

Tárolókezelés szoftver- rétegből

IBM Spectrum Storage™ Software Defined Storage

Grósz Attila (attila_grosz@hu.ibm.com)



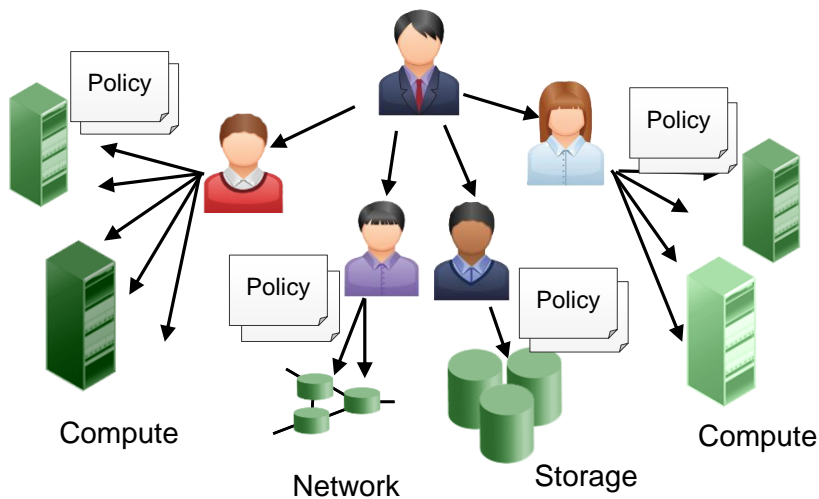
Google, Yahoo, Softlayer,... hogyan csinálják?



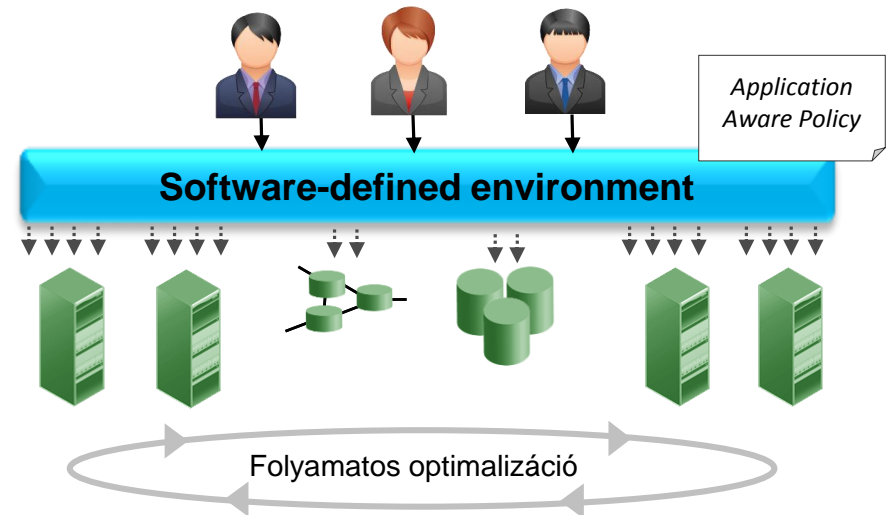
Miképp üzemelnek ezek az adatközpontok..?



**Alacsony standardizáció,
magas üzemeltetési költség
(OpEx)**



**Magas standardizáció,
alacsony üzemeltetési költség
(OpEx)**

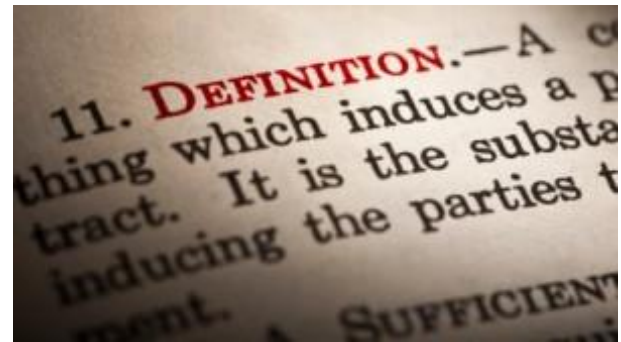


Nincsenek dedikált eszközök, nincs külön hálózat,
nincsenek „hw-offload” megoldások stb.

Szoftver-alapú tárolómegoldás



- *Szoftver-alapú tárolómegoldás bármely adattárolást biztosító szoftver csomag, amely telepíthető standard, általánosan elterjedt (x86 alapú szerver, virtualizációs réteg vagy felhő) és/vagy szabadforgalmú hardver eszközre. Teljes értékű tárolószolgáltatást és kapcsolódást nyújt a menedzselt fizikai adattároló erőforrások közt, lehetővé téve az adatok szabad mozgását a felhasználók számára.*
- *Software-defined storage is any storage software stack that can be installed on any commodity (x86 hardware, hypervisors, or cloud) and/or off-the-shelf computing hardware and used to offer a full suite of storage services and federation between the underlying persistent data placement resources to enable data mobility of its tenants between these resources*



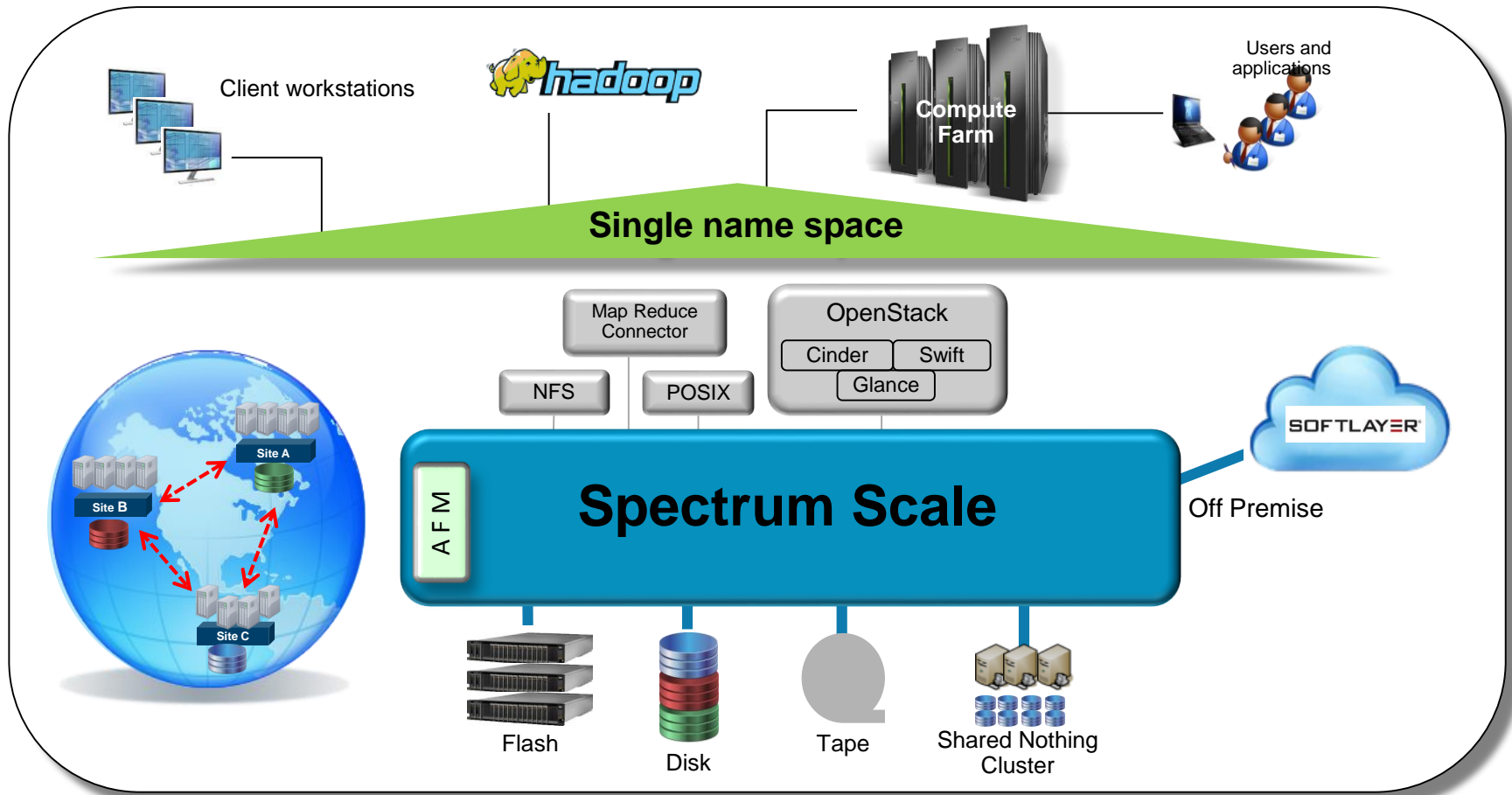
IBM Spectrum Scale



IBM Spectrum Scale - SDS



Skálázható, nagy-terjesztményű adat- és fájlmenedzsment megoldás IBM GPFS alapokon.



IBM Spectrum Scale



Skálázhatóság

Max. fájlrendszer méret:

2^{99} Bytes

(2^{63} files per file system)

Max. file méret egyenlő a fájlrendszer méretével

Meglévő ügyfelek 18 PB fájlrendszerrel

IPv6

„virtually no capacity limits”

Teljesítmény

Párhuzamos fájl hozzáférés

Disztributált, skálázható, nagy teljesítményű meta-adat

Flash gyorsítás (LROC Linux)

Automatikus ‚tier’ -elés (ILM)

Több mint 400 GB/s perf.

Távoli telephelyek közti aszinkron gyorsítás (AFM)

> 400GB/s

Megbízhatóság

Snapshot, replikáció

Beépített ‚heartbeat’, automatikus failover/failback

Üzem közben szerverek és diszkek be/ki konfigurációja

‚Rolling upgrades’

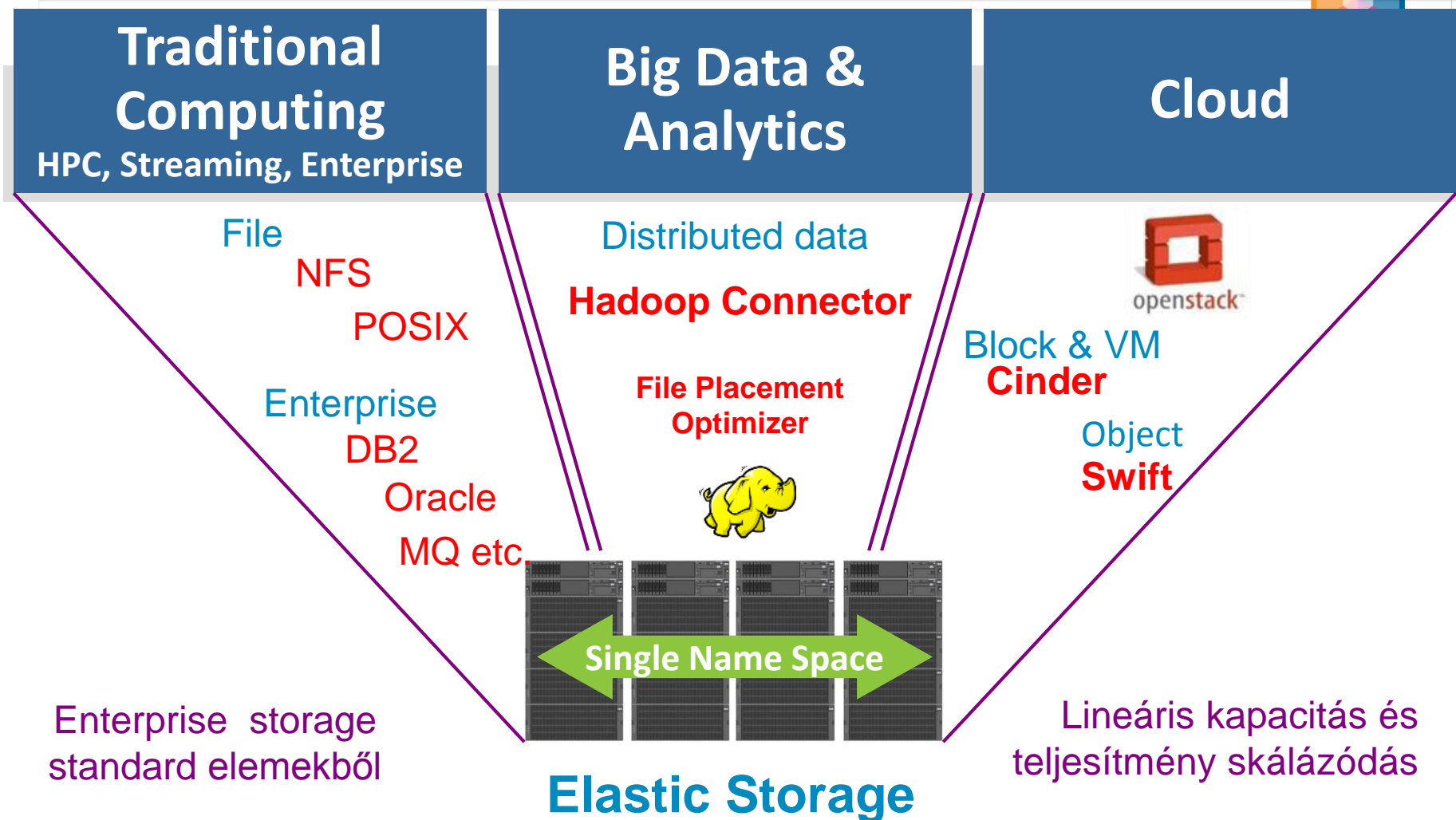
Adminisztráció bármely node - ról

Standard / disztributált mentés

15 év IBM fejlesztés



Spectrum Scale felhasználási területek



Szoftver-alapú tárolómegoldás az összes felhasználási terület számára

„Személyre szabott” megoldások



Integrált Megoldásként *IBM's Elastic Storage Server (ESS)*

Spectrum Scale Software

On Premise Infrastructure

Csupán a Szoftvert

*IBM Spectrum Scale Software, Licenc típusok –
Express, Standard, Advanced*

Platform LSF (SaaS)

Platform Symphony
(SaaS)

Spectrum Scale on Cloud

SoftLayer bare metal infrastructure

24X7 CloudOps Support

Cloud Service

Ready to use, Spectrum Scale on the Cloud

Vagy IBM SoftLayer Cloud

IBM Spectrum Scale a Felhőben

Elastic Storage Server család



- Skálázható kapacitás és teljesítmény, valódi SDS
 - nincs hagyományos storage controller
 - Mix ESS, disk, Flash, SSD, 3rd party storage stb.

- Modellek:

- **GL = High Capacity**

- Analytics, Cloud Serving, Technical, Media
 - 60 disk drawers: 2TB, 4TB disks

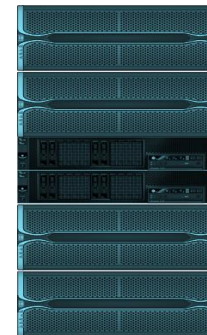
- **GS = High IOPS**

- **Hot** data and/or Metadata
 - 24 slot disk/SSD drawers: SSD (400/800), SAS 1.2TB

- Kapcsolódás: 10GbE, 40 GbE, Infiniband



Model GL2
Analytics Focused
2 Enclosures, 12U
116 NL-SAS, 2 SSD
5+ GB/Sec



Model GL4
Analytics and Cloud
4 Enclosures, 20U
232 NL-SAS, 2 SSD
10+ GB/Sec



Model GL6
PetaScale Storage
6 Enclosures, 28U
348 NL-SAS, 2 SSD
12+ GB/sec



Model GS1
24 SSD
6 GB/Sec



Model GS2
46 SAS + 2 SSD or
48 SSD Drives
2 GB/Sec SAS
12 GB/Sec SSD



Model GS4
94 SAS + 2 SSD or
96 SSD Drives
5 GB/Sec SAS
16 GB/Sec SSD



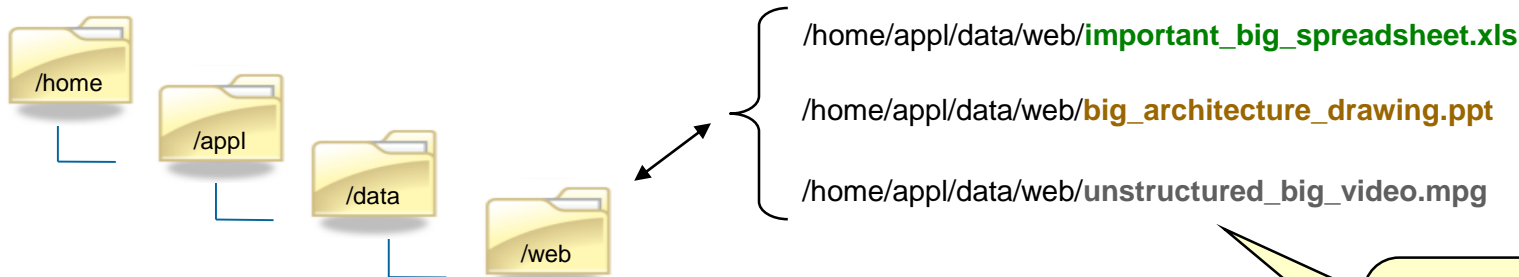
Model GS6
142 SAS + 2 SSD
7 GB/Sec

IBM Spectrum Scale

FUNKCIONALITÁS



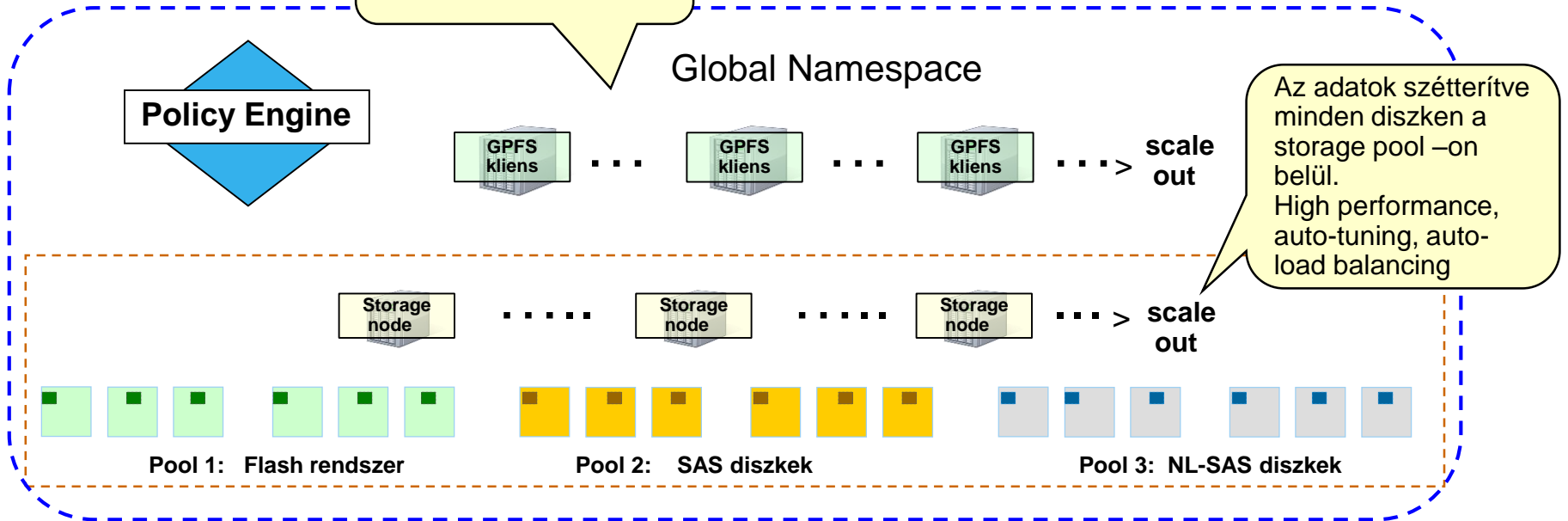
Spectrum Scale – elosztott működés



Logical

Az IO terhelés automatikusan eloszlik a node-ok közt.

Mindhárom file ugyanabban a könyvtárban, viszont más-más storage-pool –ba allokálva

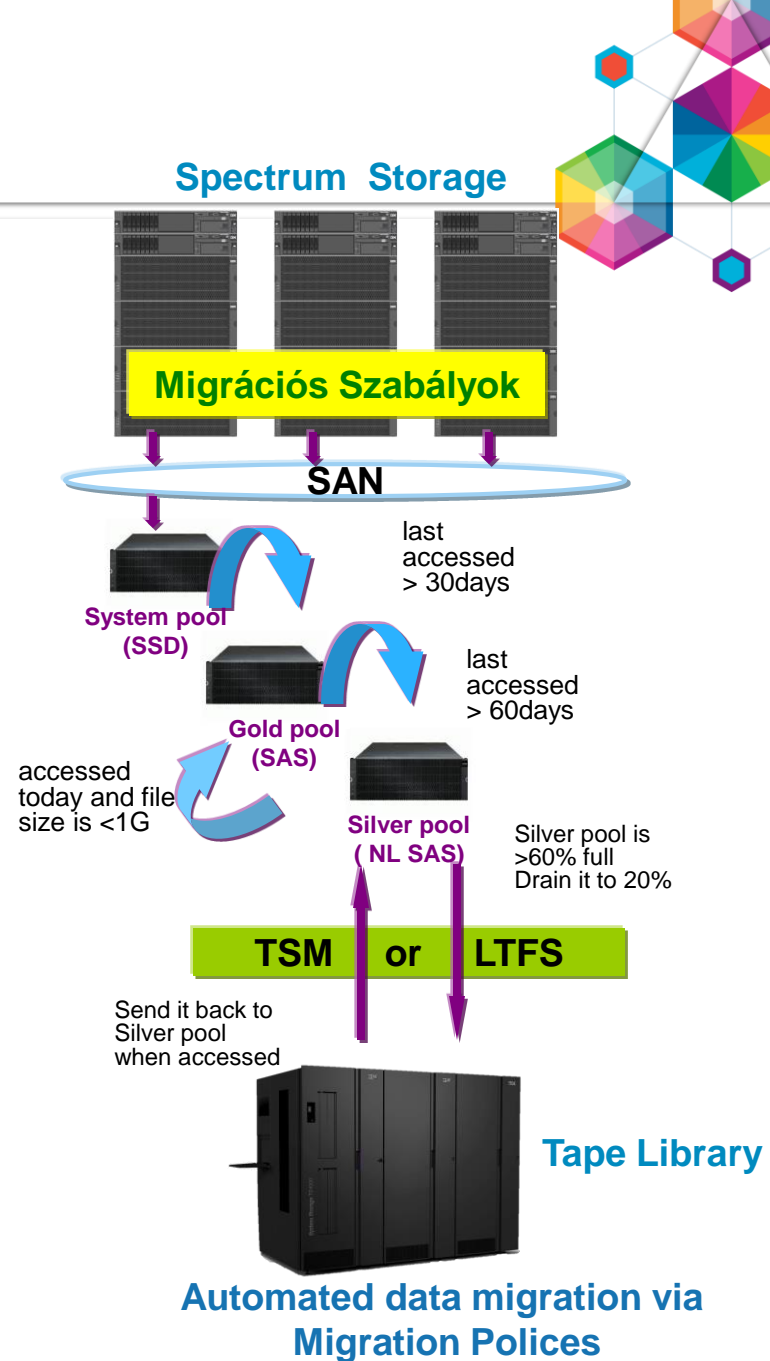


Az adatok szétterítve minden diszken a storage pool –on belül. High performance, auto-tuning, auto-load balancing

Physical

Szabály-alapú ILM és HSM

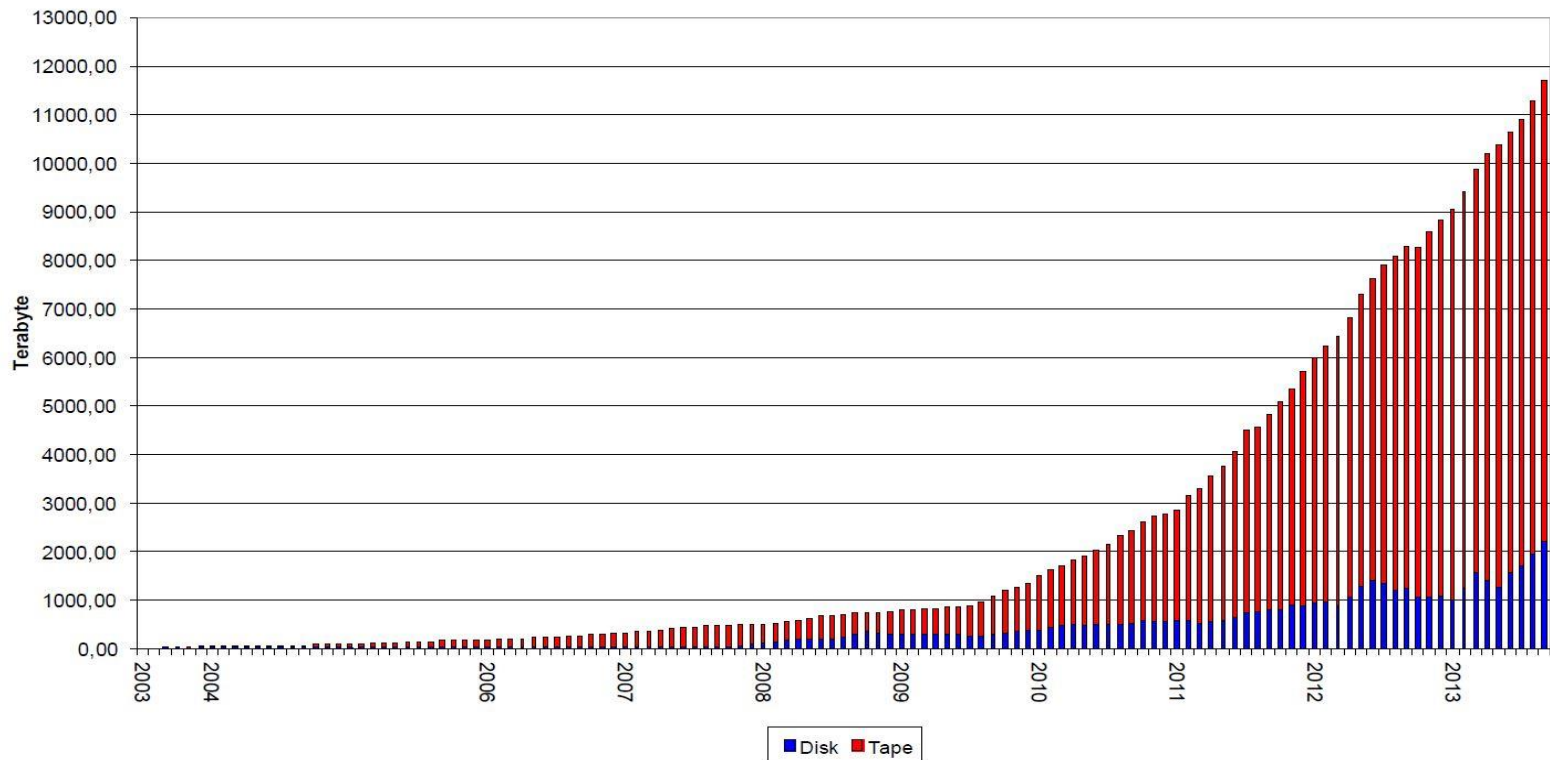
- Nagy fájlrendszerek egyszerű menedzsmentje szabály alapú, automatikus adat-életciklus kezeléssel
- Fájlok monitorozása és mozgatása küszöbérték alapján külső szalagos tárolóra
 - Példa: Online tároló 90% -os telítettséget ér el, ekkor minden 1GB -nál nagyobb és 90 napnál öregebb fájl szalagra kerül
- Adatmigráció ütemezése produkciós időn kívülre
- Egyszerű adat-menedzsment
 - „Policy engine” fájl-listát készít az adminisztrátor által beállított szabályok alapján.
 - Ezen a listán végrehajthatóak a szükséges migrációs lépések
- SQL szintakszisú script



Adatmennyiségek a Jülich –i HPC központban



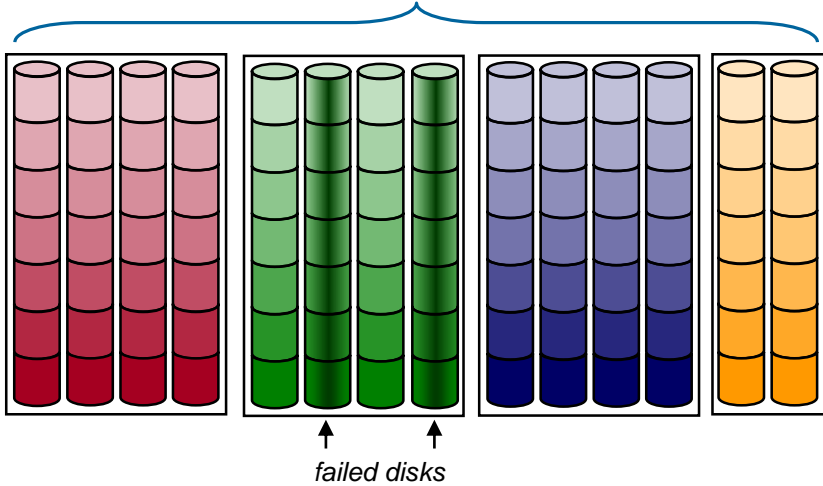
Data Growth at JSC – User Data GPFS



De-clustered RAID = gyors adat-visszaállítás (több diszk kiesése esetén is)

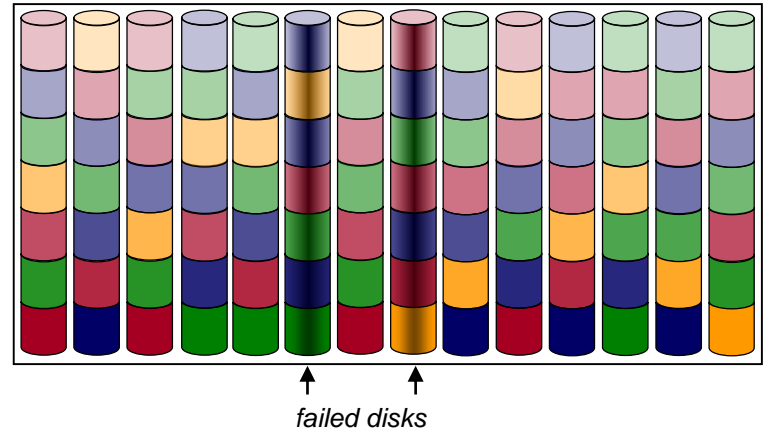


14 physical disks / 3 traditional RAID6 arrays / 2 spares



14 physical disks / 1 declustered RAID6 array / 2 spares

Decuster
data,
parity
and
spare



failed disks

| Number of faults per stripe | | |
|-----------------------------|-------|------|
| Red | Green | Blue |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |

Number of stripes with 2 faults = 7

failed disks

| Number of faults per stripe | | |
|-----------------------------|-------|------|
| Red | Green | Blue |
| 1 | 0 | 1 |
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 2 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 0 | 1 | 0 |

Number of stripes with 2 faults = 1

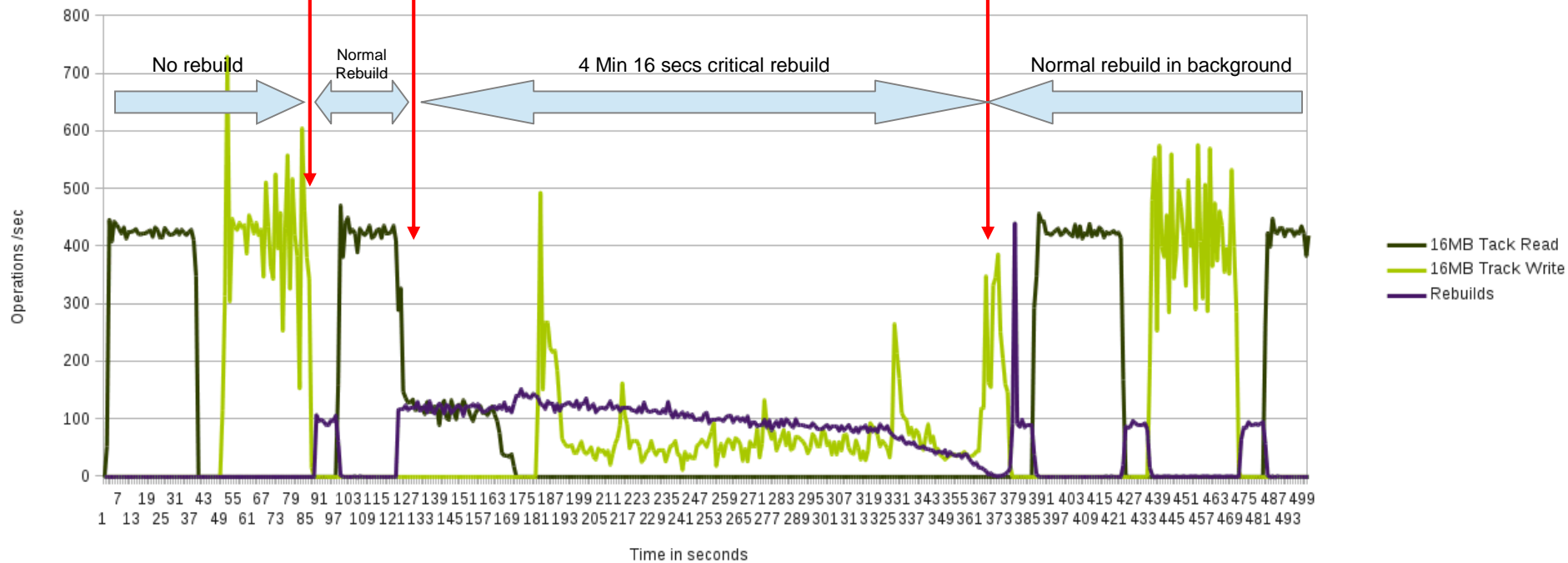
ESS 8+3P Rebuild Teszt



1st and 2nd disk failures

3rd disk failure - start of critical rebuild

Critical rebuild finished, continue normal rebuild



Test 8+3p on a GSS26 with 2TB NL-SAS, Read-Write-Read pattern

A kritikus szakaszon (a 3. diszk hiba után) a lehető legrövidebb idő alatt lefut a visszaépítés.

Amint a paritás-védelem biztosított a terhelés visszaáll kevesebb mint 5% -ra.

IBM Spectrum Scale natív RAID



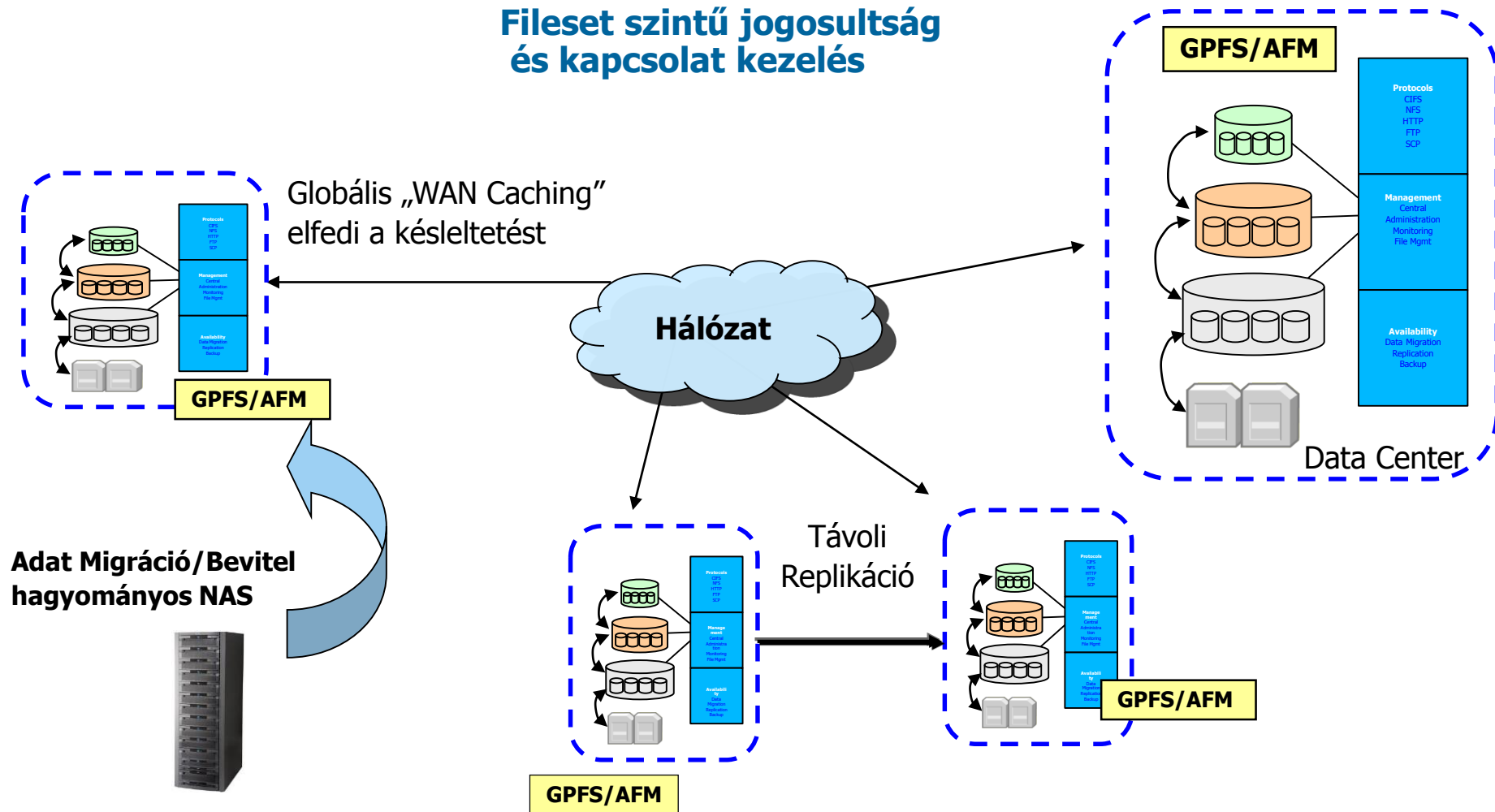
- **Elastic Storage Server - GPFS natív RAID (De-clustered RAID)**
 - Adat, paritás és spare területek egységesen partícionálva és disztributálva kerülnek elhelyezésre a tömbön
 - Rebuilds that take days on other systems, take minutes on Elastic Storage
- **Védelem kettős vagy hármás diszk hiba esetén is**
 - Reed-Solomon parity encoding; 2-fault or 3-fault tolerant
 - 3 –as vagy akár 4-es tükrözés
- **„End-to-end’ checksum & dropped write detection**
 - Adatellenőrzés a disztől a kliensig
 - Detektálja és javítja az ún. „off-track’ és „lost / dropped’ diszk írásokat
- **Aszinkron hiba-diagnózis az I/O folyamat folytatásával**
 - If media error: verify and restore if possible
 - If path problem: attempt alternate paths
- **Diszk csere támogatása üzem közben**

Active File Management (AFM) – Globális Adatkezelés

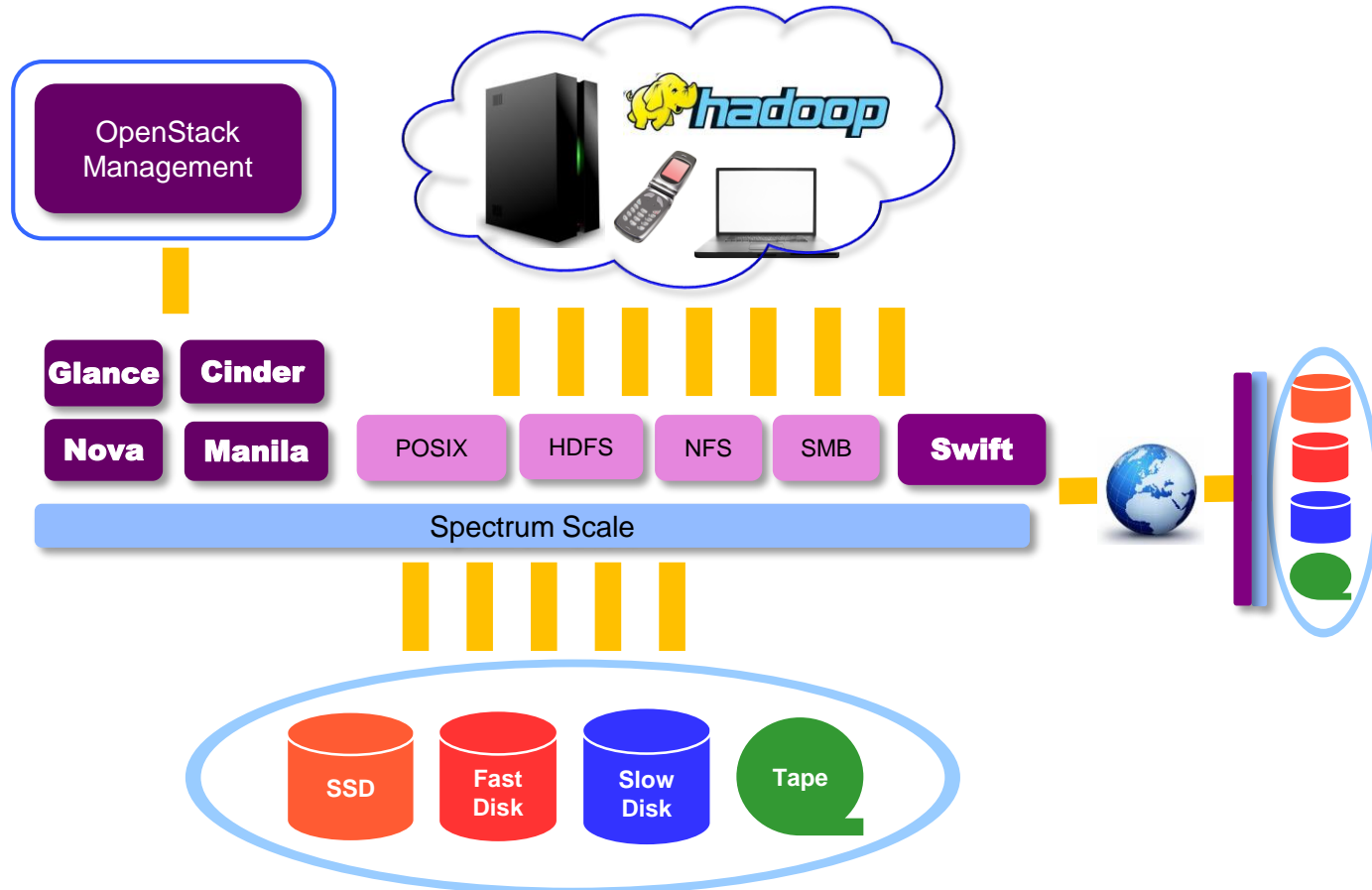


Globális névtértartomány,
nem csak „egy” névtér

Fileset szintű jogosultság és kapcsolat kezelés



IBM Spectrum Scale és OpenStack

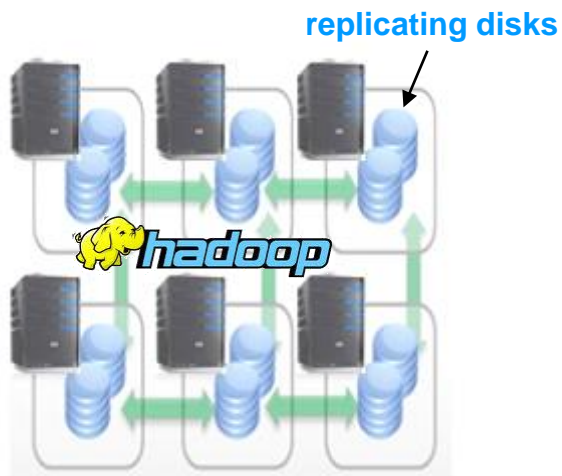


Unified Storage – Spectrum Scale & Openstack
Block (Cinder), File (Manila) és Object (swift)

Hadoop & IBM Spectrum Scale Storage

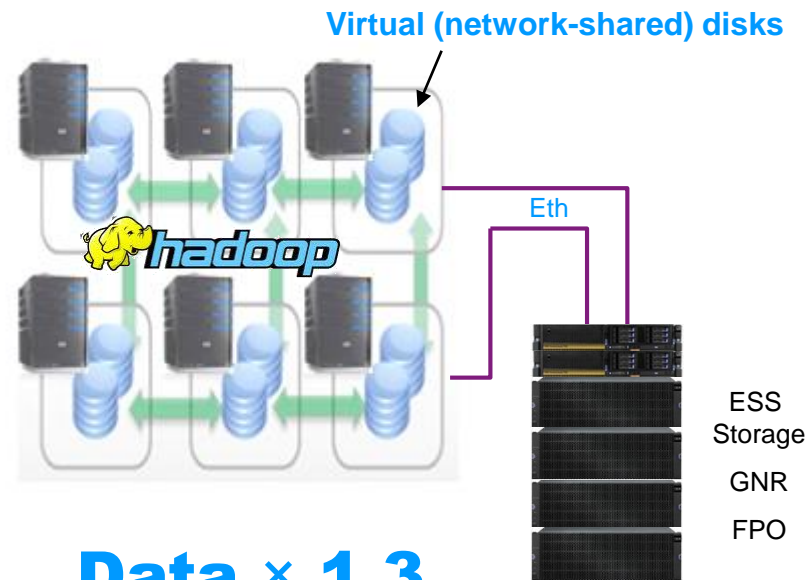


"Customers who bought this were also interested in..."



Data × 3

Klasszikus HDFS-alapú



Data × 1.3

IBM Spectrum Scale

IBM ESS vs. HDFS



| | | IBM ESS | HDFS |
|-----------------------|--------------------------|---------|------|
| Performance | Terasort: large reads | ✓ | ✓ |
| | Hbase: small write | ✓ | ✓ |
| | Metadata intensive | ✓ | ✓ |
| Enterprise readiness | POSIX compliance | ✓ | |
| | Meta-data replication | ✓ | |
| Protection & Recovery | Snapshot | ✓ | |
| | Asynchronous Replication | ✓ | |
| | Backup | ✓ | |
| Security & Integrity | Access Control Lists | ✓ | |
| Ease of Use | Policy-based Ingest | ✓ | |

Az SDS előnyei – IBM Spectrum Storage



- **Alacsony költségű** standard szerverekből, tárolóelemekből épül fel drága speciális céleszközök helyett
- **Magasfokú skálazhatóság és teljesítmény** jellemzi a lineárisan bővíthető építőelemek által (scale-out)
- **Nő az erőforrások kihasználtsága** az egyes izolált egységek csoportosíthatósága és optimalizációja miatt
- **Rugalmasságot és dinamizmust** teremt, lehetővé téve a megváltozott igényekre történő gyors reakciót az erőforrások konfigurálhatósága ill. újrallokációja segítségével.
- **Csökken az üzemeltetési költség** az automatizáció és a szabály-alapú tároló menedzsment miatt.



Kérdések..?



Köszönöm!



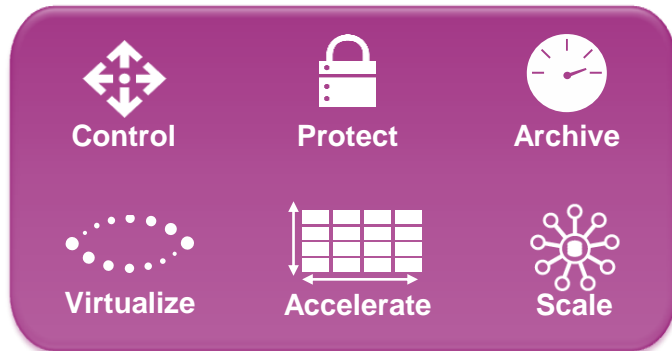
Backup slides



IBM Szoftver-alapú tárolási megoldások



Tároló Menedzsment és Optimizációs Szoftver Család



Any Storage



Flash Systems



Private, Public
or Hybrid Cloud

| | |
|---------------------|--|
| Spectrum Control | Analytics-driven data management to reduce costs by up to 50 percent |
| Spectrum Protect | Optimized data protection to reduce backup costs by up to 38 percent |
| Spectrum Archive | Fast data retention that reduces TCO for archive data by up to 90% |
| Spectrum Virtualize | Virtualization of mixed environments stores up to 5x more data |
| Spectrum Accelerate | Enterprise storage for cloud deployed in minutes instead of months |
| Spectrum Scale | High-performance, highly scalable storage for unstructured data |

Declustered RAID Example

