

CURSUL 4

STOCARE ȘI VIRTUALIZAREA STOCĂRII ÎN SISTEME INFORMATICE

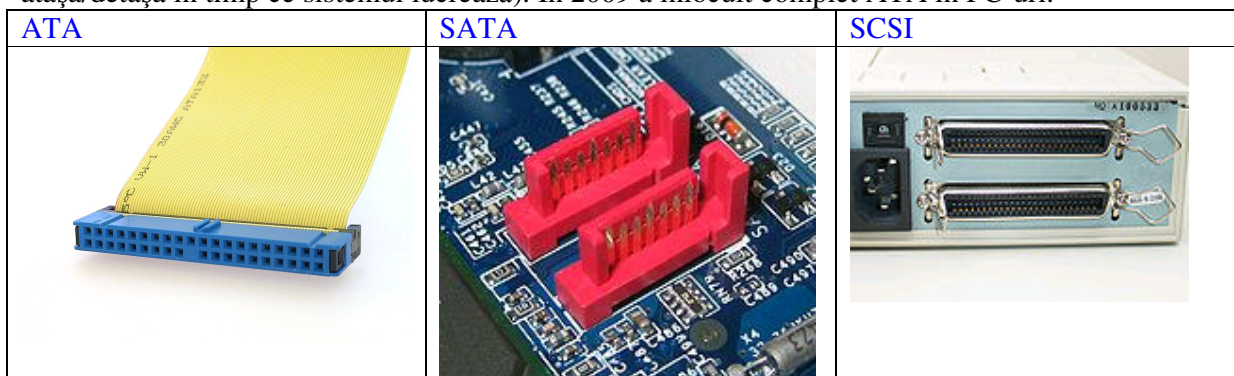
4. 1. DAS - DIRECT-ATTACHED STORAGE

Direct-attached storage (DAS) reprezintă „stocare atașată direct” și se referă la sistemul de stocare digital atașat direct la un server sau stație, [fără ca o rețea de stocare să fie interpusă](#).

Protocoalele principale utilizate în DAS sunt:

a) ATA

b) SATA (Serial ATA) . Oferă câteva avantaje asupra interfețelor ATA: 8 pini (în loc de 40), transferuri de date mai rapide și mai eficiente, capacitate „hot swapping” (dispozitivele se pot atașa/detașa în timp ce sistemul lucrează). În 2009 a înlocuit complet ATA în PC-uri.



c) SCSI

SCSI = set de standarde pentru conectarea fizică și transferul de date între calculatoare și dispozitive periferice. Standardele SCSI definesc comenzi, protocoale și interfețe electrice și optice. SCSI se folosește uzual pentru HDD și drivere de disc, dar poate fi conectat la o gamă mare de alte dispozitive, incluzând scanere și drivere CD. [Standardul SCSI definește seturi de comenzi pentru anumite tipuri de dispozitive periferice](#).

SCSI = interfață inteligentă, „ascunzând” complexitatea formatului fizic. Modalitățile de atașare a dispozitivelor la magistrala SCSI sunt similare.

SCSI = interfață periferică de tip peer to peer. **Se pot atașa max 8 respectiv 16 dispozitive la o singură magistrală**. Numărul de host/dispozitive periferice poate varia dar trebuie să existe minim un host.

Protocolul SCSI definește comunicarea (host<->host), (host<->dispozitiv periferic), (dispozitiv periferic<->dispozitiv periferic).

Majoritatea dispozitivelor periferice sunt exclusiv „SCSI targets” (ținte) neputând să se comporte ca „SCSI initiators” (nu pot iniția tranzacții către ele însele).

În contextul stocării, un **inițiator SCSI**=punctul final (endpoint) care inițiază o sesiune SCSI – adică trimite o comandă SCSI. Inițiatorul uzual nu furnizează niciun număr logic de unitate (Logical Unit Number LUN).

Un „**SCSI target**” = punctul final (endpoint) care nu inițiază o sesiune SCSI, în schimb așteaptă comenzile inițiatorului și furnizează acestuia transferurile de date I/O solicitate. Uzual target-ul furnizează inițiatorilor unul sau mai multe LUN-uri.

Uzual un calculator este inițiatorul și un DSD (**Data Storage Device** = dispozitiv de stocare date) este target. Făcând analogie cu modelul client-server, inițiatorul este clientul și targetul este serverul. Fiecare **adresă SCSI** (adică fiecare identificator de pe o magistrală SCSI) manifestă o comportare de inițiator, de target sau (rar) de ambele tipuri simultan.

d) **SAS - Serial Attached SCSI**, tehnologie pentru transferul datelor de la dispozitive de stocare ca HDD și drivere de bandă. Depinde de un protocol serial point-to-point care înlocuiește tehnologia de magistrală paralelă SCSI, folosind setul standard de comenzi SCSI. În 2009 este estimată dublarea vitezei de operare oferită de această tehnologie (la 6 Gbit/s).

Un **sistem SAS tipic** constă din următoarele **componente de bază**:

- un **inițiator** = cel care lansează **cererile de tip „device-service” și respectiv „task-management”**

care urmează a fi procesate de un disp. target și primește răspunsurile de la target-uri. Inițiatorii pot fi furnizați ca și o componentă „on-board” de pe motherboard (cazul multor motherboard-uri de tip „server-oriented”) sau ca un host bus adapter (HBA) add-on.

- un **target** = dispozitiv care conține unități logice și porturi target care recepționează cereri pentru servicii de dispozitiv și respectiv cereri de tip „task management” care urmează a fi procesate, răspunsurile la cereri urmând să se trimită inițiatorului lor. Un target poate fi un HD sau un sistem de tip „disk array” (Fig. 1).

- un subsistem de tip „**Service Delivery Subsystem**” (subsistem de furnizare a serviciilor) = partea dintr-un sistem I/O care transmite informații între un inițiator și un target. Este format în mod tipic din cablurile care realizează conexiunea dintre un inițiator și un target cu sau fără „expander”-i și „backplane”-uri .

- **expander**-i = dispozitive folosite pentru a expanda magistrala de date pentru portul de comunicație. Facilitează conexiunea mai multor dispozitive de tip „SAS End” la un singur port inițiator.



Hewlett-Packard Disk-Arrays:
HASS (sus) și NIKE

Fig.1. Disk-array

(**Backplane** – fig.2) . Backplane-urile se utilizează uzual pentru conectarea mai multor HDD la un singur controller de disk-array și se regăsesc uzual în „disk-array”-uri, servere și „disk enclosure”-uri (carcase cu mai multe discuri).

Un **domeniu SAS** reprezintă un sistem I/O constând dintr-un set de dispozitive SAS care comunică între ele prin intermediul unui subsistem de furnizare a serviciilor. Fiecare dispozitiv SAS dintr-un domeniu SAS are un **identificator unic la nivel global** (asignat de către producătorul dispozitivului – similar cu adresa MAC), denumit „**World Wide Name**” (WWN) sau adresă SAS.

Un WWN identifică în mod unic dispozitivul în domeniul SAS la fel cum un **SCSI ID** identifică un dispozitiv de pe o magistrală paralelă SCSI.

Un **WWN** identifică în mod unic dispozitivul în domeniul SAS la fel cum un **SCSI ID** identifică un dispozitiv de pe o magistrală paralelă SCSI.

Un **domeniu SAS** poate conține max. 16 256 dispoz.

Comparație SAS vs. SCSI paralel

- SAS operează point-to-point pe când magistrala SCSI este multidrop => avantaj SAS. Fiecare disp. SAS este conectat printr-o legătură dedicată la inițiator (în afara cazului când se folosește un expander). Dacă un inițiator este conectat la un target , nu este posibil să apară „conflicte” , pe când la SCSI paralel, însăși conectarea poate provoca unul.
- La SAS nu se impune utilizarea de „terminator” de magistrală (spre deosebire de SCSI -> fig.3).
- SAS elimină fenomenul de „clock skew” („clock skew”=semnalul de clock ajunge la componente diferite la momente diferite!)
- Pe un singur canal, SAS suportă max. 16 384 dispoz. (dacă se utilizează expanderi) , pe când SCSI suportă maxim 8 sau 16.
- SAS oferă viteze de transfer mai mari (1.5 sau 3.0 Gbit/s , în feb. 2009 este estimată atingerea performanței de 6 Gbit/s) comparativ cu majoritatea standardelor SCSI. Viteza este atinsă pe fiecare conexiune inițiator-target => performanțe mai bune, pt. că un disp. conectat cu SCSI paralel partajează viteza cu toate celelalte disp. de pe magistrală.
- Controllerele SAS pot suporta conectarea de dispozitive SATA, aceste putând fi conectate fie direct folosind protocolul SASA, fie prin expanderi SAS folosind protocolul SATA Tunneled



Fig.2



fig.3. Terminator SCSI

Protocol (STP).

- Și SAS și SCSI paralel folosesc setul de comenzi SCSI.

Comparație SAS vs SATA

- Sistemele identifică dispozitivele SATA după numărul de port prin care sunt conectate la HBA, pe când dispoz. SAS sunt identificate în mod unic prin WWN.
- Protocolul SAS suportă mai mulți inițiatori într-un domeniu SAS, SATA nu .
- Majoritatea driverelor SAS furnizează „tagged command queuing”, pe când cele mai noi drivere SATA furnizează „native command queuing”, fiecare cu avantajele și dezavantajele sale.
(„Tagged command queuing” (TCQ) = tehnologie cu care sunt echipate anumite HD-uri ATA și SCSI. Permite SO-ului să trimită simultan mai multe cereri de citire/scriere către un HDD.
Native Command Queuing (NCQ) = tehnologie proiectată pentru mărirea performanțelor HD-urilor SATA în anumite situații, permițând optimizarea operațiilor interne prin stabilirea ordinii în care se execută operațiile de R/W. Se reduc mișcărilor capetelor discului).
- SATA folosește setul de comenzi ATA => suportă doar HDD și drivere CD/DVD.
Teoretic, SAS suportă numeroase alte dispoz. (inclusiv scanere și imprimante). Avantajul este însă chestionabil , pt. că multe dintre aceste dispoz. folosesc căi alternative prin magistrale ca USB, IEEE 1394 (FireWire) și Ethernet.
- Hardware-ul SAS permite „multipath I/O” pentru dispozitive, pe când SATA anterior lui SATA II nu. SATA II utilizează „port multipliers” pentru a obține expandarea porturilor, a.î. anumiți producători de „port multiplier”-i au implementat „multipath I/O” folosind hardware-ul de la port multiplier-i.
- SATA a fost desemnat pe piață ca succesori de uz general al parallel ATA și se utilizează intensiv, pe când SAS-ul, mai scump, se utilizează în special în aplicații „server critice”.
- SAS folosește comenzi SCSI pentru „error-recovery and – reporting” (restabilirea în urma erorilor și raportarea acestora) , aceste comenzi având o funcționalitate mai mare decât comenzile SMART de la ATA folosite de driverele SATA.
- SAS folosește tensiuni de semnal mai mare decât SATA => crește potențialul lor de utilizare în backplane-uri din servere. SAS poate folosi cabluri de până la 8 m, SATA de max. 1 m.

Detalii tehnice

Standardul SAS definește **câteva straturi** (fig.4): Aplicație, Transport, Port, Link, PHY, fizic precum și 3 **protocoale de transport**:

- Protocolul serial SCSI = Serial SCSI Protocol (**SSP**) -> suportă drivere de disc SAS.
- Protocolul serial de tunelare ATA =Serial ATA Tunneling Protocol (**STP**) - suportă drivere SATA
- Protocolul serial de management = Serial Management Protocol (**SMP**) – pentru administrarea expanderilor SAS

Pentru straturile Link și PHY, SAS definește protocolul său unic.

Pentru statul fizic, standardul SAS definește conectori și nivele de tensiune. Cel mai probabil SAS și SATA vor progresa împreună în același ritm, vitezele oferite fiind 3.0 Gbit/s , iar cele anticipate pentru 2009 de 6.0 Gbit/s, și respectiv pentru 2012 de 12.0 Gbit/s .

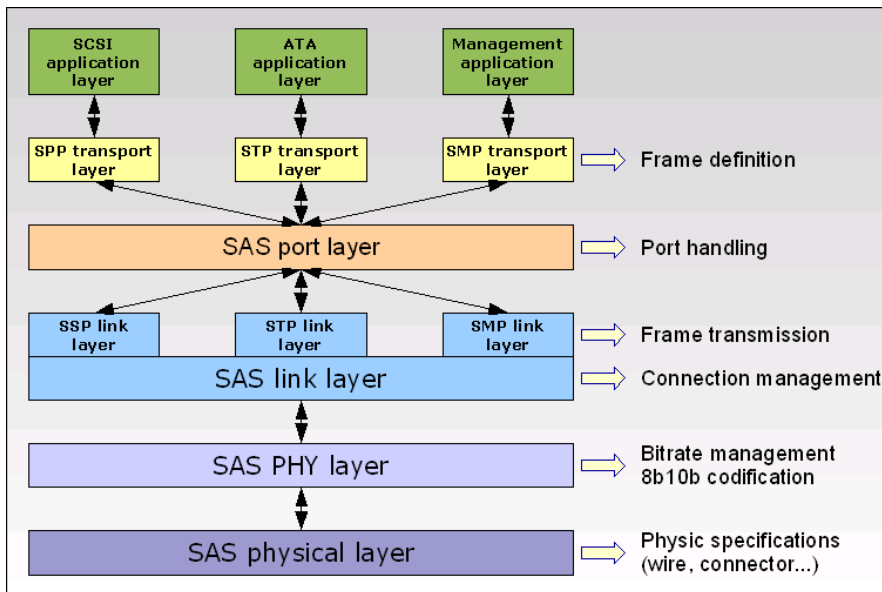


Fig. 4 Arhitectura în straturi a standardului SAS





Arhitectura în straturi a standardului SAS este formată din:

a) Stratul fizic, care definește caracteristicile electrice și fizice, modul de transmisiune diferențială a semnalelor și respectiv mai multe tipuri de conectori, ca de ex.:

SFF 8482– compatibil SATA

SFF 8484 – până la 4 dispoz.

SFF 8470 – conector extern (conector InfiniBand), până la 4 dispoz. etc. (Fig.5)

SFF 8482 (conector SATA)	SFF 8484	SFF 8470 (Infiniband connector)
		
SFF 8087 (Internal mini-SAS)	SFF 8088 (External mini-SAS)	
		

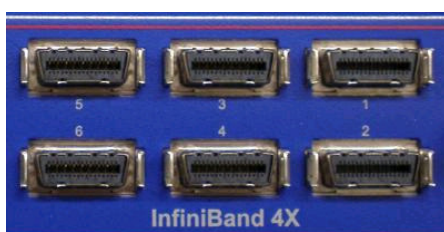


Fig. 5 Panel-ul unui switch **InfiniBand**

Specificația pentru arhitectura InfiniBand definește o conexiune serială (de tip point-to-point bidirecțională) între noduri de tip „processor” și noduri de tip „high performance I/O” (pentru DS). Această arhitectură face parte din așa-zisa „Virtual Interface Architecture”. Mai multe legături se pot conecta împreună pentru a se obține lățimi de bandă suplimentare.

b) Stratul PHY . Realizează:

- codificarea datelor 8b/10b ;
- inițializarea legăturii;
- negocierea vitezei și secvențele de reset.

c) Stratul Link (legătură), responsabil cu:

- inserarea și ștergerea primitivelor pentru „clock-speed disparity matching”
- codarea primitivelor
- realizarea operațiilor aferente procesului de „data scrambling” (codificare analogică a datelor pentru transport ???) pentru influențe electromagnetice (EMI) reduse
- determină și închide **conexiunile native** între ținte și inițiatori **SAS**
- determină și închide **conexiuni tunelate** între **inițiatori SAS** și **ținte SATA** conectate la expanderi SAS.
- gestionarea puterii (propusă pentru SAS-2.1)

d) Stratul „Port”:

Combină mai multe PHY-uri cu aceleași adrese în porturi mari.

e) Stratul transport suportă 3 protocoale de transport: SSP, STP -> suportă drivere SATA atașate la expanderi SAS și respectiv SMP ->furnizează configurația expanderilor SAS.

f) Stratul aplicație

Topologie

Un inițiator se poate conecta direct la o țintă prin unul sau mai multe PHY (o astfel de conexiune este denumită „port” indiferent dacă se utilizează unul sau mai multe PHY, deși uneori se folosește termenul de „wide port” pentru o conexiune multi-PHY).

Expander-i SAS

Componentele cunoscute ca și „SAS Expanders” facilitează comunicația ce implică un număr mare de dispoz. SAS.

Expanderii conțin **2 sau mai multe expander-port-uri externe**.

Fiecare dispozitiv expander conține **cel puțin un port de tip „ SAS Management Protocol target port” pentru administrare , putând uneori conține chiar dispozitivul**. De ex. un expander poate include un „Serial SCSI Protocol target port” pentru accesul la un dispoz. periferic.

Nu este obligatorie utilizarea unui expander pentru a realiza interfața între un inițiator SAS și o țintă, dar el permite unui singur inițiator să comunice cu una sau mai multe ținte SAS/SATA.

Există 2 tipuri diferite de expanderi: Edge Expander-i și Fanout Expander-i.

>>>

http://www.pc-pitstop.com/sas_expanders/index.asp



15-bay Hotswap SAS/SATA Tower Enclosure w/expander

Now 15 drives in one box using only a single cable connection without extra adapters or wires, for maximum compatibility and flexibility. Use SAS drives for excellent speed or SATA drives for heaps of storage space. Multiple enclosures can be daisy-chained for even more space. Requires expander-compatible SAS controller - see the description page for options.

\$1,395.00

[Add to Cart](#)

Part#: SC-SAS15



16-bay 3U Hotswap SAS/SATA Rackmount JBOD w/Redundant 650W PSU and SAS expander

Now with built in SAS-Expander! For maximum storage in minimum space, our sturdy and robust SAS/SATA JBOD provides 16 hot-swappable bays in 3U form factor. Connection to your SAS/SATA controller is through a single SFF-8088 multilane connector with 1200MB/sec bandwidth. Daisy-chaining up to eight units for a total of 128 devices! Dual 120x38mm fans provide excellent yet quiet cooling with fan failure alarm and high-temperature alarm. 650W + 350W redundant power supply and sliding rails built-in.

\$2,279.00

[Add to Cart](#)

Part#: SC-SASE16

>>>>

Un **edge expander** permite inițiatorului SAS să comunice cu până la 128 adrese SAS.

Edge expander-ii pot realiza în mod direct rutări de tip „table routing” și respectiv „subtractive routing”. În absența unui **fanout expander** se pot utiliza cel mult 2 edge expander-i în subsistemul „delivery subsystem” (pentru că se conectează împreună subtractive routing port-urile lor => nu se mai pot conecta și alți expanderi).

Un **fanout expander** poate conecta până la 128 seturi de **edge expander**-i (“edge expander device set”), permițând adresarea unui număr și mai mare de dispozitive SAS.

Portul „subtractive routing port” de la fiecare „edge expander” se va conecta la punctul de conectare denumit „PHY” de la fanout expander.

Un fanout expander nu poate realiza „subtractive routing”, poate doar să înainteze cererile aferente acestora către edge-expander-ii conectați.

Observație Standardul SAS-2.0 a renunțat la clasificarea edge/fanout expander, pentru a elimina limitările topologice inutile.

Rutarea directă permite unui dispozitiv să identifice dispozitivele care sunt conectate direct la el (în fig. 6 expanderul 2 accesează direct dispoz. A).

„**Table routing**”-ul identifică dispozitivele conectate la expanderii conectați la propriul PHY al dispozitivului (în fig.6 expanderul 2 accesează prin table routing dispoz. B).

„**Subtractive routing**”-ul este folosit atunci când dispozitivele dorite nu se află în sub-ramura de care aparținem. „Subtractive routing”-ul va pasa cererea altei ramuri (în fig. 6 expanderul 2 accesează prin subtractive routing dispoz. C).

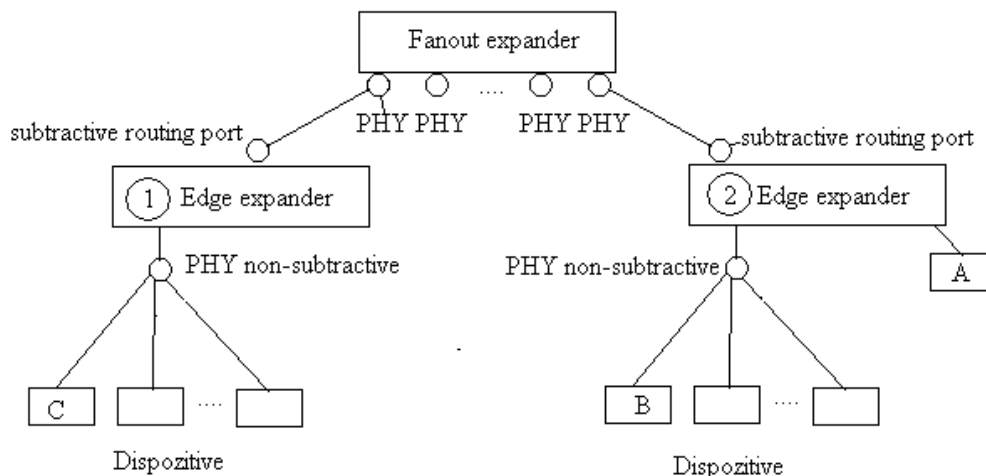


Fig. 6

Expander-ii permit **realizarea de topologii de interconectare mai complexe**. Ei:

- asistă așa-zisele „end-devices” (inițiatori sau ținte) relativ la „link-switching” (comutare de linie, care este opusă packet-switching-ului);
- pot localiza un end-device:
 - fie în mod direct (atunci când acesta este conectat direct la expander) ;
 - fir prin intermediul unei table de rutare (care realizează o mapare a ID-urilor end-device-urilor) , expanderul trebuind să realizeze rutarea către ID-ul respectiv (legătura sa este comutată pe „downstream”);
 - fie prin intermediul „subtractive routing” , caz în care legătura este rutată către un singur expander conectat la un „subtractive routing port” (în cazul când nu există un astfel de expander, end-device-ul este inaccesibil).

Expanderii pentru care nici un PHY nu este configurat ca și subtractive se comportă ca și „fanout expanders” și se pot conecta la orice număr de alți expanderi.

Expanderii cu PHY-uri de tip „subtractive” se pot conecta la cel mult 2 expanderi, caz în care trebuie să se conecteze la un expander printr-un port de tip subtractive și la celălalt printr-un port non-subtractive.

Într-un domeniu SAS există un „nod root” (cel mai de sus - the most "upstream") . Acest nod este expanderul care nu este conectat la nici un alt expander printr-un port subtractive. De aceea, dacă există un fanout expander în configurație, acesta trebuie să fie nodul root, care are informații despre toate end-device-urile conectate la domeniu.

e) Fibre Channel (Fibră optică) -> vezi cursul de Rețele.

În mod tradițional, un sistem DAS permite extinderea capacității de stocare a unui server, simultan cu asigurarea unor valori mari pentru rata de acces și respectiv lățimea de bandă „data bandwidth” (cantitate de date transportată pe unitate de timp).

Un **sistem DAS tipic** este format din una sau mai multe „enclosure”-uri (enclosure=carcasă, unitate?) în care se află **dispozitive de stocare** (vom folosi prescurtarea DS) ca de exemplu hard disk-uri și unul sau mai multe controllere. În figura 7 -> hard disk extern de tip Buffalo asamblat în fabrică într-o carcasă de tip „disk-enclosure” .



Fig.7

Interfața cu serverul sau stația de lucru se realizează prin intermediul unui HBA.

Un sistem DAS uzual este prevăzut prin proiectare cu **capacitatea „fault-tolerant design”**, relativ la mai multe aspecte:

- controller redundancy (de ex. două controllere , unul de siguranță);
- cooling (răcire) redundancy;
- pattern-uri pentru „storage fault tolerance” cunoscute ca și RAID.

>>>> <http://www.pc-pitstop.com/das/>



Arena JanusRAID 6640s U320-to-SATA RAID

Supports 16 SATA hard drives in a compact 3u chassis. Dual U320 SCSI host channels and an upgraded CPU result in increased speed. The RAID controller, fans, and power supplies are modularized to enable the least amount of operating downtime for maintenance. The scalable, high performance, and highly reliable features make it an ideal choice for DAS.

\$4,495.00

Add to Cart

Part#:

DS-ARENA6640S

>>>>

Sistemele DAS „entry-level” (ieftine, cu performanțe mai slabe) sunt cel mai adesea făcute dintr-o carcasă fără componente active (ex. controllere), la acestea logica de acces și pattern-urile pentru toleranța la defecte fiind asigurate de către HBA-ul serverului.

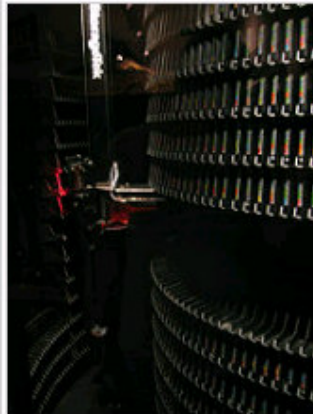
Sistemele DAS din clasele middle și top sunt prevăzute cu controllere încorporate. La acestea este încorporat managementul RAID-ului, putându-se utiliza HBA-ere simple, de tip non-RAID => costuri scăzute. Controllerele DAS-urilor pot oferi și **accesul partajat la stocare** , care permite mai multor server (uzual max. 4) să acceseze aceeași unitate logică de stocare, trăsătură utilizată în mod special pentru clustering. Din acest punct de vedere, sisteme DAS cele mai performante sunt similare cu sisteme SAN „entry-level”.

Termenul „DAS” se mai poate utiliza și pentru dispozitive comune de stocare care au conexiune de tip „unu-la-unu” între calculator și dispozitiv.

Exemple de DAS

Exemple de DAS care au conexiune de tip „unu-la-unu” între calculator și dispozitiv:

- Dispozitive de stocare portabile USB
- CD-ROM
- HD-uri ATA, SATA, SCSI, SAS
- Tape drives (drivere de bandă) și autoloader-e (Fig. 8)



Large StorageTek Powderhorn tape library, tape cartridges with barcodes packed on shelves in the front, robot arm moving in the back. Visible height of the library is about 180 cm.



Small ADIC Scalar 100 tape library, robot visible on the bottom, two IBM LTO2 tape drives behind it. Empty room for another four drives in the center. Tape cartridges on the left column, right column empty. Visible height is about 70 cm.

Fig. 8

- **J.B.O.D.** (Just a Bunch of Disks) = metodă mult utilizată pentru combinarea mai multor drivere fizice de disc într-unul singur, virtual (inversul partiționării). Controllerul tratează fiecare driver de acest tip ca și un disc independent, de aceea fiecare driver este un driver logic independent. Concatenarea nu furnizează și redundanța datelor.

De exemplu JBOD ar putea combina driverele aferente unor discuri de 3 GB, 15 GB, 5,5 GB, și respectiv 12 GB, obținând un singur driver logic de 35.5 GB. În exemplul din fig. 9, discul 0 are 63 de partiții (numerotate A1...A63), discul 1 are partițiile (A64...A91), etc. Datele sunt concatenate a.î. A64 urmează după A63, A92 după A91.

Multe distribuții Linux folosesc pentru JBOD termeni ca „linear mode” sau „append mode”.

Concatenarea este una dintre facilitățile oferite de „Logical Volume Manager”-ul de la Linux, tehnica putându-se folosi pentru a crea drivere virtuale care gestionează mai multe drivere și/sau partiții fizice.

Windows Home Server utilizează **tehnologia „drive extender”** prin care un HD-urile dintr-un JBOD (văzut aici ca vector de drivere independente) sunt combinate de către SO pentru a forma un singur „bazin” de stocare („storage pool). Acest bazin este prezentat utilizatorului ca un singur set de „network shares” (resurse partajate în rețea). Tehnologia „drive extender” adaugă facilității de concatenare și redundanța datelor prin software (un director partajat poate fi marcat în vederea duplicării, acest lucru semnificând pentru SO că trebuie păstrată o copie a datelor pe mai multe discuri fizice, deși utilizatorul vede doar o singură instanță a datelor sale).

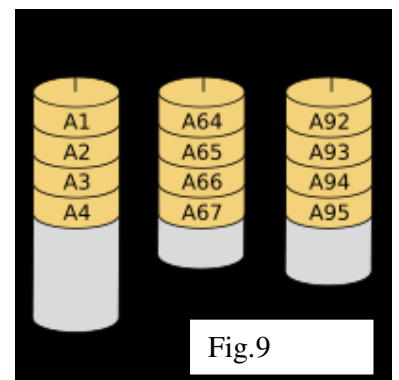


Fig.9

>>>> http://www.ex-ds.co.uk/products/products_overlandstorage_expansion.html



Sanbloc S50

Modular, scalable expansion chassis for Adaptec Snap Server systems. Each JBOD supports 12 additional hot-swappable hard drives for 12TB SATA storage or 3.6TB SAS storage. The chassis can also accommodate a mix of both SAS and SATA drives. Note that the JBOD is NOT compatible with some of the smaller storage systems from Adaptec

>>>>

- Unități RAID -> vezi cursul I/O systems

- Server Storage Expansion

>>>>http://www.overlandstorage.com/pdfs/SSS50_data.pdf

„The Overland Storage® Snap Expansion S50 provides cost-effective, flexible expansion for scalable Snap Server® storage systems. The Snap Expansion S50 comes as an unpopulated JBOD chassis, which can be filled with up to 12 SAS or SATA drives for a customized mix of price and performance that fits your specific storage needs. Use the **Snap Expansion S50** to add storage to the Adaptec Snap Server 520, 620 or 650 systems.”



Supports SAS or SATA drives for a customized blend of price and performance to fit your specific storage needs
Enables administrators easily add storage to Snap Servers as their needs grow – up to 84 TB.
Provides point-to-point connectivity to isolate drive failures, thereby maximizing reliability and performance

Dezavantaje

DAS-urile au fost supranumite "Insule de informații", deoarece nu permit partajarea datelor sau a resurselor neutilizate cu alte servere.

Atât arhitecturile NAS (network-attached storage) cât și cele SAN (storage area network) au fost concepute pentru eliminarea acestui neajuns, dar introduc aspecte noi ca: prețuri inițiale mari, administrare, securitate și arbitrarea resurselor.

4.2. Virtualizarea stocării

4.2.1. Generalități

Un **DISC LOGIC** = dispozitiv care furnizează o zonă de capacitate de stocare utilizabilă pe unul sau mai multe componente de tip driver de disc fizic într-un sistem de calcul. Se mai utilizează și alți termeni pentru a desemna un disc logic: partiție, volum logic și uneori “virtual disk” (vdisk).

Discul este descris ca *logic* deoarece nu există ca și entitate fizică de sine stătătoare. Există mai multe modalități pentru a defini un disc (sau volum) logic. Majoritatea SO-urilor moderne furnizează aceeași formă de „**logical volume management**” care permite crearea și gestionarea volumelor logice.

Discurile logice sunt de asemenea definite la nivele diferite în stiva infrastructurii de stocare (**storage infrastructure stack**). Astfel, de sus în jos există:

- SO-ul, care definește partițiile discurilor asupra cărora are permisiuni (vizibilitate), aceste discuri putând fi la rândul lor logice;
- SAN (Storage Area Network): dacă un SAN este virtualizat, un dispozitiv din acesta reprezintă din punctul de vedere al SO-ului gazdă un disc logic (vdisk)

Subsistemul de stocare furnizează uzual o anumită formă de RAID, în care discurile logice (partițiile) sunt prezentate SAN-ului ca reprezentând vectorii RAID propriu-ziși. Vectorii RAID conțin de fapt discuri fizice.

LOGICAL VOLUME MANAGER (LVM) reprezintă o metodă de alocare de spațiu pe dispozitive masive de stocare, care este mult mai flexibilă decât schemele de partiționare convenționale. În particular, un LVM poate să concateneze, să utilizeze simultan strip-uri (“strip together”) sau să combine în alte moduri partițiile obținând una mai mare, virtuală, care să poată fi redimensionată sau mutată, poate chiar în timp ce este utilizată.

LVM-ul reprezintă una dintre numeroasele modalități de virtualizare a stocării, anume cea care este implementată sub formă de strat în stiva driverului de disc al SO-ului.

REMAPAREA SPAȚIULUI DE ADRESĂ

Virtualizarea stocării sprijină obținerea **independenței de locație** (location independence) prin abstractizarea localizării fizice a datelor. **Sistemul de virtualizare** prezintă utilizatorului un **spațiu logic pentru data storage** și el însuși gestionează procesul de mapare în locația reală fizică. Forma exactă de mapare va depinde de implementarea aleasă. Anumite implementări pot limita **granularitatea mapării** care la rândul său poate limita capacitățile (capabilities) dispozitivului.

Uzual granularitatea are un domeniu de aplicabilitate care pleacă de la un singur disc fizic și poate ajunge la mici subseturi (multipli de megabytes sau gigabytes) de spațiu disc.

Într-un mediu de stocare format din dispozitive orientate pe bloc, un singur bloc este adresat folosind un “logical unit identifier” (LUN) împreună cu un offset în interiorul respectivului LUN - cunoscut ca și “adresă logică de bloc” **Logical Block Address” (LBA)**.

Un **LUN** este chiar numărul asignat unei **unități logice**. O unitate logică (**LOGICAL UNIT**) este o entitate ce folosește protocolul SCSI, singura care poate fi adresată prin operații de I/O. Fiecare țintă SCSI furnizează una sau mai multe unități. Ținta nu realizează I/O decât în contul unei unități logice specificate.

Exemplul 1. Un disk array are mai multe porturi SCSI fizice, pentru fiecare asignându-se anterior câte o “SCSI target address”. Disk array-ul este formatat ca și RAID și apoi RAID-ul este partiționat în câteva **volume de stocare separate**. Pentru a reprezenta fiecare volum, o țintă SAN este configurată pentru a furniza câte un LUN/volum. Fiecare target SCSI poate furniza mai multe LUN=> reprezintă volume multiple, acest lucru însă neînsemnând că aceste volume sunt concatenate.

Exemplul 2. Un singur driver de disc cu un port fizic SCSI. Uzual el furnizează doar un singur target, care la rândul său furnizează doar un singur LUN (cu numărul 0). Acest LUN reprezintă tot spațiul de stocare al driverului de disc.

Aplicabilitatea termenului de LUN nu este limitată la cea tradițională (SCSI paralel), acesta putând fi utilizat și de descendenții săi, cum ar fi Fibre Channel Protocol (FCP), iSCSI, HyperSCS etc.

Din perspectiva sistemului de calcul, SCSI LUN este doar o parte a adresei SCSI complete.

Adresa completă a dispozitivului (Fig. 10) este formată din:

controller ID aferent host bus adapter-ului,

target ID care identifică target-ul SCSI pe magistrală,

disk ID care identifică un LUN din target-ul respectiv,

un *slice ID* (opțional și uzual depășit) , care identifică un anumit slice (“felie”) din disk.

În familia UNIX, aceste ID-uri sunt adesea combinate într-un singur nume (de ex. */dev/dsk/ct12d3s4* se referă la controllerul 2, targetul 2, discul 3, slice-ul 4). Actual, distribuții de UNIX ca Solaris, HP-UX, NCR, ș.a., continuă să utilizeze nomenclatura “cXtXdXsX”, pe când AIX a abandonat-o în favoarea unor nume mai familiare.

Termenul de LUN se mai aplică de asemenea și unui canal de tipul “I/O access channel” în interiorul anumitor limbaje de programare (Ex. La FORTRAN: WRITE (5,32) : 5 = LUN-ul fișierului sau dispozitivului target, 32 = eticheta instrucțiunii FORMAT aferentă write-ului.)

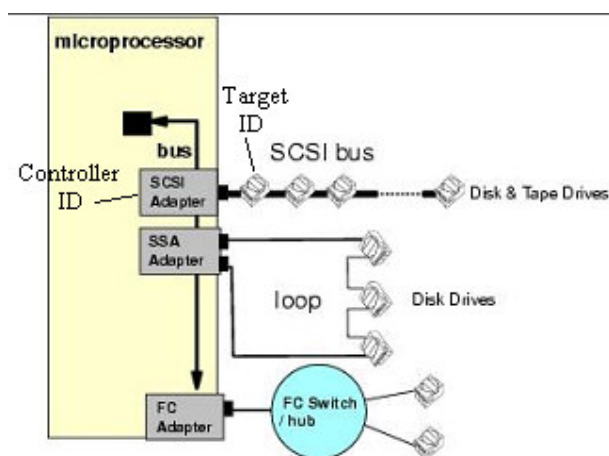


Fig. 11

Tehnica **LOGICAL UNIT NUMBER MASKING (LUN masking)** reprezintă un proces de autorizare care face ca LUN-ul să fie disponibil anumitor host-uri și indisponibile altora. Beneficiile aduse securității de această tehnică sunt limitate deoarece multe HBA-uri pot să falsifice adresele sursă (WWN-urile/MAC-urile/IP-urile). Oricum, această tehnică este implementată în special nu ca și măsură de securitate, ci mai degrabă ca o metodă de protecție împotriva unor servere care au o comportare incorectă (produc coruperea datelor gestionate de alte servere). De ex. serverele Windows atașate la un SAN pot să corupă în anumite condiții informațiile de pe volume non-Windows (Unix, Linux, NetWare) din același SAN încercând să scrie pe acestea etichete de volum de tip Windows. “Ascunzând” celelalte LUN-uri de serverul Windows, acest lucru poate fi împiedicat, deoarece serverul Windows nici măcar nu realizează faptul că există alte LUN-uri.

LOGICAL BLOCK ADDRESSING (LBA) reprezintă o schemă comună folosită pentru specificarea localizării blocurilor de date stocate pe dispozitive de stocare (uzual HD). Termenul LBA poate semnifica fie adresa blocului la care se referă adresa, fie chiar blocul.

Uzual blocurile logice au dimensiuni de 512 sau 1025 octeți, iar CD-urile ISO 9660 (și imaginile lor) folosesc blocuri de 2048 octeți.

LBA este o schemă de adresare foarte simplă: blocurile sunt localizate de un index, primul bloc având indexul 0 LBA=0, al 2-lea LBA=1 etc.

SCSI-ul a introdus LBA ca și o abstractizare. Apelurile sistem care necesită operații de I/O cu blocuri pasează LBA-urile către driverul dispozitivului de stocare. Pentru cazurile simple (când volumul mapează un singur driver fizic) controllerul driverului de disc primește direct LBA-urile blocurilor implicate în transfer și realizează mapările corespunzătoare în adrese interne de tip CHS (tupli cilindru-cap-sector).

MAPAREA LBA ȘI VIRTUALIZAREA LUN

Pentru cazurile mai complexe (în special dispozitive RAID și SAN-uri unde LUN-urile sunt compuse prin virtualizarea și agregarea LUN-urilor), LBA-urile sunt translatate din modelul aplicație al discului în cele utilizate în realitate de dispozitivul de stocare.

În dezvoltările complexe, unde se utilizează un “[storage fabric](#)” (mediu de stocare distribuit în rețea), pot apărea mai multe astfel de translații de LBA între aplicația dispecer și discul final, aflat la distanță.

În continuare se folosesc notațiile: C=nr. Cilindru, H=nr. Cap, S=nr. Sector.

O [conversie CHS->LBA](#) folosește formula:

$$LBA(C,H,S)=((C * număr_capete)+H) * sectoare_per_pistă+S-1$$

O mapare [LBA -> CHS](#) :

$$C = \frac{LBA + 1 - S}{nr_capete \times sec\ toare_per_pista}$$

$$H = \frac{(LBA + 1 - S) \bmod (nr_capete \times sec\ toare_per_pista)}{sec\ toare_per_pista}$$

$$S = LBA \bmod sectoare_per_pista + 1.$$

Adresă sau port?

În majoritatea protocoalelor, o adresă (indiferent dacă se referă la un inițiator sau la un target) este echivalentă într-un accept mai larg cu portul fizic al dispozitivului. La nivelul anului 2008 nu este comună situația în care un singur port fizic găzduiește mai multe adrese sau când o singură adresă este accesibilă din mai multe porturi ale dispozitivului. Chiar și atunci când se folosește “multipath I/O” pentru a obține toleranța la defecte, driverul de dispozitiv mai degrabă comută între diferite targeturi și inițiatori, legați în mod static la porturi fizice decât să partajeze o adresă statică între mai multe porturi fizice.

În contextul virtualizării stocării, un **DISC VIRTUAL (VDISK)** reprezintă un disc logic sau volum virtual asupra căruia un host sau o aplicație realizează operații de I/O. Este în general furnizat de un dispozitiv de virtualizare a stocării și este aplicabil în special în infrastructuri de tip FC SAN (Fibre Channel storage area network).

Actual termenul de *vdisk* este folosit la disk array-urile HP EVA și de către IBM SAN Volume Controller.

Discul este catalogat ca și virtual datorită modalității în care este mapat în capacitatea fizică de stocare pe care o reprezintă. În majoritatea sistemelor de stocare virtuale se menține un **meta-data mapping table** (tabelă de mapare a metadatelor) care translatează identificatorul de disc și LBA-urile (virtuale) în identificatorul de disc și LBA-urile fizice. Spațiul de adrese poate fi limitat de capacitatea solicitată pentru memorarea tabelii de mapare, aceasta fiind direct influențată de granularitatea informațiilor de mapare.

Granularitatea virtualizării depinde de implementare. Anumite sisteme virtualizate furnizează pur și simplu ceea ce ar putea fi considerat ca “agregarea discului” (combinarea mai multor drivere de discuri fizice/logice într-unul singur logic) a.î. granularitatea constă din chiar discul fizic.

Sistemele de virtualizare reală însă divid discurile fizice în mai multe 'chunks' (bucăți compacte) sau extensii mai mici. Astfel se poate obține un beneficiu relativ la performanță prin dispersarea unui singur disc virtual peste mai multe discuri fizice (se obține un **acces concurrent îmbunătățit** față de sistemul nevirtualizat).

Maparea spațiului de adresă se realizează între vdisk și o unitate logică prezentată de unul sau mai multe controllere de stocare. Chiar LUN-ul însuși poate fi un produs al virtualizării într-un strat diferit.

În contextul stocării, termenul de **MULTIPATH I/O** se referă la o tehnică de îmbunătățire a performanțelor pentru situațiile când există mai mult de o cale fizică între CPU-ul unui sistem de calcul și dispozitivele sale masive de stocare prin intermediul magistralelor, controllerelor, switch-urilor și bridge-urilor care le conectează.

Exemplu: un disc SCSI conectat la două controllere SCSI pe același calc. sau un disc conectat la două porturi FC. Dacă se defectează un controller, port sau switch, SO poate să ruteze I/O-ul prin controllerul valid într-un mod transparent pentru aplicație, fără modificări vizibile pentru aplicație (exceptând poate doar o oarecare latență).

Straturile de SW multipath pot ajuta căile redundante relativ la oferirea unor trăsături de îmbunătățire a performanțelor, ca:

- echilibrarea dinamică a sarcinilor (dynamic load balancing)

- traffic shaping (vezi curs rețele);
- managementul automat al căilor (Automatic path management);
- reconfigurarea dinamică.

Recent s-au făcut mari progrese în ceea ce privește tehnologiile multipathing native din SO ca Windows, Linux și HP-UX (11.31) (care sunt în majoritate implementate ca pachete SW gratuite), acestea devenind din ce în ce mai populare.

REDIRECTAREA I/O

SW-ul de virtualizare sau dispozitivul utilizează meta-date pentru a redirecta cereri de I/O. El va primi o cerere I/O care conține informații despre localizarea datelor în termenii discului logic (vdisk) și o translatează într-o cerere I/O nouă către locația discului fizic.

De ex. un dispozitiv de virtualizare poate realiza următorul set de operațiuni:

- primește o cerere de citire de la vdisk cu parametrii LUN ID=1, LBA=32;
- realizează o operație de tip “ meta-data look up “ pentru parametrii mai sus menționați și găsește maparea corespunzătoare ca având parametrii fizici LUN ID=7, LBA0 ;
- primește datele de la LUN-ul fizic;
- trimite datele înapoi la inițiator, ca și cum ar fi provenit de la vdisk din partea LUN ID=1, LBA32

4.2.2. SAN (Storage Area Networks)

4.2.2.1. Noțiuni introductive

În contextul acestui curs, SAN (*Storage Area Network*) reprezintă o *rețea dedicată*, aparte de rețelele LAN și WAN. Este adesea denumită „rețeaua din fața serverelor”.

SNIA (*The Storage Network Industry Association*) definește SAN ca „o rețea al cărui scop principal este transferul datelor între sistemele de calcul și elementele de stocare”.

Un SAN poate fi privit ca un ansamblu format din:

- infrastructura de comunicare, care furnizează conexiunile fizice;
- un strat de management, care organizează conexiunile, elementele de stocare și sistemele de calcul a.î. transferul de date să fie sigur și robust.

Termenul SAN este identificat în mod obișnuit (dar acest lucru nu este obligatoriu) cu *serviciu pentru operații de I/O cu blocuri și nu cu servicii de acces la fișiere*.

Un SAN mai poate fi privit și ca un *sistem de stocare* care constă din elemente de stocare, dispozitive de stocare (vom folosi pentru acestea abrevierea neconsacrată DS), sisteme de calcul și/sau echipamentele auxiliare aferente, plus întreg soft-ul pentru control, care comunică prin rețea.

O rețea tip Ethernet al cărui principal scop este furnizarea accesului la elemente de stocare poate fi considerată un SAN. Termenul de SAN se utilizează uneori și pentru *interconectarea sistemelor în clustere*.

Un SAN permite conexiuni de tip „any-to-any” (oricine cu oricine) prin rețea, folosind elemente de interconectare cum sunt routerele, gateway-urile, hub-urile, switch-urile și directoarele (vom nota această colecție de elemente de interconectare prin MC). Dacă se utilizează pentru interconectare canale pe fibră optică se utilizează termenul de *fabric* pentru elementele din MC.

Un SAN elimină conexiunile tradiționale dedicate între un server și un DS, precum și conceptul conform căruia serverul deținea și administra efectiv DS-urile. De asemenea se elimină orice restricție relativ la cantitatea de date pe care un server o poate accesa, limitată în mod obișnuit de numărul de DS-uri care pot fi atașate la servere individuale.

În schimb, un SAN introduce *flexibilitatea lucrului în rețea* pentru ca unul (sau mai multe) server(e) să poată partaja un *echipament comun de stocare* (common storage utility) care poate fi format din mai multe DS-uri incluzând discuri, benzi și DS-uri optice. Echipamentul de stocare poate fi localizat la distanță de serverele care îl utilizează. SAN face posibilă utilizarea a noi arhitecturi de rețea, în care gazde multiple accesează DS-uri multiple conectate la aceeași rețea.

SAN-ul poate fi privit ca o *extensie a conceptului de stocare*, extensie care permite ca DS-urile și serverele să fie interconectate utilizând elemente similare, ca în cazul LAN și WAN, adică elemente din MC. Un SAN poate fi partajat între servere și/sau poate fi dedicat unuia singur. Poate fi local sau poate fi extins pe distanțe mari (geografice).

În fig. 12 se prezintă o vedere stratificată a unui SAN (cu servere și DS de producție IBM) care conectează mai multe servere la mai multe sisteme de stocare.

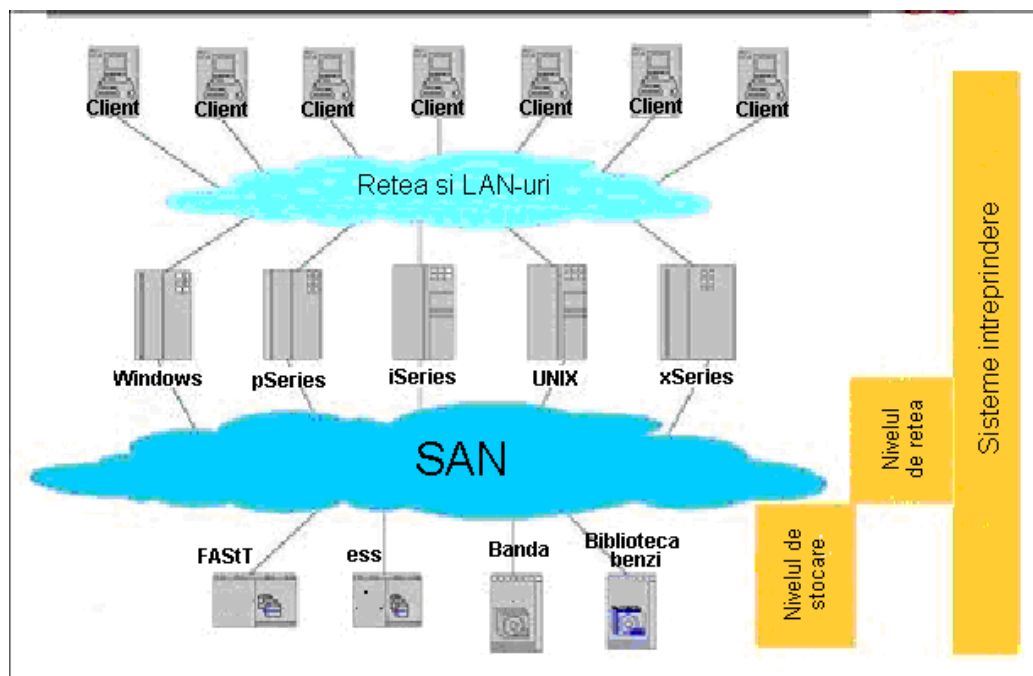


Fig. 12 Vedere stratificată a unui SAN care conectează mai multe servere la mai multe sisteme de stocare (IBM)

Se poate vorbi de o **simplificare a infrastructurii**, pentru că SAN poate conduce la **consolidarea mediului de stocare**. Mediile de stocare consolidate (*consolidated storage environment*) au mai puține elemente de gestionat. Astfel se obține o creștere a utilizării resurselor, se poate simplifica managementul stocării, se pot partaja servere de stocare aflate la distanțe mari, se pot obține *economies of scale* pentru că se dețin servere de stocare cu mai multe discuri. Aceste medii pot fi mai flexibile și pot furniza infrastructura necesară virtualizării și automatizării.

Există **4 metode de consolidare**: centralizarea centrelor de date, consolidarea fizică, integrarea datelor și integrarea aplicațiilor.

Congestionările tradiționale ale rețelei se pot evita atunci când se utilizează un SAN, deoarece acesta facilitează **transferurile de date de mare viteză în care sunt implicate serverele și DS-urile**, în oricare dintre modurile:

- Server către DS. Acesta este modelul tradițional de interacțiune cu DS. Avantajul constă în faptul că același DS poate fi accesat în mod serial sau în mod concurent de către mai multe servere.
- Server la server. Se poate utiliza un SAN pentru comunicații de mare viteză în care se vehiculează un volum mare de date comunicate între servere.
- DS către DS. Datele se mută fără intervenția serverului. Exemple: dispozitiv de tip disc care realizează backup pe bandă fără intervenția serverului, sau dispozitive aflate la distanță care realizează mirror (copie în oglindă) prin SAN.

SAN se bazează pe câteva (combinații de) tipuri diferite de interfețe de mare viteză. Actual canalele cu fibră optică sunt standard *de facto* în majoritatea SAN. De asemenea se utilizează interfețe SCSI pe post de subinterfețe între componentele interne ale membrilor SAN (spre ex. între discurile propriu-zise de stocare și controllerele RAID aferente lor).

Alte avantaje importante ale utilizării SAN sunt: (a) performanțe îmbunătățite relativ la capacitatea de stocare; (b) fiabilitate mai bună; (c) scalabilitate mărită.

(a) *Performanțele de stocare* pot fi mult mai bune decât cele ale unui DS atașat direct, datorită ratelor de transfer foarte ridicate aferente conectării dispozitivelor la un SAN (ex. canal cu fibră). Alte câștiguri de performanță provin din oportunitățile furnizate de arhitectura flexibilă a SAN – ca de ex. echilibrarea încărcării și back-up-ul independent de LAN.

(b) *Fiabilitatea stocării* poate fi mult îmbunătățită prin trăsăturile speciale ale SAN. Disponibilitatea datelor și aplicațiilor este garantată de opțiuni ca de ex: căi I/O redundante (redundant I/O paths), gruparea serverelor (server clustering), replicarea (locală sau la distanță) a datelor chiar în timpul lucrului (run-time data replication) (fig.13).

(c) Operațiuni ca mărirea capacității de stocare sau adăugarea de alte DS-uri se pot realiza ușor cu SAN, adesea fără a fi necesară oprirea sistemului. E posibil ca serverul sau clienții săi de rețea să nu trebuiască să fie aduși în stare de inactivitate. Aceste trăsături pot conduce

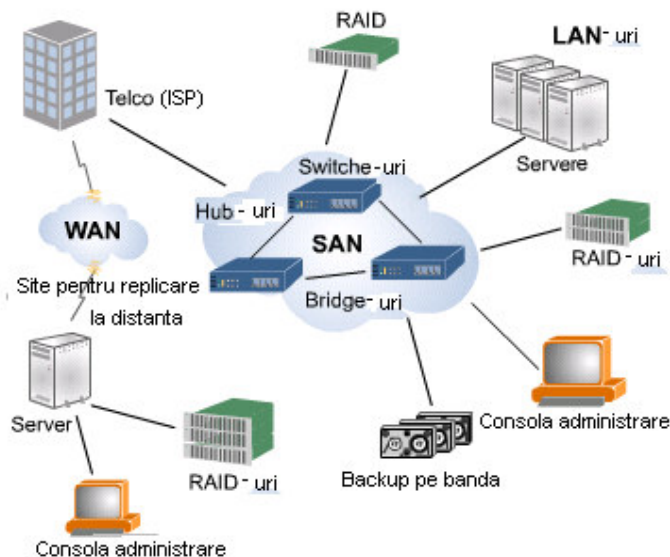


Fig. 13. Facilități speciale oferite de SAN

la reducerea drastică a costurilor, respectiv a timpilor morți (outages) în rețea, mărirea spațiului de stocare fără probleme și reducerea încărcării rețelei.

4.2.2.2. Protocoale

Majoritatea rețelelor de stocare folosesc protocolul SCSI pentru comunicarea între servere și dispozitivele disc. Ele nu folosesc interfața fizică SCSI de nivel scăzut (cablurile), deoarece topologia de magistrală a acestora este nepotrivită pentru lucrul în rețea. De aceea se folosește un [strat de mapare către alte protocoale de nivel scăzut](#) pentru a forma o rețea, după cum urmează:

- ATA over Ethernet (AoE), care realizează maparea ATA în Ethernet,
- Fibre Channel Protocol (FCP), mapează SCSI în Fibre Channel (FC)
- Fibre Channel over Ethernet (FCoE), mapează FICON în FC, folosit de calculatoare mainframe
- HyperSCSI, mapează SCSI peste Ethernet
- iFCP sau SANoIP mapează FCP peste IP.
- iSCSI Extensions for RDMA (iSER), mapează iSCSI peste InfiniBand (IB),
- iSCSI, mapează SCSI peste TCP/IP.

4.2.2.3. Partajarea stocării

La început centrele de date au creat „insule” de disk-array-uri SCSI. Fiecare insulă era dedicată unei aplicații și era vizibilă ca și un număr oarecare de HD-uri virtuale (LUN-uri). În esență, un SAN conectează laolaltă insule de stocare folosind rețele de mare viteză.

SO-urile încă văd un SAN ca pe o colecție de LUN-uri, dar uzual mențin propriile lor sisteme de fișiere pe acestea. Aceste sisteme de fișiere „locale” care nu pot fi share-ate între mai multe SO/host-uri sunt cele mai fiabile și se utilizează cel mai mult.

Dacă două sisteme de fișiere locale independente se află la același LUN, ele nu vor ști acest lucru, nu vor avea niciun fel de sincronizare a cache-ului și eventual se vor corupe reciproc. Astfel, partajarea datelor între computere printr-un SAN necesită soluții avansate, cum ar fi „[sisteme de fișiere SAN](#)” sau „clustered computing”.

În ciuda acestor aspecte care indică o creștere a complexității, SAN ajută mărirea gradului de utilizare a capacității de stocare, deoarece mai multe servere partajează același spațiu de stocare pe disc-array-uri. O aplicație comună a unui SAN este pentru accesarea tranzacțională a datelor care necesită acces la nivel de bloc de mare viteză pentru HD-uri (cazul serverelor de email, a bazelor de date și a serverelor de fișiere utilizate intens).

Prin contrast, NAS permite mai multor calculatoare să acceseze același sistem de fișiere prin rețea și să sincronizeze accesul lor. Mai târziu, introducerea de „NAS heads” a permis conversia ușoară din stocare SAN în stocare NAS. În figura 14 se prezintă modalități de stocare prin DAS, NAS și SAN.

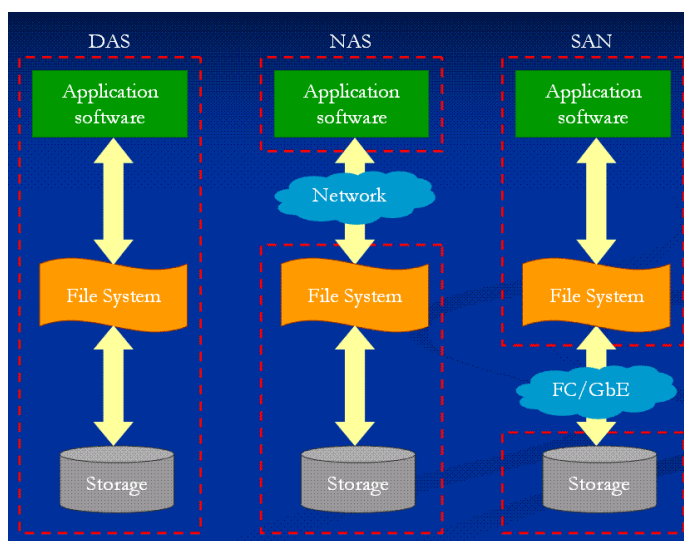


Fig.14 Modalități de stocare prin DAS, NAS și SAN.

În ciuda diferențelor dintre NAS și SAN, se pot implementa și soluții hibride, ca în fig. 15.

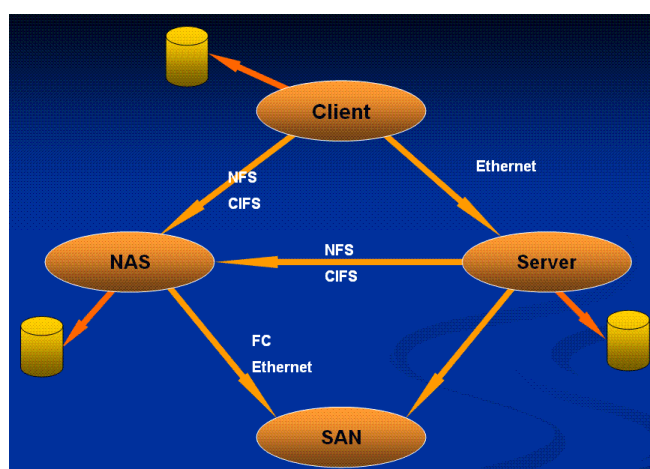


Fig. 15 Soluție hibridă NAS-SAN

4.2.3.NAS (Network Attached Storage)

NAS= stocare de date la nivel de fișier (file-level computer data storage) conectată la o rețea care furnizează acces la date unor clienți heterogeni ai rețelei.

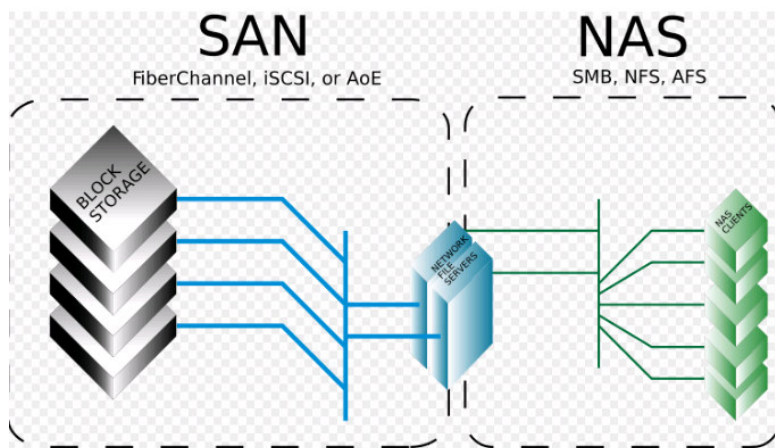


Fig. 16 SAN vs. NAS

O unitate NAS este în esență un calculator autoconținut („self-contained”) conectat la o rețea, cu singurul scop de a furniza servicii de stocare a datelor bazate pe fișiere altor dispozitive din rețea. SO-ul și alte SW-uri de pe unitatea NAS furnizează funcționalitatea stocării de date și a sistemelor de fișiere precum și accesul la fișiere, simultan cu managementul acestor funcționalități.

>>> <http://www.pc-pitstop.com/nas/>



\$3,575.00

R208DSS 2U 8-bay NAS w/Integrated iSCSI SAN

2U 8-bay NAS using Open-E Embedded OS, Core 2 Duo 2.66GHz CPU, 2GB DDR2 Memory, 2xGbE LAN ports, 500W Redundant PSU, Recovery CD, Hardware RAID 0, 1, 5, 6. Enterprise Features and Performance!



\$5,295.00

R316DSS 3U 16-bay NAS w/Integrated iSCSI SAN

3U 16-bay NAS using Open-E Embedded OS, SAS Expansion Port for adding capacity later, Core 2 Duo 2.66GHz CPU, 2GB DDR2 Memory, 4xGbE LAN ports, 650W Redundant PSU, Recovery CD, Hardware RAID 0, 1, 5, 6. Enterprise Features and Performance!

Sistemele NAS moderne sunt „enclosure”-uri externe, care folosesc HD-uri SATA și IDE atașate la rețea=> mai multe calculatoare pot partaja același spațiu de stocare care minimizează overhead-ul prin management-ul centralizat al HD-urilor. Multe dintre aceste unități încorporează controllere RAID.

>>>

Unitatea NAS nu este proiectată pentru a realiza sarcini de calcul de uz general, deși tehnic poate fi posibilă și rularea altor softuri.

Uzual unitățile NAS nu au tastatură sau display și sunt controlate și configurate prin rețea, adesea prin conectarea unui browser la adresa lor de rețea.

O alternativă la utilizarea unei stocări NAS într-o rețea constă în utilizarea unui calculator ca și server de fișiere. În forma sa de bază, un **server de fișiere dedicat** este o **unitate NAS cu tastatură și display** și un SO care, fiind optimizat pentru a furniza servicii de fișiere, poate rula și pentru a executa alte sarcini. Actual serverele de fișiere sunt utilizate din ce în ce mai mult pentru a furniza alte funcționalități, cum ar fi servicii pentru baze de date, servicii email, etc.

Un dispozitiv NAS nu necesită SO, adesea utilizându-se SO-uri cu funcționalități minimale. Ex.: FreeNAS (=software NAS gratuit open source proiectat pentru utilizarea pe un calc. standard) este doar o versiune de UNIX FreeBSD, din acesta fiind eliminate toate funcționalitățile care nu au legătură cu stocarea.

Sistemele NAS conțin unul sau mai multe HD-uri, adesea aranjate în containere de stocare redundante, logice sau în vectori RAID. **NAS degreavează celelalte servere din rețea de responsabilitatea administrării fișierelor.**

NAS utilizează protocoale bazate pe fișiere („**file-based protocols**”) ca de ex. NFS (la Linux) sau SMB/CIFS - Server Message Block/Common Internet File System (la Windows).

Rareori unitățile NAS limitează clienții la utilizarea unui singur protocol.

NAS furnizează atât stocarea cât și sistemul de fișiere, spre deosebire de SAN, care furnizează doar stocare la nivel de bloc, lăsând în sarcina clientului aspectele legate de sistemul de fișiere .

Cu timpul granițele NAS/SAN au început să se suprapună, anumite produse sugerând clar direcția de evoluție și oferind atât protocoale bazate pe fișiere (NAS) cât și protocoale pentru blocuri (SAN) în același sistem. Exemplu: Openfiler, produs free sub Linux.

Avantaje

Disponibilitatea datelor ar putea crește mult dacă se utilizează NAS ; acesta oferă în plus și facilități incorporate pentru RAID și clustering.

NAS poate mări performanțele deoarece realizează „file serving”-ul (deservirea fișierelor) , degrevând de această sarcină un alt server care trebuie să execute și alte procesări.

Totuși performanțele dispozitivelor NAS depind puternic de viteza și traficul prin rețea și de mărimea memoriei RAM cache de pe calculatoarele sau dispoz. NAS.

De notat că NAS este efectiv un server, cu toate componentele unui PC tipic (un CPU, motherboard, RAM, etc) și fiabilitatea depinde de modul său intern de proiectare.

Un NAS fără căi de acces al datelor redundante, fără controllere și surse de putere redundante , este probabil mai puțin fiabil decât un DAS conectat la un server care oferă redundanța componentelor sale majore.

Dezavantaje

Datorită caracteristicilor multiprotocol și a stratului (CPU+SO) redus, NAS prezintă limitări comparabile cu cele aferente sistemelor DAS/FC.

Dacă NAS este ocupat cu prea mulți utilizatori, are prea multe operațiuni I/O sau puterea de procesare a CPU este prea mult solicitată, NAS își atinge limitele. Dacă un sistem server se poate ușor upgrada prin adăugarea unuia sau mai multor servere în cluster a.î. să se mărească puterea CPU, NAS-ul este limitat la propriul hardware, uzual ne-upgradabil.

Anumite dispozitive NAS nu reușesc să expună servicii bine-cunoscute care sunt tipice pentru un server de fișiere, sau o fac într-un mod ineficient. Exemple: capacitatea de a calcula gradul de utilizare a discului în directoare diferite, de a indexa (localiza) rapid fișiere, de a realiza în mod eficient mirror-ul cu **rsync**.

(**rsync** =aplicație SW Unix folosită pentru sincronizarea fișierelor și directoarelor dintr-o locație în alta simultan cu minimizarea transferului de date folosind atunci când este nevoie o codificare de tip „delta encoding”).

Se poate totuși utiliza rsync, dar printr-un client NFS sau CIFS; metoda eșuează însă în enumerarea de ierarhii masive de date la viteza nominală a driverelor locale și induce un trafic de rețea considerabil.

Diferența principală între DAS și NAS constă în aceea că DAS reprezintă doar o extensie a unui server existent și nu are trăsături de lucru în rețea, pe când NAS este amplasată în rețea ca și entitate de sine stătătoare. Partajarea fișierelor folosind NAS este mai ușoară comparativ cu DAS, presupunând mai puțin timp procesor și putere I/O.

Utilizări

NAS este util nu doar pentru că oferă o stocare centralizată clienților calculatoarelor din medii unde se vehiculează cantități mari de date. NAS face posibile sisteme mai simple cu costuri mai mici, ca de ex. servere de email și web care oferă load-balancing și fault-tolerant , furnizând servicii de stocare.

Piața potențială în formare pentru NAS o formează mediile cu mari cantități de date multi-media. Appliance-urile NAS au prețuri puțin mai mari decât HD-urile externe USB sau FireWire.