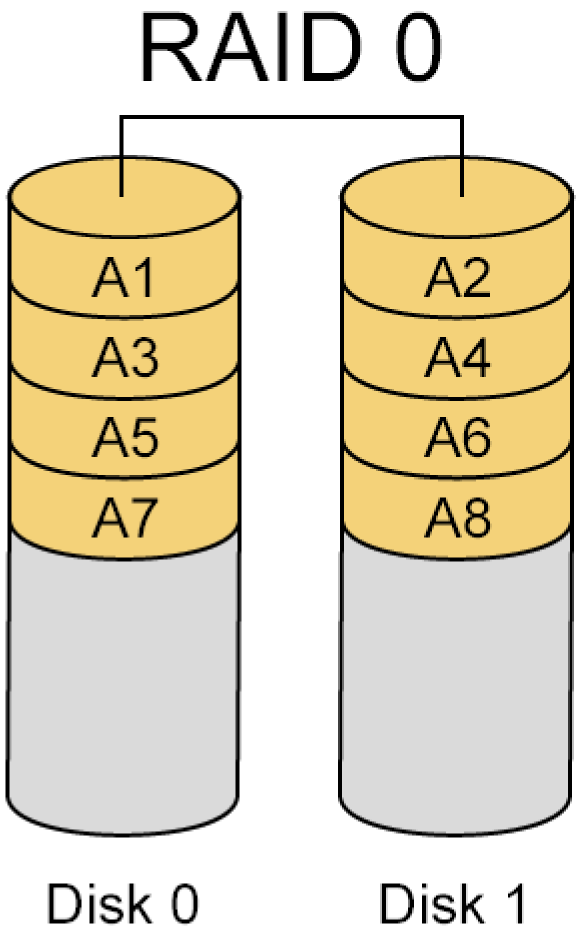
1.2.4 RAID Redundant Array of Independent (Inexpensive) Disks

- 1987 California, Berkeley University
- RAID 0, 1-5, 6
- Disks are divided into stripes



RAID 0 - striping

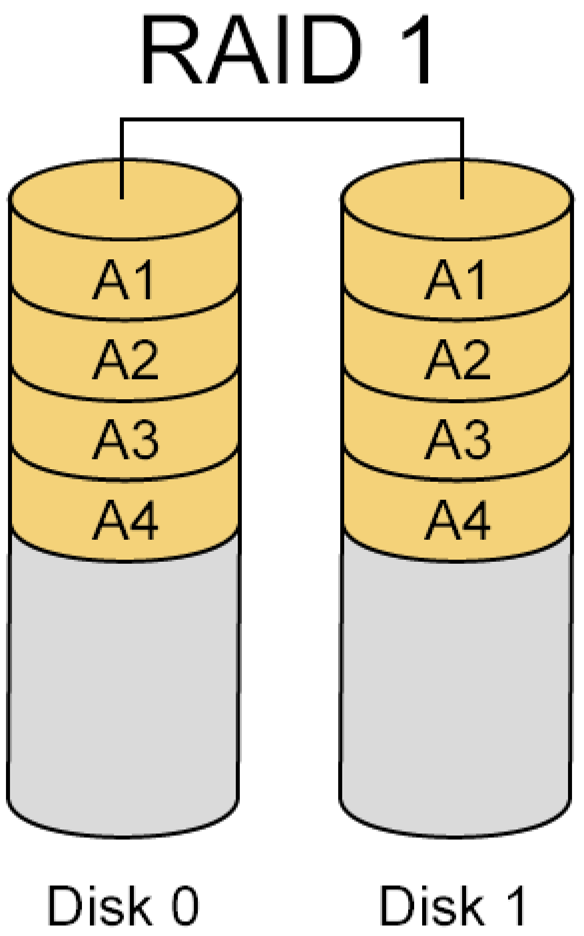
- Goal: Enlarge the capacity
  - not to increase the reliability
- Increase speed
  - (parallel read/write operations)



RAID 1 – Disk Mirroring

- Parallel operations
- High reliability
- Large overhead in size (2x!)

- Can we have approximately the same reliability with smaller overhead?

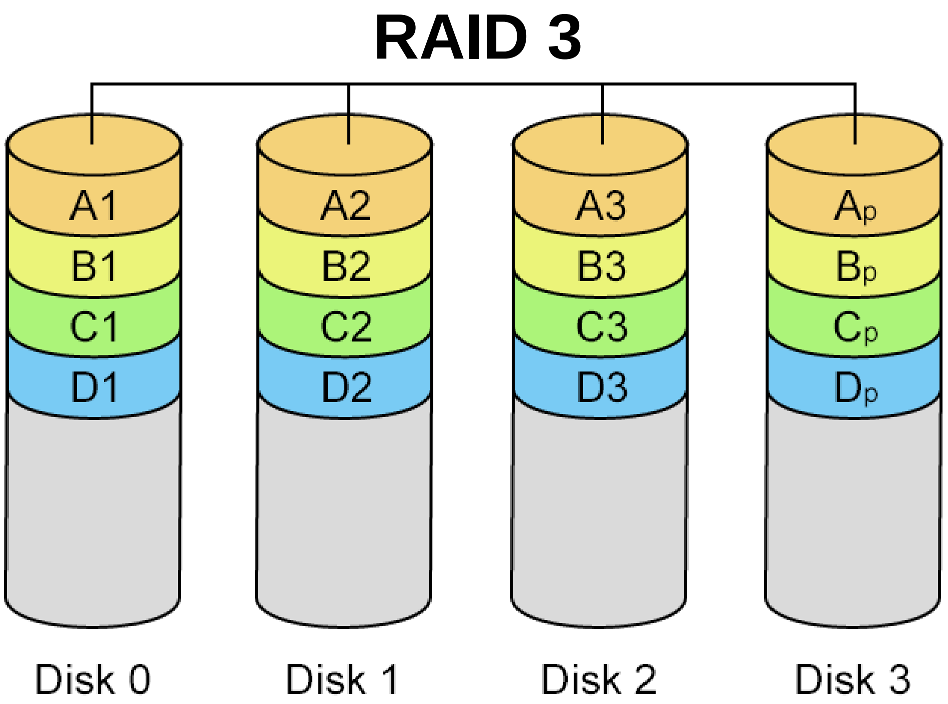


RAID 2 – Error Correcting Code

- Division into stripes
- ECC – (Error Correcting Code) are stored on certain disks
  - Suitable for detecting and correcting errors
- Not used nowadays, because ECCs are used *inside* the disks

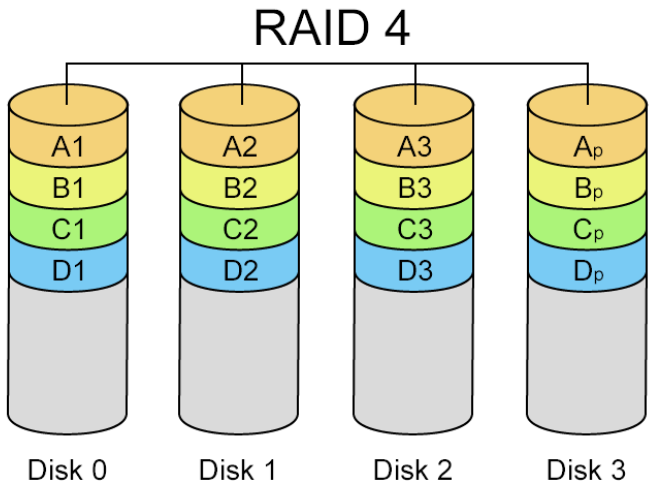
RAID 3 – Parity Disk

- +1 disk: parity (XOR)
- One disk fails: content can be reproduced from the others by XOR – time(!), slow
- Cannot detect disk errors
- Small stripes, operations always on whole stripes
  - Single user
- Typical: 2+1, 5+1, 8+1, 14+1
- Parity disk constrains
- For large files (video)



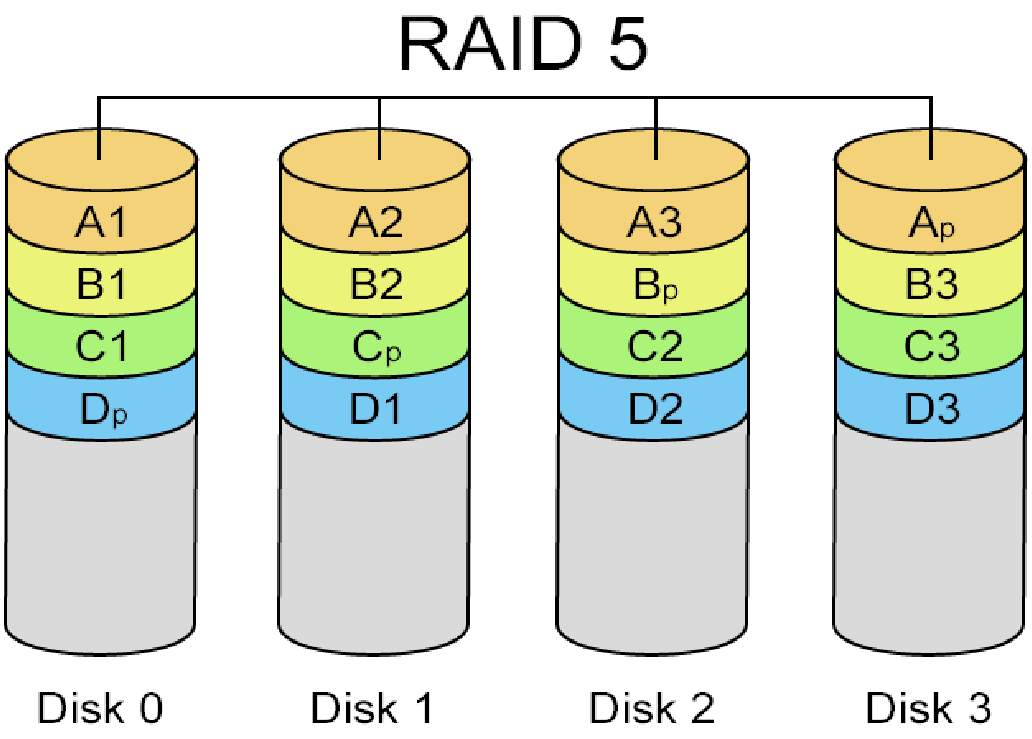
RAID 4

- Similar to RAID 3, but large stripes
- Each disk can be accessed directly
  - Allows parallel service
  - Parity disk constrains very much
- Not used in practice



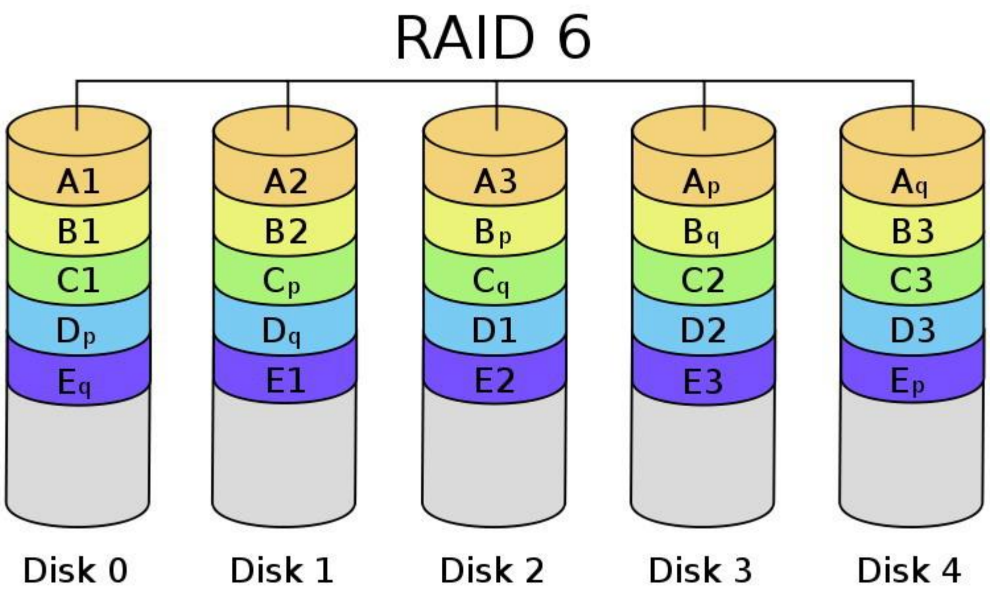
RAID 5 – Distributed Parity

- Parity distributed equally
- Eliminates the bottleneck of parity disk
- Each disk can be accessed directly
- Variable stripe size

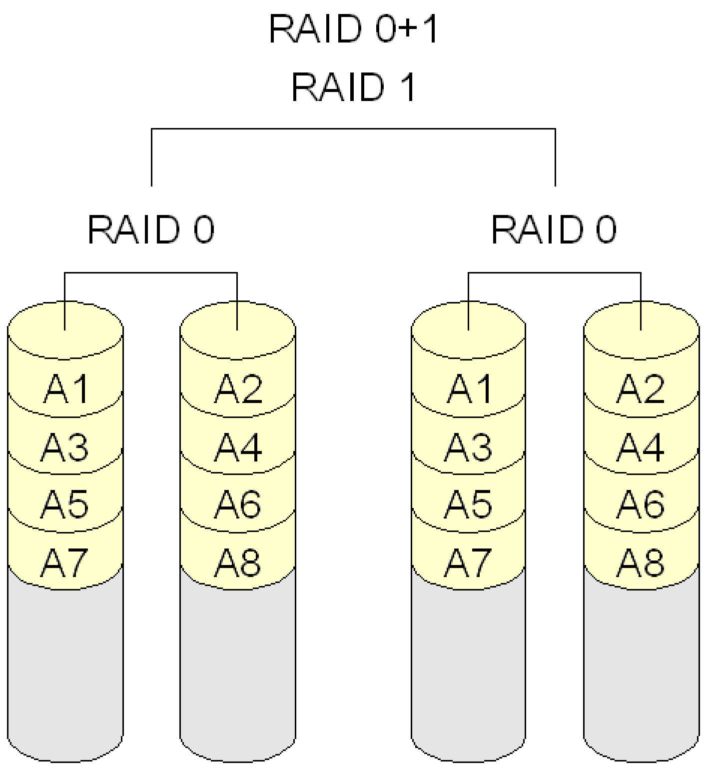


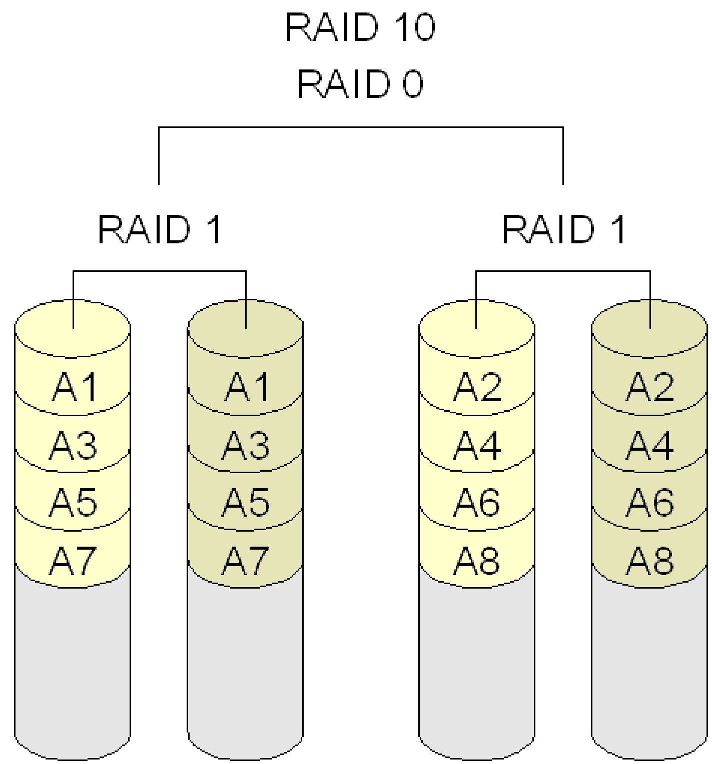
RAID 6 – Double Parity

- Row (XOR – P) and Coloumn (Reed-Solomon Code – Q) parity
  - Protects against double failures – but very slow
  - Distributed among disks



RAID 01, RAID 10





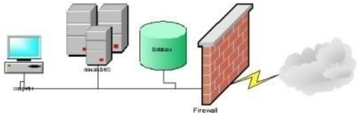
Same?

Error: whole stripe – no mirror (!!)
Restore: whole stripe(!!)
Speed: high (HW striping)

no mirror only on half
only the wrong disk
slower (SW striping)

RAID

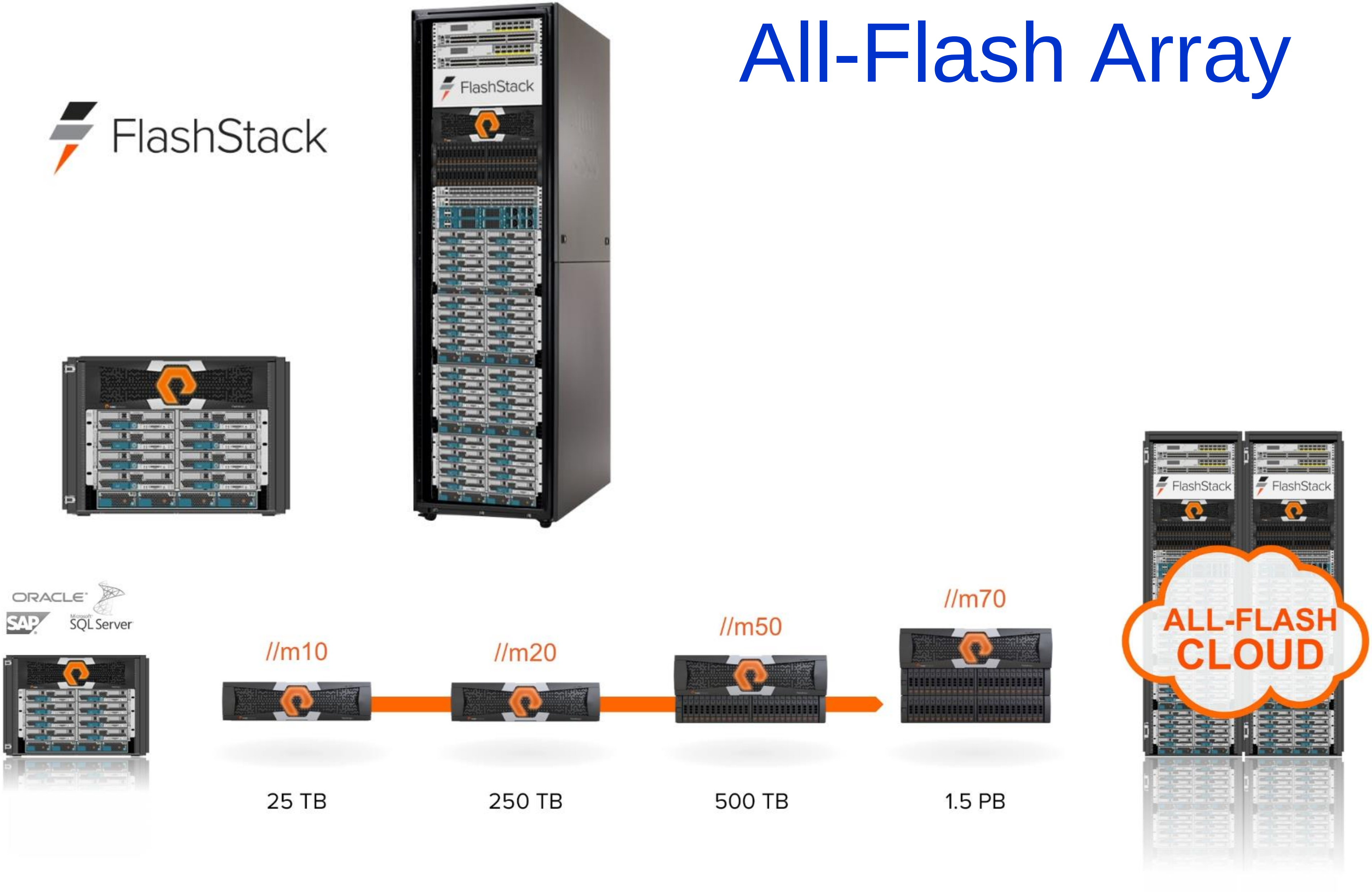
- In practice: 0,1, 5 used frequently
- RAID can protect only against PHYSICAL errors!!!!
  - – Against logical errors: back-up



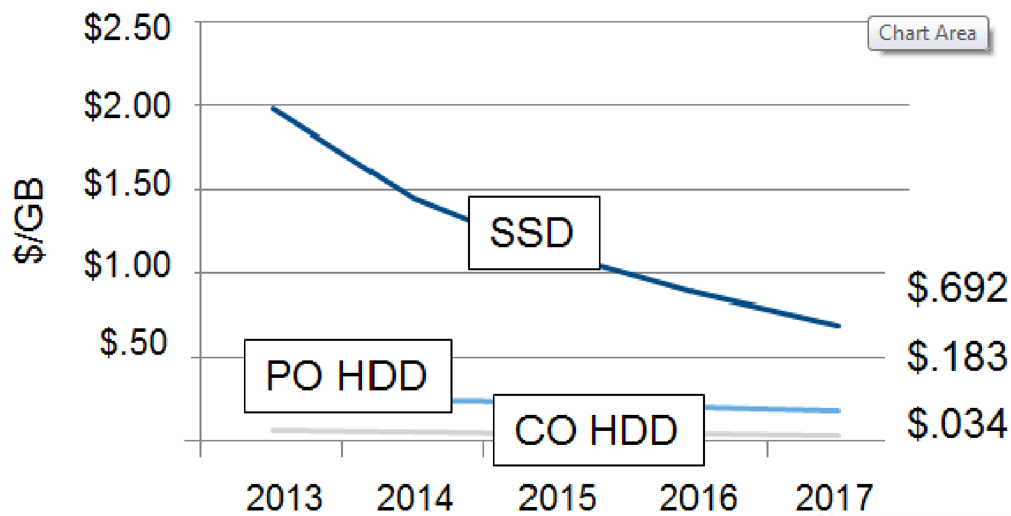
Flash memory

- Non-volatile memory (EEPROM)
  - Erasable by a special signal (‘flash’)
  - Pen-drives
  - SSD
    - Solid State Drive
  - AFA
    - All-Flash Array
  - Hybrid arrays

All-Flash Array

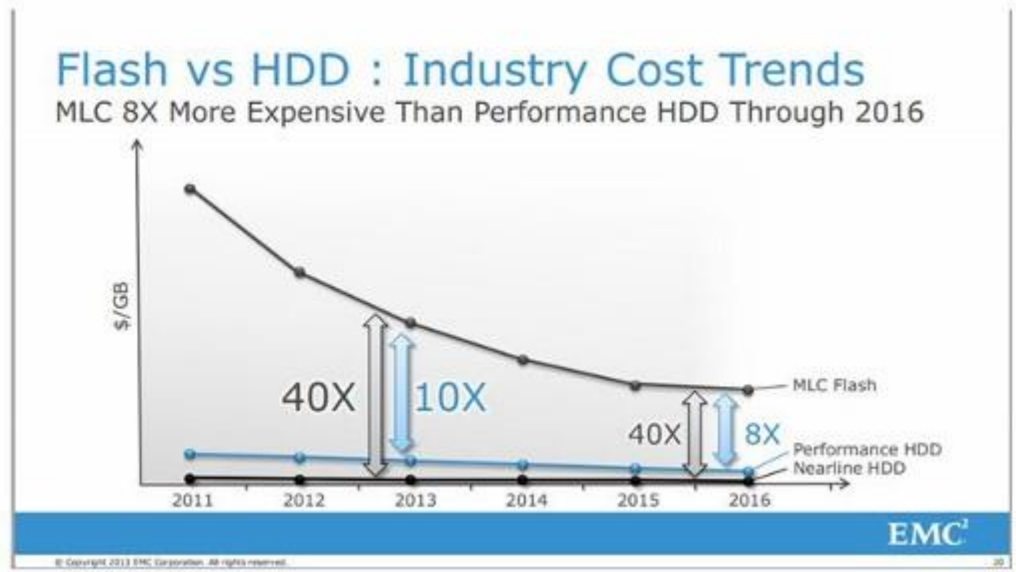


Price



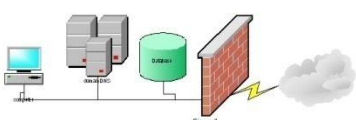
PO: Power optimised
CO: Cost optimised
TOC: Total Cost

- ~20-25% of the TOC (HDD)
- ~80% of the TOC (SSD)



Latency

- Latency 1000:1 (HDD:SSD)
  - Speed of the connections (SATA, SCSI) is the bottleneck…
  - Faster boot time
  - Data reduction technologies can be more effective
    - 6:1 (SSD) – 2:1 (HDD)
      - If it is not done by a layer above – like in VM
    - Effective usable capacity
    - In SSD typically done by controller while in HDD extra SW
  - Faster – more user requests can be served
  - More satisfied users



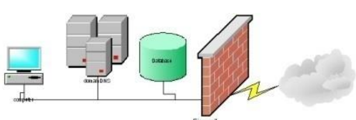
Power, cooling, size

- Power, cooling
  - No moving part
  - Less power consumption (1:10)
  - Less cooling requirement
  - 50-90% less than of HDD
- Smaller
  - Even if count with data reduction
  - 20-40* data density than of HDD
  - Up to 90% higher rack utilisation

All-Flash Array

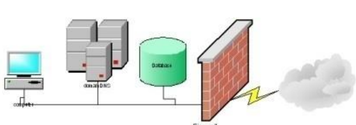
Reliability

- Less maintenance cost
- Reliability
  - 100:1 (unrecoverable bit error rate)
    - 10^{-18} : 10^{-16}
    - 10^{-18} means 1 error in several hundred years
    - In early EEPROMs writing cycles were quite limited (~1000)
    - Nowadays SSDs – for more than 200 years
  - RAID (1 in older, 5 in newer versions)



Hybrid arrays

- HDDs better in price/GB
- SSDs better in price/IOPS
  - Input/output operations per second
- Add a thin slice of flash to an HDD array
  - SSD: 2% - 5% of total capacity
  - available IOPS may double
  - read latency from 10+ ms to 3-5 ms
    - Not constant – may cause problems in some applications
  - 10% - 20% increase in array price
  - 2X performance gain
  - RAID 1



Comparision

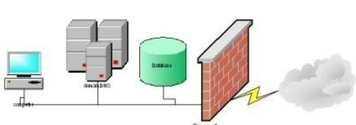
	Avg upfront per GB HW Cost	1yr avg per GB Power & Cooling	Avg upfront per usable GB Costs*	~ 3 yr per usable GB TCO
100% HDD Storage System	$0.50	$0.50	$0.63	$2.13
Hybrid Storage System	$1.60	$0.38	$0.59	$1.72
100% Flash SSD Storage Systems	$10.00	$0.17	$3.64	$4.13
100% Flash SSD Storage Systems	$5.00	$0.17	$1.82	$2.31
100% Flash SSD Storage Systems	$4.00	$0.17	$1.45	$1.95
100% Flash SSD Storage Systems	$3.00	$0.17	$1.09	$1.59

** HDD cost/usable GB goes up because of formating and RAID overhead.*

SSD based dedupe/compression increases usable GB ~ 2.75 x.

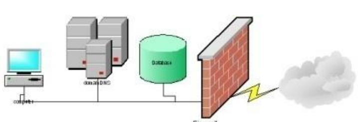
Hybrid systems a little less to account for the HDD overhead.

SSDs are typically overprovisioned from 20 to 50% to account for load balancing & garbage collection



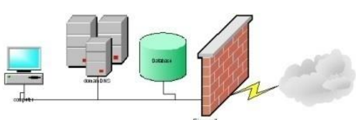
Storage Networks

- DAS – Direct-Attached Storage
- SAN – Storage Area Network
- NAS – Network-Attached Storage
- IP SAN (iSCSI)



DAS – Direct-Attached Storage

- Storage is connected directly to server
  - Block level access
  - Mainly in small(er) systems
- Two subtypes
  - Internal DAS
  - External DAS



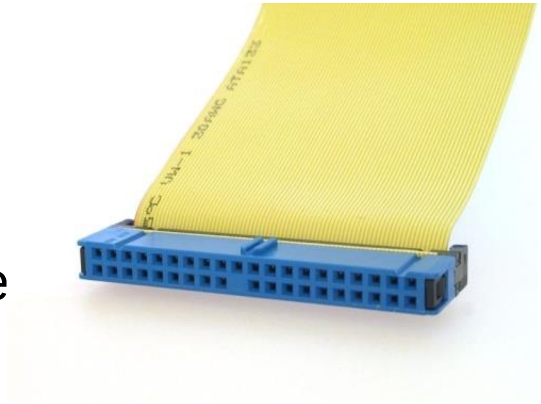
Internal DAS

- Storage is connected directly to server by an internal parallel or serial bus
  - Limited distance
  - Limited number of devices can be connected
    - (P)ATA or SATA connectors
  - Requires large space inside the server
    - Complicated maintenance

(P)ATA

- (P)ATA – Parallel Advanced Technology Attachment
  - Half duplex

- 40 & 80 wire/cable
  - 40 wire limited to UDMA 33 MB/s and below
  - 80 wire allowed for UDMA 66, 100, 133 MB/s
  - Development stopped in 2004 (because of the space requirement of the cable)

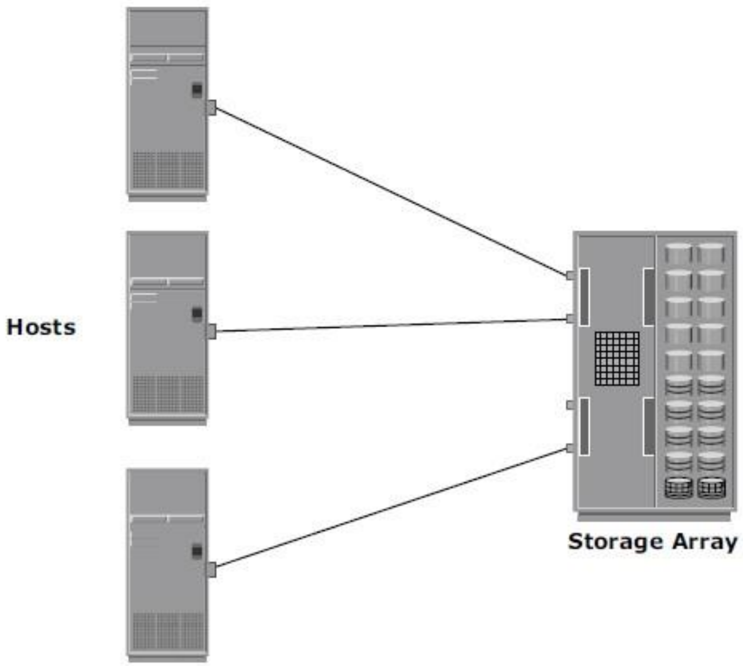


SATA

- Serial Advanced Technology Attachment (SATA)
- Point-to-point connection between the SATA host adapter and the SATA device
  - Half duplex
- New connecting interface
- Higher transmission speed
  - (P)ATA 66/100/133 MB/s
  - SATA 150/300/600 MB/s
- The connecting cable has 4 wires, max. length 1 m

External DAS

- Server is directly connected to an external storage
  - Higher distance
  - Typically not (as much) limited the number of devices
  - SCSI (or FC) connection

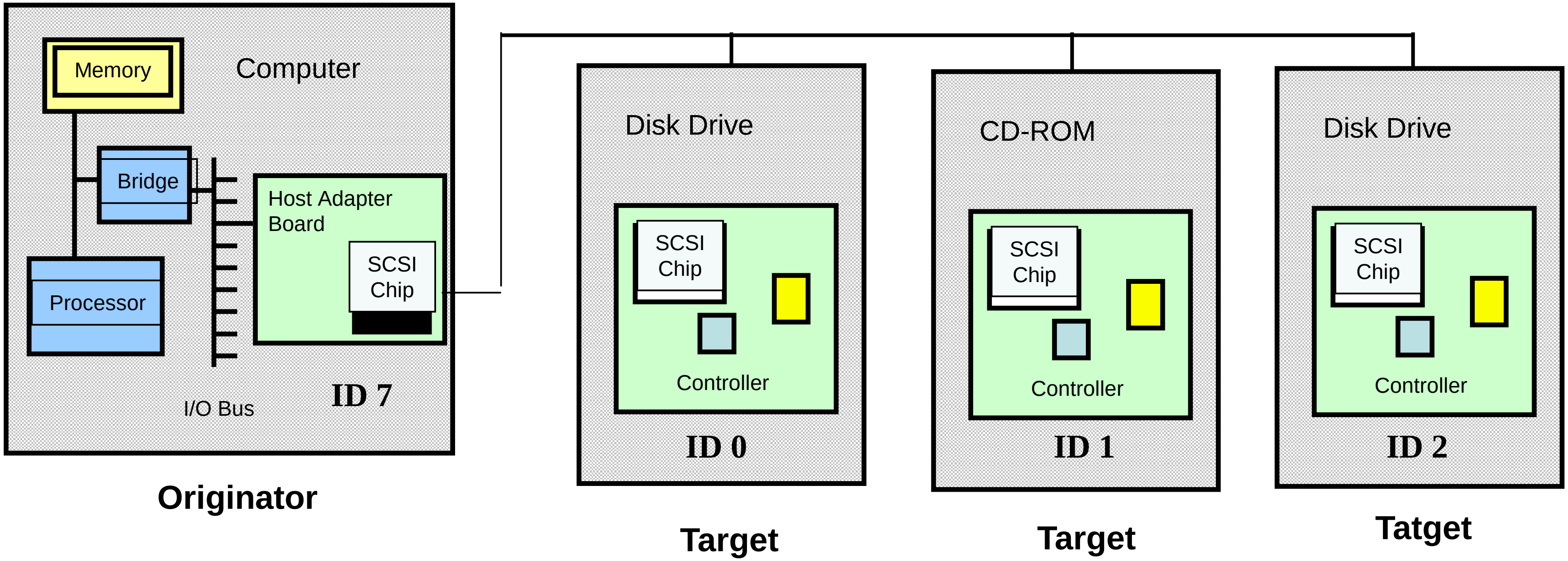


SCSI Interface

- SCSI
  - Small Computer System Interface (SCSI)
  - Standardised I/O bus
- Devices
  - Disk Drives
  - Tape Drives
  - Removable Media Drives
  - CD-ROM, CD-R/CD-RW Drives
  - Optical Memory Drives
  - Printers

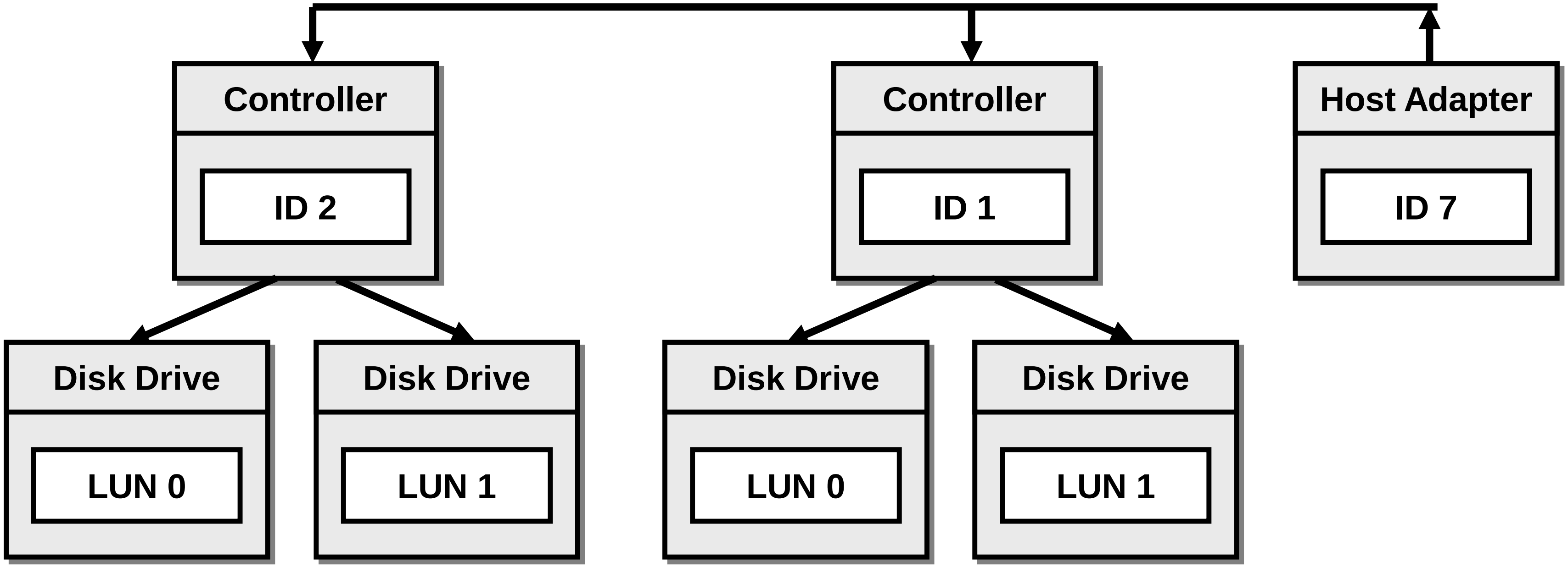
SCSI Topology

- All devices on one common bus
- Devices have their own, unique identifiers
- Roles
  - The originator - generally computer
  - The target - typically disk, CD-ROM

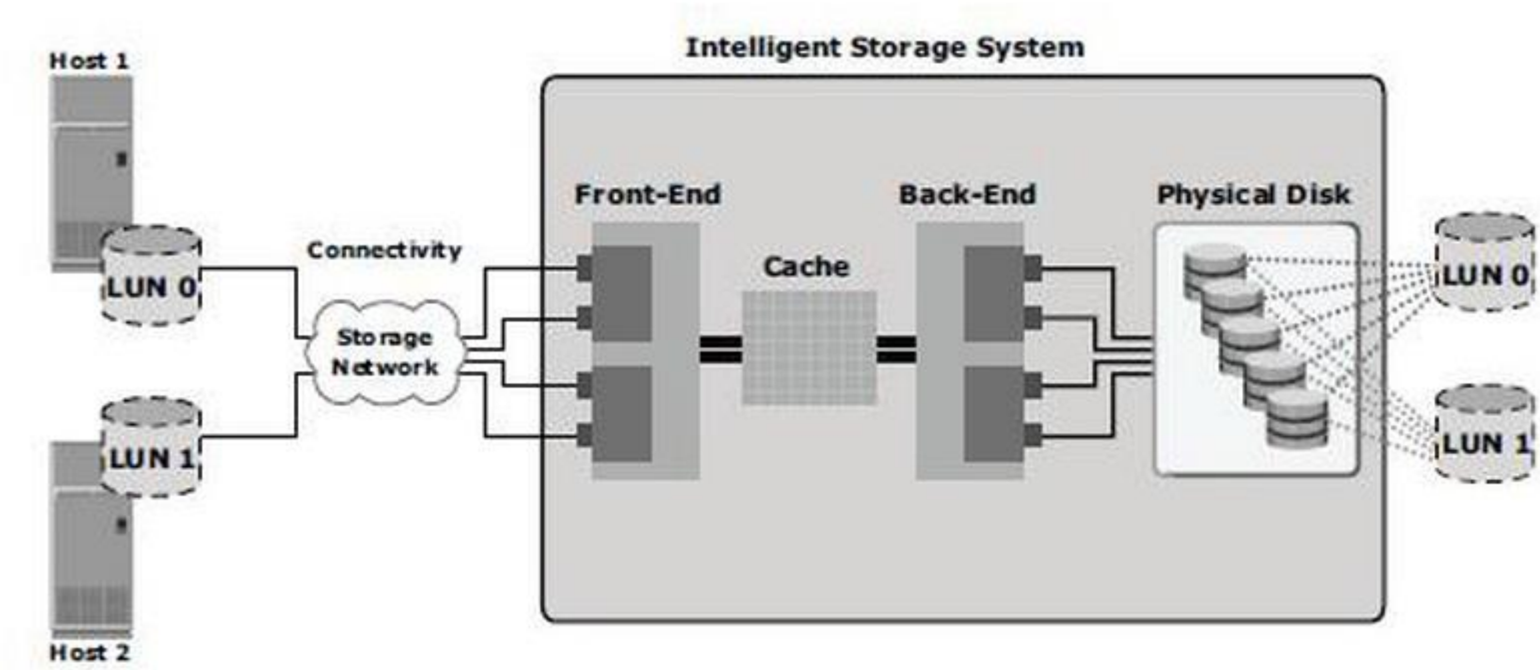


Logical Unit Number (LUN)

- NOT(!!!) a number – Group of storages
- LUN is an SCSI group address method



LUN


Intelligent Storage System
Host 1
LUN 0
LUN 1
Host 2
Connectivity
Storage
Network
Front-End
Cache
Back-End
Physical Disk
LUN 0
LUN 1

SAS - Serial Attached SCSI

- Serial Attached SCSI (SAS)
  - Improvement of the parallel SCSI adapter
  - Transmission speed 3, 6 or 12 Gbit/s
  - Full duplex, dual port drives, higher reliability
  - More drives can be addressed from one controller port

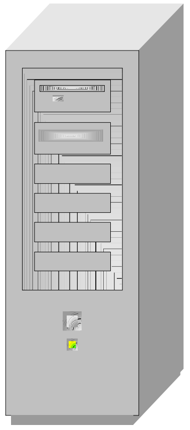
Direct-Attached Storage (DAS)

Advantages

- Better than to store data on the client
- Limited redundancy
- Low cost, simple

Disadvantages

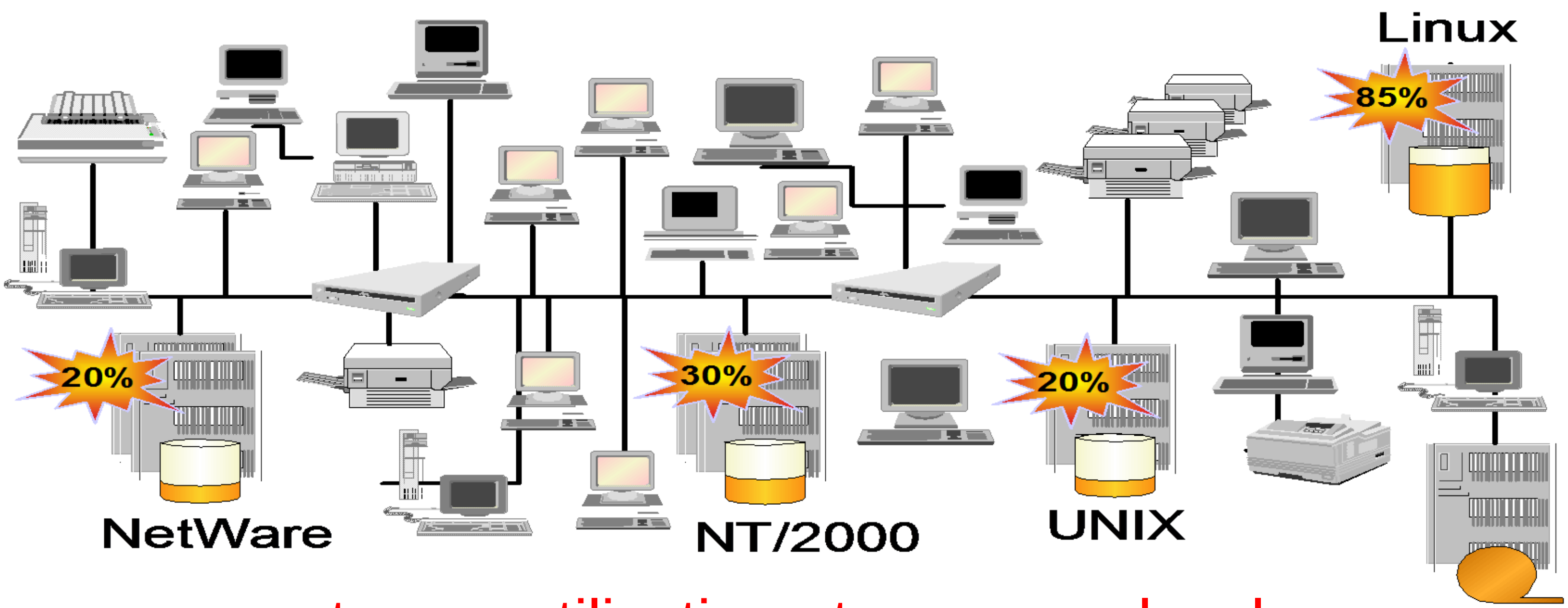
- Difficult management
- Low utilisation
- High cost of back-up
- Difficult data sharing
- Non well scalable
- Limited number of devices
- Low throughput



DAS Device

Dedicated Storage Devices

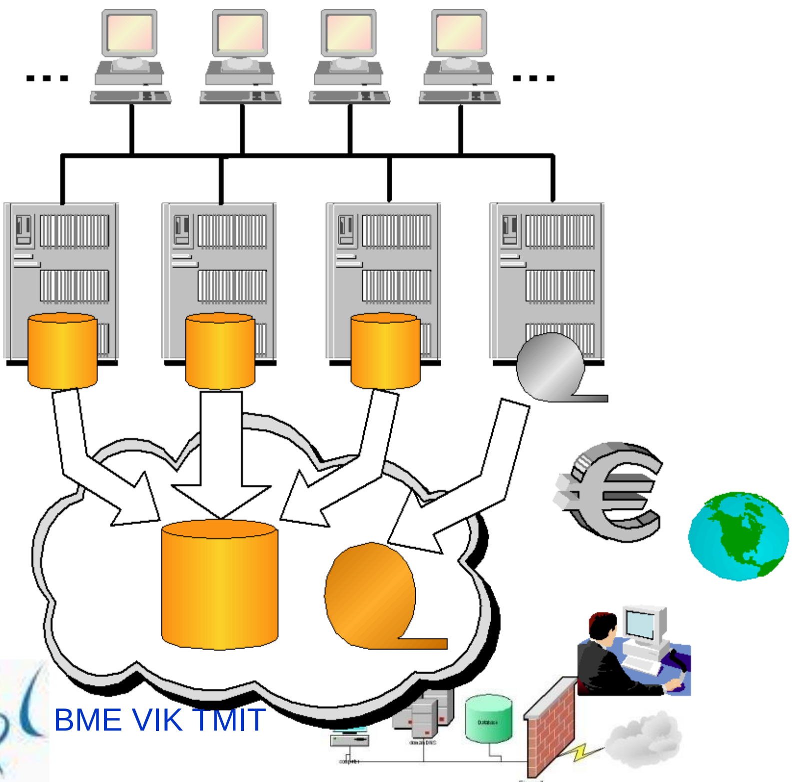
- Separated servers and storage – information islands, under separated management
- Unefficient source utilisation, high costs



The average storage utilisation at company-level is typically only 40-60%

Consolidated Storage Devices

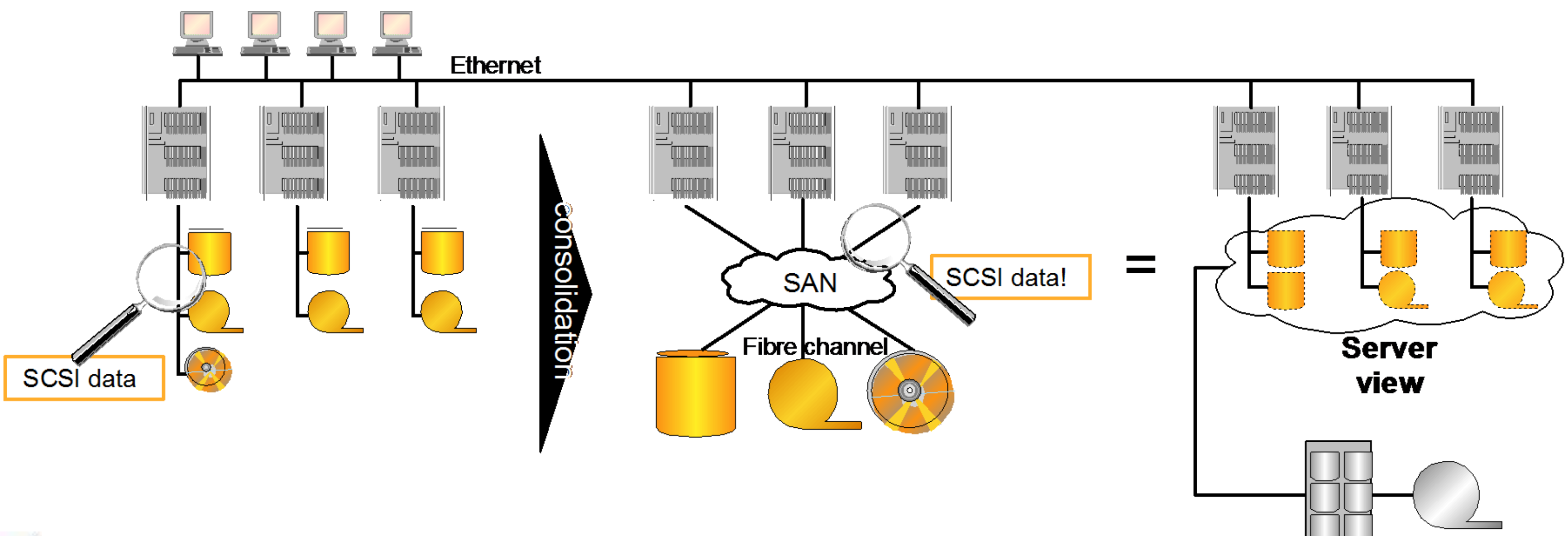
- Consolidated devices, management, data
- low complexity, lower specific cost
- High availability, scalable, disaster tolerate systems can be built



The consolidated storage unit handling can be realized in network architecture

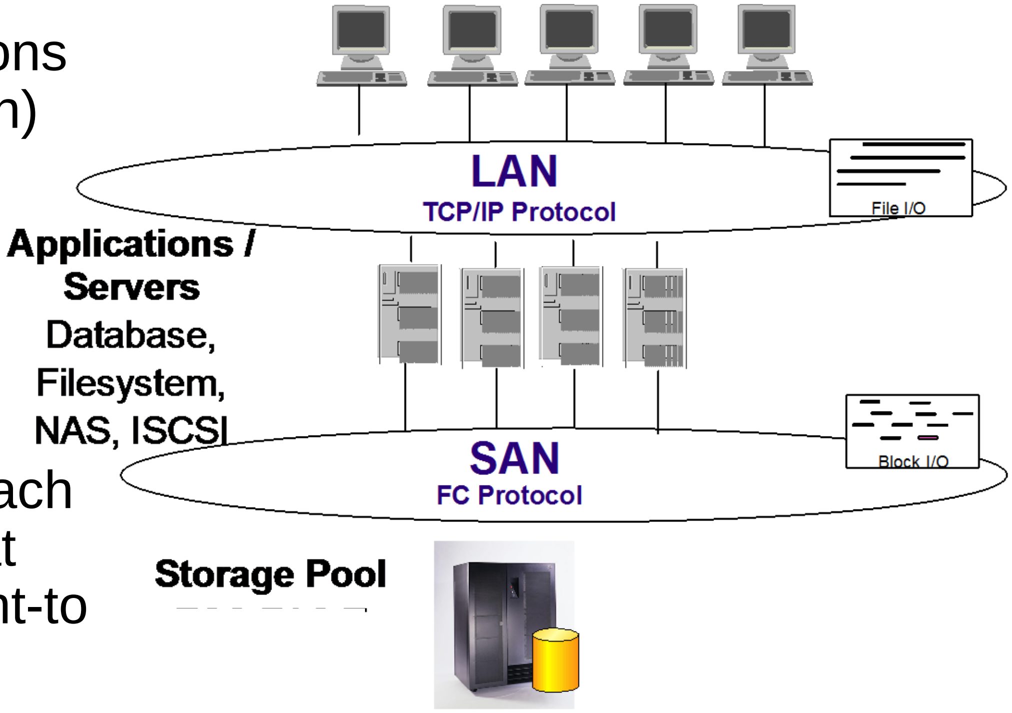
SAN - Storage Area Network

- Network dedicated to data transmission
- The storage devices are physically independent from the servers, more server can reach the same device
- Data handling protocol not changes, servers see as dedicated own storage unit



Advantages of SAN Networks

- **Resources**
  - Increase resource efficiency,
  - Scalability
  - Higher level system functions can be installed (replication)
- **Management**
  - higher efficiency
  - higher service levels
- **Information access**
  - Application servers can reach any data on the network, at any time – typically in point-to point connection

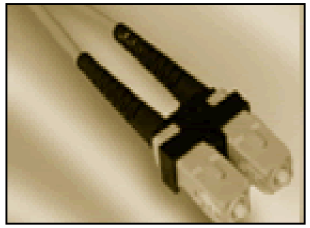


Transmission Technology: Fibre Channel Protocol

- Scalable
  - Large number of devices
  - Long distance
  - Transmission solution for numerous protocol
    - SCSI-3 (remote disks can be accessed in the same way as locals)
    - IP, ATM, …
- Point-to-point or Switched network topology
- Different devices, speed
  - Different types of copper wire, fibre optic
  - Max. speed: 128 Gb/s

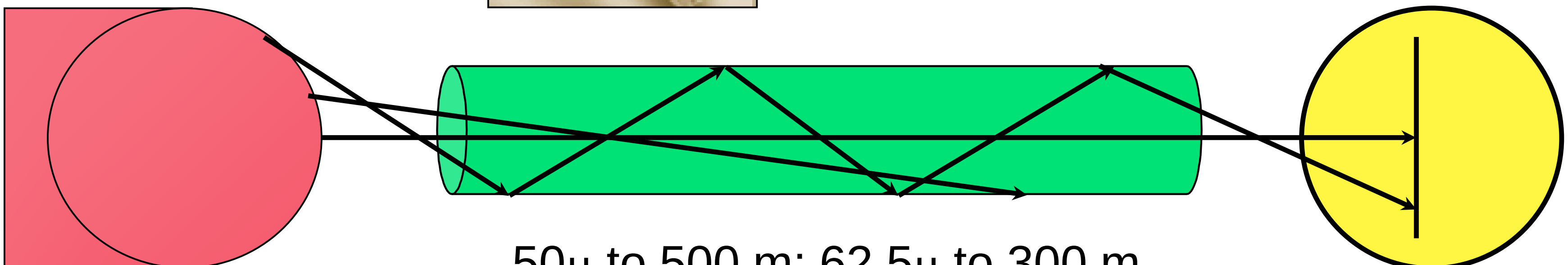
Fiber Optic

- Multimode



LED

Optical receiver

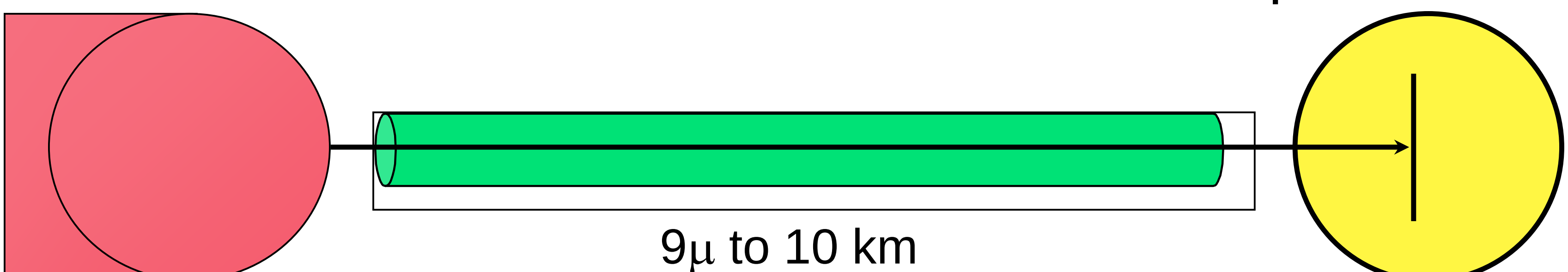


50μ to 500 m; 62,5μ to 300 m

- Single mode

LED

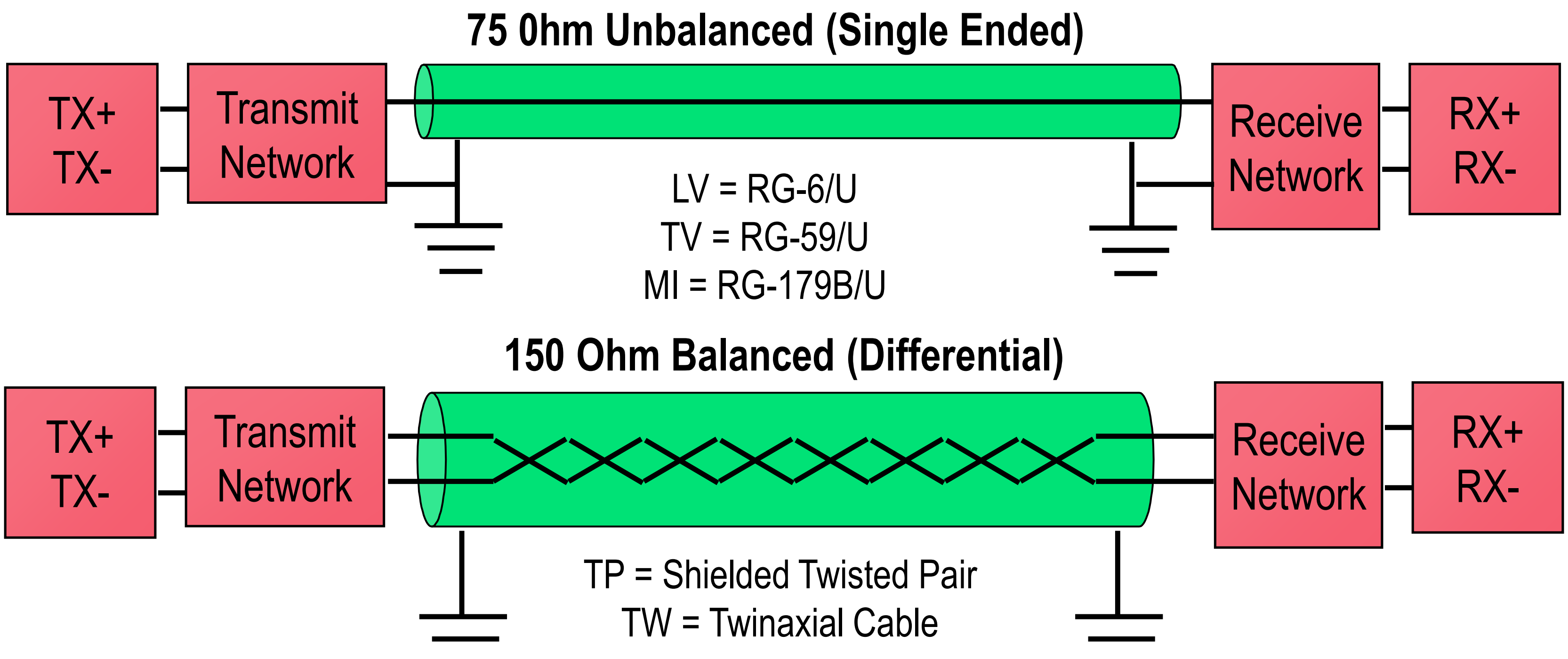
Optical receiver

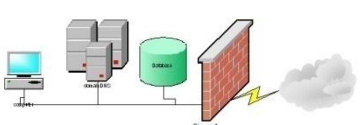


9μ to 10 km

Copper Wire

- Mainly Back-end
- Max. 15 m
  - Better signal-to-noise ratio than that of the fiber



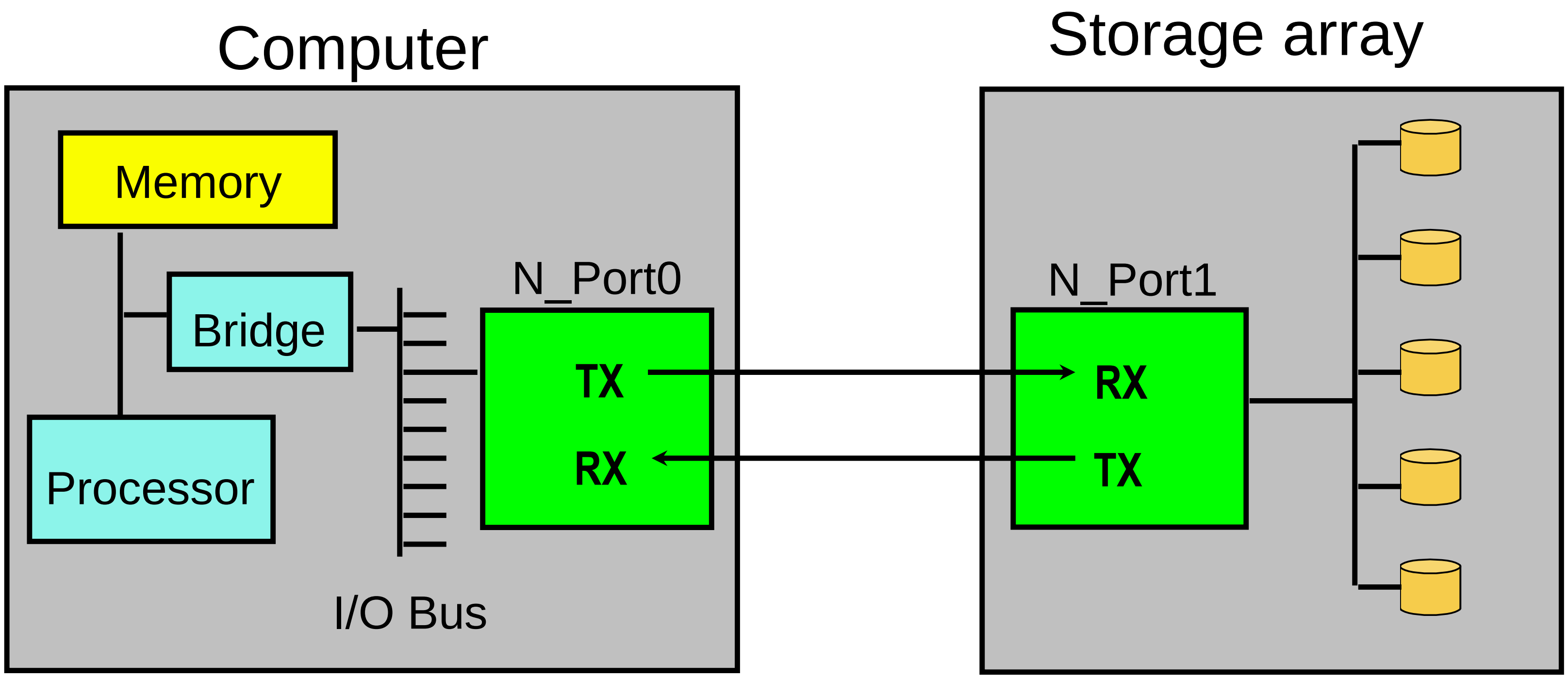


Fibre Channel Protocol

- 3 different topologies
  - Point-to-point
  - Arbitrated loop (not used)
  - Switched Fabric
- FC interconnects ports
  - Any entity communicating over FC, not a physical port
    - N ports (Node)
      - Disk
      - HBA (Host Bus Adapter) in computers
    - F ports
      - FC Switch

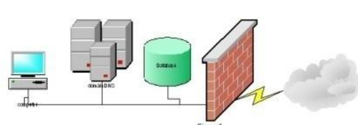
Point-to-point

- DAS
  - SCSI: max. 1,5 GB/s (12 Gb/s)
  - FC: max. 16 GB/s (128 Gb/s)

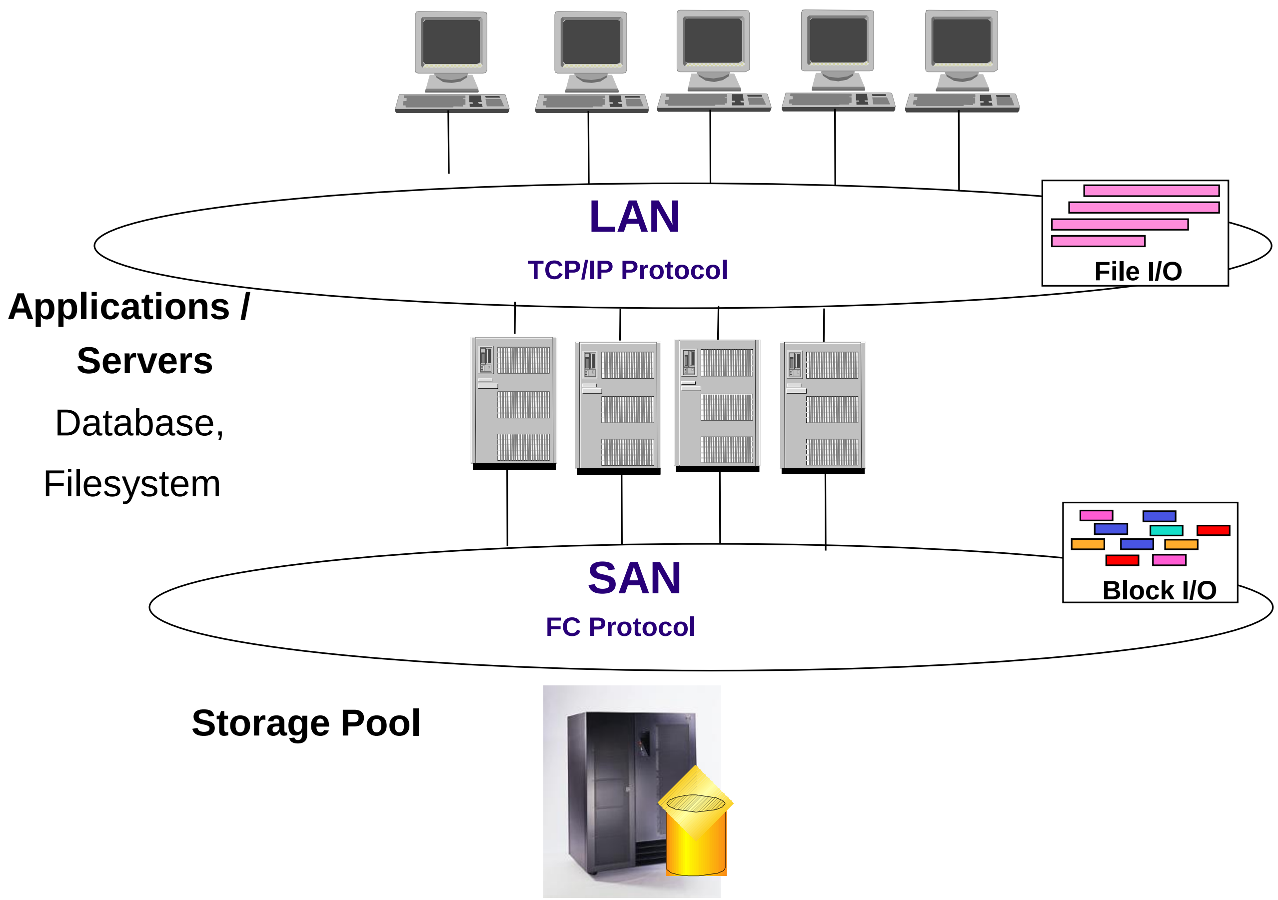


Summary of HDD Technologies

	SATA III		SAS-3		Fibre Channel
Teljesítmény	Half-duplex	≠	Full-duplex with Link Aggregation	=	Full Duplex
	6 Gb/s		12 Gb/s	≠	128 Gb/s
Interfész	1 m internal cable	≠	10 m internal and external cables		15 m copper cable 500m/10 km optic
	Multipliers 15 HDD max	≠	Expanders >128 devices	=	16 Million (Fabric)
Kialakítás	Single-port HDDs	≠	Dual-port HDDs	=	Dual-port HDDs
	Single-host	≠	Multi-initiator	=	Multi-initiator
Driver sw.	Software transparent with Parallel ATA	≠	Software transparent with Parallel SCSI	=	Software transparent with Parallel SCSI

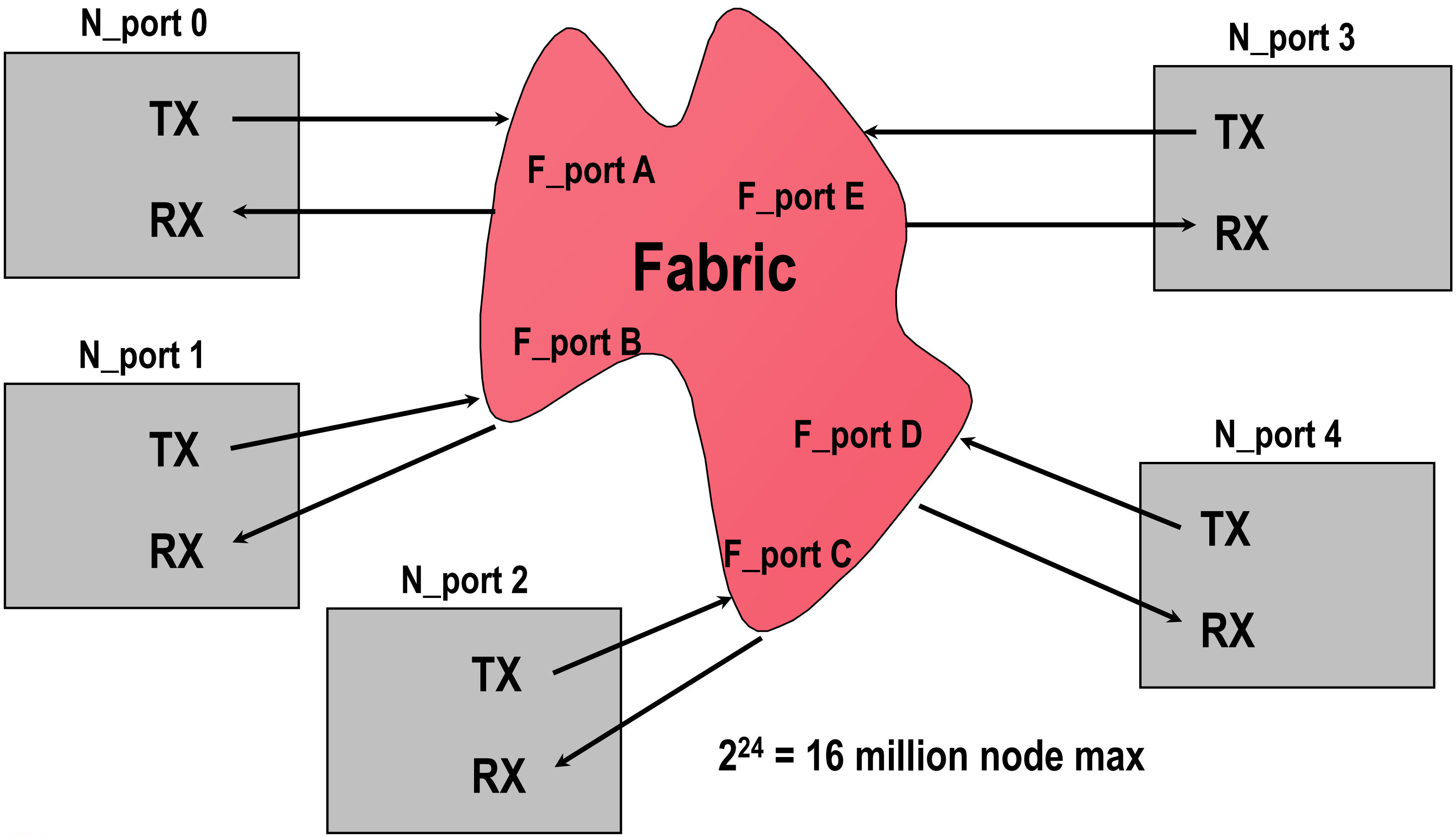


Refresh: SAN


LAN
TCP/IP Protocol
File I/O
Applications /
Servers
Database,
Filesystem
SAN
FC Protocol
Block I/O
Storage Pool

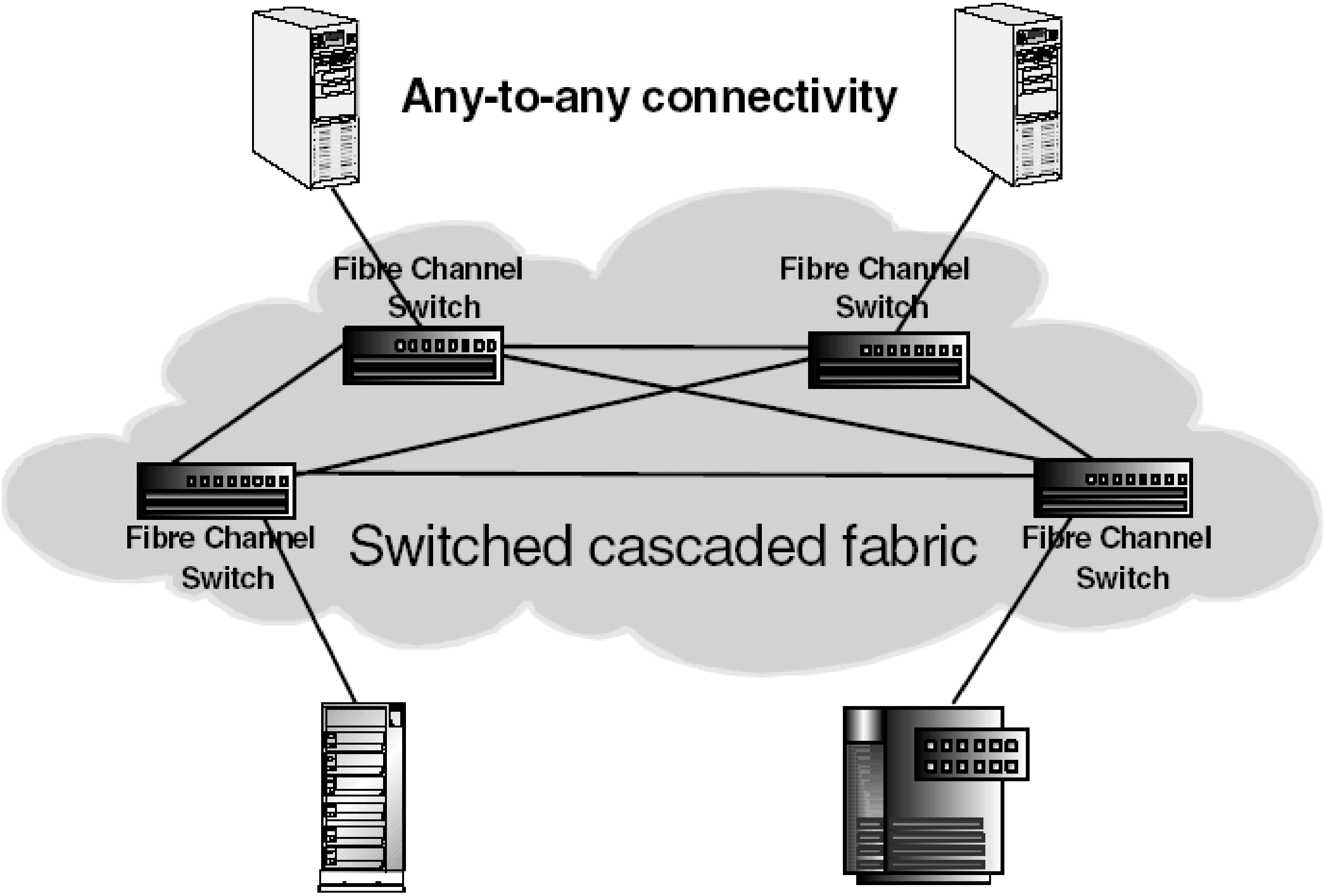
Switched SAN Network / Fabric

There are other connection types, but nowadays this is typical



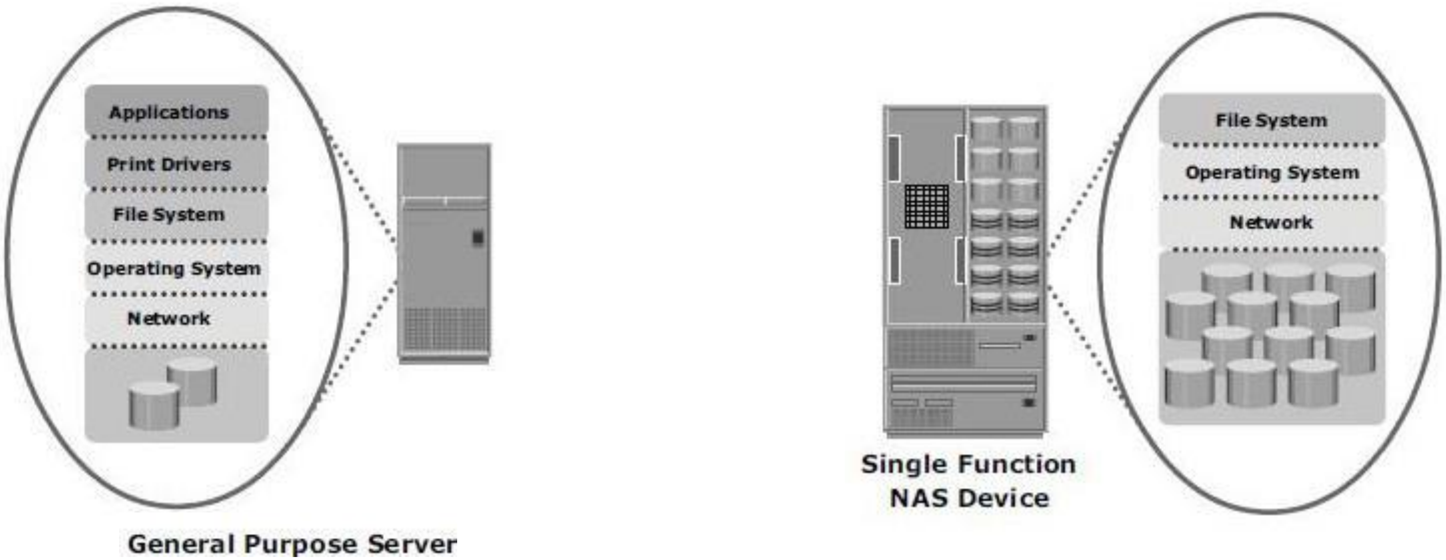
2^{24} = 16 million node max

Switched SAN Network



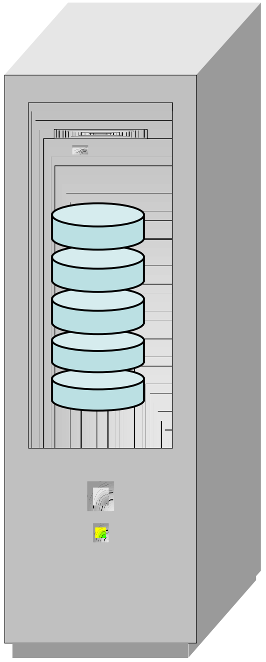
NAS – Network-Attached Storage

- Disk, which is connected to an IP network
  - Dedicated file server
    - with op. sys. optimised for I/O operations



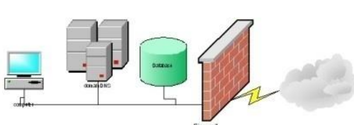
NAS (Network Attached Storage)

- IP (LAN ,WAN) connection, internal SCSI structure
- Internal RAID - error resistance
- File level access (Block access not supported!!)
- Easy installation
- Scalability inside the device
- Performance limitations (LAN bandwith, protocol overhead)



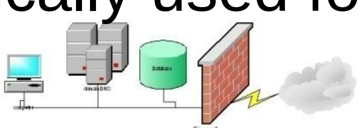
NAS Protocols

- NFS (Network File System) – UNIX
  - A protocol above UDP, specialised for file operations
- CIFS (Common Internet File System)
  - Op. sys. independent
    - Server
    - Client
  - Above TCP/IP
- FTP (File Transfer Protocol)



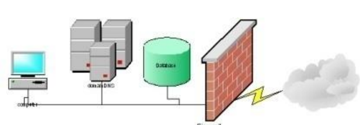
IP SAN (iFCP, iSCSI)

- SAN: block level access, FC network
- NAS: flie level access, IP network
- IP SAN: block level access, IP network
  - iFCP (Internet Fibre Channel Protocol)
    - Congestion ctrl, error detection&recovery by TCP
  - iSCSI (Internet Small Computer System Interface)
    - SCSI commands over existing LAN/VLAN/WAN
    - Emulates the SCSI bus over IP networks
    - Though any SCSI device can be connected, typically used for server <-> data storage



IP SAN

- For storage consolidation without the need of a dedicated network
  - Low-cost equivalent of Fibre Channel, but performance highly depends on ‘other’ traffic
- For disaster recovery
  - To mirror storages between (remote) data centers
    - As ‘hot stand-by’
    - Over WAN



SAN or NAS summary 1.

SAN (Storage Area Network)

Centralized, high performance network, dedicated exclusively for data storage

- Interconnects servers and data storage devices,
- Contains network and switching elementes and supporting software solutions

Advantages:

- Scalable, extendable
- High data transmission speed (4 -10 Gb/s)
- Maintenance: can be centralized; supports hierarchial storage – BUT in case of complex, heterogenous network complicated management

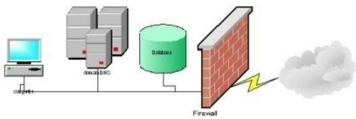
SAN or NAS summary 2.

NAS (Network Attached Storage)

Data storage device connected to the network; supports data sharing among server and clients

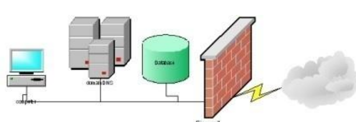
Advantages:

- Scalable, extendable, but limited bandwidth (LAN)
- Easy to install and maintain
- Typically in small/medium environments



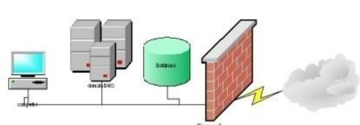
Data back-up in a disk system

- Copy (very) large amount of data
- Block level copies
- High speed, firmware-supported back-up processes
- *Volume copies* – real copies of disks



Volume Copy - Clone

- Back-up technology with firmware tools in a storage device
- Real duplicated file
  - back-up, analitic, datamining etc.
- Possibility of migration to other disk types
- Data-cosistency have to be ensured – stop application during cloning
  - Long time
  - Alternative: split mirror



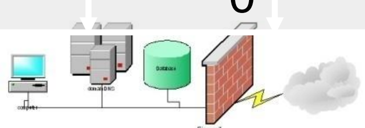
FlashCopy (Snapshot)

- If we modify a block – no overwrite – store at a different place
- FlashCopy table to register the blocks of the files
  - „Snapshot"
  - Rollback to any time
- Not suitable for back-up since no real copy generated!
  - For version handling systems

- COW: Copy On Write
  - The application must be stopped only for a while

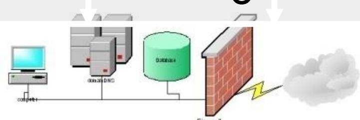
FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
	B0		B8	B0		B8
	B1	B1		B1	B1	
	B2		B9	B2		B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8			8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		



FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0		B8
	B1	B1		B1	B1	
	B2		B9	B2		B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8			8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		

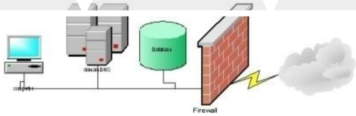


FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
	B2		B9	B2		B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8			8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		

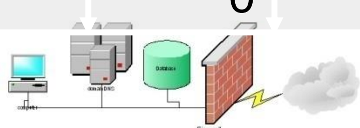
FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
Write t2	B2	**B2>B9**	B9	B2		B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8			8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		



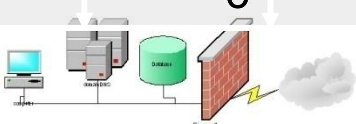
FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8			8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		



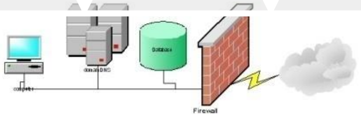
FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10		8	8	8
Delta (flashcopy inctement)	0			Virtual Volume A		



FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10		8	8	8
Delta (flashcopy inctement)	0	2		Virtual Volume A		

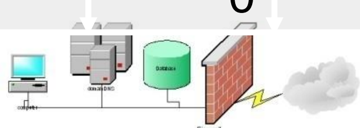


FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
	B1	B1		B1	B1	
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10		8	8	8
Delta (flashcopy inctement)	0	2		Virtual Volume A	B	

FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
Write t3	B1	B1	**B1>B10**	B1	B1	
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10		8	8	8
Delta (flashcopy inctement)	0	2		Virtual Volume A	B	



FlashCopy - Blocks

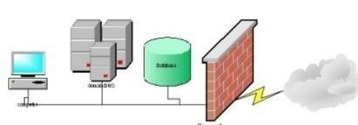
Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
Write t3	B1	B1	**B1>B10**	B1	B1	**B10**
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10		8	8	8
Delta (flashcopy inctement)	0	2		Virtual Volume A	B	

FlashCopy - Blocks

Block table				Flashcopy table		
Time	**T1**	**T2**	**T3**	**F1**	**F2**	**F3**
Write t2	B0	**B0>B8**	B8	B0	**B8**	B8
Write t3	B1	B1	**B1>B10**	B1	B1	**B10**
Write t2	B2	**B2>B9**	B9	B2	**B9**	B9
	B3	B3	B3	B3	B3	B3
	B4	B4	B4	B4	B4	B4
	B5	B5	B5	B5	B5	B5
	B6	B6	B6	B6	B6	B6
	B7	B7	B7	B7	B7	B7
Total # of blocks	8	10	11	8	8	8
Delta (flashcopy inctement)	0	2	3	Virtual Volume A	B	C

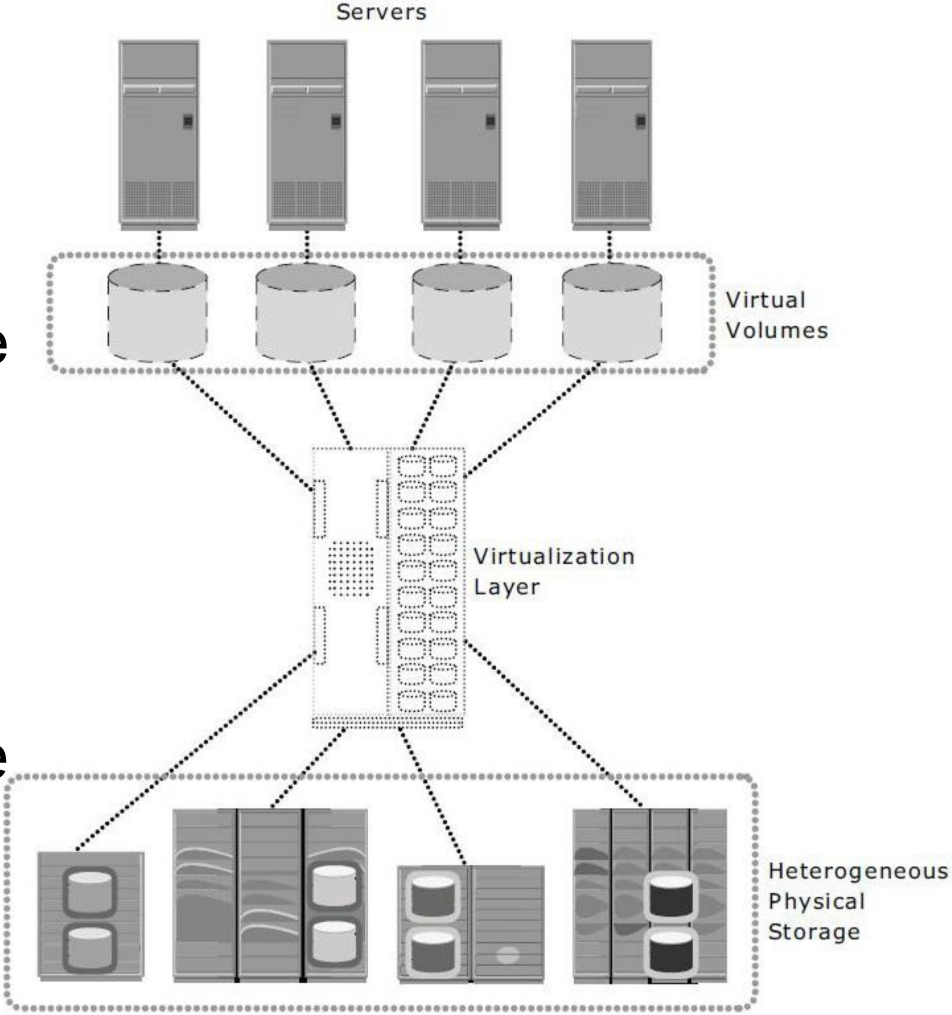
Virtualisation

- **Storage virtualisation** refers to the process of abstracting logical storage from physical storage. The term is today used to describe this abstraction at any layer in the storage software and hardware stack
- The virtualization can be realized in a different system levels



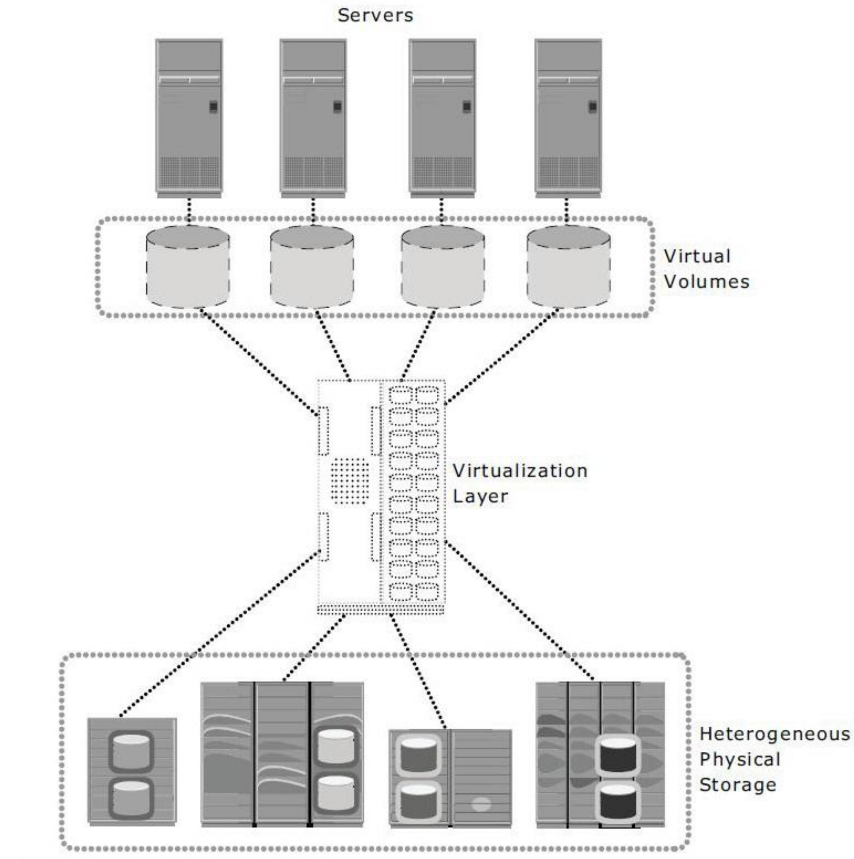
Storage Virtualisation

- Virtualisation Appliance or Virtualisation Engine hides the differencies of the disks
  - „translates” between the two formats
  - disks can be shared
    - better utilisation
  - replacement or modification of disks are not seen for the server



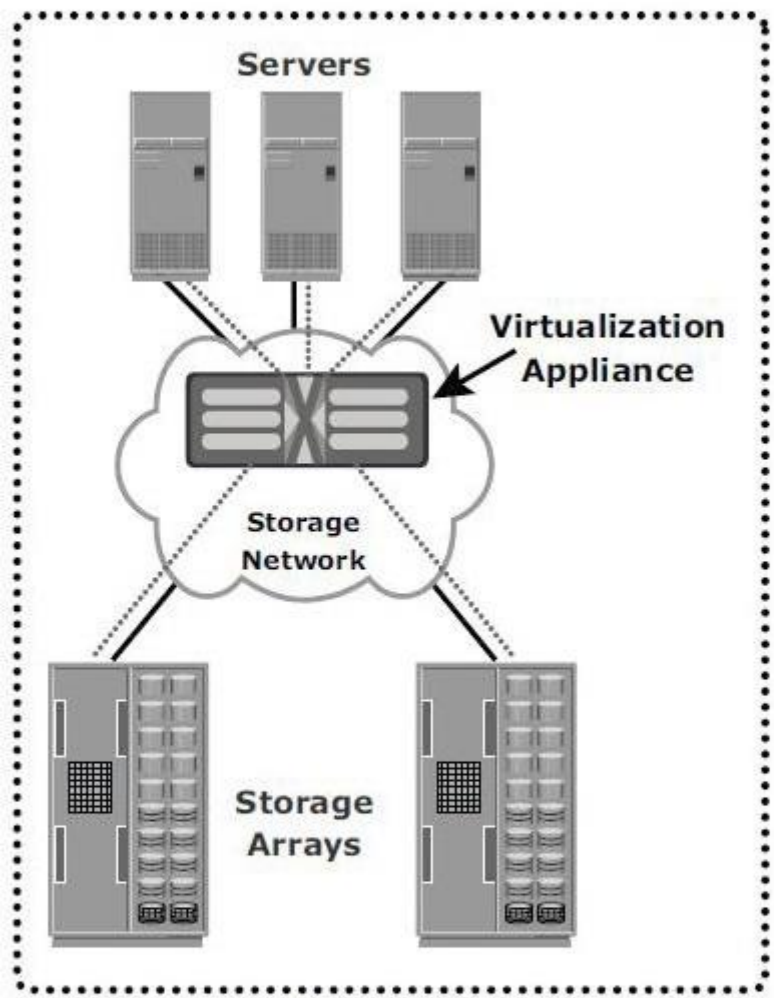
Behaviour of Virtualisation Appliance

- Meta-data
  - matching table between physical – logical addresses
- Server: LUN=1, LBA=32
  - LBA Logical Block Address
- VM: from table, this corresponds to the physical LUN=4, LBA=0
- Requests the data from physical disk
- Data is transmitted to server as if it came from LUN=1, LBA=32



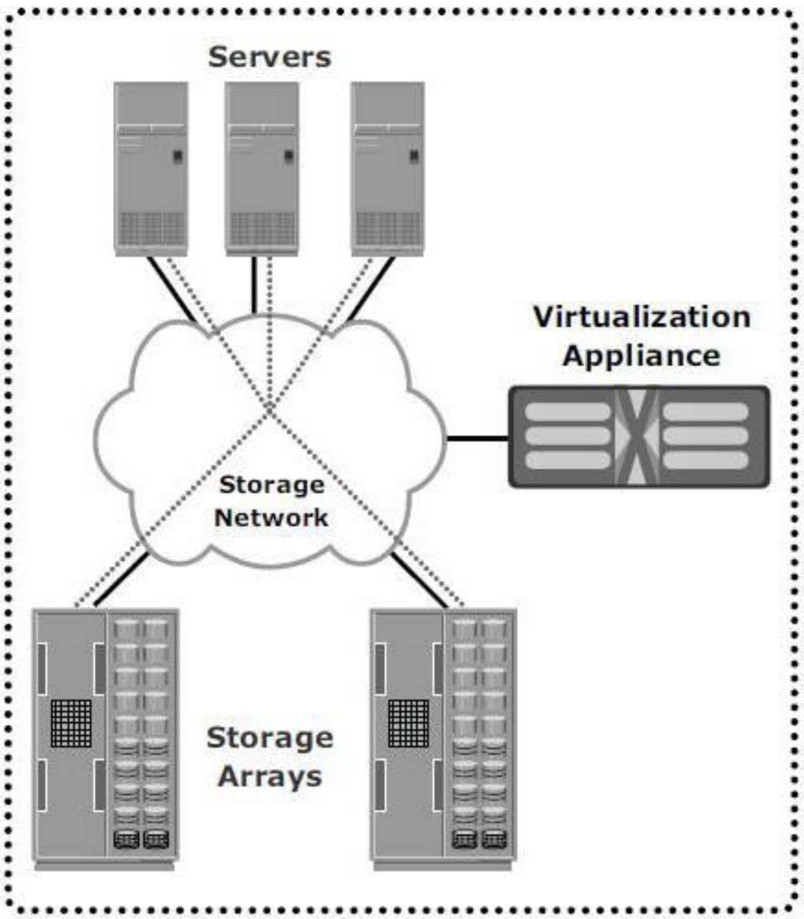
Virtualisation Configuration – in-band

- Virtualisation Appliance in data path
  - no need for a special software on server
  - but slower since data goes through the Virtualisation Appliance



Virtualisation Configuration – out-of-band

- Separate control and data paths
  - Special software on server:
    - first asks the physical location of the data from Virtualisation Appliance
    - then reaches the data directly
  - Faster data transfer since no additional layer in data path

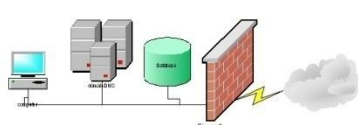


Levels of Virtualisation

- Block level virtualisation
  - discussed till this point
  - *server* wants to have an access to a *data block*
  - knows its (logical) address
- File level virtualisation
  - a *client*/host wants to have an access to a *file* on a file server
  - must know on which

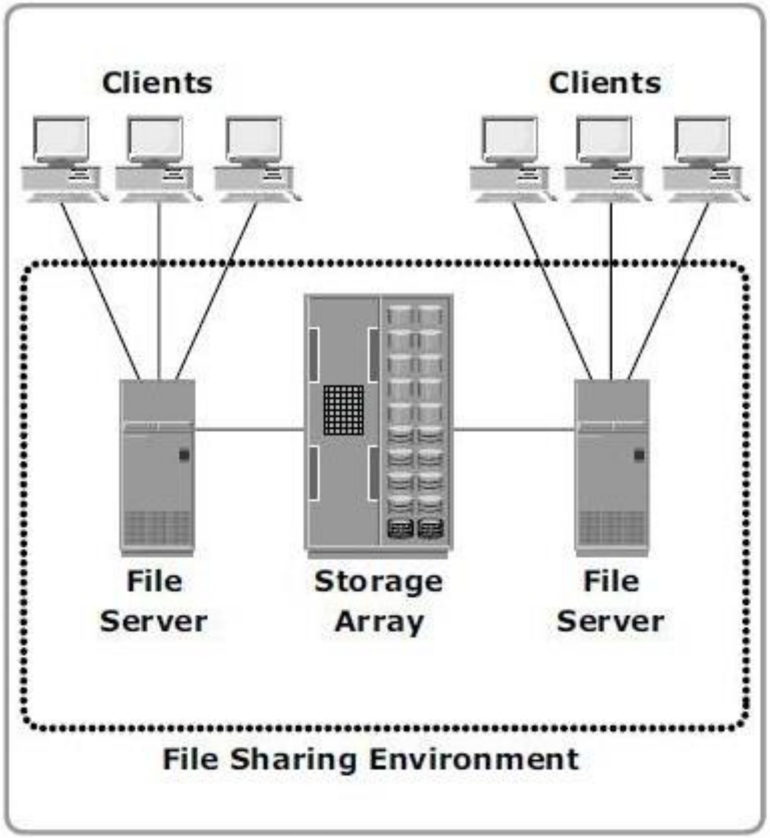
Block Level Virtualisation

Real sources …	Virtual sources…
• LUN (different types, different vendors)	• Virtual LUNs, they seem as same type of same vendor
• Tipically fix sized	• Size can be modified dinamically
• Different vendor configurations & back-up services	• Centralised management and services
• Migration to new technologies is difficult	• Migration without disturbing the applications



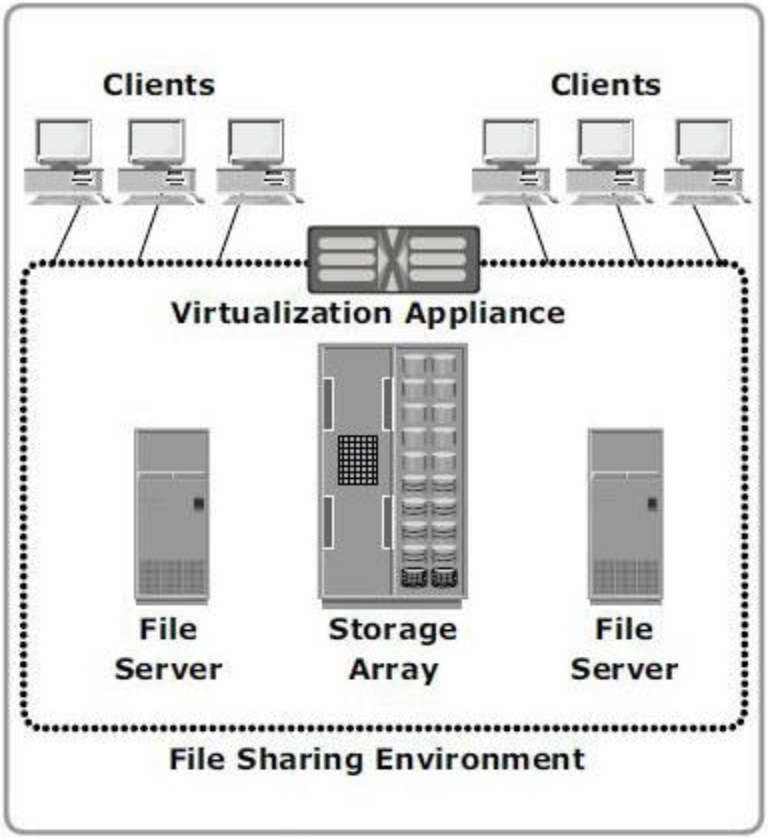
File Level Virtualisation

- Without File Level Virtualisation
  - if a client/host wants to have an access to a file on a file server
  - must know on which
  - one server may be empty while other full
  - file movement affect the client, too



File Level Virtualisation

- Virtualised file server
  - client should not know on which server the file is
  - simpler
    - load sharing
    - file movement
    - extension
- Cloud computing



Data upload to cloud

- Time…
- On media + delivery boy
- 100 petabyte
  - Film archive
  - NASA satellite pictures
  - 2000 years of mp3
  - 200x the genom of all humans
  - Via Internet up to years