Universitatea Politehnica București

Facultatea de Electronică, Telecomunicații și Tehnologia Informației

**iSCSI**

**(Internet Small Computer System Interface)**

 Grupa 443A

 Coordonator științific: Studenți:

 Conf. Dr. Ing. **Ștefan Stăncescu Rădulescu Vicențiu-Mihail**

 **Rădulescu Victor Cristian**

**Cuprins:**

**1. Introducere**

**2. iSCSI Implementare**

**3. Proiectarea unei rețele iSCSI**

**4. Exemple: Scenari de utilizare ISCSI**

**5. Alternative Ethernet la iSCSI**

**6. Concluzi**

**Rădulescu Victor:1,2,3**

**Rădulescu Mihail:4,5,6**

**iSCSI**

**(Internet Small Computer System Interface)**

Introducere

**Definiție:**

 iSCSI este un protocol al nivelului transport care descrie modul în care pachetele ale Small Computer System Interface ( SCSI ) ar trebui să fie transportate printr-o rețea TCP / IP

 iSCSI,care provine de la Internet Small Computer System Interface,lucrează deasupra protocolului TCP(Transport Control Protocol) și lasă ca comanda SCSI să fie trimisă end-to-end la rețelele locale LAN,rețele globale de internet(WAN-uri) sau Internetul.IBM a dezvoltat iSCSI ca o dovadă a concepului din 1998 și prezentată ca prima variantă a standardului iSCSI la Internet Engineering Task Force (IETF) in 2000. Protocolul a fost ratificat în anul 2003 .

**Modelul Arhitecturi iSCSI:**



**Cum funcționează ISCSI:**

iSCSI funcționează prin transportul de date la nivel de bloc între un inițiator iSCSI pe un server și un obiectiv iSCSI pe un dispozitiv de stocare . Protocolul iSCSI incapsuleaza comenzi SCSI și asamblează datele în pachete pentru stratul TCP / IP . Pachetele sunt trimise prin rețea folosind o conexiune punct la punct. La sosire , protocolul iSCSI dezasamblează pachetele , separând comenzile SCSI astfel încât sistemul de operare ( OS) va vedea spațiul de stocare ca un dispozitiv SCSI local, care pot fi formatat ca de obicei . Astăzi , popularitate iSCSI în unele dintre micile întreprinderi sau în întreprinderile mijlocii ( IMM-uri) are de a face cu modul în care virtualizarea serverului folosește spațiul de stocare. Într-un mediu virtualizat , spațiul de stocare este accesibil tuturor gazdelor din cadrul clusterului și nodurilor clusterului, nodurile care comunica cu spațiul de stocare în rețea prin utilizarea protocolului iSCSI .

****

Această imagine ilustrează modul în care un simplu mod de implementare iSCSI SAN ar putea împărtăși aceeași infrastructură ca și restul calculatoarelor din rețea . În mod ideal , cele două rețele ar fi separat , pentru a evita congestionarea rețelei .



Această imagine arată o abordare mai practică a planificării unei rețele de iSCSI . Acest design poate fi scalat pentru a permite mai multe căi de I / O prin adăugarea switch-uri Ethernet dedicate pentru rețeaua de stocare sau permițând producerea LAN a fi folosit ca o cale de rezervă în cazul în care rețeaua dedicată Ethernet nu este disponibilă.

Un inițiator iSCSI este o bucată de software sau hardware care este instalat pe un server pentru a trimite date către și de la iSCSI-suprafața de bază de stocare sau destinație iSCSI .

**Formatul unui pachet iSCSI:**



 iSCSI Implementare

**ISCSI și IP-ul de stocare a rețelei:**

Avem două protocoale primare:

1)ISCSI-Internet-SCSI-pentru a transporta SCSI CDB și data în conexiunile TCP/IP.



2)FCIP-Fibre-Channel-over-IP-pentru a transporta cadrele în interiorul conexiunilor TCP/IP-oricare cadru FC(Fibre Channel) nu doar SCSI

****

**Numirea iSCSI:**

Inițiatorul și destinatarul necesită iSCSI.Un nume este o locație indepedentă iar un nume de node iSCSI este numele dispozitivului SCSI al dispozitivului iSCSI asociat cu nodurile iSCSI,nu adaptoarelor.

Se utilizează SLP(Service Location Protocol)V1,iSNS,sau interogări destinație pentru nume(Trimitere Destinatar).

Două tipuri de nume iSCSI:

 Iqn-iSCSI qualified name

Eui-Extended Unique Indentifier(IEEE EUI-64- de asemenea folosit pentru FC WWN-uri)

**Relația dintre SCSI și iSCSI:**

Dispozitiv SCSI = node iSCSI

Port SCSI = port iSCSI

Portul rețea definit de adresa IP + portul TCP

Grup de porturi = o singură conexiune SCSI

**Sesiune iSCSI între nodul inițiator și nodul iSCSI destinație:**

****

**Sesiuni iSCSI:**

iSCSI are concepția de sesiune și ele sunt de două tipuri:

1)Discovery

 2)Normal operation

Ambele tipuri de sesiuni au câteva stagii:

a)Faza de logare inițială

b)Autentificare pentru securitate

c)Negocierea parametrilor de operație

d)Full-Featured Phase

Sesiunea poate să se ocupe de comenzile SCSI și de date după ce logarea a fost completă.

**Secvența de logare iSCSI (fără autentificare):**

****

**1)Discovery**

Rețele mici : Emițători,destinatari și configurație statică

 Comanda trimiteDestinatari face configurația mai ușoară

Rețele de lungime medii:Protocolul Serviciului de localizare(SLP multicast discovery)

Rețele de lungime marii:iSNS(Internet storage-name service)

 Include domenii soft

 Include bază de date pentru vitoare administrări

**Driverul Soft:**

****

ISCSI este un servici de rețea deschis prin folosirea unui driver software iSCSI.Multe sisteme de operare suportă prin drivere Cisco și/sau de la furnizorul OS.

Cisco oferă driver suport total pentru Solaris 2.6(EOL),7,8,9.Este bazat pe Linux pe kernel 2.6 ,Windows 2000 cu SP2(service pack 2) sau mai noul Windows XP Pro etc.

Furnizoarele OS suportă driverele iSCSI.

Descărcări TCP și opțional procesare iSCSI în hardware.Deschide gazda CPU de la:- procesare TCP-16 biți de control pe pachet.

 -procesarea iSCSI opțională,cu header de 32 biți și digestie de date(CRC32C)

**HBA iSCSI și Enginul de descărcări TCP(TOE):**

****

 **Proiectarea unei rețele iSCSI**

Pentru proiectarea unei rețele iSCSI luăm în considerare următoarele:

-cunoașterea profilului aplicației de I/O

-cunoașterea transferul aplicației tale

-determinarea disponibilități necesare

-cea mai bună perfomanță obținută (fără latență) de la IP-ul local dedicat al rețelei de stocare(fără trafic competițional).

-considerarea distanței

**IP-ul rețelei de stocare dedicat**

Separarea logică a rețelei IP dar nu neapărat separare fizică a rețelei.Poate folosi un VLAN pentru o rețea Ethernet existentă.Este recomandat folosirea unui NIC dedicat pe o gazdă pentru iSCSI.Avem nevoie de un potențial minim pentru a afirma banda.

****

Schema **unei Dedicated IP Storage Network**

**Mapare VLAN și VSAN**

\*Apartanența VSAN

-atribuite hosturilor iSCSI:switch(config)#iscsi Initiator ip-add 51.51.51.51

 switch(config-(iscsi-init)#vsan 10

-atribuite interfeței iSCSI :switch(config)#vsan database

 switch(config-vsan-db)#vsan 10 interface iscsi 2/3

-atribuite discului

\*Terminația VLAN

-sub-interfața

\*Controlul de acces

-bazat pe ISCSI:sw(config-(iscsi-tgt)#pwwn ....

 sw(config –(iscsi-tgt)#initiator ip

 add 51.51.51.51 permit

-Zone based:Sw(config)#zone name marketing vsan 10

 Sw(config-zone)#member ip-add ...

 Sw(config-zone)#member pwwwn ...



Schema unei mapări VLAN și VSAN

**Inițiatorul Proxy**

Modul default al inițiatorului Proxy este cel transparent.

Un pWWN reprezintă multipli emițători iscsi pentru scalabilitate.Sunt mai puține intrări în numele serverului FC.

Avem un managemenet ușor al zonei la inițiatorul(emițătorul) proxy și este folositor în clusterele la aplicații .

Inițiatorul mută sarcina de mascare a configuraților și operaților la IPS pentru o ușuară administrare.

**Proxy-modul inițiator:**

Sw(config)#int iscsi 2/3

Sw(config-if)#switchport proxy-initiator ...nwwn .. pwwn



Schema unui Inițiator,emițător Proxy

**Internet Storage Name Service(iSNS)**

**ISNS** un nume de serviciu care oferă spațiu pentru resursele descoperite (register/deregister/query).Este un serviciu de control automat al logării,unde starea ei schimbă serviciul de notificare,fiind un serviciu de tip client-server cu serviciu director.ISNS deschide maparea pentru FC și dispozitive ISCSI.ISNS are un protocol redus la capătul TCP-ului numit iSNSP.

**Propietăți de proiectare cu iSNS:**

-un dinamic discovery pentru o implementare de scală largă.

-o integrare fară folos cu serviciul nume FC

-server suport iSNS disponibil de la MS Windows și Linux



Schema unui iSNS

**iSCSI cu GPFS**

Transferul de fișiere de către NFS sau NAS ar putea fi vârful de perfomanță.ISCSI oferă perfomanțe mari la nivelul de blocaj la acces dar stă prost la capitolul de transfer de fișiere.

O soluție ar fi combinarea iSCSI cu GPFS ce ar duce avantajul de perfomanțe înalte și consolidare transferului de fișiere pe disc.

**GPFS** (General Parallel File System) este o tehnologie de cluster pentru perfomanțe mari de calcul.Oferă partajarea sistemului de fișiere prin discuri de către toți clienți iSCSI în cluster prin administrare de token.



Exemplu de combinare iSCSI cu GPFS

**Citirea iSCSI**

****

Există o relație unu la unu între iSCSI și cadre FCP.

Există un singură parcurgere dus-întors de la emițător la destinator pentru a retrage date.

**Scrierea iSCSI**

****

O scriere tipică SCSI are nevoie de două parcurgeri dus-întors.Trebuie să aștepte pentru „transfer gata” înainte să trimită date.

Inițiatorul și destinatarul iSCSI poate negocia „Inițial R2R = no” în timpul logării (date nesolicitată).Datele de ieșire de la iSCSI pot urma scrierea SCSI.

**Throughput-ul (lărgimea de bandă) la iSCSI**

SCSI controlează deplasarea datelor.

Graficul de mai jos reprezintă **impactul latenței asupra throughput**:



Throughput-ul reprezintă numărul total de date transmise de la un capăt la altu a latenței.Factori care contribuie la latență sunt : echipamentul ,protocolul , distanța.Cu cât e mai mare distanța cu cât e mai mare latența (Ex:1ms/300km)

Pentru a maximiza througput-ul trebuie ca distanța să fie cât mai mică.Alt factor important este impactul TCP-ului asupra throughput.Acest impact îl descriem mai jos:

-numărul de bytes în derulare = lățimea de bandă \* întârzierea (BW\*Delay)

-pentru a se recupera din erori

-opțiuni de înaltă perfomanță a rețelei precum ferestre mari (RFC1321) și SACK.

-pentru a maximiza throughput-ul:

\*goodput (iesirea = intrarea)

\*fereastra de congestie TCP (cwnd) să fie mai mare sau egală cu produsul dintre lățimea de bandă și întârziere.Dacă avem un cwnd prea mic, throughput-ul este mai mic decât capabilitățile rețelei.Dacă este prea mare (cwnd) apare congestia care duce la retransmisie,la golirea rețelei.

**Creșterea perfomanței rețelei iSCSI**

Pentru a crește perfomața trebuie să înțelegem mai întâi obiectivele precum numărul de utilizatori ai rețelei,numărul de cereri de I/O,perioada răspunsului de acceptare,througput-ul dorit.

Factori care influențează perfomanța:

* Resursele sistemului(CPU,memoria,arhitectura magistralei)
* Resursele de stocare,hard(Ex:RPM,cache,implementarea nivelului RAID pe suprafețe de disk)
* Echipamentul de rețea
* Lățimea de bandă a IP-ului disponibilă (în special la WAN)
* Distanța între inițiatori(emițatori) și destinațiile iSCSI
* Implementarea și configurarea TCP .Exemple precum:

\*creșterea dimensiuni ferestrei TCP pentru a umple pipe-ul(minimul și maximul lățimi de bandă TCP,estimarea întârzieri)

\*evitarea unei rețele îngreunate sau utilizarea Qos pentru dirijarea traficului iSCSI

\*permite suport pentru cadre-jumbo capăt la capăt

\*permite SACK (default)

\*folosirea buferului de trimitere TCP la IPS

* Dimensiunea blocului de I/O (Ex:creșterea dimensiunea blocului de I/O de la 8k la 64k sau mai mult)
* Utilizarea cardurilor TOE pe sisteme cu utilizare CPU mare

 **Scenari de utilizare ISCSI**

Avem următoarele 4 scenari:

1)Foarte puțină întârzieri datorită distanței mici dintr-o rețea de campus ce rezultă un throughput apropiat lățimi de bandă.



Exemple de aplicați:consolidarea discului și partajarea de fișiere.

În acest scenariu luăm în considerare următoarele:

* Consolidarea spațiului de stocare duce la salvarea de spațiu pe disk
* Administrare ușoară a sistemului și a memoriei în comparație cu DAS.
* Securitatea
* O mare rată de hosturi iSCSI și porturi de memorie

2)Foarte puțină întârziere sau deloc în data center



În figură avem un iSCSI ce găzduiește un server NFS, un Filer Nas și un GFPS

Exemple de aplicați:servici Web,gazde iSCSI pentru a oferi backup pentru serverele FC primare.

În acest scenariu luăm în considerare următoarele:

* „oversubscription ISL”și „fan out”
* Integrarea „Blade-server”.Exemplu figura de mai jos:



* Numărul de carduri NIC ce se folosesc
* Partajarea de fișiere

 3) Rețelele metro

Transportul pentru rețelele metro poate fi : metro Ethernet, DWDM, CWDM, Sonet și fibru negru.

Rețeaua suportă o latență mică datorită distanței de metro .De exemplu 1 ms întârziere pe o singură direcție pentru 300 km.

Toate considerațiile de proiectare în rețelele de campus și data center se aplică și aci.

Exemple de aplicați: serviciu de streaming media în distribuția de data în centre,pentru consolidarea memoriei și serviciul de partajare de fișiere.

 **Exemplu de rețea Metro:**

****

4)SAN extins transparent peste distanță pentru DR prin centru de date distribuite.

Exemplu:



Exemple de aplicați:serviciu de partajare de fișiere rapidă pentru utilizatori îndepărtați

În acest scenariu luăm în considerare următoarele:

* OS-uri dispărute și spațiu de stocare
* Impactul distanței (pe distanțe mari ):

\* impact asupra operatorilor/ utilizatorii finali din punct de vedere a experienței utilizatorului

\*impact asupra aplicaților din punct de vedere a perfomanțelor și a througput-ului

 - Aplicarea Qos asupra marginile dispozitivelor WAN

 - Securizarea datelor

 **Alternative Ethernet la iSCSI**

ISCSI este cel mai popular din punct de vedere abordări stocării transmisilor de date prin rețelele de Internet Protocol ( IP ), dar exista alternative , cum ar fi următoarele :

1. Fibre Channel over IP ( FCIP ) , de asemenea , cunoscut sub numele de Fibre Channel „tunneling”, mută date între SAN peste rețele IP pentru a facilita schimbul de date pe o întreprindere distribuită geografic .
2. Internet Fibre Channel Protocol ( iFCP ) este o alternativă la FCIP care îmbină rețelele SCSI și FC în Internet .
3. Fibre Channel over Ethernet ( FCoE ) a devenit un standard oficial în 2009. Împins de către Cisco Systems, Inc. și alți furnizori de rețele , a fost dezvoltat pentru a face Ethernet-ul mai potrivit pentru transportul pachetelor de date pentru a reduce nevoia de Fibre Channel . Dar în timp ce acesta este folosit în switch-uri cu servere Cisco , FCoE este rar utilizată în comutare SAN .
4. ATA over Ethernet ( AoE ) este un alt protocol Ethernet SAN , care a fost vândut comercial , în special prin Coraid Inc. AoE interpretează protocolul de stocare Advanced Technology Attachment ( ATA ) direct de la rețeua de Ethernet , decât să se bazeze pe un protocol de nivel înalt , precum iSCSI cu TCP / IP .

 **Concluzi**

Pentru că folosește standardul Ethernet, iSCSI nu are nevoie de scumpe și uneori complexe switch-uri și carduri care sunt necesare pentru a rula rețele Fibre Channel ( FC). Asta îl face mai ieftin să fie adoptat și mai ușor de gestionat .

O suprafață de stocare FC ( SAN ) transmite datele fără a pierde pachete și suportă în general lărgimi de bandă mai mare , dar tehnologia FC este scumpă și necesită specializare pentru a instala și configura în mod corespunzător . Un iSCSI SAN , pe de altă parte , ar putea fi implementat cu carduri obișnuite de interfață de rețea Ethernet ( NIC) și switch-uri și să ruleze pe o rețea existentă . În loc de a învăța , a construi și de a gestiona două rețele - un LAN Ethernet pentru comunicare utilizator și un FC SAN pentru memorare - o organizație ar putea folosi cunoștințele și infrastructura deja existentă pentru LAN-uri și SAN-uri .