

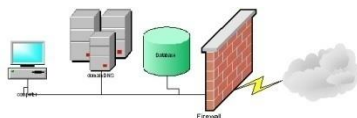
Információs rendszerek üzemeltetése

V. fejezet Adattárolás, tároló rendszerek, tároló hálózatok

BME VIK TMIT
Mérnök-informatikus alapképzés



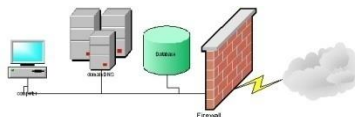
BME VIK TMIT



Információs rendszerek üzemeltetése

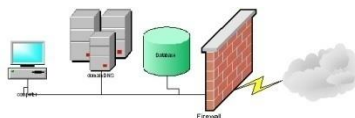
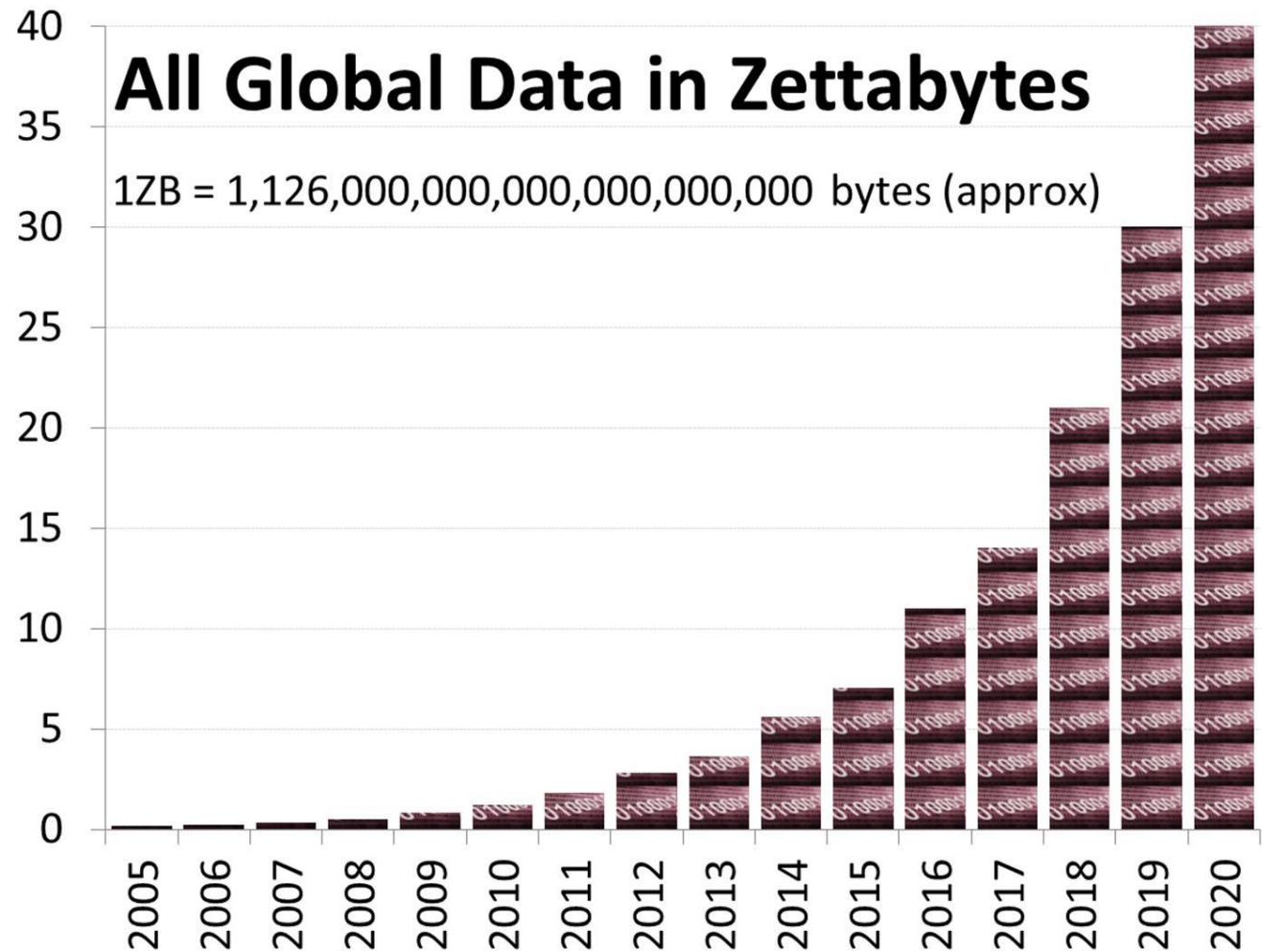
Adattárolók, adattárolás

- Adatok típusai
- Adattároló eszközök
- Diszkek megbízhatósága (RAID)
- Tárolórendszerek (DAS, SAN, NAS, IP SAN) és átviteli technológiák
- Adatmásolási eljárások (Volume/Flash copy)
- Virtualizáció
- Adattároló hálózatok menedzsmentje



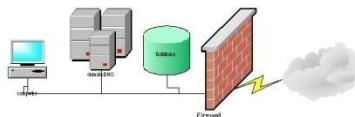
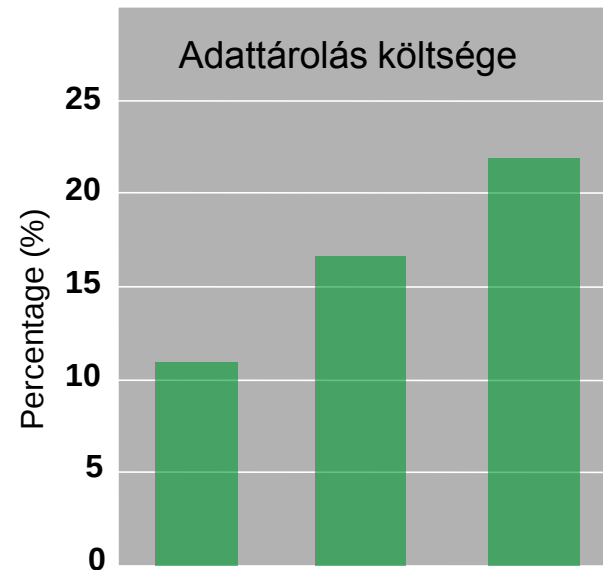
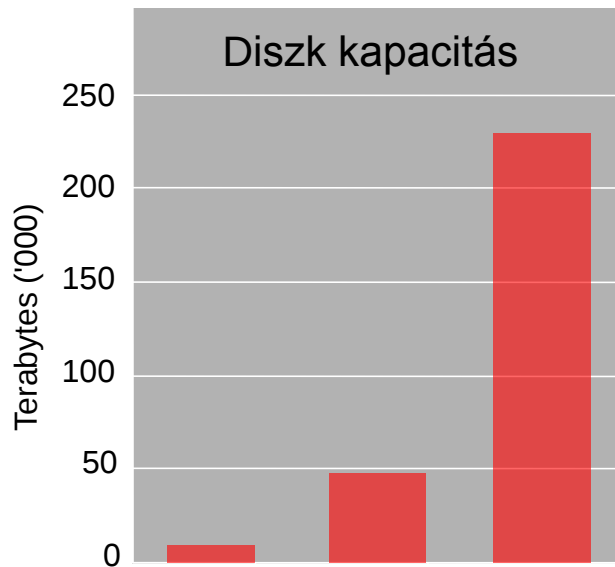
Előállított adat

- Zetta = 10^{21}
- Az elmúlt két évben kb. 10x annyi adat született, mint addig összesen

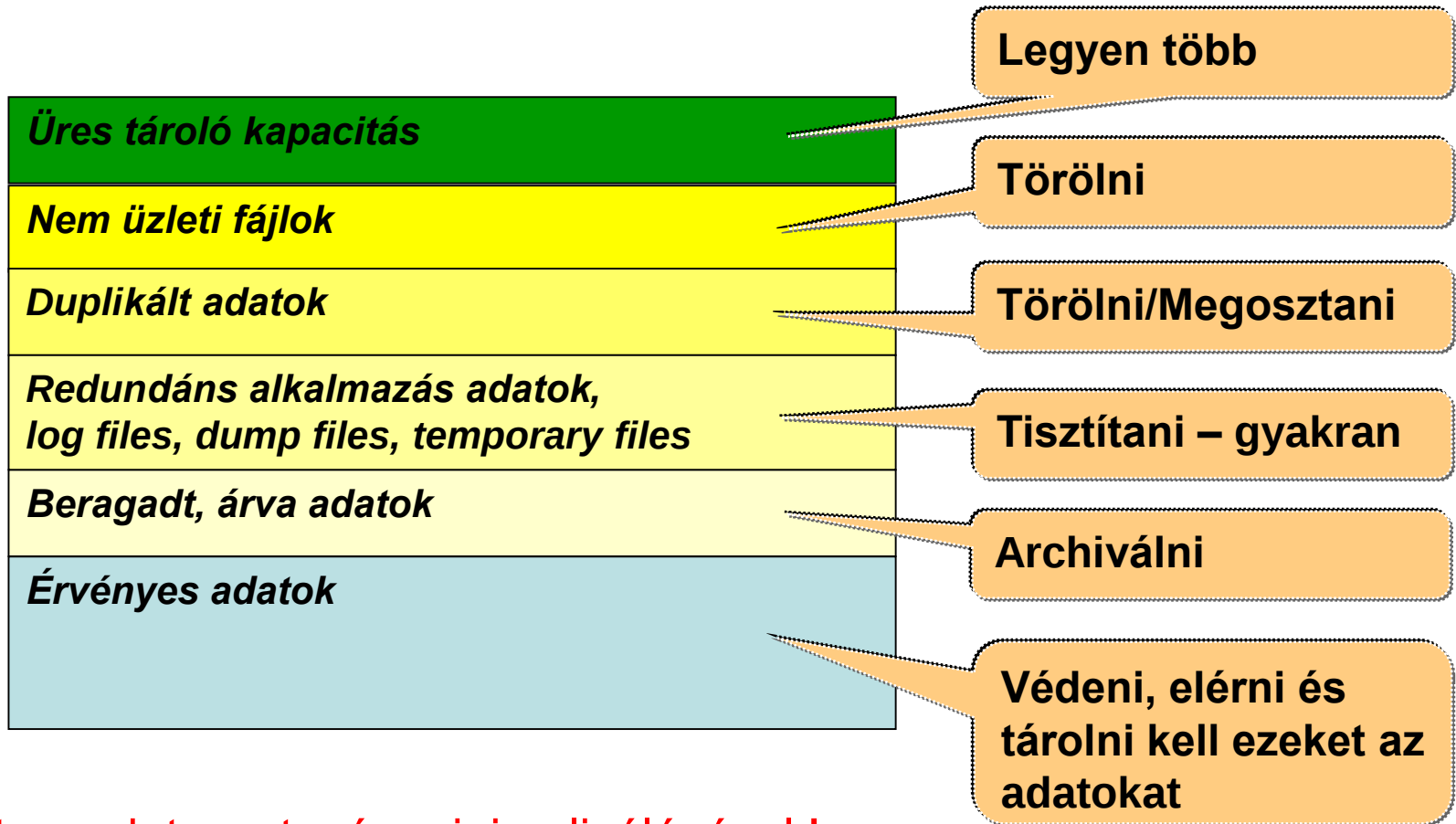


Tárkapacitás, költség

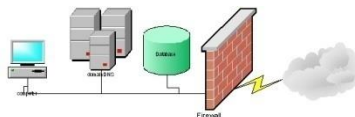
Több adat önmagában nem jelent több információt – de több költséget biztosan!



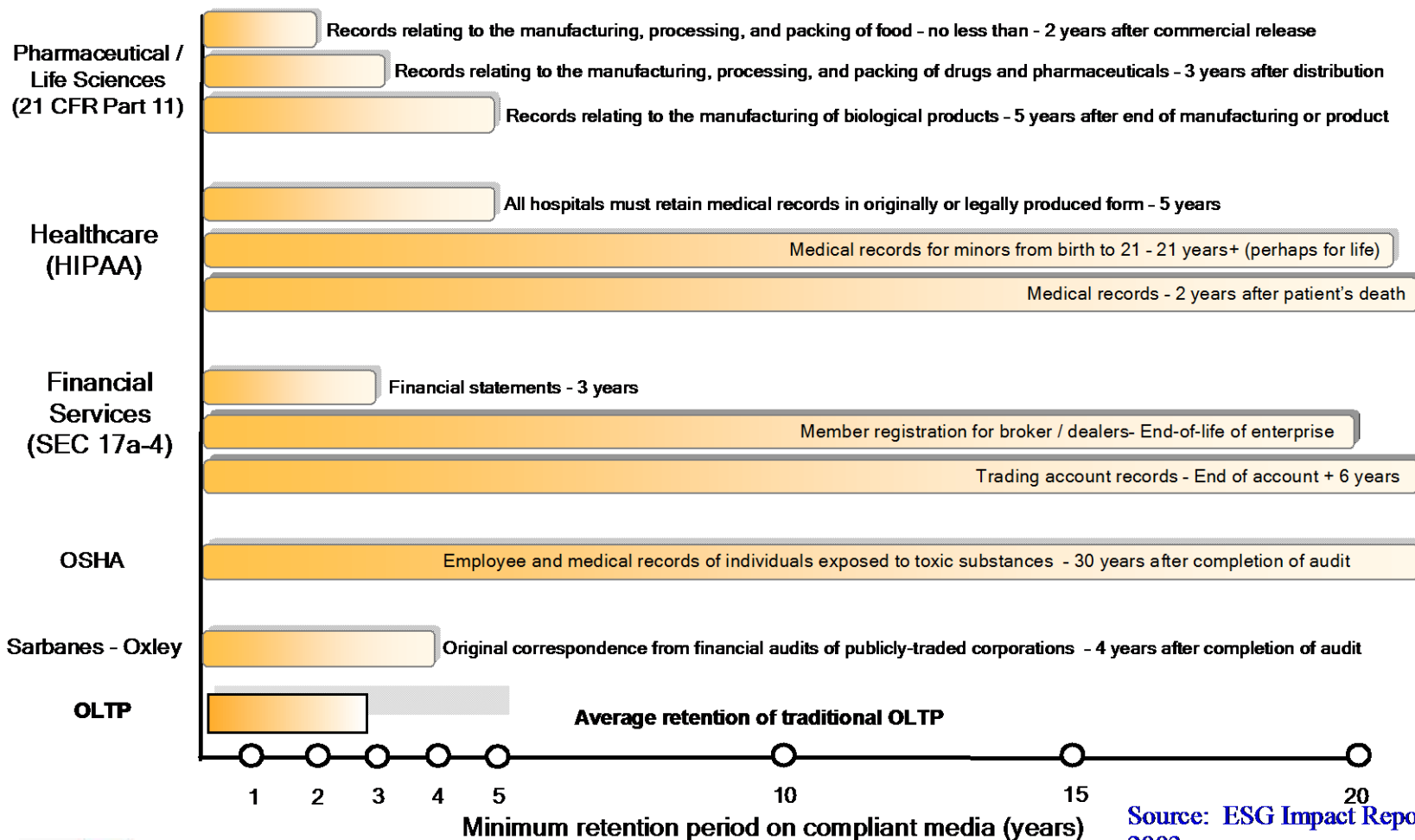
Adatok típusai



Mindezt az adatvesztés minimalizálásával !



Kötelező adatmegőrzési időkre vonatkozó törvényi előírások



Adatfolyamok

Strukturált adatfolyam

Forrása és célja ismert

Server-to-Server

Üzleti eszközök

Speciális alkalmazások

Erős biztonság

Privát hálózatok

Tervezhető

Adatbázis adatok

Monitorozható és ez alapján tervezhető

Strukturálatlan adatfolyam

Nem ismert

Client-to-client, Client-to-Server

Személyes eszközök

E-mail, Web

Gyenge biztonság

Publikus hálózatok

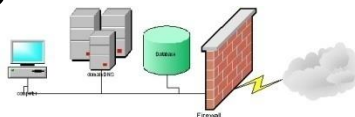
Nehezen tervezhető, statisztika

Fájlszerverek

Inkább házirendekkel szabályozható



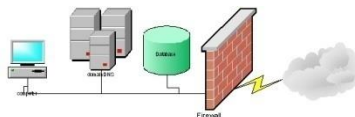
BME VIK TMIT



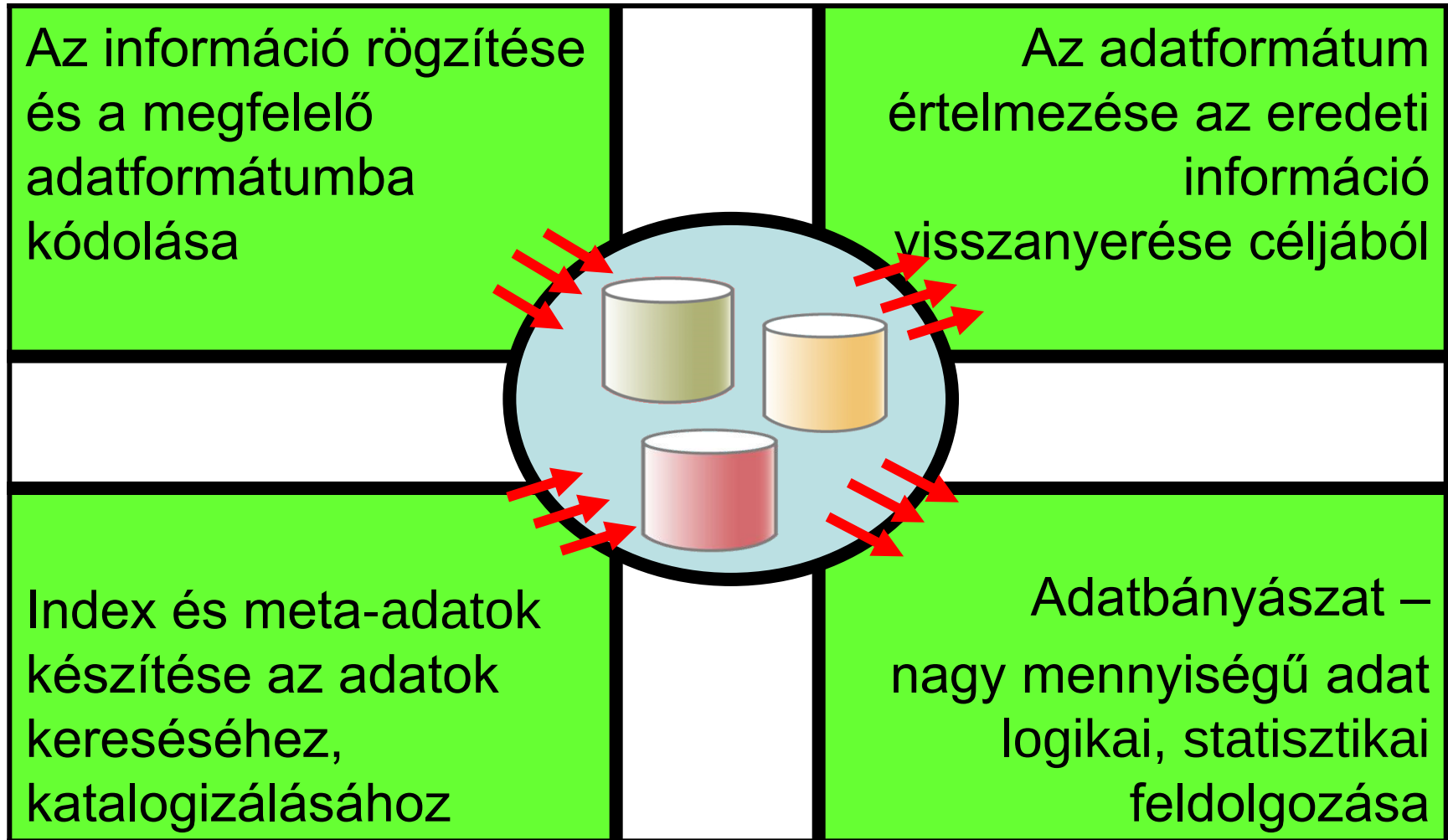
Információs rendszerek üzemeltetése

Adatgazdálkodás

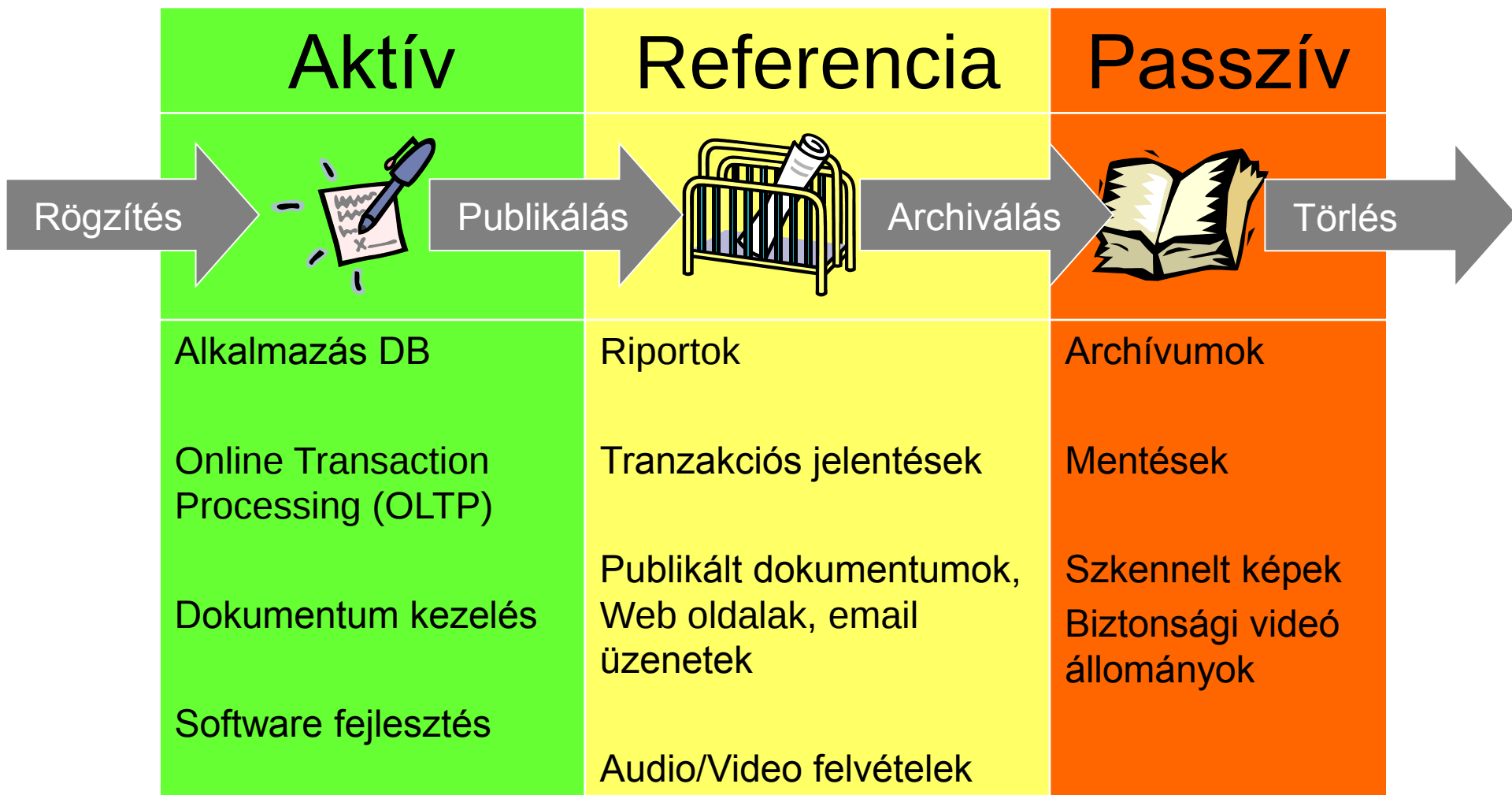
- IT infrastruktúra és menedzsmentjének egyszerűsítése, automatizálása
 - Cél: alacsonyabb élettartam-költség (TCO, *Total Cost of Ownership*), magasabb megtérülés (ROI, *Return on Investment*)
- Üzletfolytonosság, biztonság, adat-sérthetetlenség biztosítása
 - Cél: az üzlet folyamatosságának biztosítása
- Hatékony információ élettartam-menedzsment
 - Cél: az információt az értékének megfelelő tárterületen kell tárolni, az adat-megőrzési kötelezettségeket is figyelembe véve



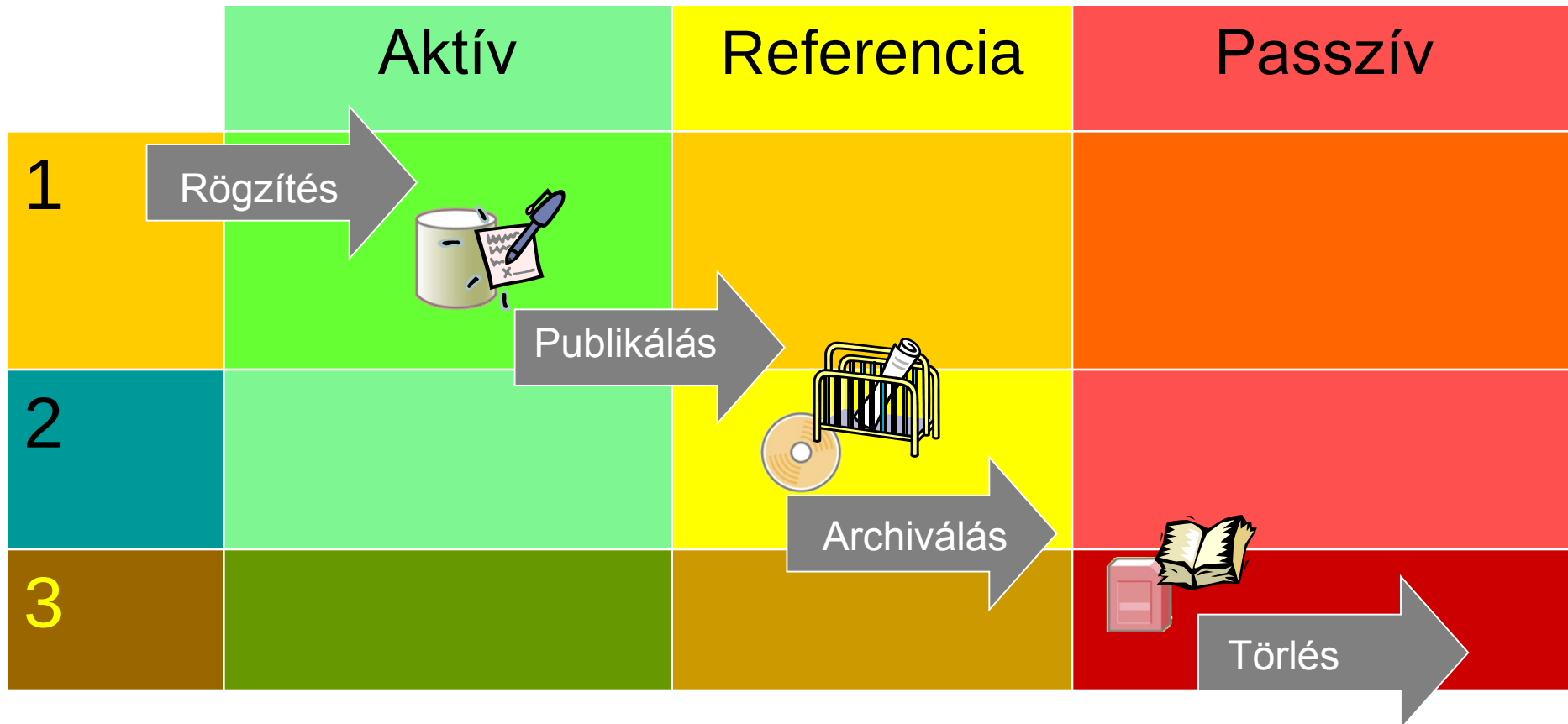
Információ és adat élelciklus, adattípusok



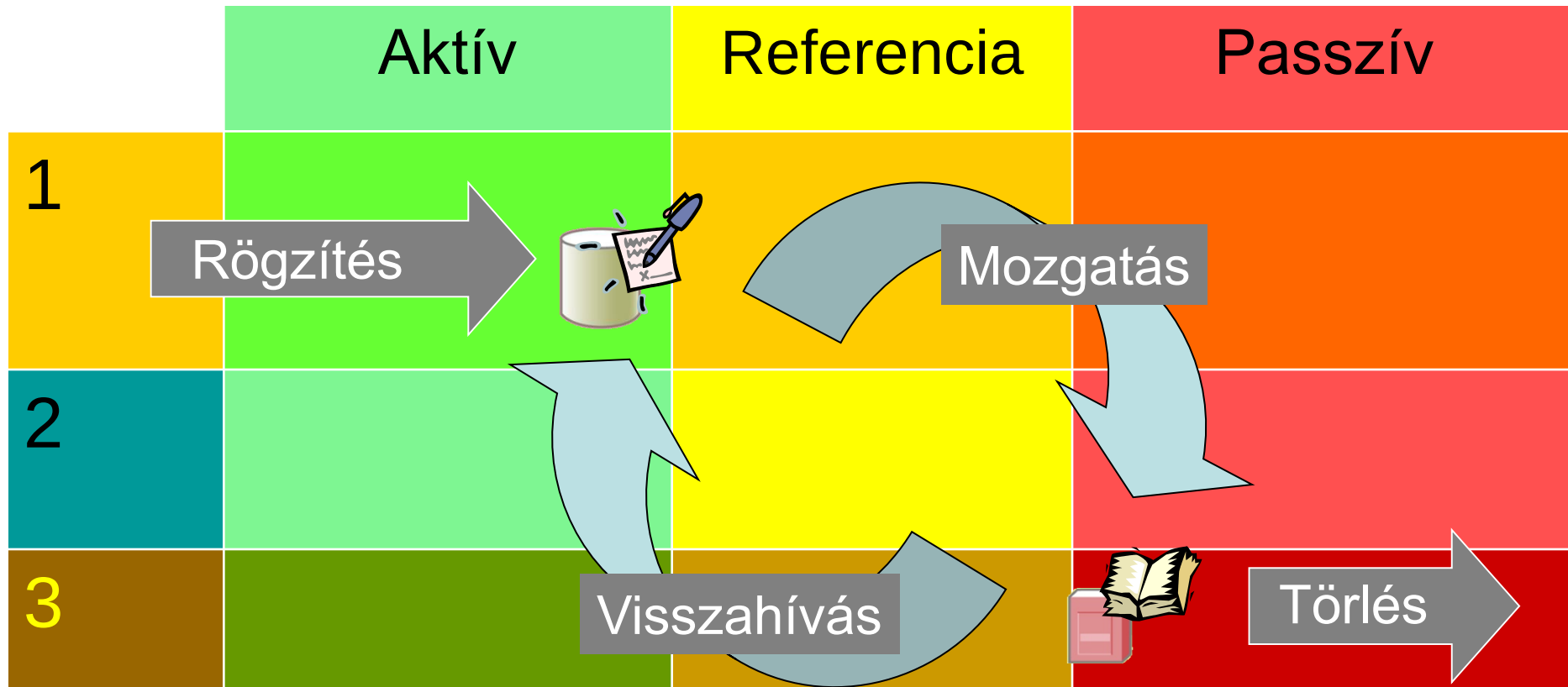
Az egyes adatok iránti igény változó



Optimális adattárolási technológia: választás az adat értékének megfelelően



Hierarchikus tároló menedzsment (HSM) (Hierarchical Storage Management)



HSM paraméterei: adat mérete, típusa és az utolsó olvasás időpontja

Adattároló eszközök

Előny

Probléma

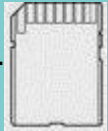
Diszk



- „Azonnali” adat elérés
- Közvetlen írás/olvasás (R/W)

- Diszk csere
- Tápellátás, hűtés
- Élettartam 3-4 év!
- Mozgó alkatrészek

Flash memó-
ria



- Nincsenek mozgó alkatrészek
- Gyors

- (Még) drága

Optikai



- Másodlagos tároló
 - automata könyvtárak
- WORM (*Write Once Read Many*)

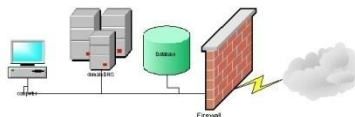
- Nem igazán tartott lépést a diszk és szalag fejlődéssel, SOHO eszköz (*Small Office Home Office*)

Szalagos



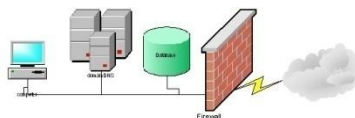
- 10-20x olcsóbb, mint a diszkes tároló
- 30 éves adatmegőrzési idő

- Nem azonnali elérés
- Soros írás/olvasás



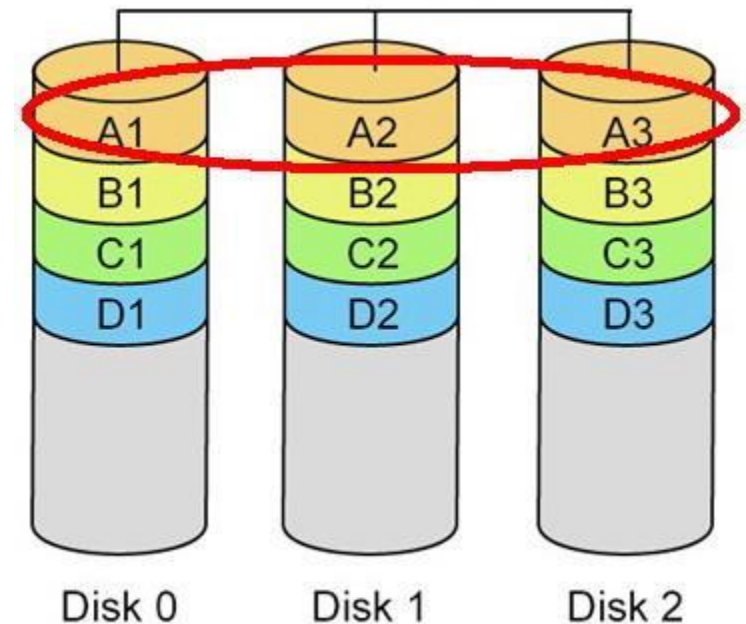
Diszkek megbízhatósága

- A szervereken tárolt adatok ne sérüljenek meg
- Eredetileg egyetlen nagy, nagymegbízhatóságú diszk
 - SLED: Single Large Expensive Drive
 - Drága...
 - MTBF (Mean Time Between Failure)
 - Kb. 750 000 óra (kb. 85 év)
 - De egy nagy(??) tárolótömb 1000 diszkkal
 - MTBF: $750\,000 \text{ óra} / 1000 = \text{kb. 1 hónap}$



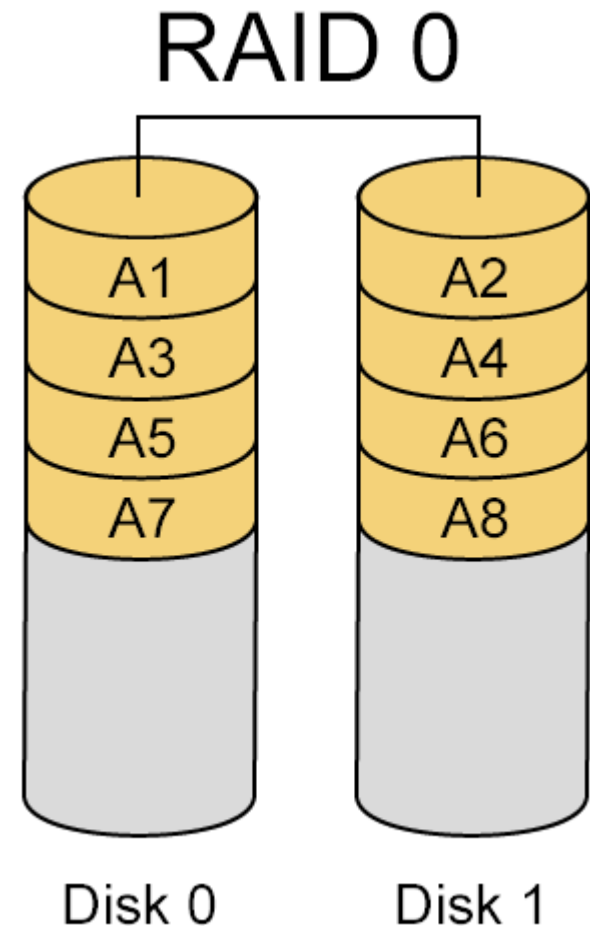
RAID Redundant Array of Independent (Inexpensive) Disks

- 1987 California, Berkeley Egyetem
- RAID 0, 1-5, 6
- Lemezek sávokra (stripes) osztása



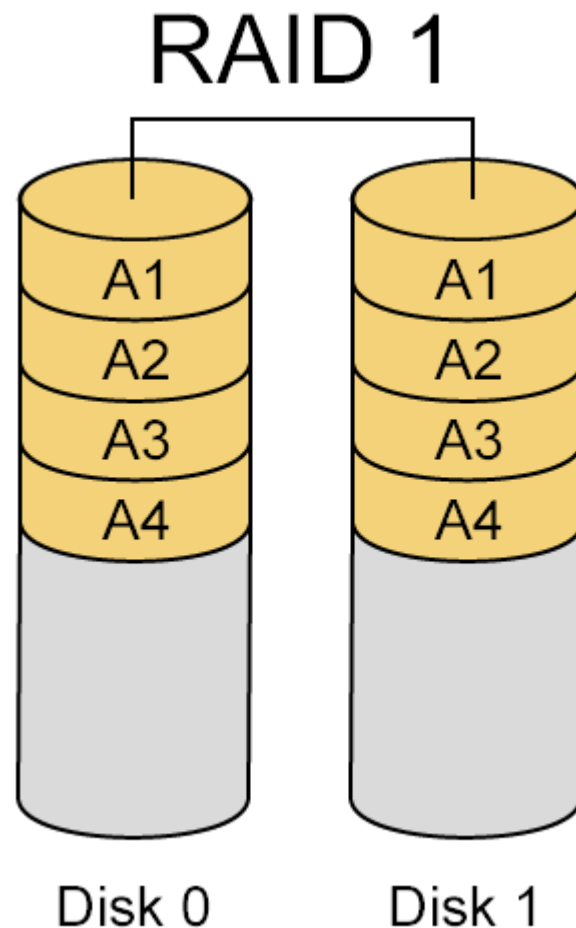
RAID 0 - striping

- Nem biztonság növelési cél
- Kapacitásnövelés
- Sebességnövelés
(párhuzamos írás/olvasás)



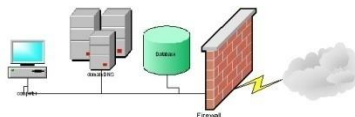
RAID 1 – diszk tükrözés (mirroring)

- Diszk duplikálás
- Nagy megbízhatóság
- Nagy (2x) méretnövekedés
- Lehet közel ekkora biztonság, kisebb veszteséggel?



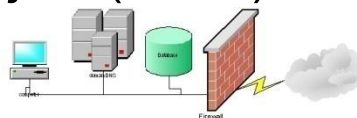
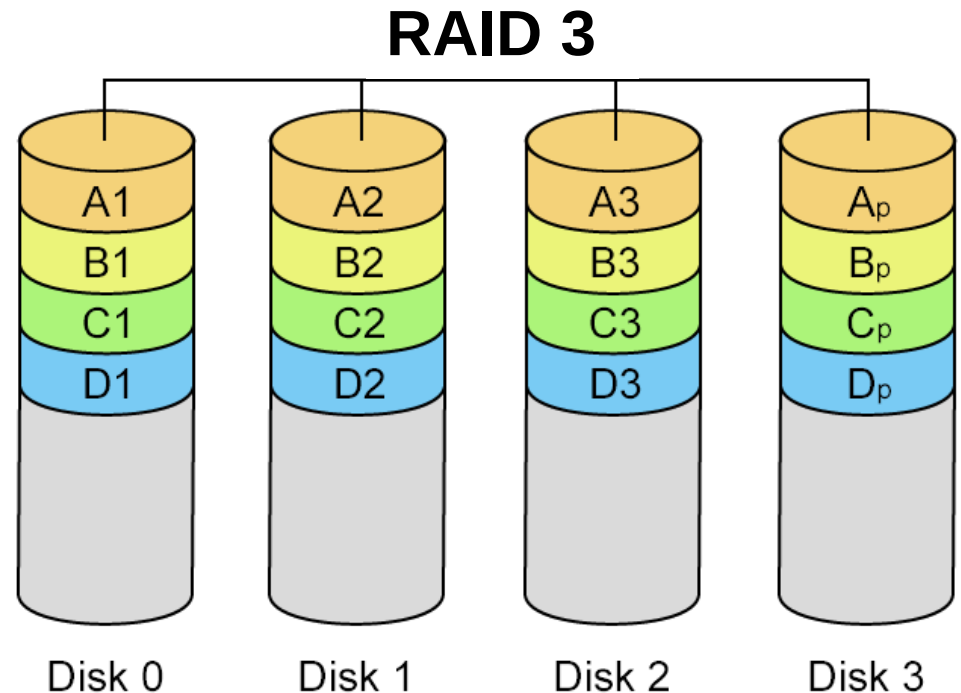
RAID 2 – hibajavító kód

- Sávokra bontás
- Egyes meghajtókon hibajavító kódokat (ECC – Error Correcting Code) tárolnak
 - Képes a diszkhiba detektálására, javítására
- Ma már nem használják, mert ma már a meghajtókon belül képeznek hibajavító kódokat



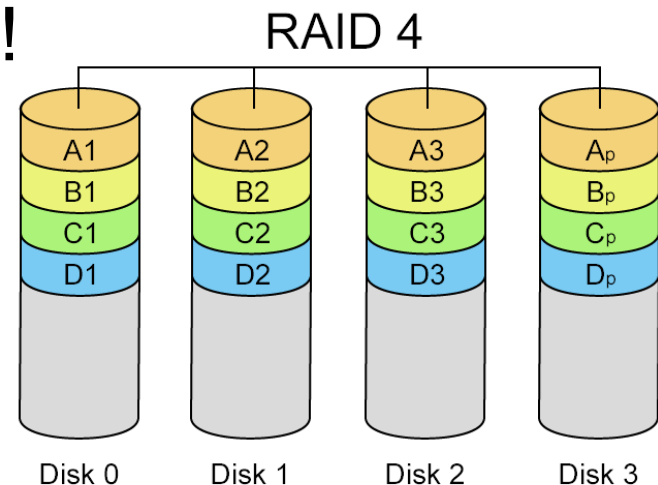
RAID 3 – paritásdiszk

- +1 diszk: paritás (XOR)
- Egy meghajtó kiesése: a többiből XOR művelettel helyreállítható – idő, lelassul
- Diszkhibát nem érzékel
- Tipikus: 2+1, 5+1, 8+1, 14+1
- Paritásdiszk korlátoz
- Kis sávok, mindig egész stripe művelet (Adatdiszkek egyszerre érhetőek el)
 - Single user
 - Nagy fájlok (video)



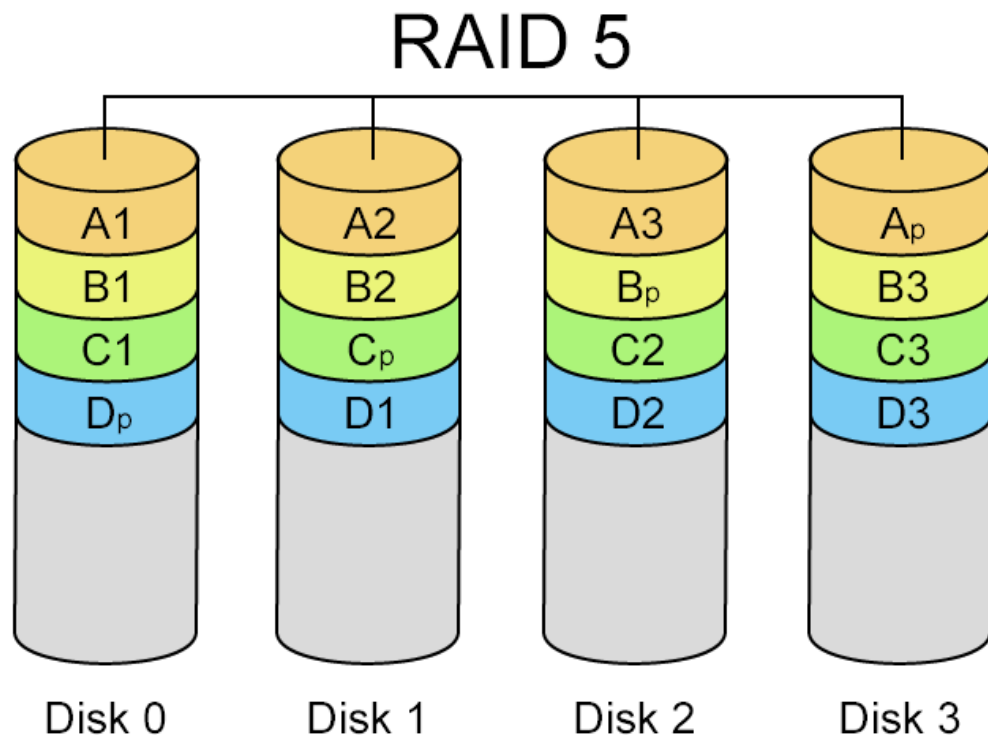
RAID 4

- Hasonló a RAID 3-hoz, de nagyméretű sávok
- Közvetlenül bármelyik diszkhez hozzáférhetünk
 - Párhuzamos kiszolgálást lehetővé tesz
 - Paritásdiszk nagyon korlátoz!
- Nem használják a gyakorlatban



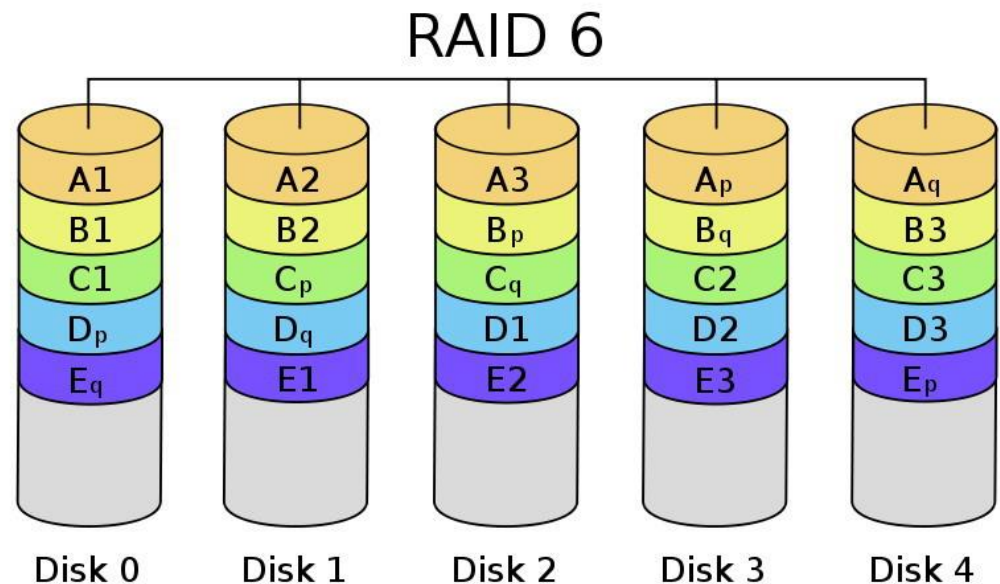
RAID 5 – elosztott paritás

- Paritás egyenletesen szétosztva
- Kiküszöböli a paritás meghajtó okozta szűk keresztmetszetet
- Közvetlenül elérhetők a diszkek
- Változtatható sávméret

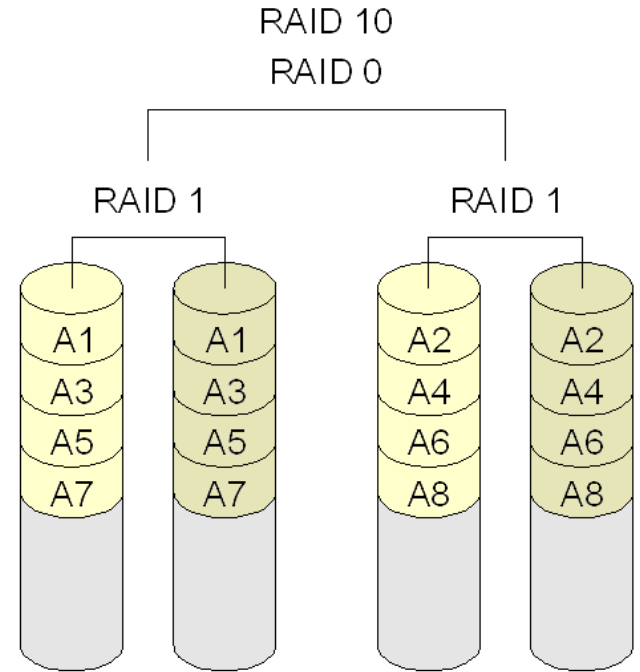
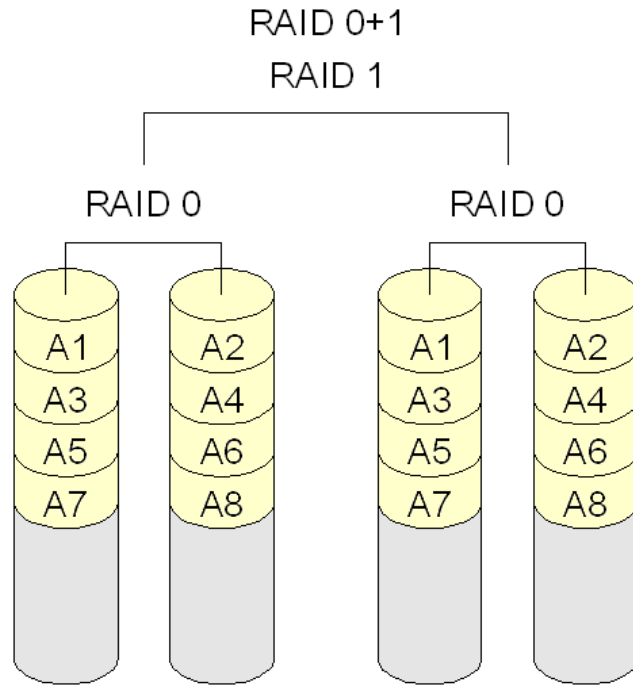


RAID 6 – kettős paritás

- Sor (XOR – P) és oszlop (Reed-Solomon kód – Q) paritás
 - kettős hiba ellen is véd – de lassú
- Diszkek között szétosztva



RAID 01, RAID 10



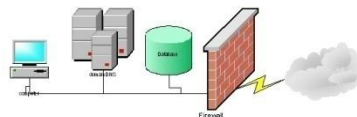
Egyformák?

Hiba: egész stripe – nincs tükrözés (!!)

Helyreállítás: egész stripe(!!)

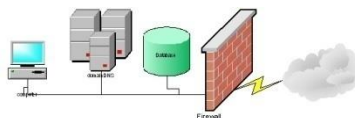
Sebesség: nagyobb (HW striping)

csak a felén nincs tükrözés
csak a rossz diszk
lassabb (SW striping)



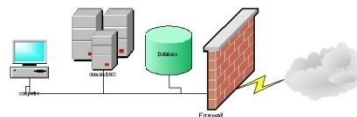
RAID

- Gyakorlatban 0,1, 5 elterjedt
- RAID csak a FIZIKAI diszk hibák ellen véd!
 - Logikai hibák ellen: mentés!



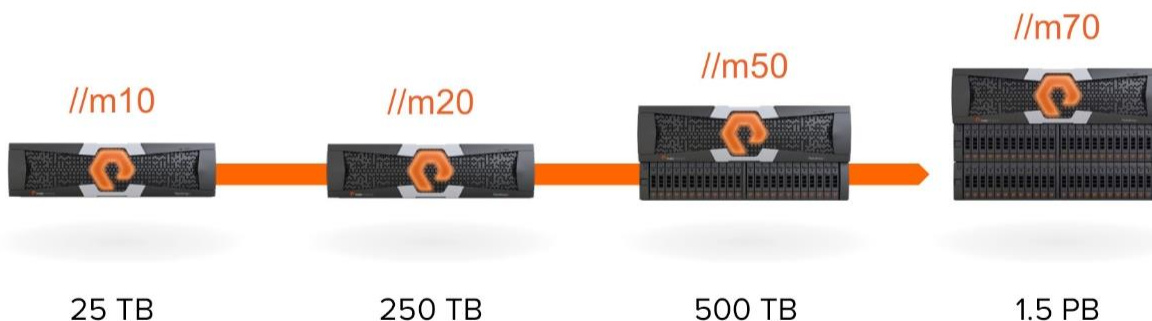
Flash memóriák

- Nonvolatile memory (EEPROM)
 - ‘Nem felejtő’
 - Speciális jellel törölhető (‘flash’)
 - Pen-drives
 - DataTraveler Ultimate GT
 - SSD
 - Solid State Drive
 - AFA
 - All-Flash Array
 - Hybrid arrays

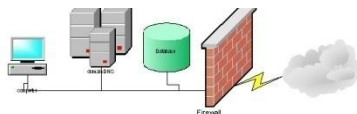




All-Flash Array

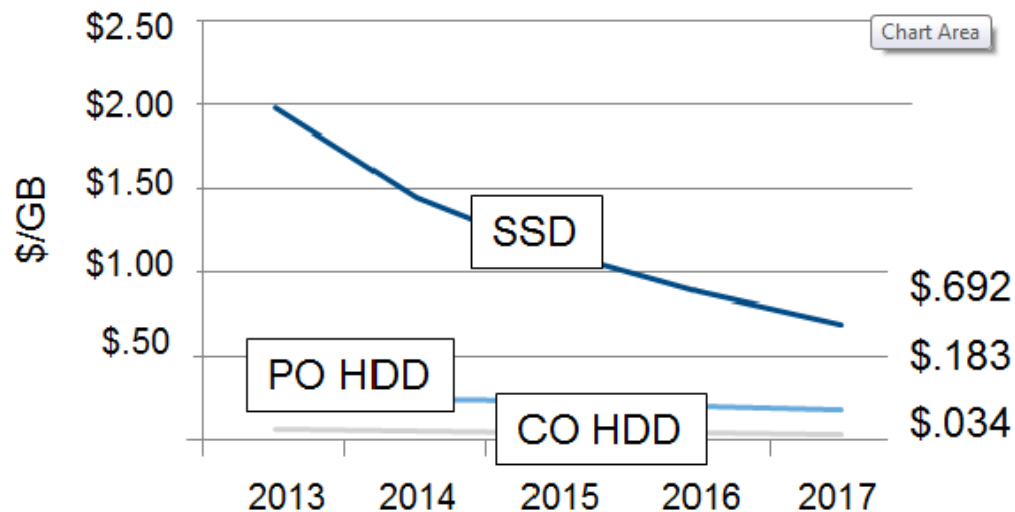


BME VIK TMIT



Információs rendszerek üzemeltetése

Ár



PO: Power optimised

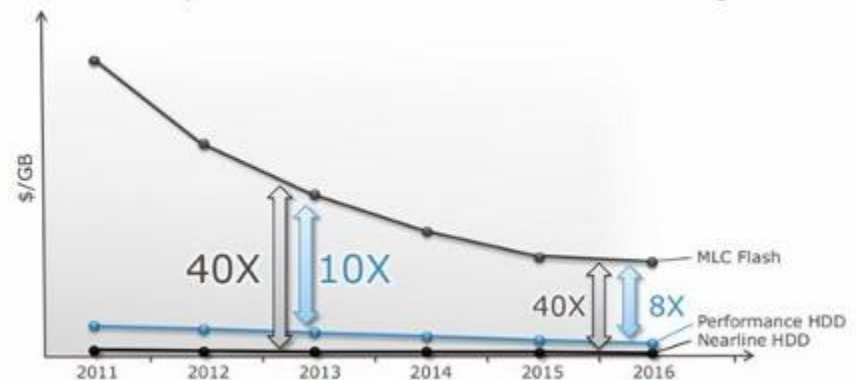
CO: Cost optimised

TOC: Total Cost

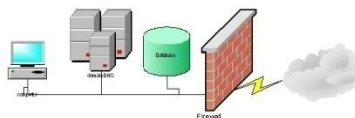
~20-25% of the
TOC (HDD)
~80% of the
TOC (SSD)

Flash vs HDD : Industry Cost Trends

MLC 8X More Expensive Than Performance HDD Through 2016

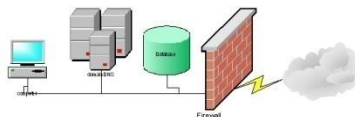


EMC²



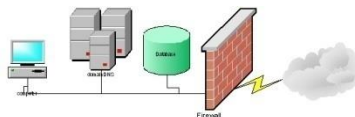
Késleltetés

- Latency 1000:1 (HDD:SSD)
 - Ma már a csatlakozó (SATA, SCSI) sebessége korlátoz
 - Gyorsabb boot-olás
 - Adat kompressziós (data reduction) technológiák sokkal hatékonyabbak
 - 6:1 (SSD) – 2:1 (HDD)
 - Ha ezt már eleve egy felső réteg nem tette meg... – VM
 - Effektív kapacitás (Effective usable capacity)
 - SSD – tipikusan a kontroller, míg HDD - extra SW
 - Gyorsabb – több használói igény elégíthető ki
 - Elégedettebb használók

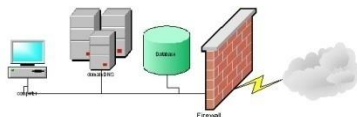
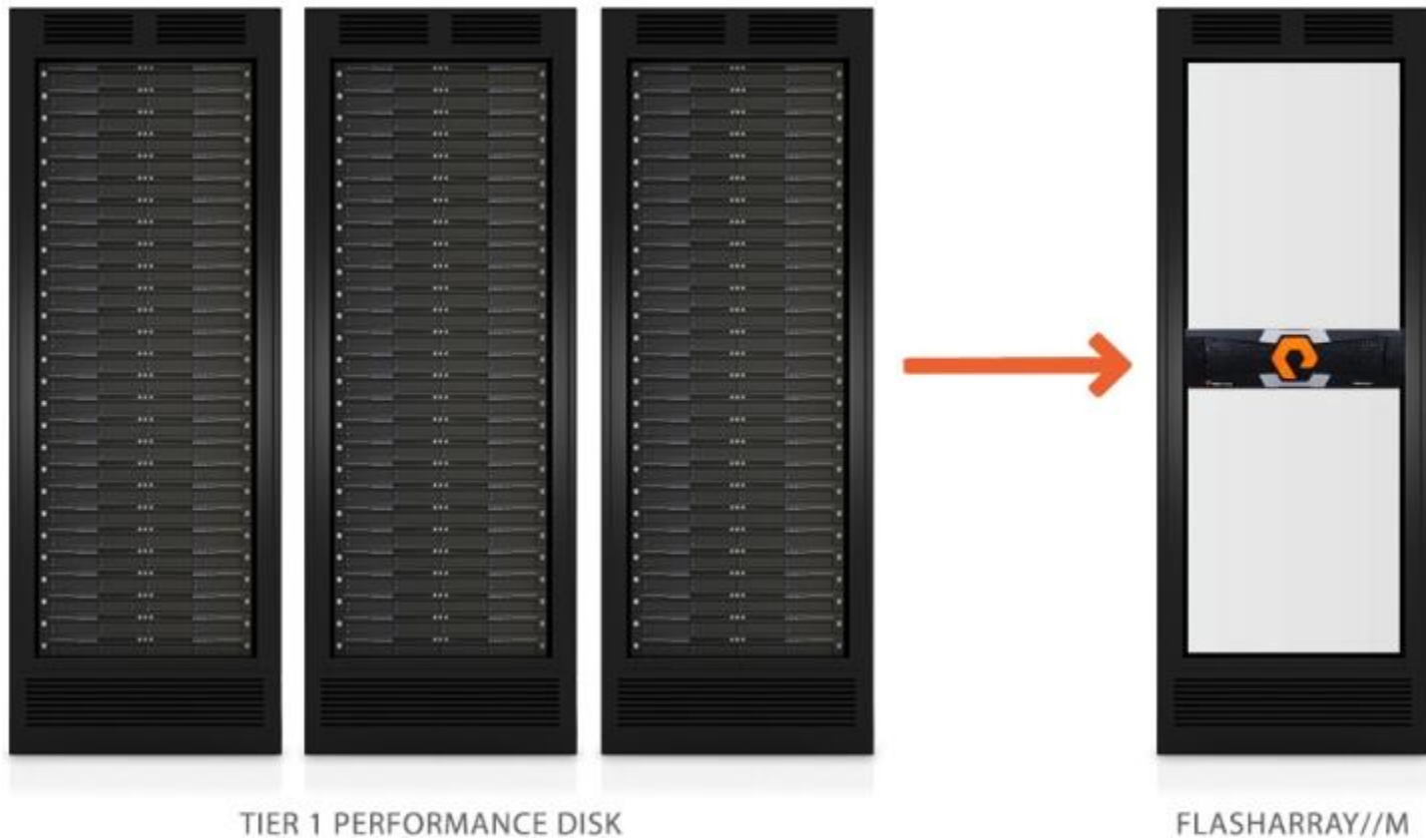


Energia, hűtés, méret

- Táp, hűtés
 - Nincs mozgó alkatrész
 - Kisebb energia-felhasználás (1:10)
 - Kisebb hűtés igény
 - Kb. 50-90%-kal kevesebb, mint a HDD esetén
- Kisebb
 - Pláne, ha az adatkompRESSZIÓ hatását is beszámítjuk
 - 20-40* adat sűrűség, mint a HDD-nél
 - Akár 90%-kal jobb rack kihasználtság

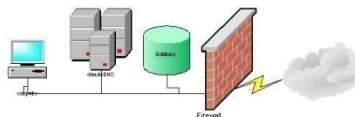


All-Flash Array



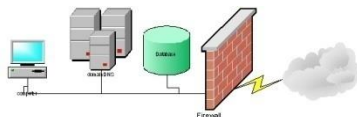
Megbízhatóság

- Kisebb fenntartási költség
- Reliability
 - 100:1 (unrecoverable bit error rate)
 - 10^{-18} : 10^{-16}
 - 10^{-18} : 1 hiba többszáz év alatt
 - A korai EEPROM-oknál az írási ciklusok limitáltak voltak (~1000)
 - Mai SSDs – több mint 200 év
 - RAID (1 a régebbiekénél, 5 az újabbaknál)



Hybrid arrays

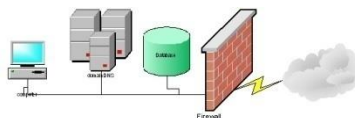
- HDD jobb ár/GB
- SSD jobb ár/IOPS
 - Input/output operations per second
- Rakjunk egy vékony SSD réteget egy HDD tömb fölé
 - SSD: 2% - 5% az összkapacitásnak
 - Elérhető IOPS duplázódik
 - Olvasási késleltetés 10+ ms helyett 3-5 ms
 - DE nem egyenletes – ez néhány alkalmazásnál problémát okozhat
 - 10% - 20% árnövekedés a teljes tömbre
 - 2X teljesítmény
 - RAID 1



Összehasonlítás

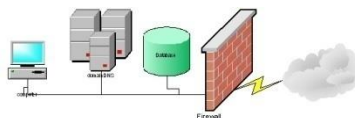
	Avg upfront per GB HW Cost	1yr avg per GB Power & Cooling	Avg upfront per usable GB Costs*	~ 3 yr per usable GB TCO
100% HDD Storage System	\$0.50	\$0.50	\$0.63	\$2.13
Hybrid Storage System	\$1.60	\$0.38	\$0.59	\$1.72
100% Flash SSD Storage Systems	\$10.00	\$0.17	\$3.64	\$4.13
100% Flash SSD Storage 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 formatting 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



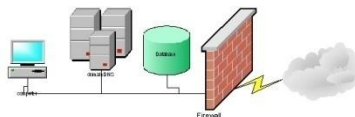
Tároló rendszerek

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



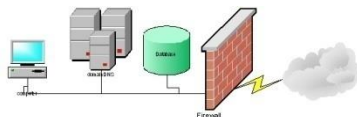
DAS – Direct-Attached Storage

- A tároló közvetlenül a szerverhez csatlakozik
 - Blokkszintű hozzáférés
 - Főképp kis rendszereknél
- Két alaptípus
 - Belső (Internal DAS)
 - Külső (External DAS)



Internal DAS

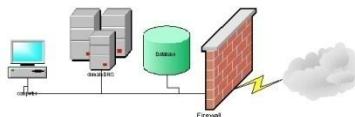
- A tároló közvetlenül a szerverhez kapcsolódik belső soros vagy párhuzamos buszon keresztül
 - Távolság korlátozott
 - Általában a buszhoz csak korlátozott darabszámú eszköz csatlakoztatható
 - (P)ATA vagy SATA csatlakozó
 - Nagy helyet foglal a szerveren belül
 - Nehezebb karbantartás



(P)ATA



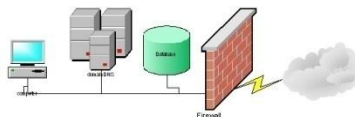
- Párhuzamos elérésű HDD interfész
 - Half duplex
- PATA – Parallel Advanced Technology Attachment
- 40 & 80 eres kábel
 - 40 eres: UDMA max. 33 MB/s
 - 80 eres: UDMA 66, 100, 133 MB/s
 - nagy helyigény => fejlesztés megállt (~2004)



SATA

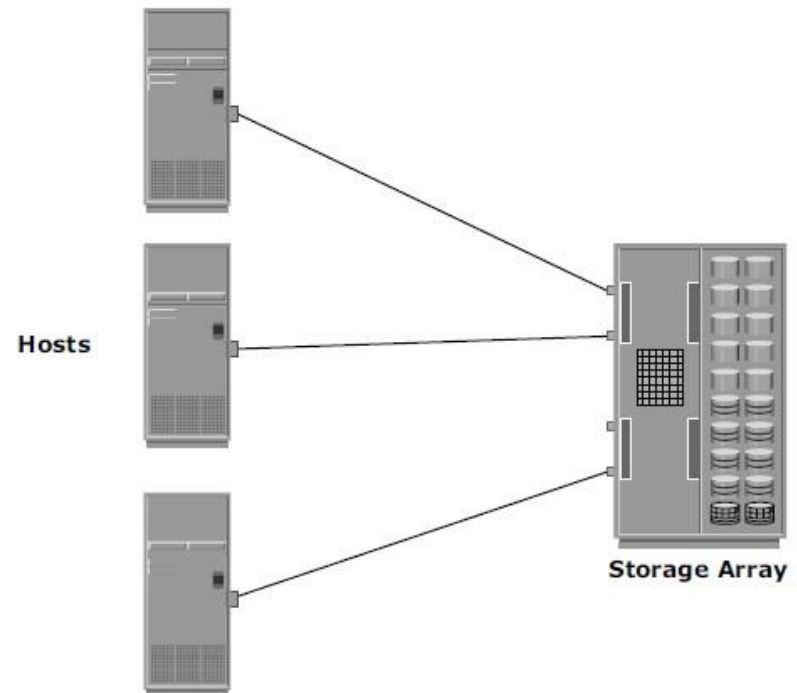


- Serial Advanced Technology Attachment (SATA)
- Pont-pont összeköttetést teremt a SATA host adapter és a SATA eszköz között
 - Új csatlakozó felület
 - Nagyobb átviteli sebesség
 - (P)ATA 66/100/133 MB/s
 - SATA 150/300/600 MB/s
 - A csatlakozó kábel 4 eres, maximális hossza 1 m
 - Half duplex



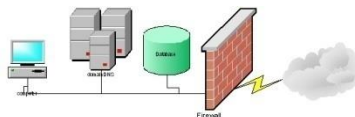
External DAS

- A szerver közvetlenül kapcsolódik egy külső tárhoz
 - Nagyobb távolság
 - Általában nem (annyira) korlátozott a csatlakoztatható eszközök száma
 - SCSI (vagy FC) csatlakozás



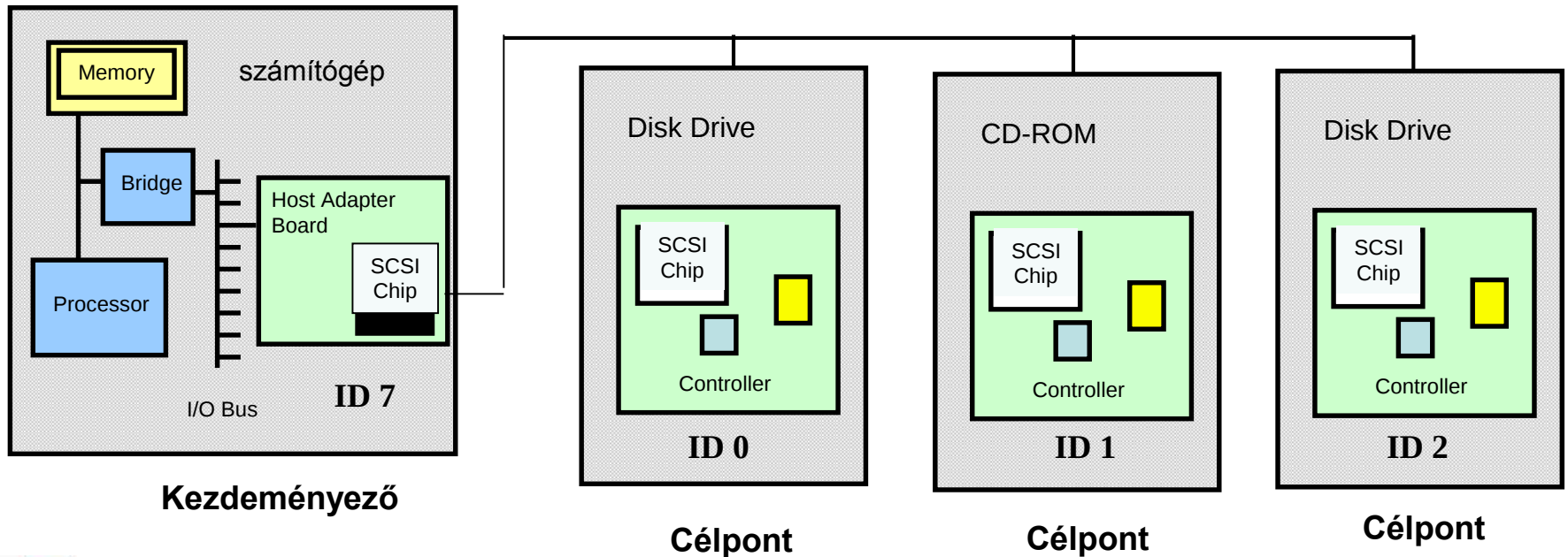
SCSI interfész

- SCSI
 - Small Computer System Interface (SCSI)
 - Szabványos I/O busz
 - Nagy teljesítményű interfész
- Eszközök
 - Disk Drives
 - Tape Drives
 - Removable Media Drives
 - CD-ROM, CD-R/CD-RW Drives
 - Optical Memory Drives
 - Printers



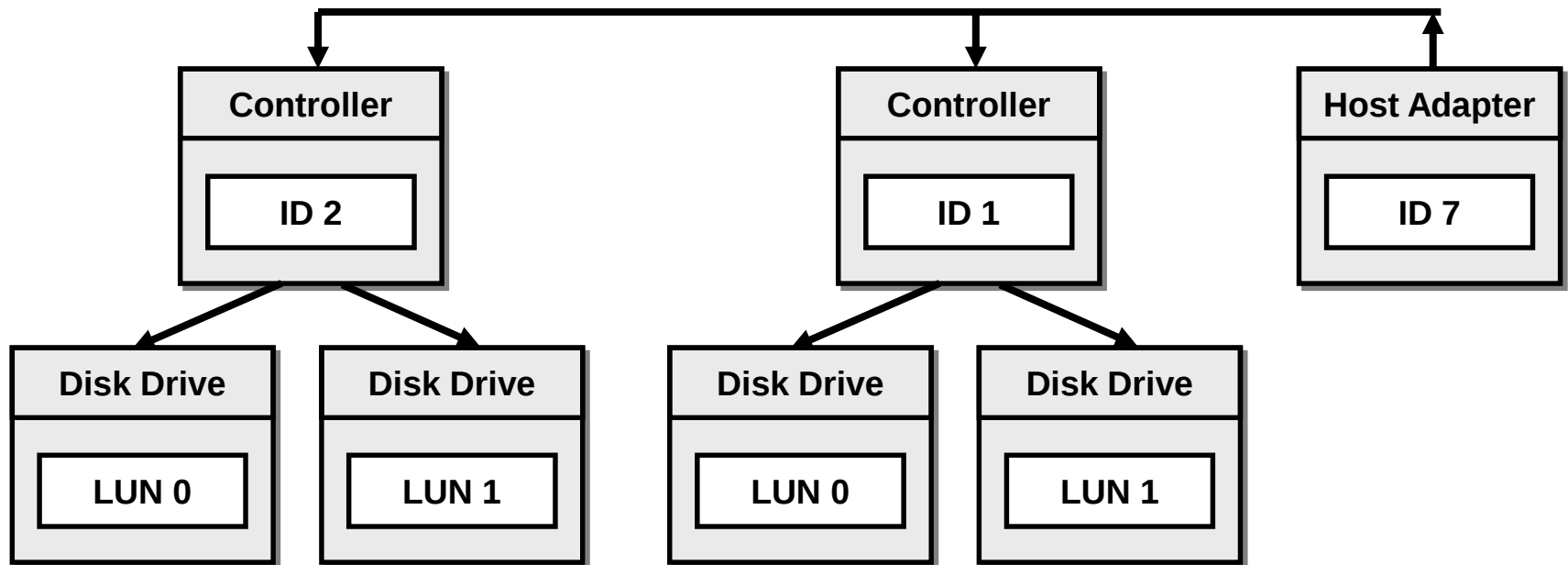
SCSI topológia

- Összes eszköz egy közös buszon osztozik
- Az eszközök egyedi azonosítóval rendelkeznek
- Csoportosításuk szerepük alapján történik
 - A kezdeményező általában számítógép
 - A célpont tipikusan merevlemez, CD-ROM

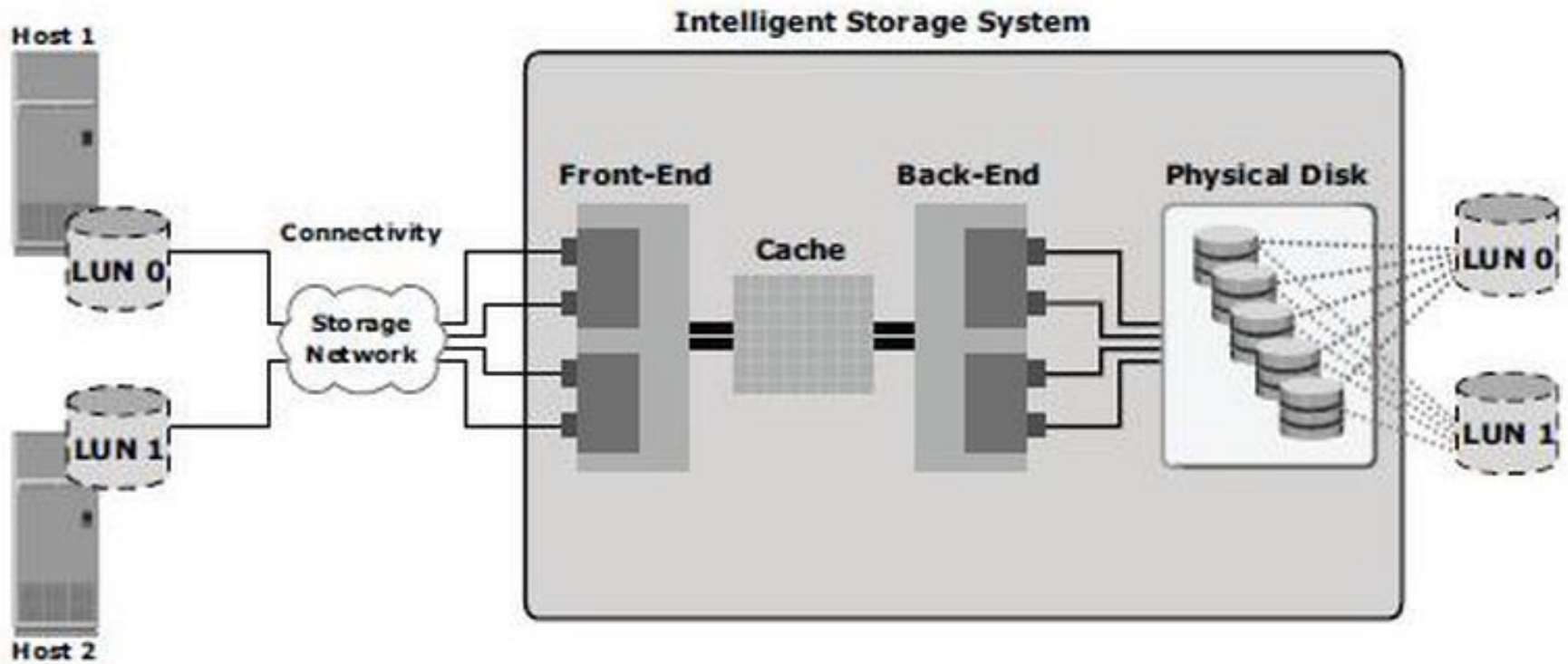


Logical Unit Number (LUN)

- **NEM (!!!) szám** – egy tárolócsoporthoz
- LUN egy SCSI részcsoporthoz címezési mód
 - Különböző fizikai eszközökön elhelyezett területek logikailag egy egységnek látszanak



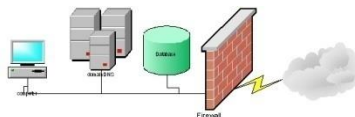
LUN



SAS - Serial Attached SCSI



- Serial Attached SCSI (SAS)
 - A párhuzamos (parallel) SCSI adapter továbbfejlesztése
 - Adatátviteli sebesség 3, 6 or 12 Gbit/s
 - Full duplex, magasabb megbízhatóság
 - Több drive címezhető egyetlen controller portról



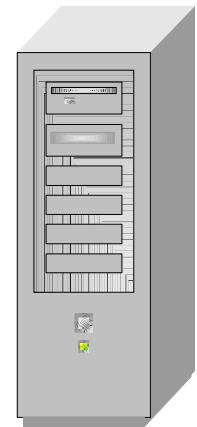
Direct-Attached Storage (DAS)

Előnyök

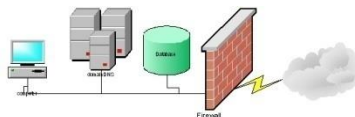
- Jobb, mint ha az adatokat a kliens tárolná
- Korlátozott redundanciát tud nyújtani
- Kis költség
- Egyszerű

Hátrányok

- Nehézkes menedzsment
- Backup költséges
- Elpocsékolott tárolóterület
- Nehézkes az adatmegosztás
- Nem (jól) skálázható
- Korlátozott darabszámú eszköz csatlakoztatható
- Kis teljesítmény, korlátozott sávszélesség

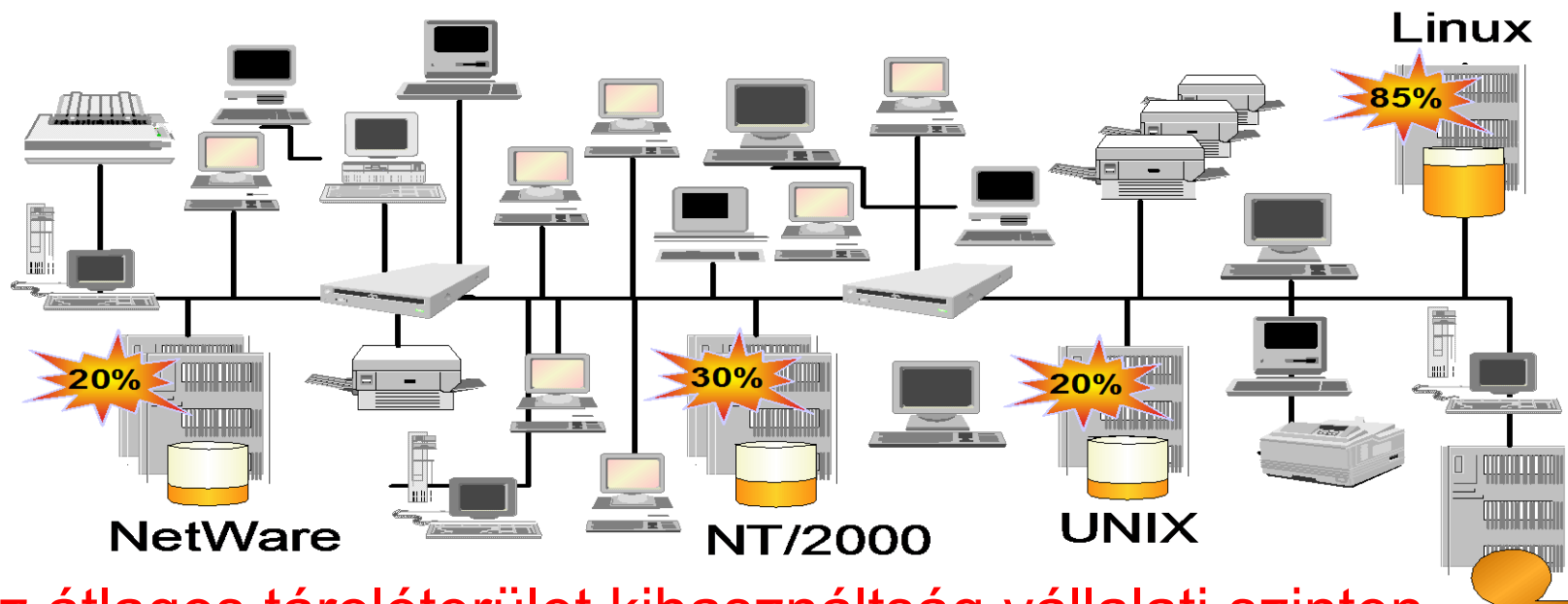


DAS Device



Dedikált tároló eszközök

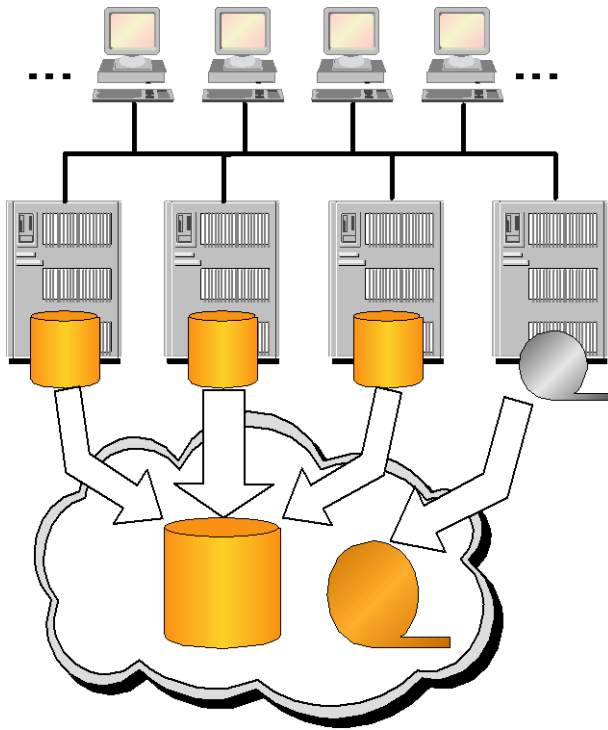
- Elosztott szerverek és tárolók – információ szigetek, különálló menedzsment alatt
- Nem hatékony az erőforrások (technikai és emberi) használata, magasak fajlagos költségek



Az átlagos tárolóterület kihasználtság vállalati szinten tipikusan 40-60%.

Konszolidált tároló eszközök

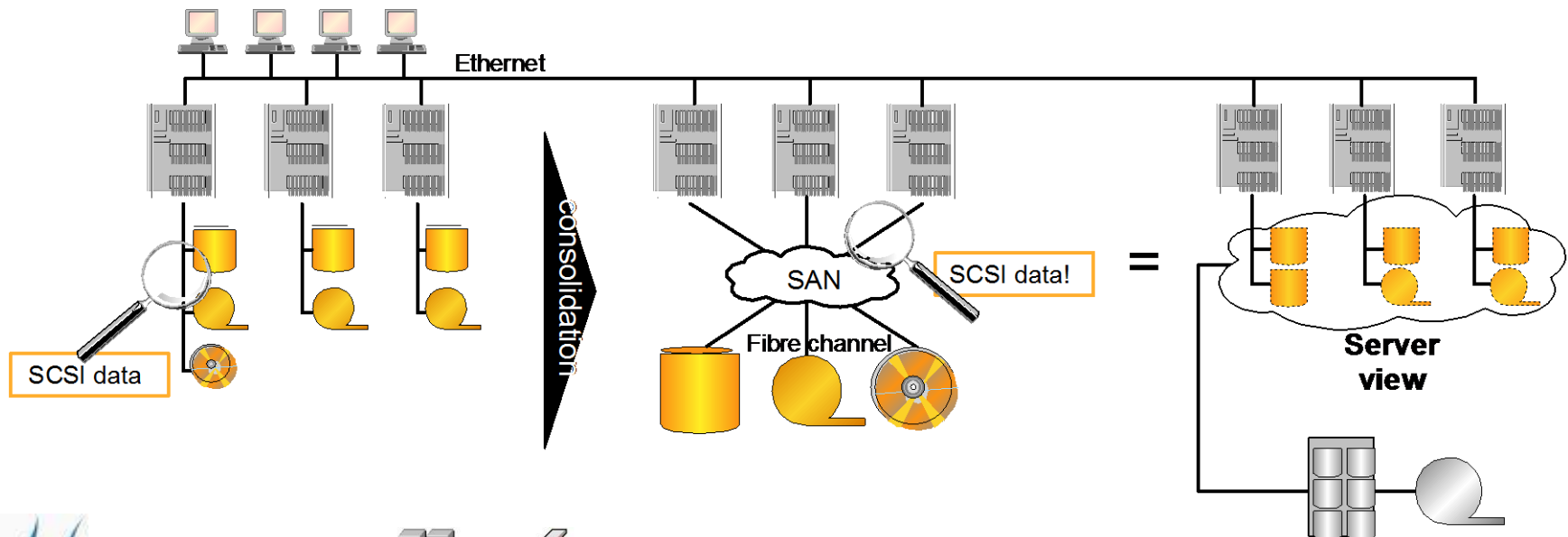
- Konszolidált eszközök, menedzsment, adat
- Kevésbé komplex, alacsonyabb fajlagos költségek
- Nagy rendelkezésre állású, skálázható, katasztrófa tűrő rendszerek alakíthatóak ki



A konszolidált
(consolidated) tárhasználat
hálózati architektúrában
valósítható meg

SAN - Storage Area Network

- Adattároló hálózat, adatforgalomra dedikált hálózat
- A tároló eszközök az egyes szerverektől fizikailag elválnak, több server képes ugyanazt az eszközt elérni
- Nem változik az adatkezelési protokoll, a szerverek dedikált, saját tárat látnak



A SAN hálózatok előnyei

- **Erőforrások**

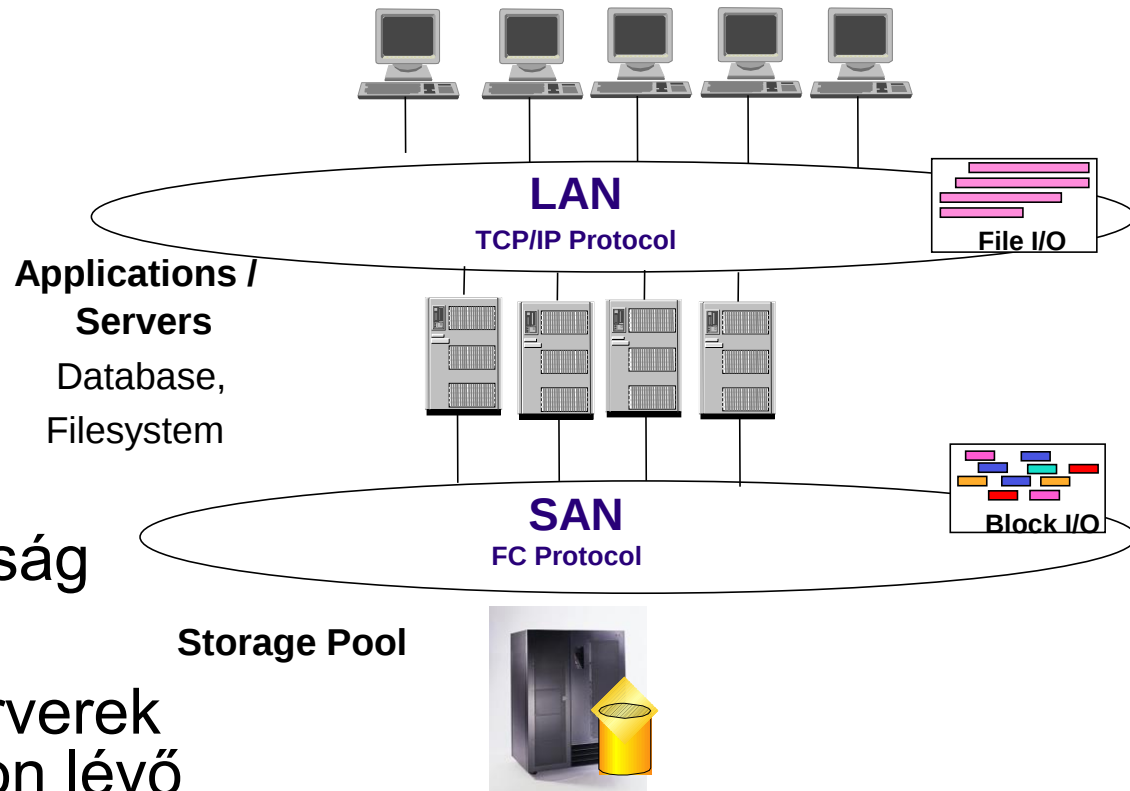
- Az erőforrások kihasználtsága nő
- Skálázhatóság
- Magasabb szintű rendszerfunkciók valósíthatók meg (**replikáció**)

- **Menedzsment**

- Nagyobb hatékonyság

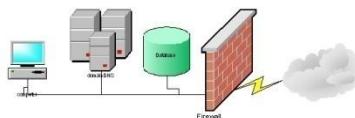
- **Információ elérés**

- Az alkalmazás szerverek bármely, a hálózaton lévő adatot bármikor elérhetnek



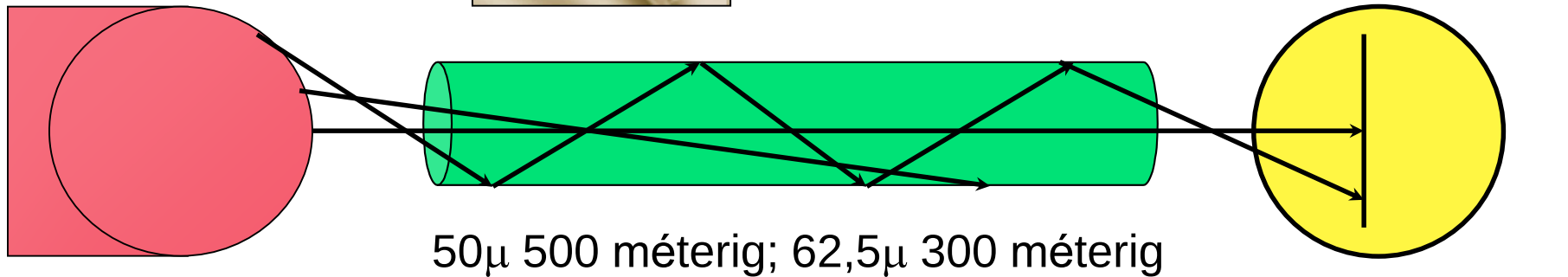
Átviteli technológia: Fibre Channel Protokoll

- Skálázható
 - Nagyszámú eszköz
 - Nagy távolság
 - Átviteli megoldás számos protokollhoz
 - SCSI-3 (=> **diszkek ugyanúgy érhetők el, mint ha helyiek lennének !!**)
 - IP, ATM, ...
- Pont-pont és kapcsolt hálózati topológia
- Sokféle eszköz, sebesség
 - Különböző típusú rézdrót, üvegszál
 - Max 128 Gb/s sebesség

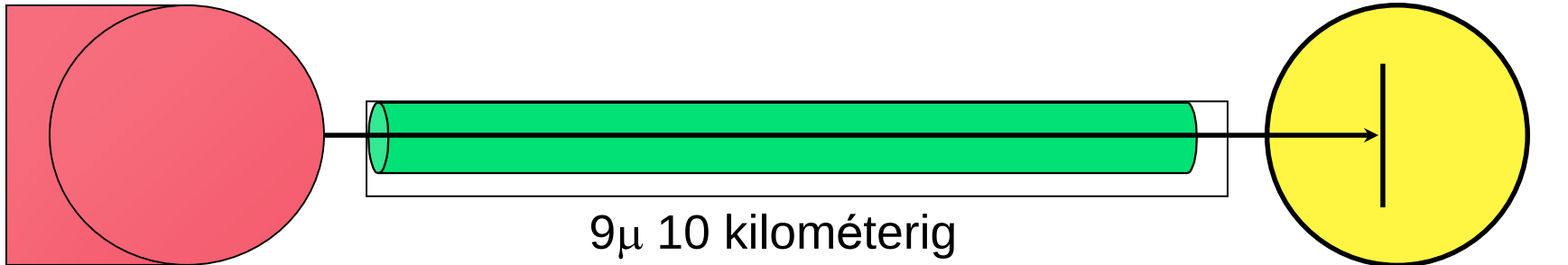


Üvegszál (Fiber Optic)

- Multimode LED



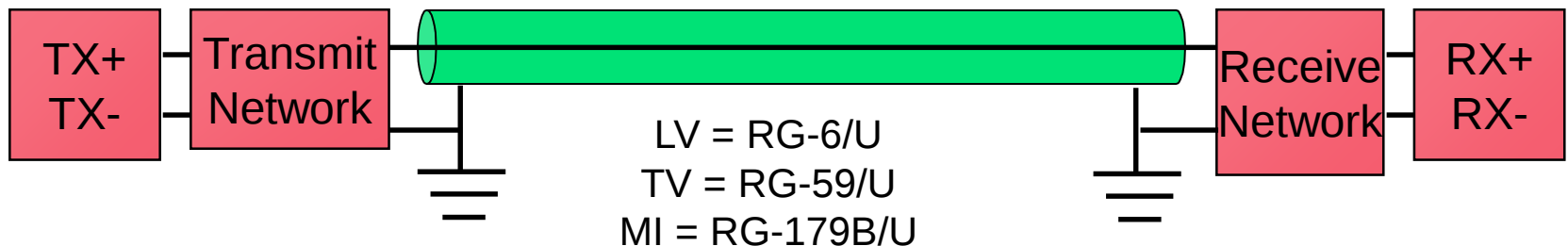
- Single mode LED



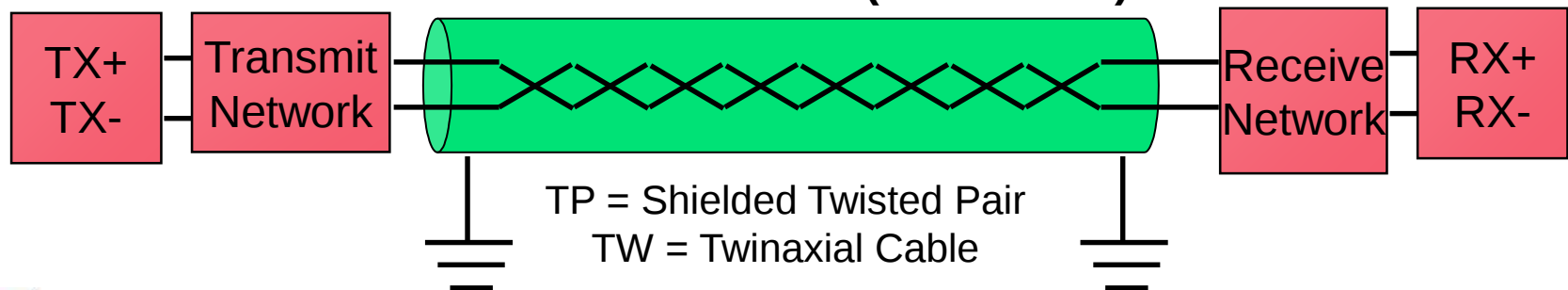
Rézvezeték

- Főleg Back-end
- Max. 15 méter
 - Jobb jel-zaj viszony, mint az üvegszálon

75 Ohm Unbalanced (Single Ended)

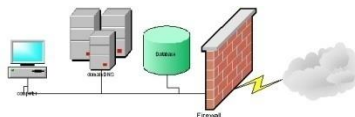


150 Ohm Balanced (Differential)



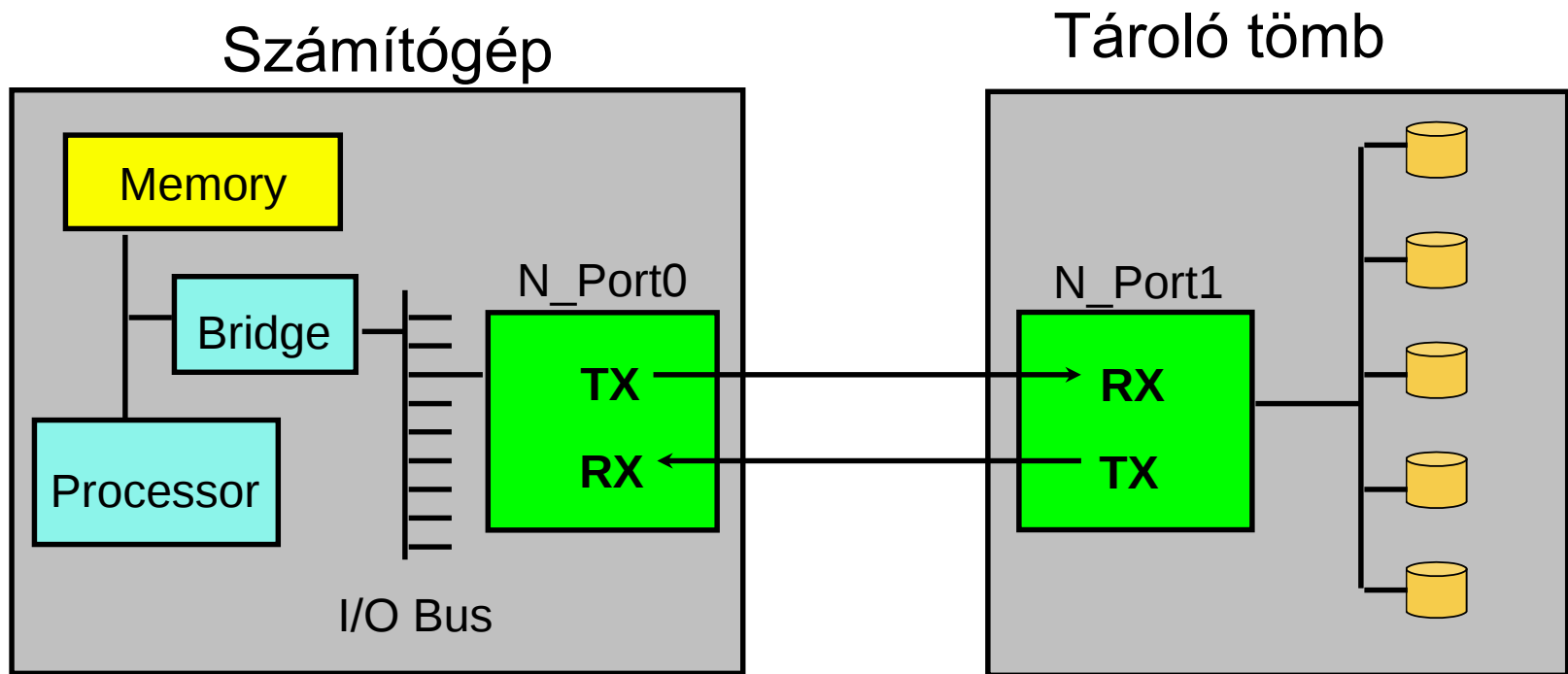
Fibre Channel Protokoll

- Háromféle topológia
 - Pont-pont
 - Arbitrált gyűrű (Arbitrated Loop) (nem használt)
 - Kapcsolt hálózat (Switched Fabric)
- FC portokat köt össze
 - Bármilyen, FC-n kommunikálni képes eszköz, nem fizikai port
 - N port (Node)
 - Disk
 - HBA (Host Bus Adapter) számítógépekben
 - F port
 - FC Switch



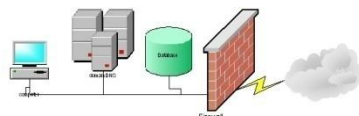
Pont-pont

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

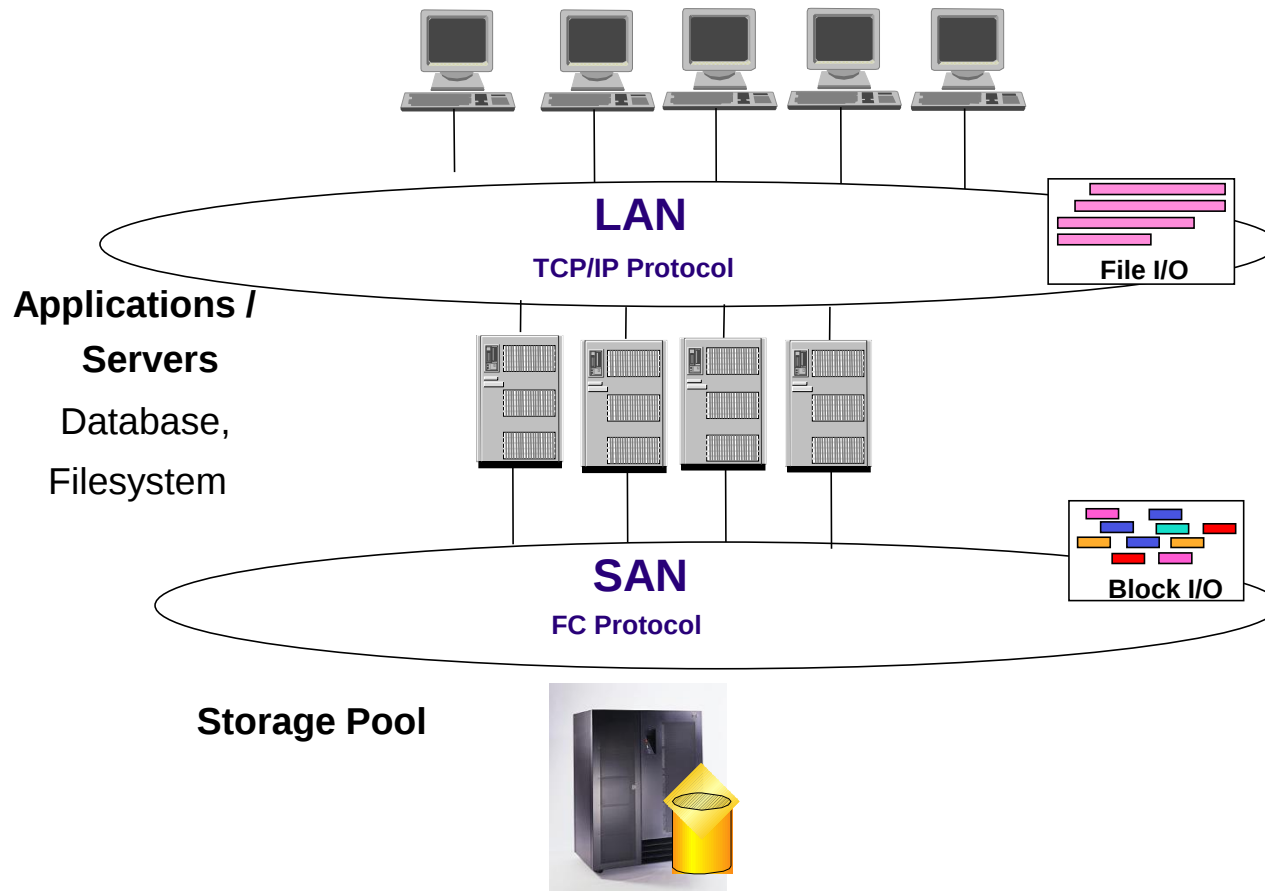


HDD Technológiák összehasonlítása

	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	≠	> 6 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

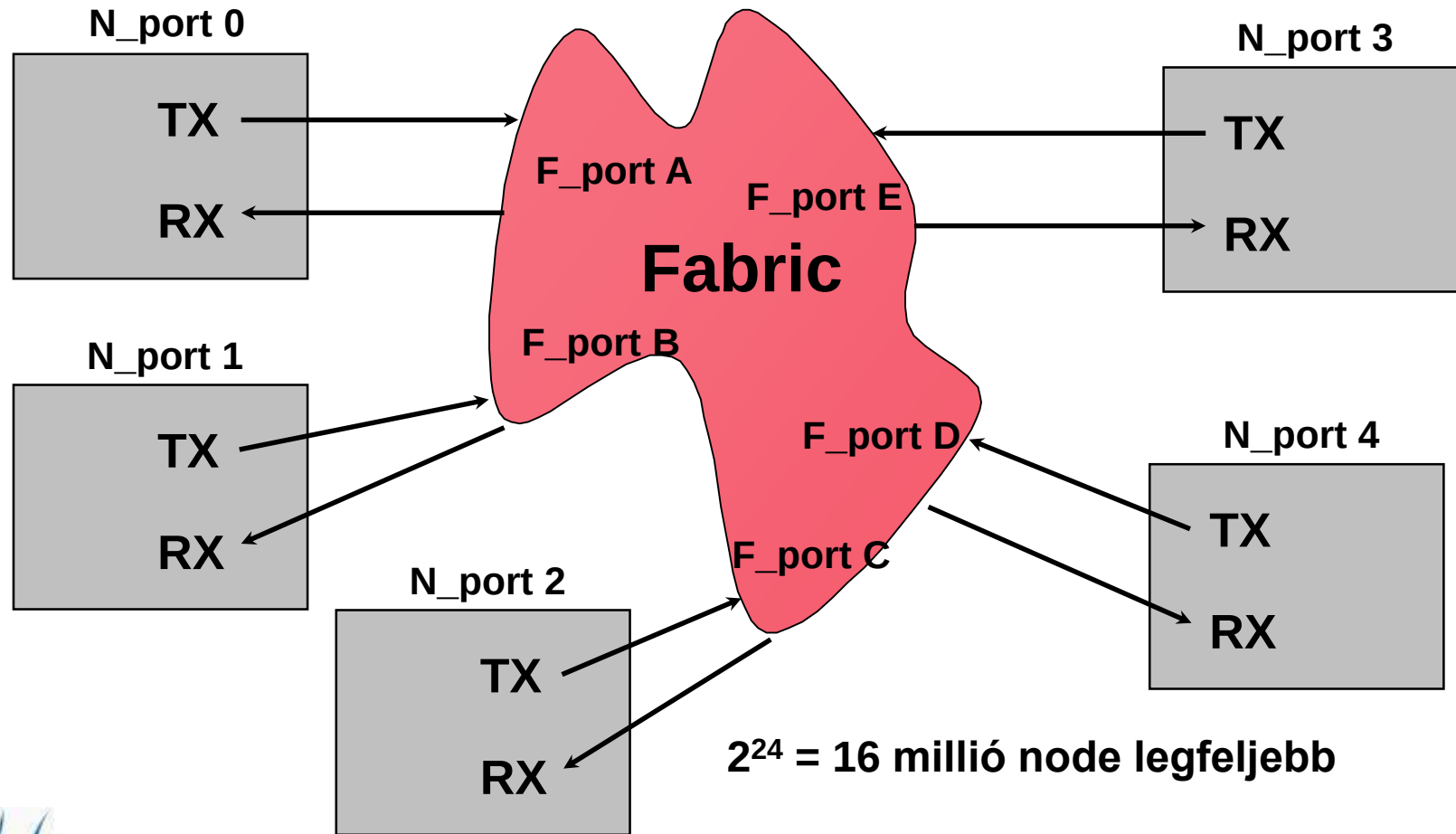


Vissza: SAN hálózatok

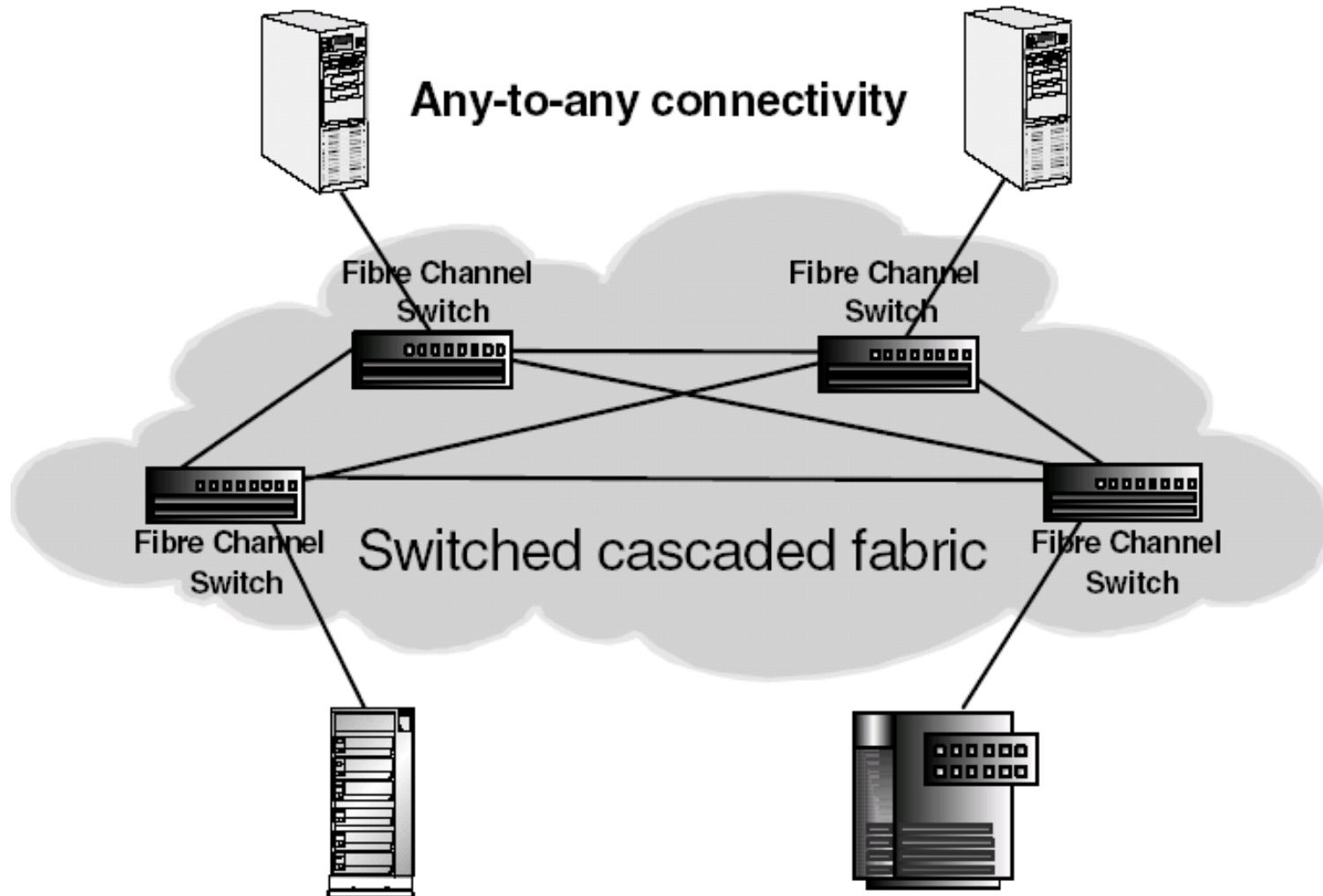


Kapcsolt hálózat / Fabric

- Vannak más elrendezések is, de ma ált. ezt használják

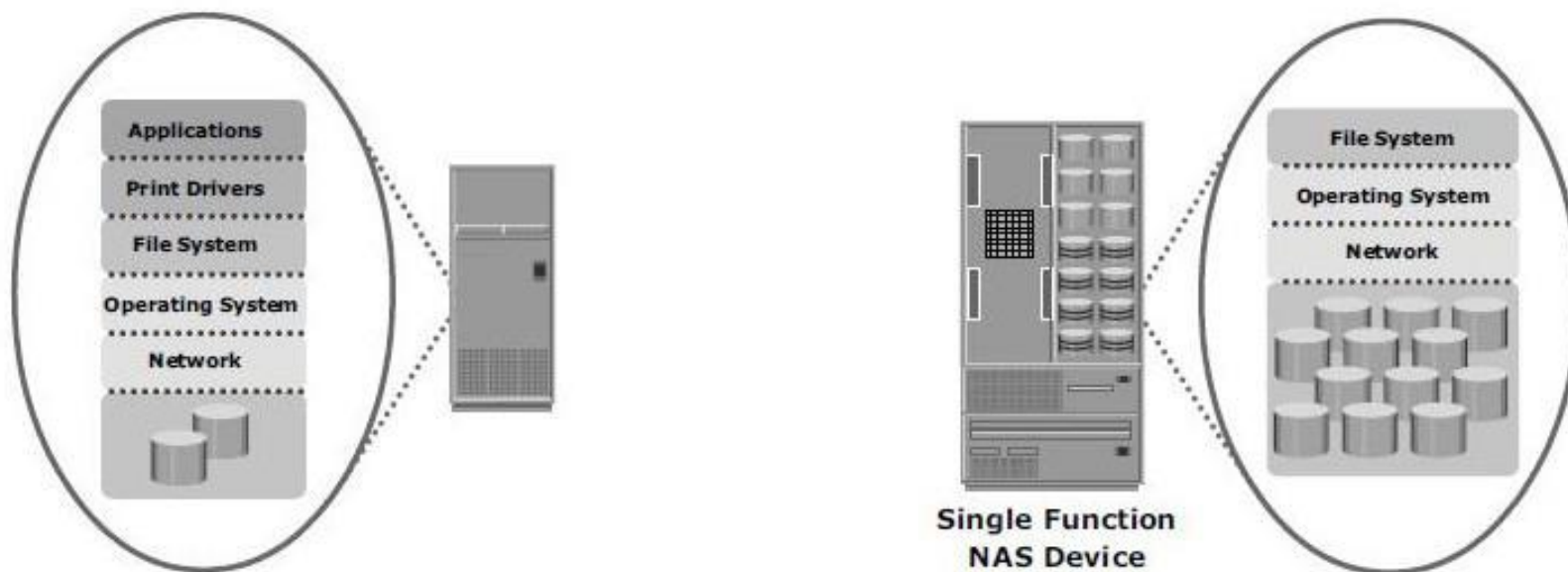


Kapcsolt hálózat



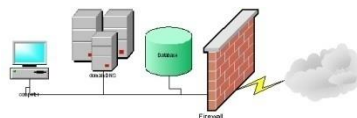
NAS – Network-Attached Storage

- Olyan tároló, amely IP hálózathoz csatlakozik
 - Dedikált célú fájlserver
 - I/O műveletekre optimalizált op. rendszerrel



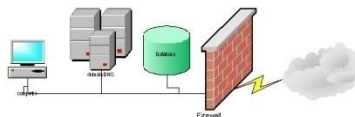
General Purpose Server

Single Function
NAS Device



NAS

- IP (LAN, WAN) kapcsolat, belső SCSI struktúra
- Belső RAID - hibatűrés
- Fájl szintű elérés (Blokk elérés nem támogatott)
- Könnyen telepíthető
- Skálázás eszközön belül
- Performancia korlátok (LAN sávszélesség, protokoll overhead)

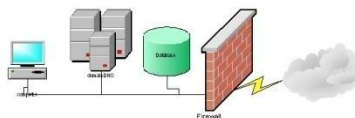


NAS protokollok

- NFS (Network File System) – UNIX
 - UDP feletti, fájlműveletekre specializált protokoll
- CIFS (Common Internet File System)
 - Operációsrendszer-független
 - Szerver
 - Kliens
 - TCP/IP feletti
- FTP (File Transfer Protocol)

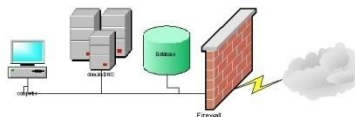
IP SAN (iFCIP, iSCSI)

- SAN: blokk szintű hozzáférés, FC hálózat
- NAS: fájl szintű hozzáférés, IP hálózat
- IP SAN: blokk szintű hozzáférés, IP hálózat
 - iFCP (Internet Fibre Channel Protocol)
 - Torlódás vez., hiba detektálás&javítás TCP-vel
 - iSCSI (Internet Small Computer System Interface)
 - SCSI parancsok meglevő LAN/VLAN/WAN-on
 - SCSI busz emulációja IP hálózaton
 - Bár tetszőleges SCSI eszköz csatlakoztatható, gyakorlatban server <-> adat tároló kommunikációra



IP SAN

- Tároló konszolidáció dedikált hálózat nélkül
 - Low-cost megfelelője a Fibre Channel-nek, de a teljesítmény nagyon függ a 'más' forgalomtól
- Vészhelyzeti helyreállítás (disaster recovery)
 - Tárolók tükrözése (távoli) adatközpontok között
 - 'Hot stand-by'
 - WAN felett



SAN vagy NAS összefoglalás 1.

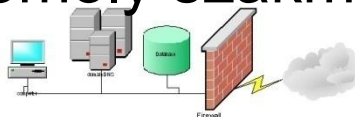
SAN (Storage Area Network)

Központosított, nagy teljesítményű, adattárolásra dedikált hálózat, amely

- Összeköti a szervereket, tároló eszközöket,
- Hálózati elemeket és kapcsoló eszközöket, valamint támogató szoftver megoldásokat tartalmaz

Előnyei:

- Skálázható, bővíthető
- Nagy adatátviteli sebesség (4 -10 Gb/s)
- Központilag felügyelhető, támogatja a hierarchikus tároló technológia kialakítását, de komplex, heterogén hálózat esetén komoly szakmai feladat a megfelelő üzemeltetés



SAN vagy NAS összefoglalás 2.

NAS (Network Attached Storage)

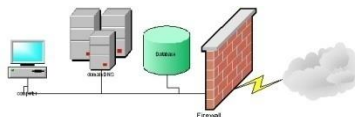
A hálózatra csatlakoztatott adattároló eszköz, amely támogatja a adatmegosztást kliensek és szerverek között.

Előnyei:

- Skálázható, bővíthető, de az adatátviteli sebesség korlátozott (LAN)
- Könnyen telepíthető, üzemeltethető eszköz
- Főleg kisebb környezetek esetén alkalmazható

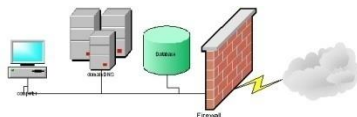
Diszkrendszeren belüli adatmásolás

- Nagymennyiségű adat másolását kell elvégezni
- Blokkszintű másolati példányokat készítünk
- Diszkrendszeren belüli nagysebességű, firmware támogatott másolási eljárások
- A felhasználó szempontjából teljes másolati kötegek – *volume copies* – jönnek létre



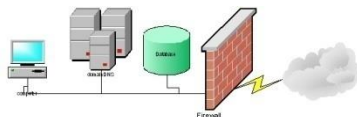
Volume Copy - Clone

- Egy adott tároló eszközön belül **firmware** eszközökkel megvalósított másolati technológia
- **Valódi duplikátum** kötet jön létre
- Teljes másolat: back-up, analitika, adatbányászat, stb. célra is alkalmazható
- Alkalmas alkalmazások más karakterisztikájú diszkekre történő **migrálás**ához



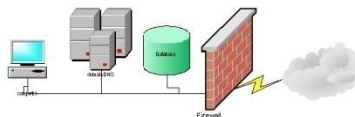
Volume Copy - Clone

- Az **adat-konzisztenciát biztosítani kell**, az alkalmazások állapotának figyelembe vétele, I/O műveletek leállítása szükséges a másolat készítésekor
- A másolás a terjedelemtől függően **időigényes**, az **alkalmazások** addig **állnak**
- Alternatíva: **Split mirror**, tükrözött állományok közötti kapcsolat felbontása után az állományok külön kezelése. Az alkalmazások futása folyamatos, de külön idő, amíg az újabb szinkronizáció létrehozható



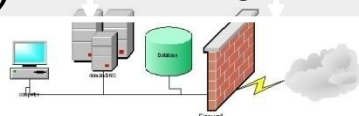
FlashCopy (Snapshot)

- Ha egy **blokkot módosítunk**, nem írjuk felül az eredetit, hanem **máshová tesszük**
- **FlashCopy tábla**, mely nyilvántartja az egyes fájlok blokkjait
 - „Pillanatfelvétel”
 - **Tetszőleges** időpontbeli állapot helyreállítható
- COW: Copy On Write
- A futó alkalmazást csak egy pillanatra kell leállítani, hogy koherens legyen
- **Nem alkalmas back-up célra, mivel nem keletkezik valódi másolt állomány!**



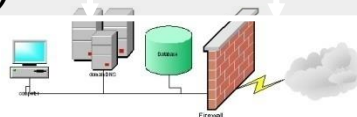
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	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
Összes blokkszám	8			8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



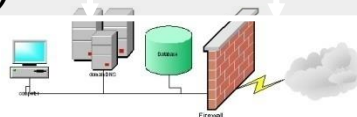
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás 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
Összes blokkszám	8			8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



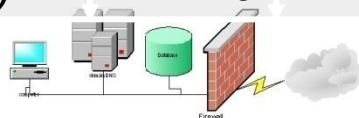
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás 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
Összes blokkszám	8			8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



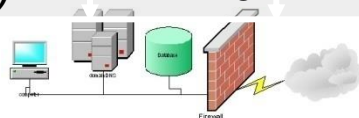
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
	B1	B1		B1	B1	
Írás 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
Összes blokkszám	8			8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



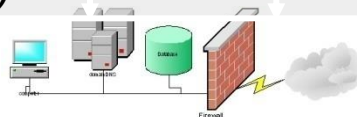
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
	B1	B1		B1	B1	
Írás 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
Összes blokkszám	8			8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



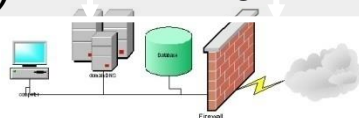
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
	B1	B1		B1	B1	
Írás 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
Összes blokkszám	8	10		8	8	8
Delta (flashcopy inkrementum)	0			Látszólagos Volume A		



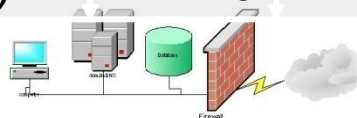
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
	B1	B1		B1	B1	
Írás 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
Összes blokkszám	8	10		8	8	8
Delta (flashcopy inkrementum)	0	2		Látszólagos Volume A		



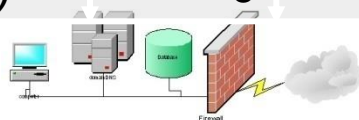
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
	B1	B1		B1	B1	
Írás 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
Összes blokkszám	8	10		8	8	8
Delta (flashcopy inkrementum)	0	2		Látszólagos Volume A	B	



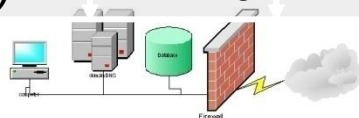
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
Írás t3	B1	B1	B1>B10	B1	B1	
Írás 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
Összes blokkszám	8	10		8	8	8
Delta (flashcopy inkrementum)	0	2		Látszólagos Volume A	B	



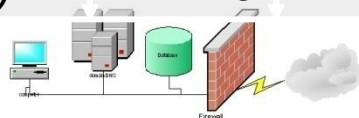
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
Írás t3	B1	B1	B1>B10	B1	B1	B10
Írás 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
Összes blokkszám	8	10		8	8	8
Delta (flashcopy inkrementum)	0	2		Látszólagos Volume A	B	



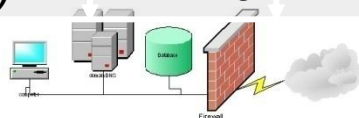
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
Írás t3	B1	B1	B1>B10	B1	B1	B10
Írás 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
Összes blokkszám	8	10	11	8	8	8
Delta (flashcopy inkrementum)	0	2	3	Látszólagos Volume A	B	



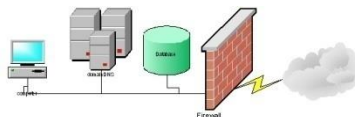
FlashCopy - blokkok

Blokk tábla				Flashcopy tábla		
Időpontok	T1	T2	T3	F1	F2	F3
Írás t2	B0	B0>B8	B8	B0	B8	B8
Írás t3	B1	B1	B1>B10	B1	B1	B10
Írás 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
Összes blokkszám	8	10	11	8	8	8
Delta (flashcopy inkrementum)	0	2	3	Látszólagos Volume A	B	C



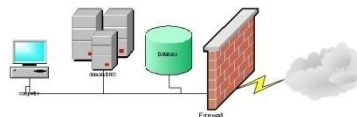
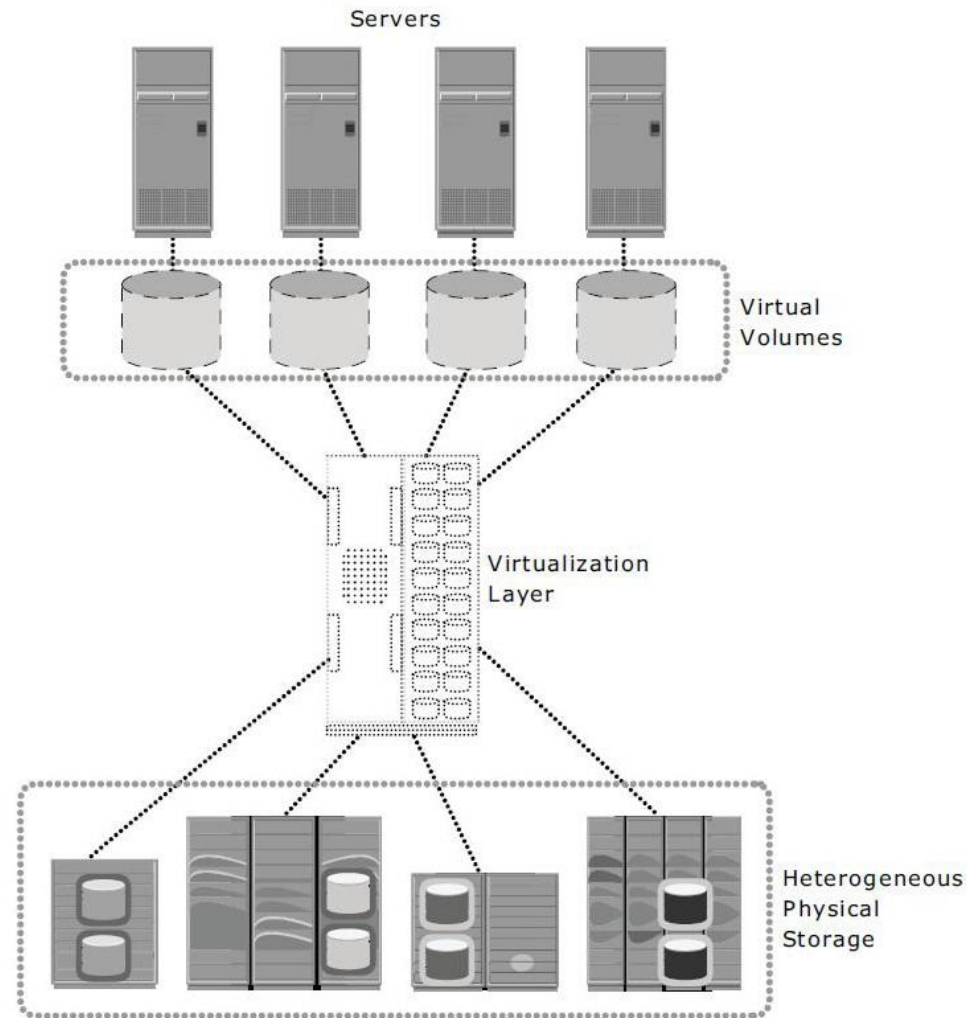
Tárterület virtualizáció

- A virtualizáció olyan technológia, amely lehetővé teszi, hogy bizonyos erőforrások **más erőforrásoknak tűnjenek, lehetőleg kedvezőbb tulajdonságokkal**
- A virtualizáció jellemzően **elfedi** a háttér-rendszer **komplexitását**, és új, hatékonyabb funkciókat biztosít a háttér-rendszer szolgáltatására építve
- A virtualizáció a rendszer **különböző rétegeiben** valósítható meg



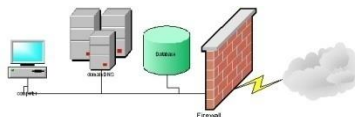
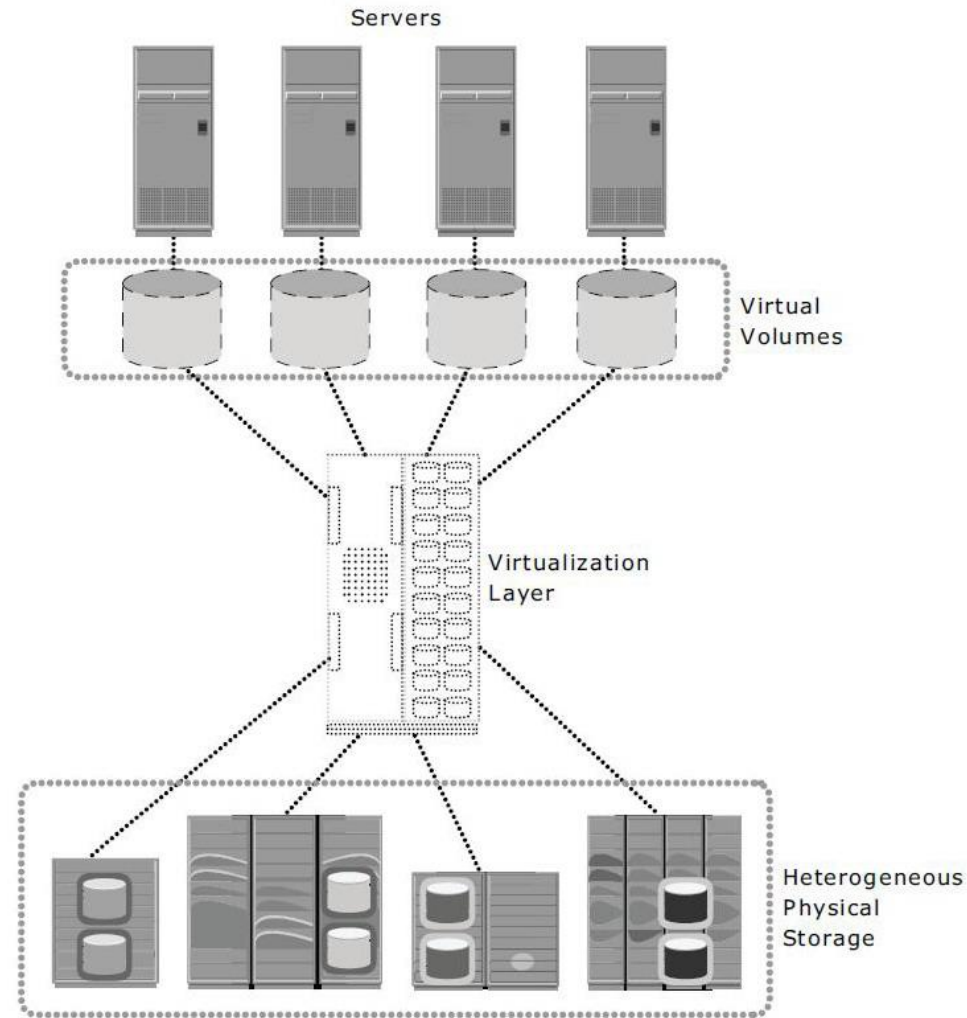
Tároló virtualizáció

- A virtualizáló motor (Virtual Engine or Virtualization Appliance) elfedi a diszkek különbözőségét
 - „fordít” a két formátum között
 - diszkek megoszthatók
 - nő a kihasználás
 - diszkek cseréje, módosítása nem látszik a szerver számára



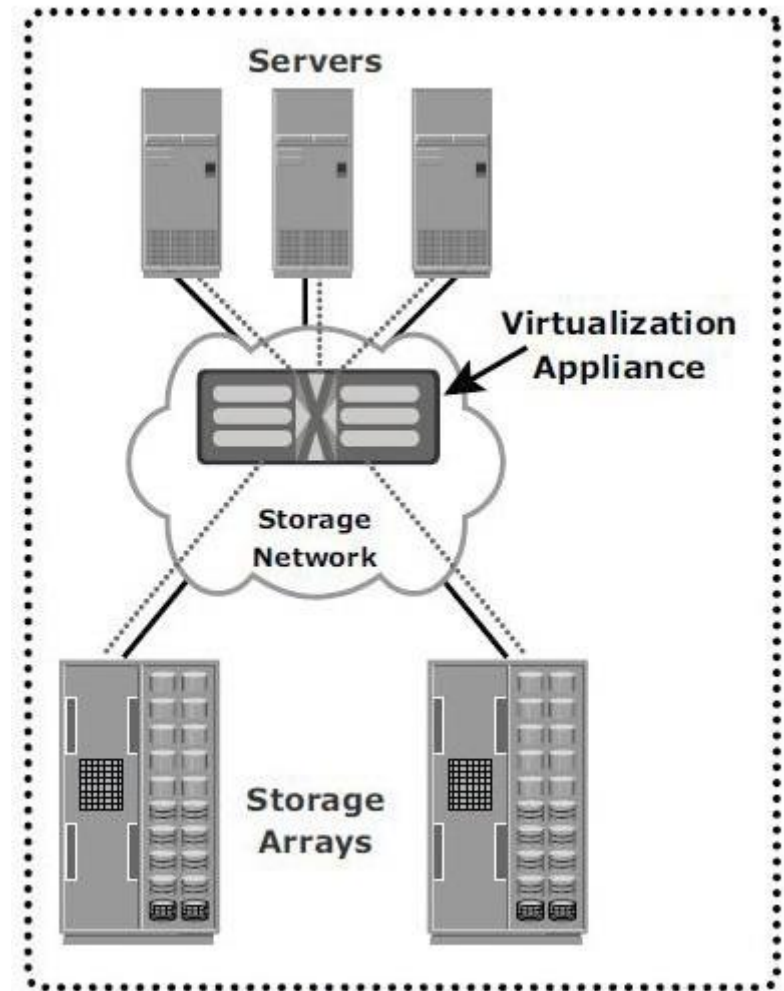
Virtualizáló motor működése

- Meta-data
 - tkp. egy összerendelési táblázat a logikai – fizikai cím között
- Szerver: LUN=1, LBA=32
 - LBA Logical Block Address
- VM: táblázatból, ez megfelel a fizikai LUN=4, LBA=0 címnek
- Elkéri az adatot a fizikai diszktől
- A megkapott adatot úgy továbbítja a servernek, mintha az a LUN=1, LBA=32 címről érkezett volna



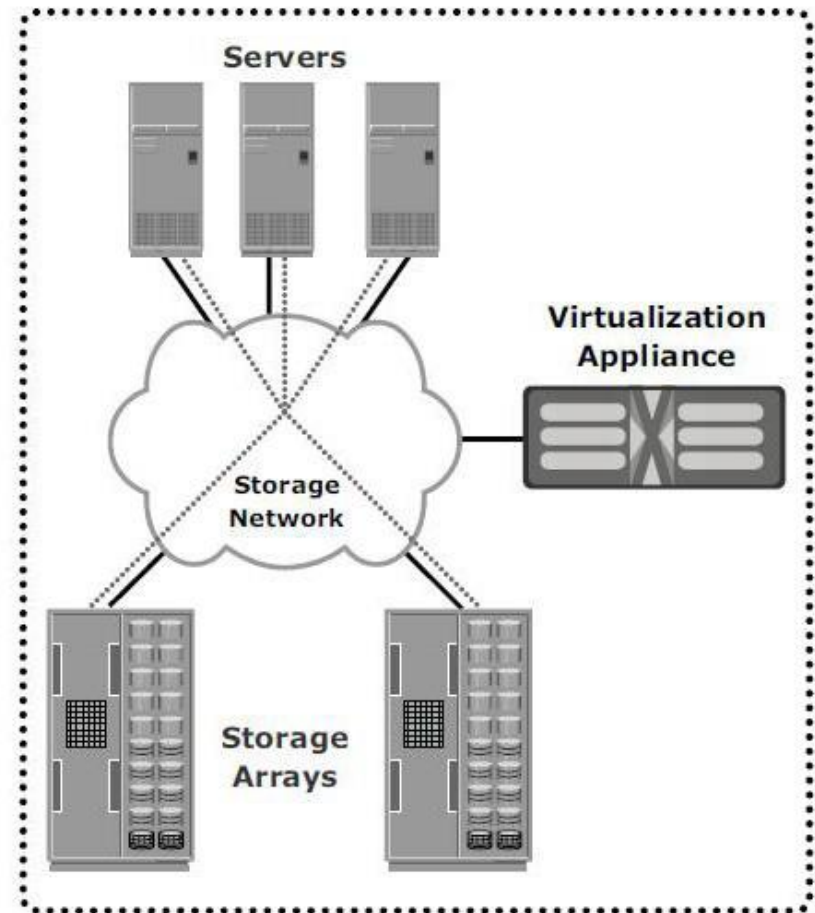
Virtualizáció konfiguráció – in-band

- Virtualizáló motor az adatútban
 - nincs szükség külön szoftverre a serveren
 - de lassabb, mert az adat átmegy a VM-en



Virtualizáció konfiguráció – out-of-band

- A vezérlés és az adatút elválik
 - a szerveren külön szoftver kell:
 - először lekérdezi az adat fizikai helyét a VM-től
 - majd közvetlenül eléri az adatot
 - gyorsabb az adatátvitel, mert nincs az adatútban semmi



Blokkszintű virtualizáció

Ezek az erőforrások ...

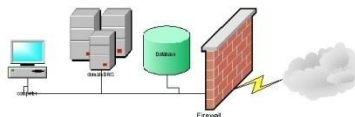
- LUN (különböző gyártók, különböző típusai)
- Tipikusan fix méretű
- Különböző szállítói felületek konfiguráció, és copy szolgáltatások
- Nehéz a migráció az új technológiákra

ilyennek tűnnek ...

- Virtuális LUN-ok, amelyek úgy tűnnek, mintha egy gyártó egy típusa lennének
- A mérete dinamikusan csökkenthető és növelhető
- Központi menedzsment, egységes felületek és szolgáltatások
- Migráció az alkalmazások zavarása nélkül

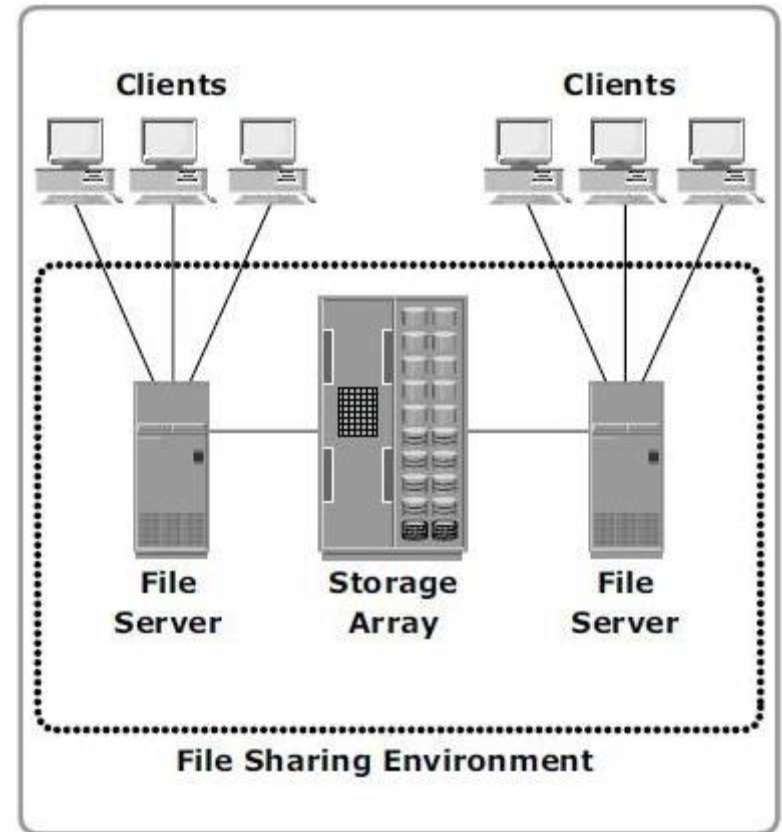
Virtualizáció szintjei

- Blokkszintű virtualizáció
 - eddig erről volt szó
 - a *szerver* egy adott *adatblokk*hoz akar hozzáférni
 - aminek ismeri a (logikai) címét
- Fájl szintű virtualizáció
 - egy *kliens*/host egy *fájlt* szeretne elérni egy adott fájlserveren
 - tudnia kell, hogy melyiken



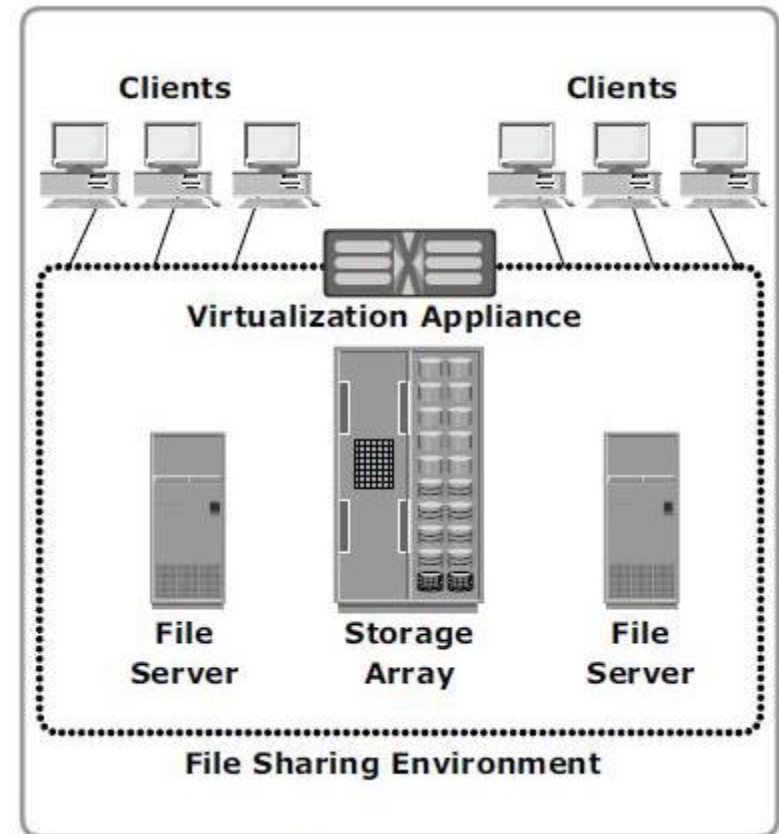
Fájlszintű virtualizáció

- Fájlszintű virtualizáció nélkül
 - egy kliens/host egy fájlt szeretne elérni egy adott fájlszerveren
 - tudnia kell, hogy melyiken
 - lehet, hogy az egyik szerver tele, a másik üres
 - fájl mozgatás érinti a klienst is



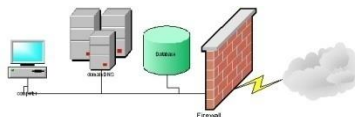
Fájlszintű virtualizáció

- Virtualizált fájlserver
 - a kliensnek nem kell tudni, melyik fizikai serveren van a fájl
 - egyszerűbb
 - terhelésmegosztás
 - fájlmozgatás
 - bővítés
- Cloud computing



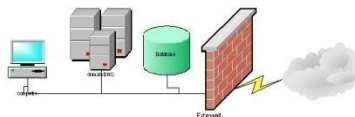
Adatátvitel a felhőbe

- Idő...
- Adathordozón
- 100 petabájt
 - Filmarchívum
 - NASA archívum
 - 2000 év mp3
 - 200x az emberiség génállománya
 - Akár évekig tart a feltöltés neten keresztül

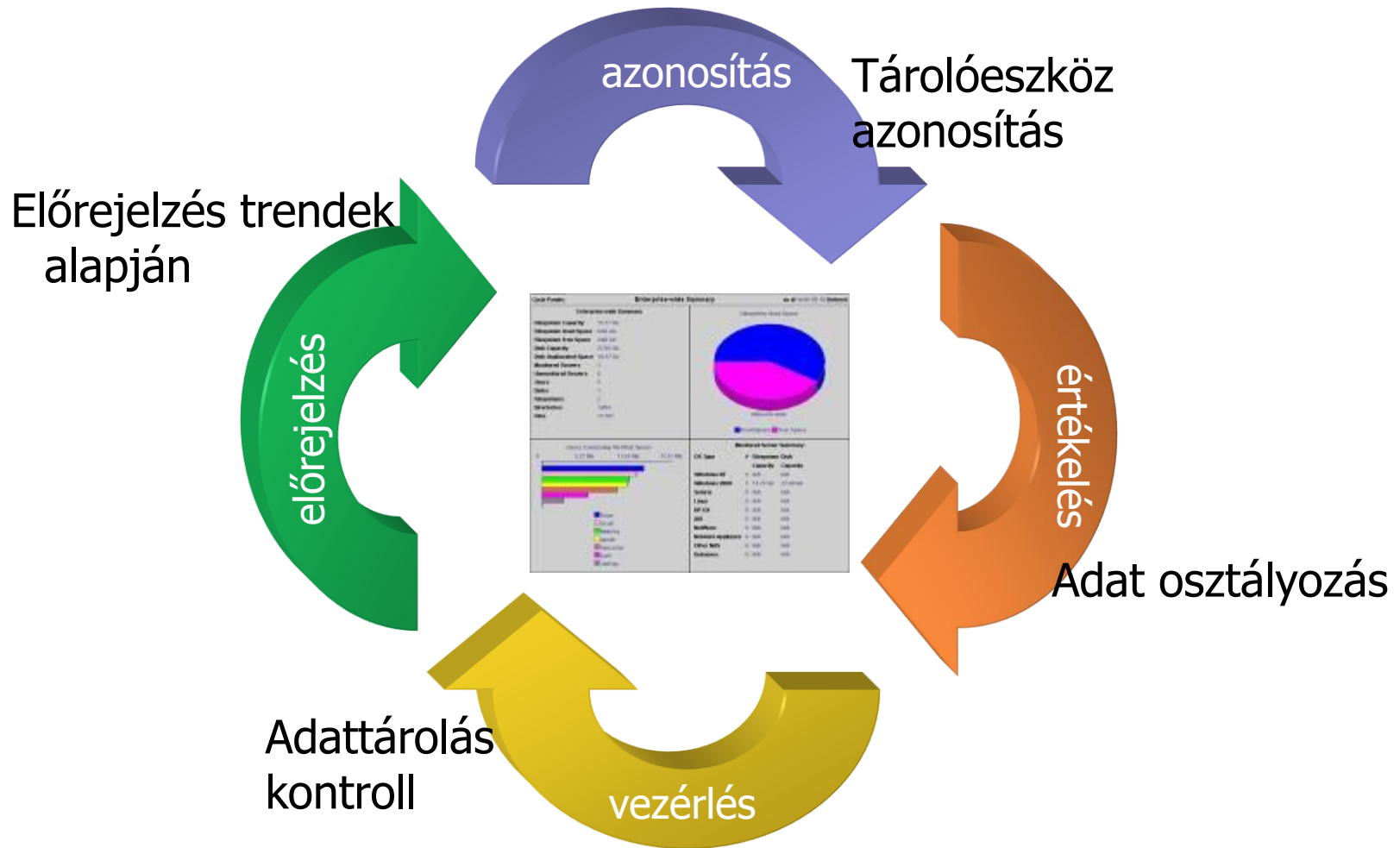


Adattároló hálózatok menedzsmentje

- Néhány kérdés, amit nem egyszerű megválaszolni:
 - Milyen az aktuális tárterület foglaltság?
 - Mi miatt növekedik legjobban a felhasznált terület?
 - Hogy lehet a tárterület növekedést prognosztizálni?
 - A tárolt adatok hányadrésze értéktelen?
 - Mely rendszereket kell először migrálni?
 - Tárterület, tároló eszköz leltár hogyan valósítható meg?
 - A leállások közül mi vezethető vissza tárolási okokra?
 - Tárterület használati szabályok hogyan definiálhatóak, és hogyan tartathatóak be?

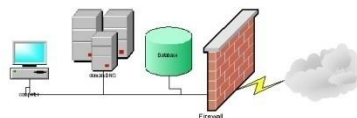
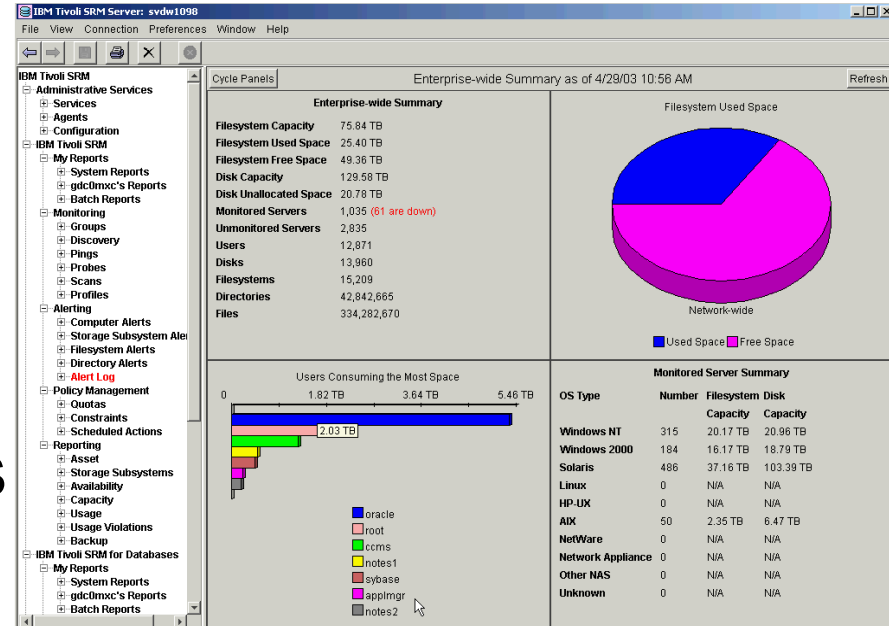


A tároló erőforrás menedzsment főbb lépései (Storage Resource Management)



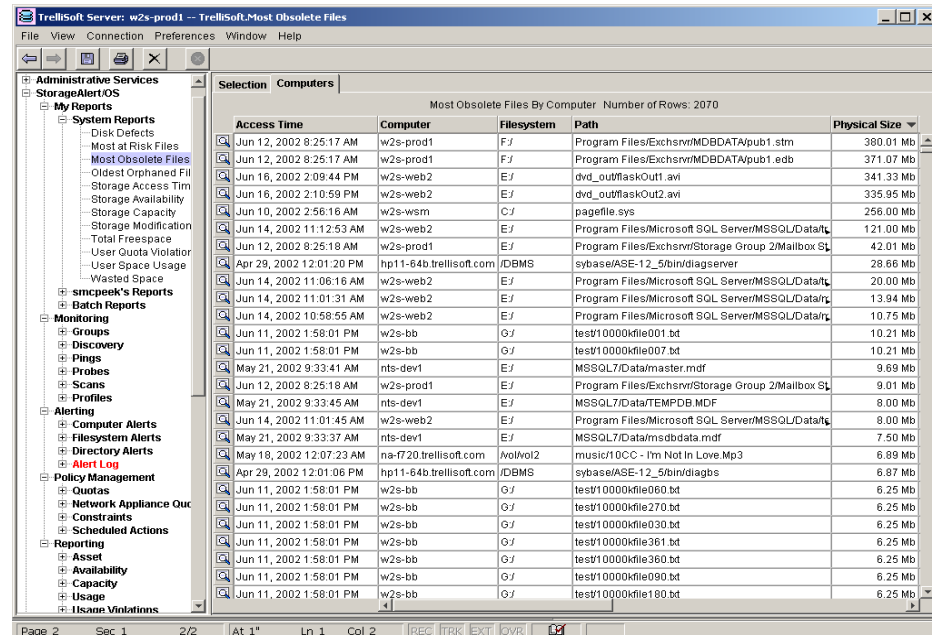
SRM megoldás (1)

- Azonosítás
 - Eszközleltár
 - Lefoglalt területek
 - Mi az aktuális kapacitás kihasználás
 - Van-e kritikus fájlrendszer?
 - Allokált, de nem használt területek



SRM megoldás (2)

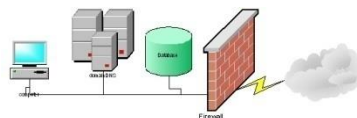
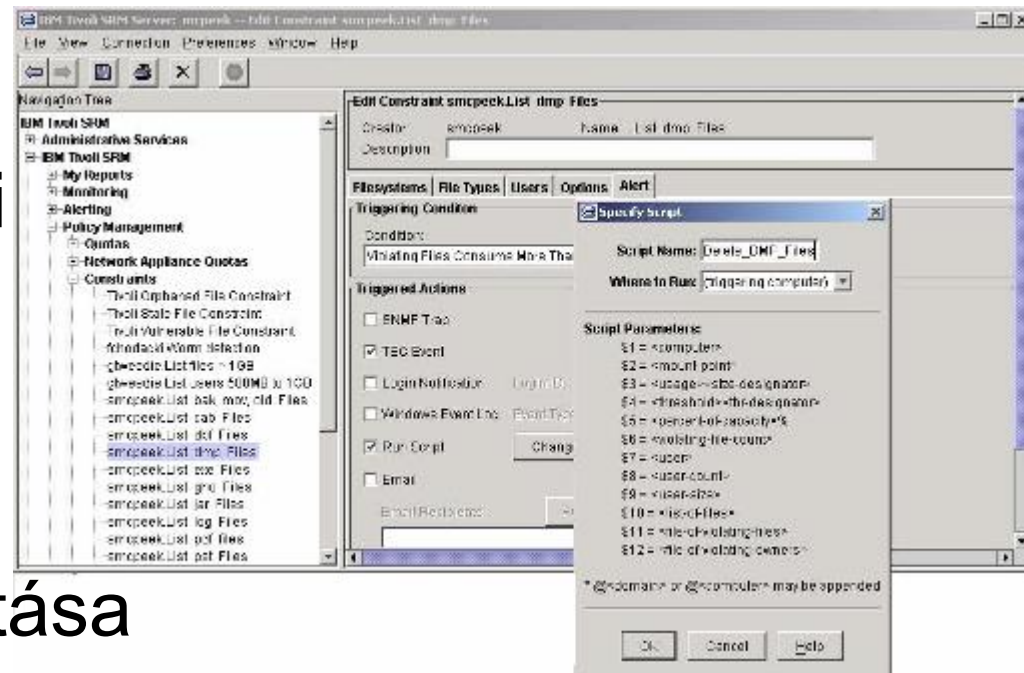
- Értékelés (adat osztályozás)
 - Fájl és könyvtár szintű analízis
 - Feleslegesen foglalt tárterületek azonosítása
 - Árva, régi, nem használt, duplikált állományok azonosítása



Access Time	Computer	Filesystem	Path	Physical Size
Jun 12, 2002 8:25:17 AM	w2s-prod1	F:\	Program Files\Exchsrvr\MDBDATA\pub1.stm	360.01 Mb
Jun 12, 2002 8:25:17 AM	w2s-prod1	F:\	Program Files\Exchsrvr\MDBDATA\pub1.edb	371.07 Mb
Jun 16, 2002 2:09:44 PM	w2s-web2	E:\	dvd_outflaskOut1.avi	341.33 Mb
Jun 16, 2002 2:10:59 PM	w2s-web2	E:\	dvd_outflaskOut2.avi	335.95 Mb
Jun 10, 2002 2:58:16 AM	w2s-wsm	C:\	pagefile.sys	256.00 Mb
Jun 14, 2002 11:12:53 AM	w2s-web2	E:\	Program Files\Microsoft SQL Server\MSSQL\Data\...	121.00 Mb
Jun 12, 2002 8:25:18 AM	w2s-prod1	E:\	Program Files\Exchsrvr\Storage Group 2\Mailbox SL...	42.01 Mb
Apr 29, 2002 12:01:20 PM	hp11-64b.trellissoft.com	/DBMS	sybase\ASE-12_5\bin\diagserver	28.66 Mb
Jun 14, 2002 11:06:16 AM	w2s-web2	E:\	Program Files\Microsoft SQL Server\MSSQL\Data\...	20.00 Mb
Jun 14, 2002 11:01:31 AM	w2s-web2	E:\	Program Files\Microsoft SQL Server\MSSQL>Data\...	13.94 Mb
Jun 14, 2002 10:58:55 AM	w2s-web2	E:\	Program Files\Microsoft SQL Server\MSSQL>Data\...	10.75 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile001.bt	10.21 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile007.bt	10.21 Mb
May 21, 2002 9:33:41 AM	nts-dev1	E:\	MSSQL7\Data\master.mdf	9.69 Mb
Jun 12, 2002 8:25:18 AM	w2s-prod1	E:\	Program Files\Exchsrvr\Storage Group 2\Mailbox SL...	9.01 Mb
May 21, 2002 9:33:45 AM	nts-dev1	E:\	MSSQL7\Data\TEMPDB.MDF	8.00 Mb
Jun 14, 2002 11:01:45 AM	w2s-web2	E:\	Program Files\Microsoft SQL Server\MSSQL>Data\...	8.00 Mb
May 21, 2002 9:33:37 AM	nts-dev1	E:\	MSSQL7\Data\msdbdata.mdf	7.50 Mb
May 18, 2002 12:07:23 AM	na-f720.trellissoft.com	/vol1vol2	music\10CC - I'm Not In Love.Mp3	6.89 Mb
Apr 29, 2002 12:01:06 PM	hp11-64b.trellissoft.com	/DBMS	sybase\ASE-12_5\bin\diagbs	6.87 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile060.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile070.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile030.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile361.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile360.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile090.bt	6.25 Mb
Jun 11, 2002 1:58:01 PM	w2s-bb	G:\	test\10000kfile180.bt	6.25 Mb

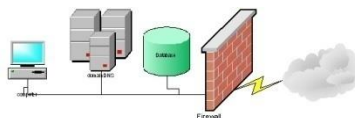
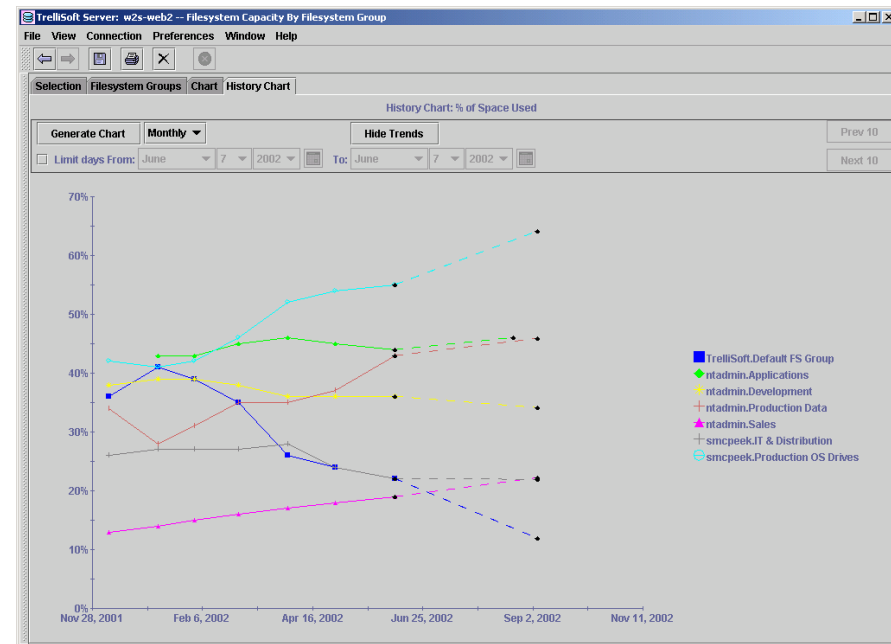
SRM megoldás (3)

- Vezérlés
 - Központi riasztási rendszer
 - Kvótakezelés
 - Automatikus válaszakciók indítása
 - A menedzsment funkciók automatizálhatóak



SRM megoldás (4)

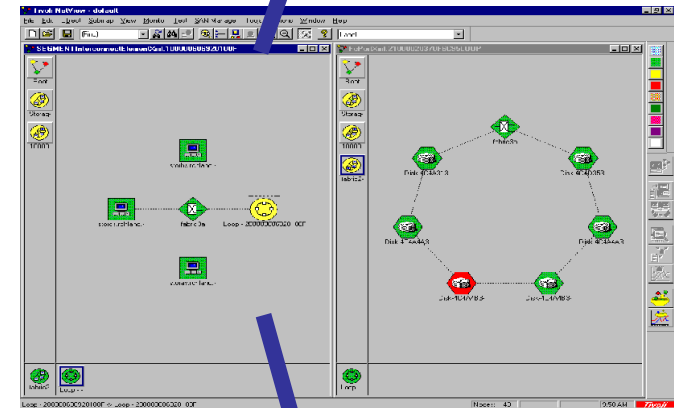
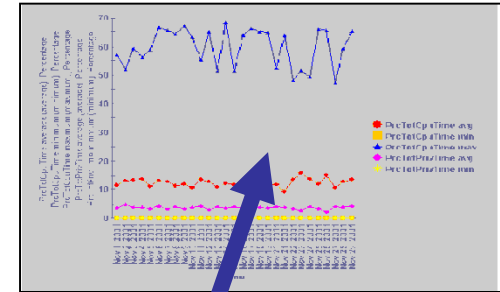
- Előrejelzés
 - Leggyorsabban növekvő felhasználók, fájlrendszerek, adatbázis táblák.
 - Trendek azonosítása, és előrejelzés



SAN menedzsment eszköz

6 fő funkció

- Topológia térkép (redundancia)
- Monitoring
- Központi esemény-felügyelet és riasztás
- Központi alkalmazás indítás
- Hibadetektálás és előrejelzés, kiküszöbölés
- Zóna kontroll



Time	Source	Event	Severity	Status	Message
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001
2008-11-10 10:00:00	TS0_0001	TS0_0001	Warning	Open	TS0_0001

