

Les baies de stockage

Maintenant que l'on a défini notre « unité de base », prenons un peu de hauteur et parlons de nos baies de stockage.

Il existe différents types de baies de stockage :

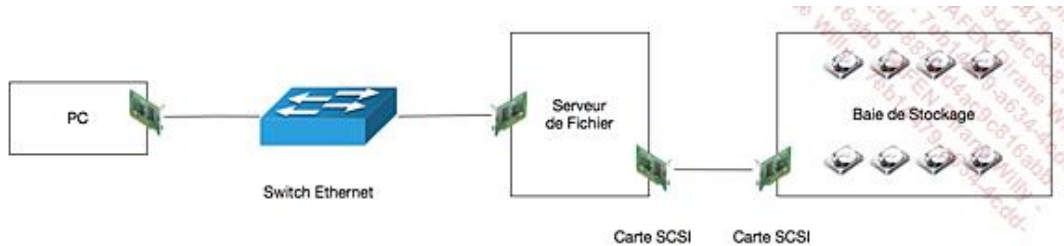
- Les baies de stockage de type Direct Attached Storage ou DAS.
- Les baies de stockage de type Network Attached Storage ou NAS.
- Les baies de stockage de type Storage Area Network.

Quel que soit le type de la baie que l'on utilise, son rôle est de fournir une grande quantité d'espace de stockage, un accès à cet espace rapide.

1. Direct Attached Storage - DAS

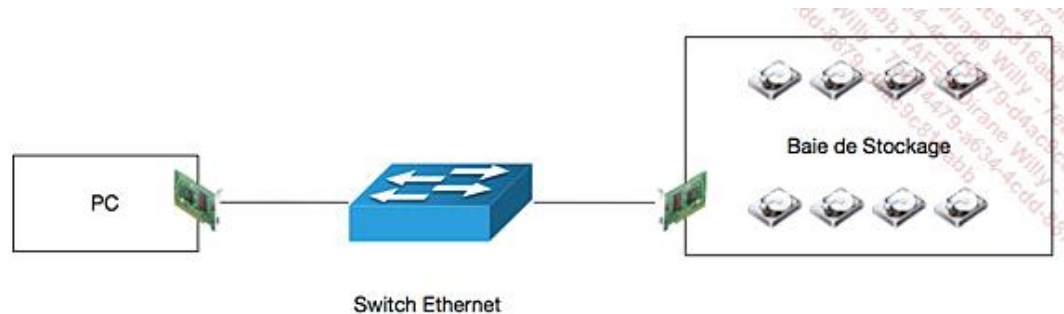
Le DAS est une baie de stockage qui, comme son nom l'indique, est directement connectée à un serveur. Elle est connectée à son serveur par le biais de connectiques diverses et variées allant de l'USB au Fibre Channel en passant par le SATA, la gestion des périphériques de stockage présent dans la baie DAS, se fait à partir du serveur. Il faut aussi inclure un disque dur interne connecté en SATA sur la carte mère comme un DAS.

Le coût est faible mais la gestion peut rapidement devenir « compliquée ».

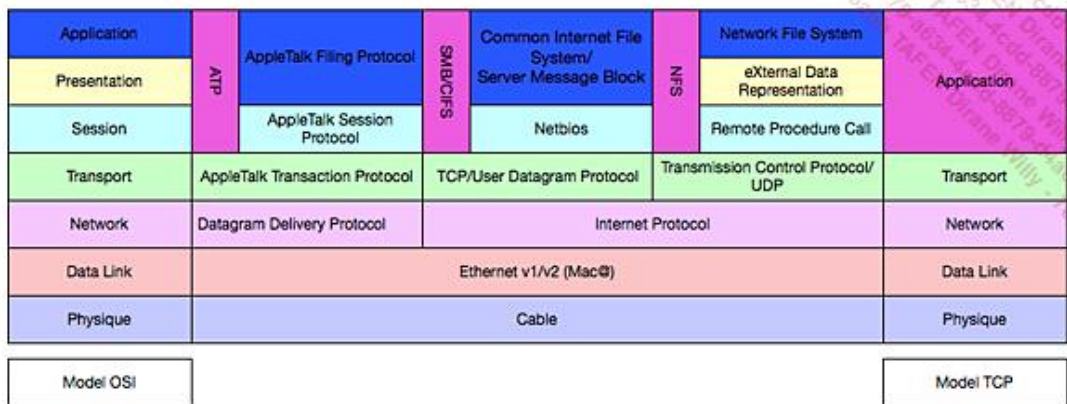


2. Network Attached Storage - NAS

Le NAS est un stockage accessible à travers le réseau, donc accessible par plusieurs machines simultanément. La connectivité des NAS est actuellement basée principalement sur l'Ethernet dont le connecteur est l'adaptateur réseau, ou NIC pour Network Interface Card. Il utilise des protocoles d'interconnexion basés sur TCP/IP. Le NAS est un fournisseur d'accès à des fichiers car il présente un système de fichiers, et s'appuie sur les protocoles applicatifs tels que le CIFS/SMB (Microsoft), l'AFP (Apple) et NFS (Unix/Linux), pour le partager. Le NAS est un espace de stockage dit en mode fichier, cela permet d'avoir accès aux fichiers une fois les droits d'accès configurés.



Comparaison entre les modèles OSI et TCP incluant les protocoles d'accès aux fichiers.



a. Le Network File System

Le NFS est un protocole de fichier distribué initialement conçu par SUN Microsystems en 1984. Il permet à un utilisateur (sur un ordinateur client) d'avoir accès à des fichiers à travers le réseau. Les spécifications du NFS sont disponibles via les RFC suivantes :

- NFSv2 : RFC1094 (1989)
- NFSv3 : RFC1813 (1995)
- NFSv4 : RFC3010 avec les révisions suivantes 3530, 7530 (2000, puis 2003, et 2015)
- NFSv4.1 : RFC 5661 (2010)

Comme nous pouvons le voir dans le schéma précédent, le NFS s'appuie sur les couches de sessions et de présentation du modèle OSI. Ces couches sont respectivement représentées par :

- Le RPC (*Remote Procedure Call*) pour la couche de session.
- L'XDR (*eXternal Data Representation*) pour la couche de présentation.

Le RPC est un ensemble de programmes rassemblant des procédures d'exécution. Il se définit par la combinaison de l'adresse du serveur hôte, le Program Identifier (PID) et le numéro de la procédure spécifique.

Le XDR est un langage utilisé uniquement pour décrire les données.

Maintenant que nous avons vu les composants sous-jacents du NFS, regardons-le du point de vue du modèle TCP.

Le NFS est aussi une application créée et pensée sous UNIX, qui voit un système de fichiers comme un système hiérarchique sous forme d'arbre. Pour ce faire, le NFS s'appuie sur le *Virtual File System* ou VFS qui agrège les différents types de systèmes de fichiers distants comme étant vus comme un système de fichiers local. Il est possible de lui greffer (continuons l'analogie avec notre arbre) des dossiers provenant d'autres serveurs via le protocole MOUNT (et non la commande Unix mount). Le protocole mount permet au serveur NFS d'avoir un accès à des ressources (fichiers et répertoires) d'un serveur distant comme une ressource locale. Pour ce faire, le NFSv2 se basait principalement sur les UID (*User Identifier*) et GID (*Group Identifier*) des permissions du système de fichiers.

Le NFS est un protocole dit « stateless », c'est-à-dire que la gestion de l'information se fait au niveau du client

(historique des requêtes). Un client renverra autant de fois que nécessaire la même requête jusqu'à avoir un résultat de la part du serveur NFS.

Il est aussi dit « idempotent », c'est-à-dire que quel que soit le nombre de fois que la même requête est envoyée, tant que le résultat de la requête ne change pas, il n'y a pas d'effet négatif.

Le NFSv3 apporte principalement les évolutions suivantes :

- Support des fichiers de plus de 2 Go
- Les autres évolutions concernant des modifications des procédures (<https://tools.ietf.org/html/rfc1813#page-11>).
 - La gestion de la taille des fichiers (64 bits)
 - Le client et le serveur négocient la taille maximale du transfert (max transfer size) qui était auparavant à 8 ko

Le serveur NFS se base sur les démons ou services suivants (https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/ch-nfs.html#s2-nfs-how-daemons) :

- Nfsd : service qui gère les requêtes des clients NFS.
- Mountd : service utilisé par le nfsd pour fournir au client la liste des systèmes de fichiers disponibles.
- Lockd : qui est présent tant sur le client que sur le serveur, gère les requêtes de verrouillage depuis le lockd du client vers le lockd du serveur. Il fait partie du protocole NLM (*Network Lock Manager*)
- Statd : qui travaille de pair avec lockd et qui note que le client a fait une demande de réservation et qui surveille le client. Il fait partie du NSM (*Network Status Monitor*) du protocole RPC.

Le NFSv4.x est lié à :

- Un besoin d'utilisation du NFS via Internet. Cela signifie plus de contrôle de sécurité (utilisation du Kerberos et des listes de contrôles d'accès - ACL).
- Une plus grande interopérabilité avec les environnements Windows (réservation -locking).

Pour avoir plus d'informations sur le NFS, nous vous recommandons de lire les RFCs ou Managing NFS and NIS, 2nd Edition de Mike EISLER, Ricardo LABIAGA, Hal STERN, ou bien encore le document de 15 pages (http://www.snia.org/sites/default/files/ESF/FINAL_SNIA_An_Overview_of_NFSv4-4_20Oct2015.pdf) du SNIA "An Updated Overview of NFSv4".

Avantages/inconvénients

Avantages du stockage NFS

- NFS est très stable et fiable.
- Peut être agrandi à tout moment sans répercussions côté VMware.
- Souvent déjà existant dans la plupart des organisations.
- Stockage optimal (Thin provisioning : n'est alloué que ce qui est consommé).
- Stockage centralisé pouvant s'adapter à toute structure (Ethernet).
- Peut rivaliser avec les performances de la fibre optique avec l'arrivée de NFSv4 et l'Ethernet 10 Gbit.
- Un partage NFS peut bénéficier directement de la déduplication.

Inconvénients du stockage NFS

- L'overhead au niveau CPU est conséquent bien qu'inférieur au protocole iSCSI.
- Seule une session de donnée est active. Ainsi, l'agrégation de lien n'a aucun effet et ne permet pas de dépasser 1 Gbit (il faut une carte 10 Gbit pour augmenter la bande passante).
- Sécurité en retrait (dans la plupart des implémentations)

Le détail du processus de connexion via le NFS est disponible à l'adresse suivante : http://www.eventhelix.com/realtimemantra/networking/NFS_Protocol_Sequence_Diagram_Summary.pdf

3. Storage Area Network - SAN

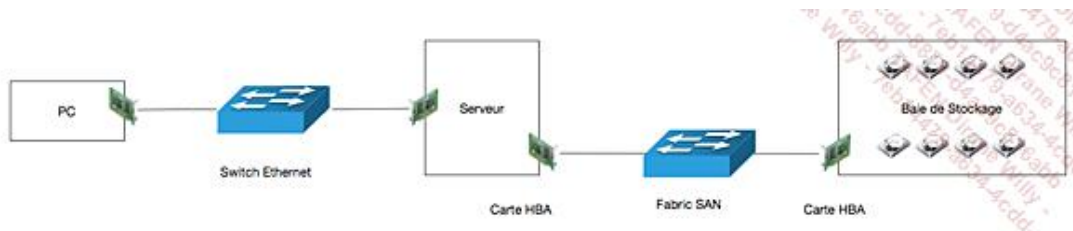
Le SAN contrairement au NAS s'appuie historiquement sur un réseau dédié entre les serveurs et la baie de stockage. Il est à très grande vitesse et faible latence (pour le FC).

a. Fibre Channel

Le Fibre Channel ou FC est un standard défini comme le SCSI par l'INCITS avec le groupe de travail T11, et est suivi par le FCIA (*Fibre Channel Industry Association*).

L'interface pour ce type de réseau est le *FC-Host Bus Adapter* ou FC-HBA, et les modules d'interconnexion sont des Switch Fibre Channel. Lorsque l'on agrège plusieurs switches Fibre Channel entre eux, nous obtenons une « fabrique » (*Fabric*) FC. Historiquement, un réseau FC est en fibre optique et nécessite aussi des connecteurs spécifiques nommés SFP (*Small Form-factor Pluggable* transceiver) ou QSFP (*Quad Small Form-factor Pluggable*). De nos jours, ils sont aussi disponibles en cuivre.

Attention : ne pas confondre (en anglais) Fibre Channel (les protocoles de transport de données : SAN) et Fiber (qui concerne le câblage optique).



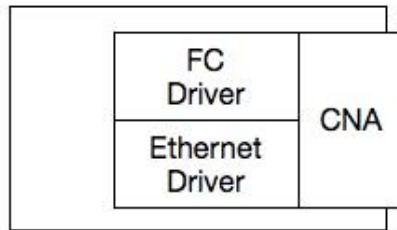
b. Composant du SAN et architecture physique du SAN

Interface de connexion

Dans une architecture Fibre Channel, nous sommes dans une communication de type client/serveur. Le client se nomme Initiateur et peut se trouver sous deux formes physiques :

- FC HBA (*FC Host Bus Adapter*)
- CNA (*Converged Network Adapter*)

Là où la carte FC-HBA est dédiée au réseau FC, la carte CNA permet de connecter deux types d'interfaces FC et Ethernet.



Les CNAs dans le cadre de l'utilisation du FCoE nécessite des équipements réseaux (Switch) compatibles FCoE, qui sont eux-mêmes connectés au SAN via du FC.

Le câblage peut être soit de la fibre optique, soit du cuivre, les vitesses atteintes sont les suivantes (<http://fibrenchannel.org/fc-roadmaps>) :

Product Naming	Throughput (MBps)	Line Rate (GBaud) [†]	T11 Spec Technically Completed (Year) [*]	Market Availability (Year) [*]
1GFC	200	1.0625	1996	1997
2GFC	400	2.125	2000	2001
4GFC	800	4.25	2003	2005
8GFC	1600	8.5	2006	2008
16GFC	3200	14.025	2009	2011
32GFC	6400	28.05	2013	2016
128GFC	25600	4X28.05	2014	2016
64GFC	12800	56.1	2017	2019
256GFC	51200	4X56.1	2017	2019
128GFC	25600	TBD	2020	Market Demand
256GFC	51200	TBD	2023	Market Demand
512GFC	102400	TBD	2026	Market Demand
1TFC	204800	TBD	2029	Market Demand

Les WWNN et WWPN

L'architecture physique liée au Fibre Channel est classique, c'est-à-dire qu'elle est comparable à l'architecture d'un réseau Ethernet. Là où il y a une interface réseau Ethernet, dans un réseau Fibre Channel, nous avons les initiateurs logiciels ou matériels. Les baies de stockages qui hébergent les données auxquelles on cherche à accéder sont nommées cibles (target).

Comme les cartes réseau, qui ont des adresses MAC pour les distinguer les unes des autres, tous les équipements FC ou nœuds (nodes) ont des WWN (*World Wide Name*) ou WWNN (*World Wide Node Name*). C'est une adresse codée sur 64 bits. Il existe deux nomenclatures, dont la structure est la suivante :

Première version de la norme :

10:00:XX:XX:XX:YY:YY:YY ou 20:00:XX:XX:XX:YY:YY:YY

Seconde version de la norme :

5X:XX:XX:XY:YY:YY:YY ou 6X:XX:XX:XY:YY:YY:YY

X : représente l'identifiant de la compagnie

Y : représente les informations spécifiques fournies par la compagnie

Une carte HBA comporte un WWNN, et son ou ses ports comportent des WWPN (*World Wide Port Name*). Pour un switch FC, le switch aura un unique WWNN et de multiple WWPN. Pour les baies de stockages de milieu de gamme, les contrôleurs auront chacun leur WWNN.

Pour les baies de stockage haut de gamme, le WWNN est global et propre à la baie.

En fonction du rôle qu'il joue, et de sa localisation (HBA, Hub, Switch), le port n'a pas le même nom.

Type de port FC	Nom complet	Description	Architecture
N-Port	Node Port	Se connecte à d'autres N-Port ou F-Port	FC-P2P ou FC-Switch
NL-Port	Node Loop Port	N-Port lorsqu'ils sont connectés dans un anneau	FC-AL (Arbitrated Loop)
F-Port	Fabric Port	Port du switch connecté aux N-N-port	FC-P2P ou FC-Switch
FL-Port	Fabric Loop Port	Port du switch connecté aux NL-Port dans un anneau	FC-AL (Arbitrated Loop)
E-Port	Expansion Port	Switch Port connecté à un autre switch Port, formant un Inter Switch Link (ISL)	FC-Switch

Il existe un type de N_port particulier le N_Port ID Virtualization ou NPIV. Le NPIV est un standard du groupe T11.org. Ce standard permet à une carte unique de type FC-HBA de pouvoir enregistrer auprès de la Fabric FC, plusieurs WWPN (*World Wide Port Name*). Chacun de ses WWPN est vu comme un étant un N_port lié à une carte fc-HBA par la Fabric FC.

Chacun de ses WWPN peut être assigné à une machine virtuelle, comme nous le ferions en ajoutant une carte physique dans un serveur. Une fois que cela est fait, nous pouvons directement assigner des LUNs à la VM. Ce mode ne fonctionne qu'avec les RDM ou Raw Device Mapping.

Lorsqu'une machine virtuelle reçoit une carte HBA virtuelle via l'intermédiaire d'une carte NPIV, le fichier VMX de la machine virtuelle est modifié afin d'inclure l'association WorldWide Node Name et WorldWide Port Name. Du point de vue de l'hyperviseur, le VMKernel crée un virtual port (vPORT) - au démarrage de la machine virtuelle-, sur la carte physique FC-HBA de l'hyperviseur se crée un tunnel d'accès à la LUN associée via la Fabric. Ce vPORT est détruit lorsque la machine virtuelle est éteinte. Dans le cas d'une migration (vMotion), le vPORT est clos sur le premier hyperviseur, et ouvert sur le second. Attention, le NPIV ne supporte pas les migrations Storage vMotion.

La baie

La baie de stockage est composée des éléments suivants :

- Un ou des contrôleurs
- De mémoire additionnelle dédiée
- Des tiroirs (de disques)

Le contrôleur est un ensemble de processeurs et de mémoires dédié à la gestion de l'ensemble des périphériques de stockages auxquels il a accès (RAID, LUN, Vol...), ainsi qu'à la gestion des chemins pris par les I/O. Par souci de résilience, il y a souvent deux contrôleurs sur une baie. Ils sont nommés SP ou Storage Processors. Historiquement

les SP n'étaient pourvus que d'un seul type de connectique (FC). Les équipements actuels permettent la mutualisation des couches FC et Ethernet.

Les tiroirs embarquent les périphériques de stockage de données.

Disk Array Extension - DAE

Le DAE que l'on nomme plus couramment tiroir ou shelf est une extension pour une baie de stockage ne contenant que des périphériques de stockage (disque dur et/ou SSD). Il est relié à la baie par de la fibre optique ou du cuivre.

c. Le Fibre Channel

Le modèle Fibre Channel

Le fonctionnement de Fibre Channel est complètement différent des modèles OSI et TCP. Le modèle FC est composé de cinq niveaux que l'on nomme de FC 0 à FC4.

Chaque niveau à un rôle bien spécifique :

- FC0 : ce niveau correspond au niveau physique du modèle OSI. Il correspond aux câbles et connecteurs permettant le transit des informations sur un réseau FC.
- FC1 : ce niveau correspond au niveau Data Link du modèle OSI. Il gère le codage/décodage de la donnée,
- FC2 : ce niveau gère la transmission des trames, les créations de sessions, le zoning.

Ces trois premiers niveaux représentent le FC-PH ou *Fibre Channel Physical*

- FC3 : toujours en développement, et n'est pas utilisé actuellement.
- FC4 : correspond aux protocoles de haut niveau (SCSI, ATM, IP...).

Les types d'architecture Fibre Channel

Il existe trois types d'architecture Fibre Channel :

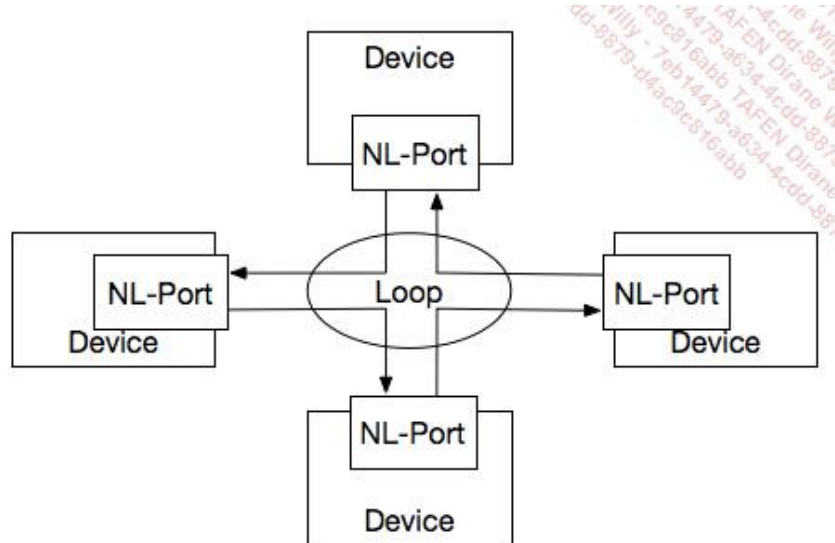
- FC Point à point FC P2P.

Ce type d'architecture peut se rencontrer dans les DAS.

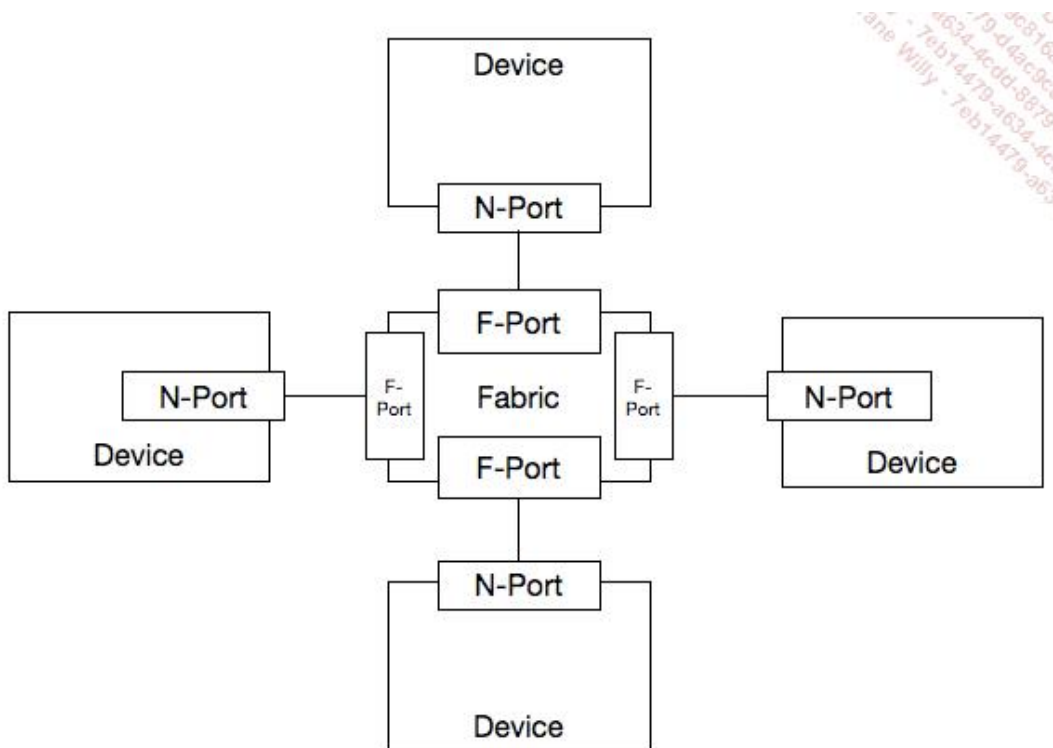


- FC en boucle - FC AL (*Arbitrated loop*)

Ce type d'architecture est similaire à l'architecture Token Ring d'IBM, où chaque élément attend son tour pour discuter.



- FC Switched FC SW



Sécurisation

Le Lun Masking et le Zoning permettent de limiter les accès et connexions à certaines parties de l'infrastructure SAN au périphérique connecté au réseau FC. Les deux techniques sont utilisées de manière complémentaire en général.

Lun Masking

Le Lun Masking est une technique permettant de « cacher » via une gestion des droits des HBA, certaines LUN du stockage. Il s'agit donc d'une technique d'autorisation/révocation d'accès à une LUN.

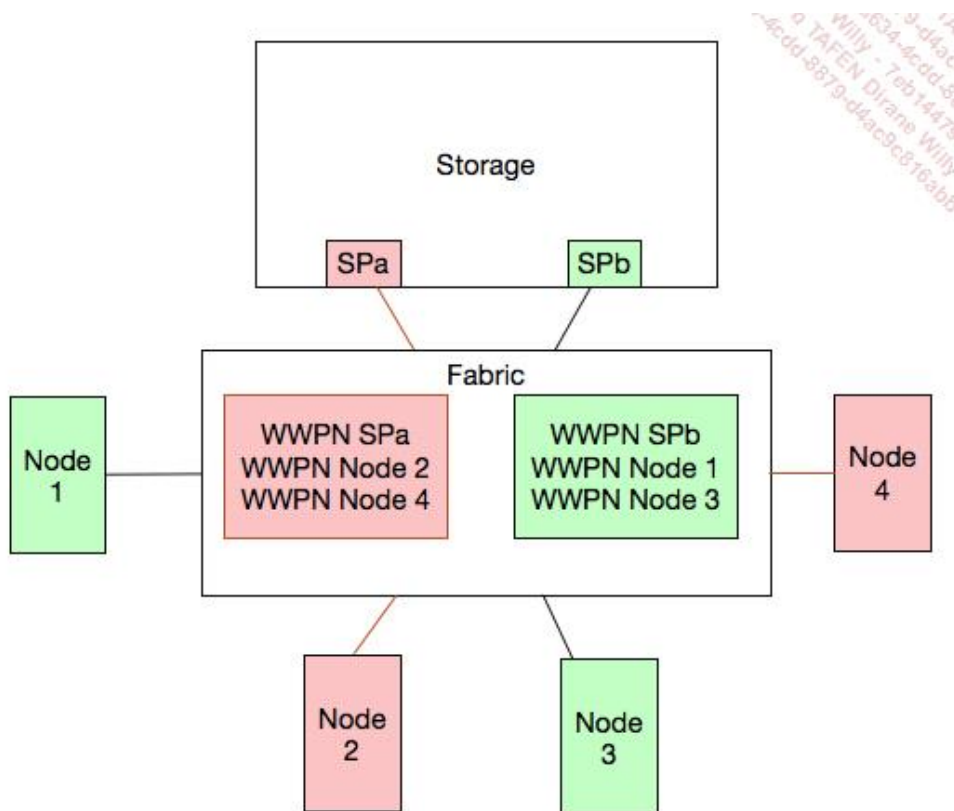
Zoning

Le zoning représente des zones de communication entre des initiateurs et des cibles (targets). Tous les éléments présents dans une même zone peuvent communiquer entre eux et sont indexés dans un annuaire de ressources : le Name Server. Le zoning est l'équivalent d'un Vlan dans un réseau Ethernet.

Au niveau du zoning, on distingue deux types :

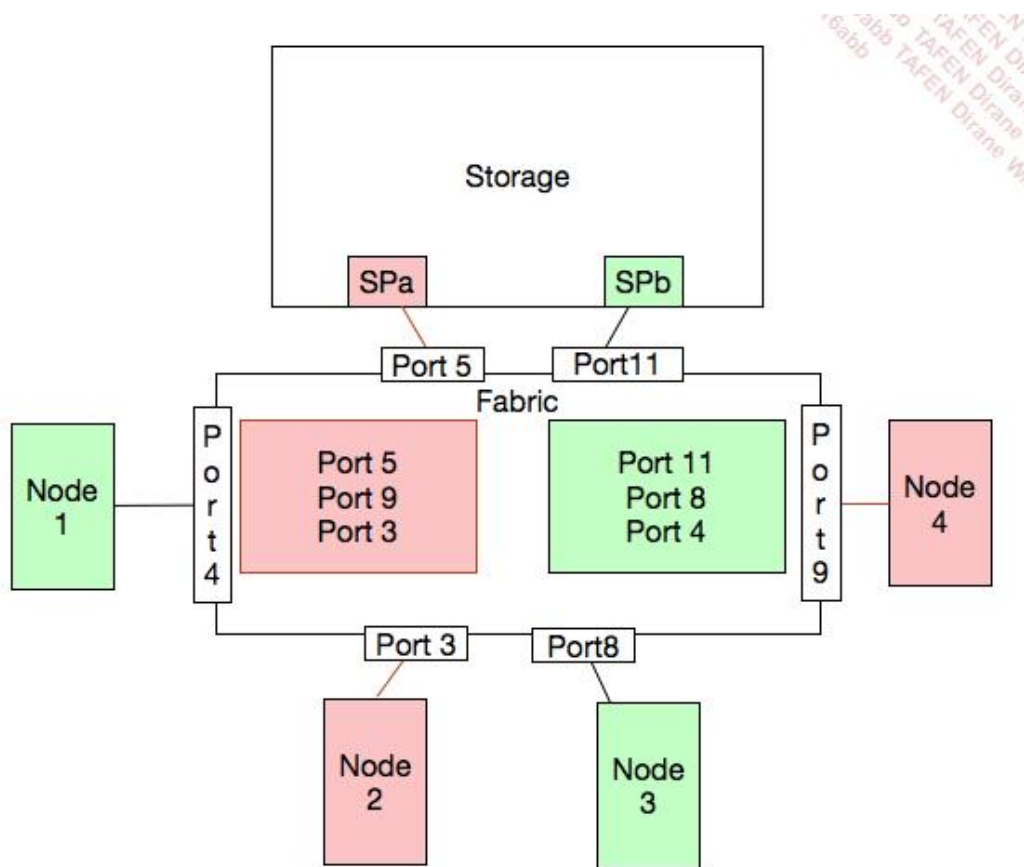
- Soft Zoning

Le Soft Zoning est basé sur l'association entre WWPN dans une même zone. Il est possible d'utiliser des noms symboliques (ou Alias) afin que l'association soit plus aisée à configurer.



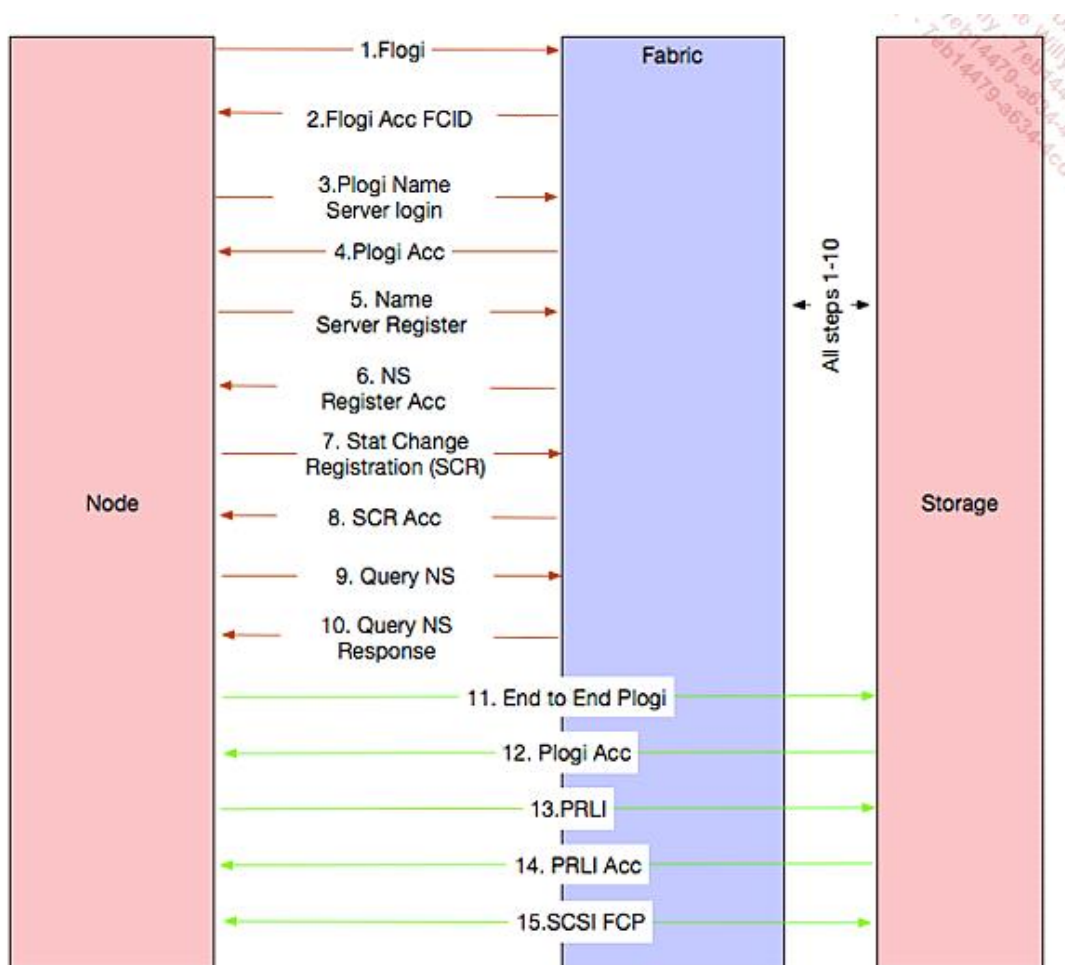
- Hard Zoning

Le hard Zoning se base sur les port_ID (les ports de connexion au niveau du switch FC) auxquels sont connectés les N_Port des équipements de stockage et/ou des serveurs (nodes).



La gestion de la connectivité dans le monde Fibre Channel

Dans un environnement SAN, la connexion allant de la carte FC HBA au périphérique de stockage via la Fabric, se fait en plusieurs étapes.



La première étape est le Flogi correspondant au Fabric Login, où le N_Port cherche à se déclarer au niveau de la Fabric.

La seconde étape est la validation du Flogi par la Fabric. La Fabric renvoie aussi le FCID (*Fibre Channel Identifier*).

Le FCID est un adressage en 24 bits.

23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0								
Domain								Area								Port/Node AL															
Fabric Identifier																Port Physical Address															
Adresse 24 Bits																															

Domain correspond à l'identifiant unique du switch ou de la Fabric.

Area est utilisé pour identifier le port sur la Fabric ou un groupe de ports.

Port correspond à l'adresse attachée au N-Port ou NL-Port.

À ce niveau, la carte HBA (N_Port) est authentifiée au niveau de la Fabric (F_Port), mais la connexion n'est pas fonctionnelle.

L'étape trois correspond à la seconde authentification, le Port Login : le Plogi. Le port de la carte HBA (N_Port) cherche à s'enregistrer au prêt du serveur de noms de la Fabric.

L'étape 4 est la prise en compte par le port de la Fabric (F_Port) de la demande et la génération de la réponse

(acceptation) contenant les informations concernant la Fabric et le F_Port.

L'étape 5 est l'enregistrement de la carte HBA (N_Port) dans l'annuaire de serveur de noms.

L'étape 6 correspond à l'acceptation de l'enregistrement par la Fabric.

L'étape 7, le port de la carte HBA (N_Port), demande à la Fabric l'ensemble du contenu concernant l'état tous les enregistrements (*Registered State Change Notification* - RSCN) à chaque changement au sein de la Fabric (Zoning soft et hard, ajout de domaine ou de switch).

L'étape 8 est l'acceptation de l'enregistrement du changement d'état (*State Change Registration* - SCR).

L'étape 9 : à ce moment-là, le serveur de noms est contacté par la carte HBA (N_Port) afin de connaître l'ensemble des ressources accessibles.

L'étape 10, la Fabric, renvoie les informations du serveur de nom ainsi que les informations de zone (hard et soft zoning, membres de la zone - WWPN, WWNN).

Les étapes 11, et 12 correspondent à une connexion Port Login de point à point, entre un client (Node) et le stockage (Storage).

Les étapes 13 et 14 correspondent au Process Login, il est utilisé dans la configuration de l'environnement des processus entre le client et la target.

Avantage/inconvénient

Avantages du stockage SAN fibre

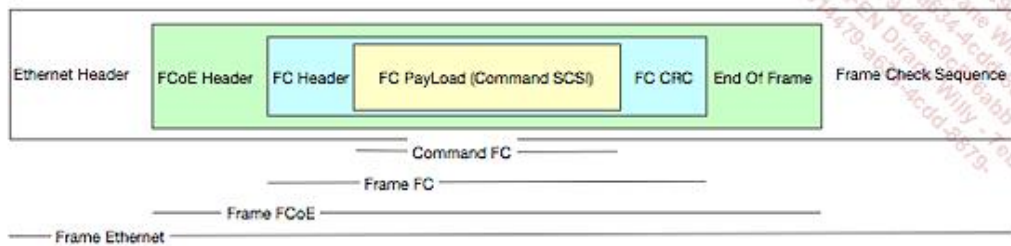
- Utilisé depuis longtemps en production et donc très mature.
- Performances excellentes dans des environnements multiserveurs.
- Sécurité accrue (LUN Masking + Zoning) et réseau dédié.
- Accès au niveau bloc.
- Peu de consommation de ressources au niveau des serveurs (hyperviseurs).
- Latence faible (particulièrement utile pour certaines applications critiques).

Inconvénients du stockage SAN fibre

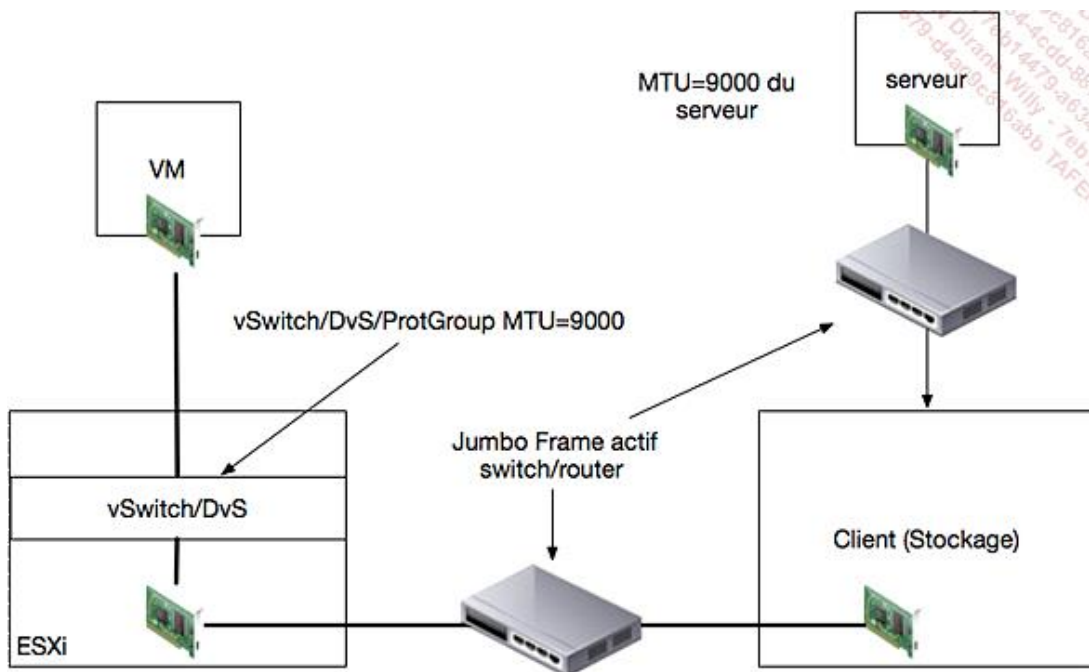
- Le prix.
- Nécessite une infrastructure SAN, donc Switch SAN, fibre optique, etc.
- Complexité grandissante lorsque l'environnement s'étend (notion de directeurs, fabrics, etc.).
- adapté aux applications générant de nombreuses E/S (type messagerie, ERP, BDD).

d. Fibre Channel over Ethernet

Le FCoE est une évolution du FC. Cette évolution permet de réduire les coûts en s'appuyant sur les infrastructures Ethernet déjà en place dans les entreprises, tout en s'inscrivant dans le cadre de la convergence des centres de données (DCB - *Data Center Bridging*) et d'un réseau Ethernet sans perte de paquets (LossLess). Le FCoE permet d'encapsuler des commandes FC dans les réseaux IP.

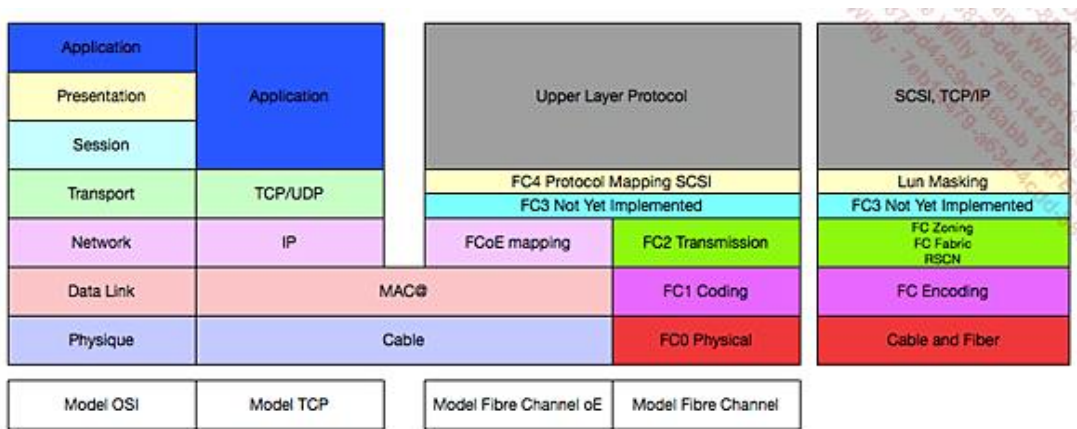


Cette mise en place de type « poupées russes » autour de la commande SCSI, a un impact conséquent sur la taille du paquet Ethernet. Il dépasse la taille standard du MTU (*Maximum Transmission Unit*) qui est de 1 500 par défaut. Ces paquets réseau peuvent atteindre une taille allant jusqu'à 9 000. Ce sont ce que l'on appelle les Jumbo Frames. Lorsque l'on décide d'utiliser des Jumbo Frames, il est obligatoire que l'ensemble des éléments réseau constituant la chaîne soient configurés avec un MTU de 9 000 (switch - physique et virtuel - et carte Ethernet).



L'infrastructure nécessite alors des équipements compatibles FCoE permettant l'interopérabilité entre les technologies Ethernet (IP) et Fibre Channel (FC). Une fois encore il n'y a pas de comparaison possible entre les modèles OSI et TCP par rapport au modèle Fibre Channel.

Les équipements FCoE nécessitent une compatibilité DCB switch sans perte de paquet, (Data Center Bridging - LossLess Ethernet Switch) et FC. Ces switches sont appelés FCF (*FCoE Forwarders*).



Au niveau VMware, il faut absolument relier l'adaptateur software à une carte physique. Pas besoin de préciser le VLAN puisque l'adaptateur est relié à un vSwitch, et FIP (*FCoE Initialization Protocol*) découvrira automatiquement le VLAN ID.

Avantage/inconvénients

Avantages du stockage FCoE

- Possibilité de converger vers le tout Ethernet.
- Permet de s'intégrer facilement afin de continuer à utiliser des baies FC ou Ethernet.
- Bénéficie des capacités montantes de l'Ethernet (100 Gbit prévus d'ici peu...).
- Supporté par Cisco (notamment sur l'UCS).
- Initiateur Software possible également.
- Moins d'Overhead généré par rapport à iSCSI.

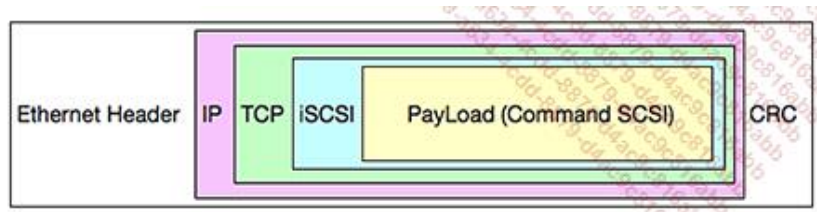
Inconvénients du stockage FCoE

- Jeune
- Nécessite un réseau Ethernet Loseless, ce qui induit des transformations complexes sur le réseau (mais peut être fait sur un périmètre restreint).
- Complexité totalement déportée côté réseau.
- L'investissement côté équipements réseau non négligeable.
- Les standards sont en préparation (entre FCoE - DCB et TRILL). Il faudra attendre pour que le marché soit totalement mature.
- FCoE n'est pas routable contrairement à iSCSI qui l'est nativement (FCIP et iFCP arrivent...).

e. Internet Small Computer System Interface

L'iSCSI est un protocole lié au SAN, comme nous l'avons vu dans le modèle architectural du SCSI. Il s'appuie sur Internet (IP) pour transmettre des commandes SCSI. Il est lié aux spécifications définies par le groupe du T10.

Schéma de la trame réseau iSCSI :

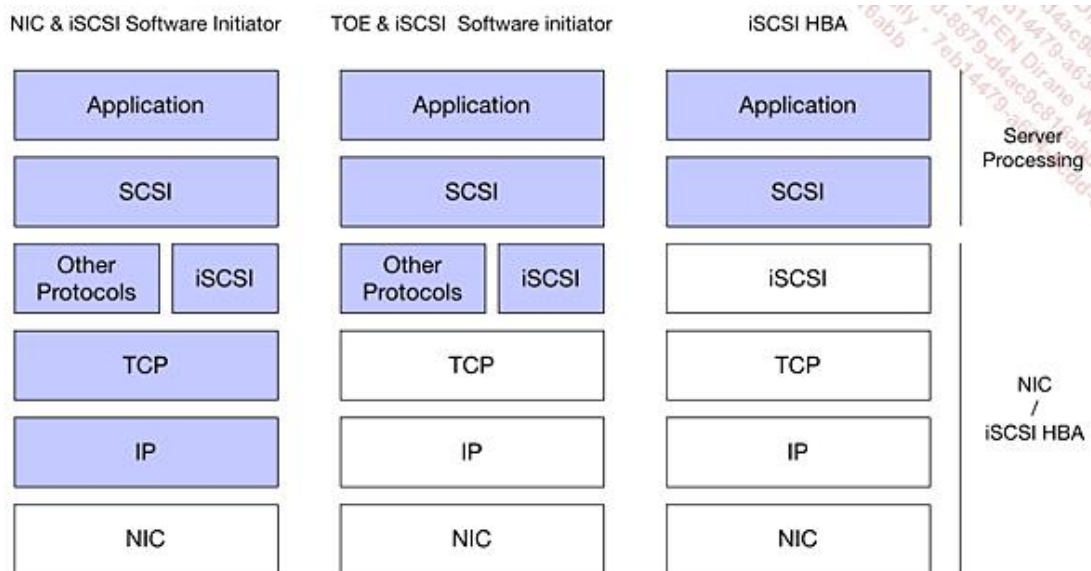


Les types d'interfaces pour l'iSCSI

Dans le monde iSCSI, nous sommes dans une architecture de type client/server. Le client est l'initiateur et le serveur est la cible (target) qui correspond aux stockages.

L'initiateur peut se trouver sous deux formes :

- Physique (hardware-dependent ou hardware-independent) : iSCSI Host Bus Adapter (iSCSI HBA).
- Logiciel (software initiator), qui se trouve dans le système embarqué dans le système et s'appuie sur une carte réseau Ethernet.



IQN, NAA et EUI

Comme pour les cartes Ethernet avec leurs adresses MAC ou les FC HBA avec les WWPN et WWNN, les HBA iSCSI ont leur équivalent : des noms. Les noms sont associés au nœud et non à la carte HBA elle-même. Pour un même nœud, il est possible d'avoir trois noms différents en fonction du standard de nom utilisé.

- iSCSI Qualified Name ou IQN

Pour générer des IQN, il « suffit » de générer un nom de domaine. Attention, nous ne parlons pas obligatoirement d'un nom de domaine DNS, l'utilisation de l'identifiant IANA (*Internet Assigned Numbers Authority*) de l'entreprise est possible.

La structure de l'IQN est la suivante : IQN.[Date (yyyy-mm)].[Autorité de Nom en sens inverse]:[nom unique]

- Extended Unique Identifier ou EUI

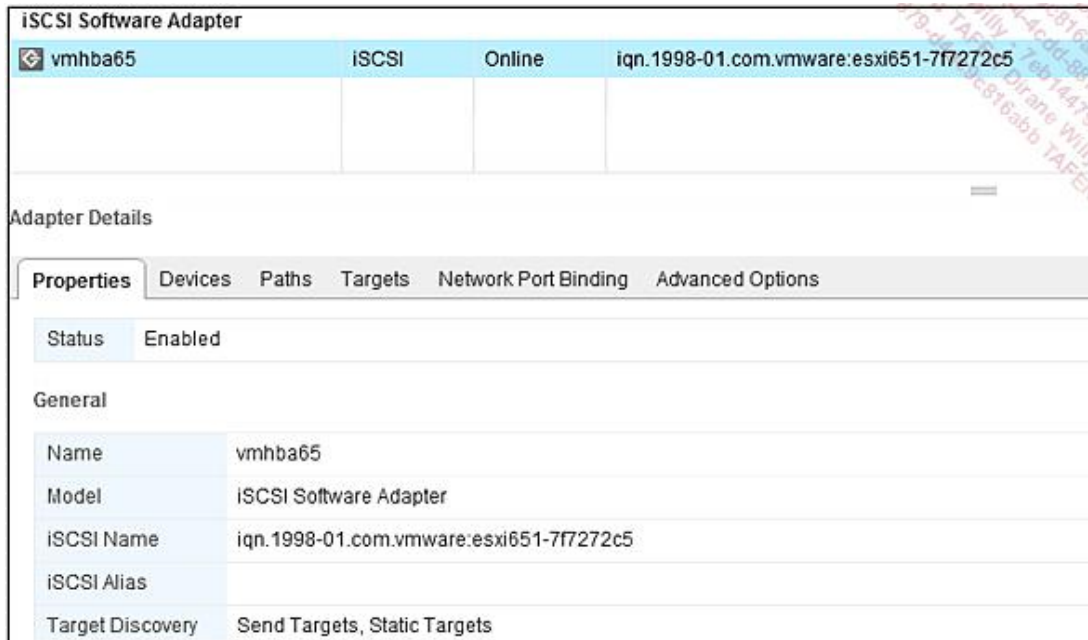
La structure de l'EUI est la suivante : EUI.[Identifiant sur 64 bits]. Sur ces 64 bits, les 24 premiers bits représentent l'identifiant IEEE de la société qui a conçu l'interface. Les 40 autres bits sont assignés par cette même société. Ils sont uniques.

- Network Address Authority ou NAA

Cette norme est issue du T11. Sa structure est la suivante : NAA.[Adresse sur 64 ou 128 bits].

Dans le cadre de l'utilisation de l'iSCSI, VMware utilise les formats IQN et EUI pour les cartes HBA.

En sélectionnant dans la liste des adaptateurs de stockage, lorsque l'on clique sur la carte iSCSI logique, nous avons les informations de nom et de type de carte.



Fonctionnement de l'iSCSI

Comme nous l'avons vu précédemment, l'iSCSI se base sur l'IP. Les composants nécessaires sont les suivants :

- L'initiateur
- La cible
- La session iSCSI et la connexion iSCSI
- Le portail iSCSI

L'initiateur et la cible, ou target, sont des sujets que nous avons traités précédemment. Mais pour faire un rappel, rapide, l'initiateur correspond à une interface iSCSI (matériel ou logiciel). La cible représente la baie de stockage accessible via une connexion réseau.

Avant de pouvoir agir sur les LUNs, on les « découvre » en configurant l'initiateur.

Lors de la session de découverte, il est possible d'utiliser soit :

- Une session de découverte dynamique qui utilise le nom DNS ou l'IP de la baie. L'initiateur envoie une requête de type SendTarget à la baie. La baie lui répond en envoyant la liste des IQNs et LUNs accessibles.
- Une session de découverte statique qui s'appuie sur le nom DNS ou l'IP de la baie ainsi que sur le nom IQN de la LUN.

Une fois la LUN découverte, une session de connexion iSCSI est établie en s'appuyant sur le TCP/IP. Une session

de connexion iSCSI a un identifiant : ISID (*iSCSI Session ID*).

Un ISID peut contenir plusieurs sessions TCP. Cela permet de faire de la répartition de charge ou de l'agrégation de bande passante au niveau des connexions TCP.

Le portail est le nom que l'on donne à un composant réseau ayant une adresse IP et pouvant être utilisée par un nœud iSCSI (initiateur/target) afin d'accéder à une session iSCSI. Dans le cas de l'initiateur, le portail est défini par son adresse IP. Dans le cas de la cible, le portail est défini tant par son adresse IP que par le port d'écoute (3260).

Avantage/inconvénient

Avantages du stockage iSCSI

- Excellent rapport qualité/prix.
- Stockage centralisé pouvant s'adapter à toute structure (Ethernet).
- Accès au niveau bloc.
- Pourra aller au-delà des performances fibres dans un avenir proche.
- Peut être routé donc très facile à utiliser dans du multisite.
- Souvent utilisé pour du PRA (la réplication iSCSI est très simple et efficace).
- Nécessite peu de compétences stockage mais plutôt réseau.

Inconvénients du stockage iSCSI

- Performances en retrait par rapport à la fibre.
- Le passage à 10 Gbit coûte relativement cher.
- La sécurité : il n'y a plus de ségrégation Stockage (SAN FC/réseau Ethernet).
- Peut nécessiter une isolation physique (non prévue au départ) des réseaux iSCSI pour des performances maximales.
- Très impactant en overhead (consommation de ressources nécessaires au niveau hyperviseur).
- Le diagnostic est rendu beaucoup plus complexe car les performances/coupures peuvent aussi provenir du réseau.