

Pierre-Olivier BOURGEOIS - Alexis MARCOU

Storage Area Network (SAN)

26 juillet 2004

Table des matières

Table des matières	1
I Introduction	3
II Le Stockage : Un enjeu pour les entreprises	4
1 Des données de plus en plus inaccessibles	4
2 Le partage des ressources de stockage	4
3 Des ressources hétérogènes	4
4 Protection des données contre les sinistres	5
III Le stockage en réseau : deux technologies	6
5 Le NAS : Une vraie solution ?	6
6 Le SAN : Ses promesses	6
6.1 Intégrité et disponibilité	7
6.2 Espace d'échange de données	7
6.3 Souplesse d'évolution	8
6.4 Performance	8
6.5 Indépendance du stockage	8
6.6 La virtualisation du stockage	8
IV La technologie SAN	9
7 le réseau SAN	9
7.1 L'utilité d'un réseau dédié	9
7.2 Une technologie encore assez jeune	9
7.3 Les topologies du réseau NAS	9
8 Les interfaces	10
8.1 Le SCSI Parallèle	10
8.2 Le SCSI Série	11
9 Fibre Channel	12
9.1 La topologie FC	13
9.2 Les couches de FC	14
10 la virtualisation	18
10.1 La technologie RAID	18
10.2 La virtualisation dans les reseaux SAN	18
11 Internet Fibre Channel Protocol (iFCP)	19
11.1 La description de iFCP	19
11.2 Architecture d'une passerelle iFCP	20
12 le iSCSI (ou le SCSI over IP)	20
V Conclusion	23

Première partie

Introduction

Le système d'information d'une entreprise est considéré comme un capital. L'entreprise qui saura l'exploiter le plus rapidement et de la manière la plus efficace sera gagnante.

Mais, le problème qui est posé aux entreprises, c'est la croissance exponentielle de la taille de l'information. Des questions se posent :

- Comment accéder à des bases de données de plus en plus volumineuses avec des temps d'accès acceptables ?
- Comment réussir à faire les sauvegardes des données, dont la taille augmente tout les jours, dans un délai de sauvegarde qui lui ne change pas ?
- Comment assurer l'interopérabilité des données dans un environnement de systèmes hétérogènes ?

De plus, l'adoption de l'outil informatique dans de nombreux domaines professionnels et l'ouverture des réseaux (Internet) sont des sources d'accroissement des données.

Et afin de tirer le meilleur des données, il faut les analyser, et donc les enregistrer au préalable, sans oublier qu'elles doivent être toujours disponibles.

Dans ce dossier, nous allons vous présenter un réseau dédiée au stockage, le *Storage Area Network* (le SAN).

Deuxième partie

Le Stockage : Un enjeu pour les entreprises

1 Des données de plus en plus inaccessibles

Les entreprises collectent au fur et à mesure des années un volume très important de données. Le fait de récupérer ces données coûte à l'entreprise. Ces données prennent donc de la valeur. Si elles étaient perdues, le fait de ne plus les avoir ou tout simplement de les retranscrire une nouvelle fois dans le système d'information fait perdre de l'argent à l'entreprise.

D'où, l'intérêt pour l'entreprise d'investir dans des solutions de stockages et de sauvegardes performantes.

Mais, on ne résoud pas le problème en ajoutant des nouveaux périphériques de stockage sur le serveur de stockage, l'estimation de la taille des données est très difficile à prévoir. Et l'indisponibilité de données ou le manque de place, provoque un manque de productivité qui fait virtuellement perdre de l'argent à l'entreprise.

Il faut noter que si on augmente la taille des données, la taille des sauvegardes augmentent nettement aussi. Et le problème, c'est que le temps qui est nécessaire pour faire la sauvegarde des données prend de plus en plus de temps. Et la fenêtre de temps de sauvegarde, souvent la nuit lorsque les bureaux sont vides, elle ne change pas. Elle va même jusqu'à se réduire, car le traitement des données, s'effectue souvent avant les sauvegardes et en plus, avec Internet, on doit pouvoir accéder aux données 24h/24h.

Le réseau de l'entreprise n'est pas du tout dédié au transfert. Même avec des réseaux Gigabit Ethernet, il est impossible d'attendre les 100Go par heure. Les données sont freinées par les couches logicielles de communication. Car, on ne parle plus en Giga octet (Go) , mais en Téra octet (To) pour le système d'information des entreprises.

2 Le partage des ressources de stockage

Dans une entreprise, il peut exister plusieurs domaines d'applications spécialisées (ressources humaines, marketing, production,...). Donc, chaque service possède ses propres applications et ses propres données. Les besoins en stockage, de chaque service augmentent continuellement. Comment faire pour que lorsqu'un service, qui est saturé en stockage, se développe beaucoup plus vite que les autres ?

On ne peut pas imaginer, ne pas utiliser l'espace disponible des serveurs d'autres services qui n'utilisent pas entièrement leurs capacités de stockage. Mais, alors, le serveur surchargé sera obligé d'aller chercher l'espace qui lui est nécessaire à travers le réseau de l'entreprise. Ce qui provoquera une très nette dégradation des performances du système applicatif. Ajuster constamment l'espace de stockage en fonction du besoin est un vrai défi.

3 Des ressources hétérogènes

Avec tout les différents systèmes qui existent, on peut facilement s'apercevoir qu'au sein d'une même entreprise, il est très difficile, même impossible d'avoir une infrastructure informatique homogène. Les différents applicatifs ont besoin de tel ou tel environnement. Que ce soit de l'UNIX, du Windows, du Mac ou du Linux, fournir les mêmes données à toutes les plates-formes relève pratiquement du miracle. De plus, administrer des ressources hétérogènes et dispersées à un coût qui peut être jusqu'à 4 fois plus élevé, que dans un environnement homogène.

4 Protection des données contre les sinistres

Les risques de sinistre d'une entreprise ne peuvent être écartés. Il est nécessaire de dupliquer les données dans un autre lieu géographique, et même aller jusqu'à un autre continent, si on veut se mettre à l'abri d'un tremblement de terre. Mais, cela a un prix et il faut prendre en compte le temps de transfert et le coût de ce choix. Car, même si certaines technologies (comme le SCSI) permettent la création d'un miroir instantanément, la longueur maximale des liens introduit une contrainte de distance (à peine quelques mètres).

Troisième partie

Le stockage en réseau : deux technologies

Dans le stockage réseau, deux technologies sont proposées : le *Network Attached Storage* (NAS) et le *Storage Area Network* (SAN)

5 Le NAS : Une vraie solution ?

Avec ce type de technologie, les systèmes de stockage viennent se connecter au réseau de l'entreprise existant. Ils se comportent alors comme un serveur de fichier quelconque. Ils s'appuient sur le mode de communication TCP/IP et s'administrent à travers à des agents implantés dans les serveurs. Ils sont en général compatibles avec plusieurs types de plate-formes. Le NAS est un service au niveau "fichier".

L'avantage, de cette solution, est qu'une unité de stockage est autonome. De plus, elle n'a pas besoin d'une infrastructure spécifique, ce qui rend la solution plutôt économique. Elle est souvent assez évolutive grâce à des possibilités de branchement de nouveaux disques à chaud. La redondance des données se fait au sein d'une seule et même unité de stockage, grâce au technologie RAID.

L'inconvénient principal, de ce type de solutions, est qu'elle vient surcharger le réseau de l'entreprise comme le montre le schéma (cf. FIG.1) . Les données vont de l'unité NAS vers le serveur d'application, puis vers le client. Ensuite, le NAS s'oriente plus dans le cadre d'un serveurs de fichiers autonome. Et le fait de passer par mode de communication non dédié comme TCP/IP ralentit le débit et les temps de réponses d'accès aux données. Pour des serveurs d'applications, avec des traitements lourds et conséquents, ce n'est peut-être pas la véritable solution.

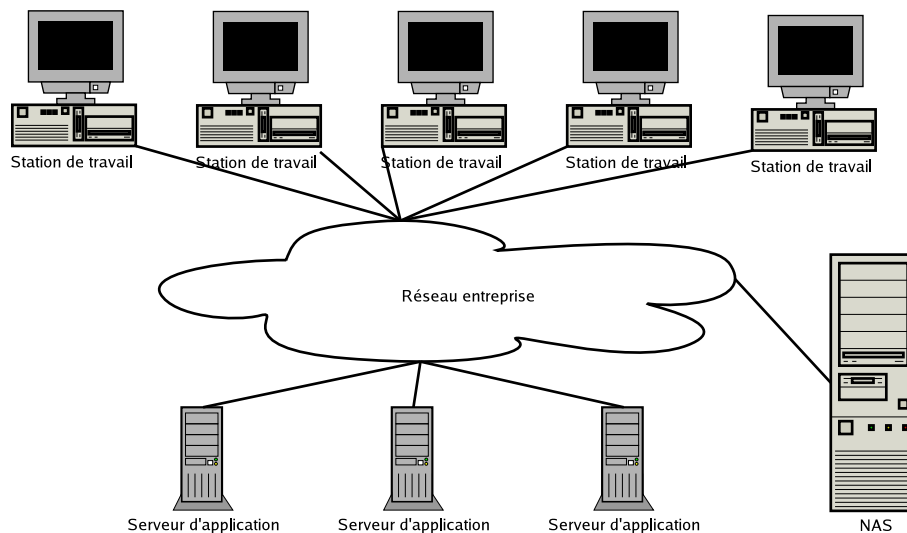


FIG. 1 – Schéma simple de l'utilisation d'un NAS

Donc, La technologie NAS, n'est pas réellement une solution miracle pour le stockage des données. Elle a peut-être l'avantage de centraliser les données, ce qui résoud certains problèmes, mais pour d'autres cas, on déplace le problème.

6 Le SAN : Ses promesses

Le SAN a pour énorme avantage de pouvoir offrir au serveur applicatif des périphériques de stockages distants de plusieurs centaines de mètres avec des temps de réponse comparables à des

liaisons privés et locales. On arrive à obtenir ceci, grâce à l'utilisation d'un réseau dédié au stockage (cf. FIG.2). De plus, le SAN peut être administré à partir d'un seul point.

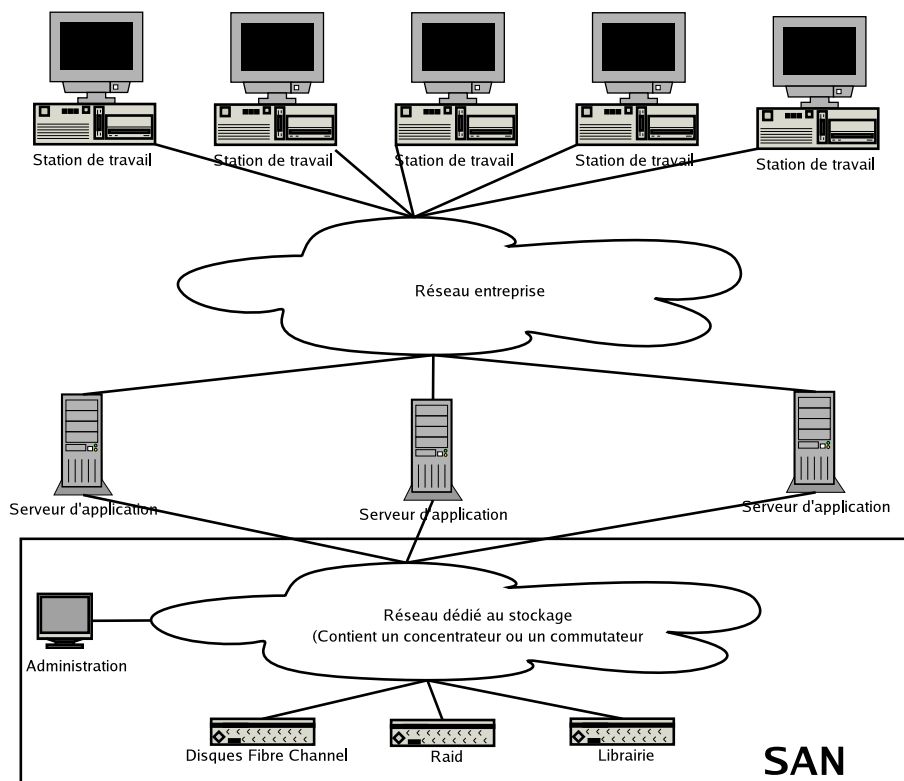


FIG. 2 – Schéma simple de l'utilisation d'un SAN

6.1 Intégrité et disponibilité

Par sa conception, le SAN offre une redondance naturelle. Ayant un réseau dédié, il suffit de multiplier les chemins d'accès aux périphériques de stockage pour avoir un simple miroir.

De plus avec des systèmes de passerelles dont dispose un SAN, il est très simple de dupliquer les données vers un centre à l'autre bout du monde, pour avoir une complète protection contre les sinistres.

6.2 Espace d'échange de données

Il est possible de partager des données à travers le SAN.

6.2.1 Partage physique

On partage un espace de stockage pour chaque serveur. Tout se passe comme si chaque serveur possédait ses propres ressources. Les serveurs peuvent être hétérogènes. Mais, un système d'isolation permet de préserver l'intégrité des données.

6.2.2 Partage en lecture

Ce deuxième niveau de partage, pour des serveurs hétérogènes, permet d'accéder en lecture aux données d'un serveur d'application, par exemple un serveur NT peut exploiter les données générées par un serveur UNIX.

6.2.3 Partage coopératif

Pour ce niveau de partage, il permet de travailler en écriture et en lecture sur des mêmes données. Malgré que le SAN soit bien conçu pour ce genre d'opérations, on comprend vite que la cohérence des données ne puisse être conservée sans qu'elle soit implémentée au niveau des serveurs d'applications. Certaines implémentations sont en cours, notamment, en ce qui concerne les serveurs en clusters.

L'avantage de ce genre de partage, est qu'il est possible, de faire des sauvegardes de manières rapides, entre deux périphériques, sans surcharger le réseau de l'entreprise.

6.3 Souplesse d'évolution

Le SAN permet une réelle souplesse d'évolution. L'architecture d'une SAN par le fait d'utilisation d'un réseau dédié, permet l'ajout d'espace de stockage, en ajoutant des disques Fibre Channel, sans avoir à rendre indisponible le reste des données. C'est aussi le cas des systèmes de stockages locaux. Mais un SAN n'est pas limité par une faible connectivité et une complexité de câblage.

6.4 Performance

Les performances du SAN ne sont pas limitées. L'utilisation d'un réseau dédié ultra-rapide en est une des causes. Les différents éléments d'un SAN autorisent une croissance modulaire de la bande passante. Même si la bande passante nominale est limitée, il est possible d'installer des dizaines, voire des centaines de chemins de données indépendants, qui conduit à une bande passante quasi illimitée.

En plus, une évolution du débit nominal s'ajoute à la démultiplication des chemins de données

6.5 Indépendance du stockage

Le SAN permet ainsi de ne pouvoir dissocier le stockage des serveurs d'applications. Mais aussi, l'administration des données est ainsi facilitée, car centralisée.

6.6 La virtualisation du stockage

On alloue un espace de stockage à chaque serveur. Mais le serveur ne sait pas exactement comment se présente physiquement le stockage. Son espace de stockage correspond à une partie de l'espace virtuel disponible.

Cette virtualisation permet une réorganisation dynamique en fonction du besoin de l'entreprise. Tout ce qui est attrait au stockage n'est plus du tout géré par leur serveur d'application, mais par le SAN (Sauvegarde, mirroring,...) (cf. FIG.3).

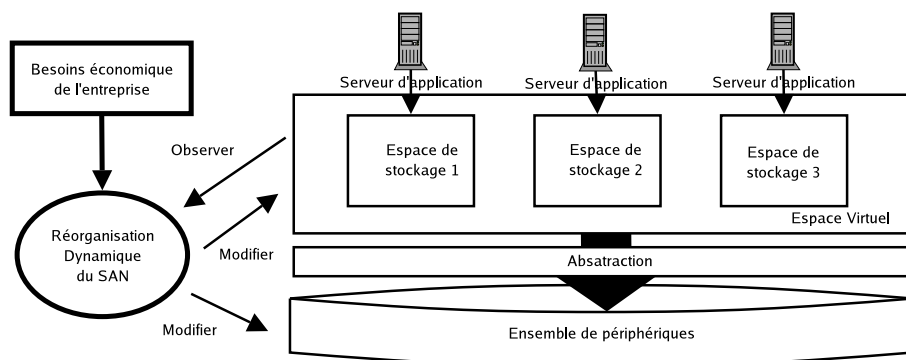


FIG. 3 – Abstraction des éléments physiques pour les serveurs d'applications

Quatrième partie

La technologie SAN

Avant le *Storage Area Network*, les architectures de stockage se reposaient sur le bus parallèle SCSI (*Small Computer Systems Interface*). Mais le principale inconvénient de ce bus était sa portée qui ne dépassait pas quelques dizaines de centimètres et ne pouvant connecter qu'un ensemble limité de périphériques.

Mais l'arrivée de la fibre optique a tout changé. On a pu ainsi externaliser le stockage en utilisant, par exemple le Fibre Channel. La mise en l'oeuvre de réseaux dédiés au stockage a alors pu débuter.

7 le réseau SAN

7.1 L'utilité d'un réseau dédié

La création d'un réseau dédié au stockage, a tout d'abord été opérée pour désengorger le réseau local, le LAN (*Local Area Network*), dépassé par l'explosion de la masse de données produite par l'ère numérique.

Le SAN peut ainsi prendre en charge le trafic nécessaire pour transférer le trafic lié aux stockages et aux sauvegardes. L'utilisation de Fibre Channel, permet d'atteindre des taux de transfert allant jusqu'à 100 Mo/s sur une centaine de kilomètres. Ce qui permet de transférer des données entre différentes filiales.

7.2 Une technologie encore assez jeune

Les travaux de standardisation se poursuivent, car encore assez simple de créer des SAN stratégiques, la gestion de ces solutions restait encore réalisable pour une équipe de spécialistes. Mais, pour concevoir une SAN au niveau de l'entreprise, cela s'avère délicat, surtout dans un contexte multiconstructeurs. Mais, les principaux fournisseurs en matériels pour les SAN commencent à atteindre un standard permettant l'interopérabilité entre les différents matériels.

7.3 Les topologies du réseau NAS

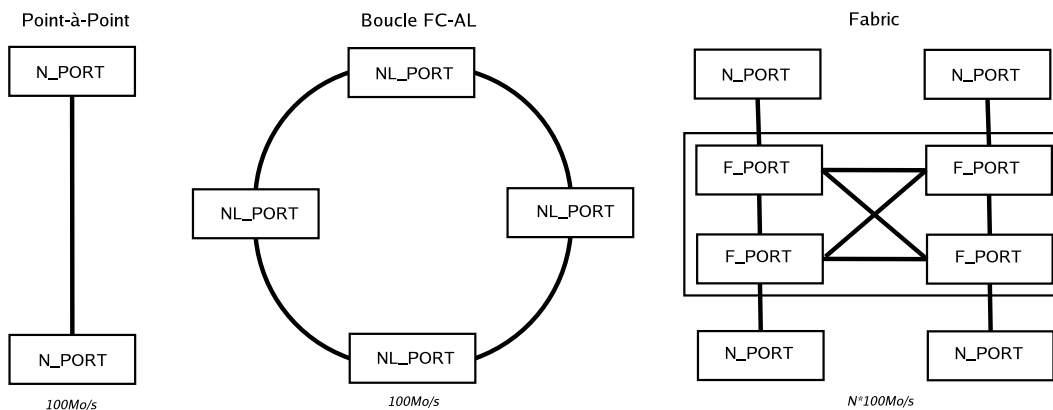


FIG. 4 – Différentes topologies utilisées par le Fibre Channel

Comme le montre la figure (cf. FIG.4), il existe trois types de topologie.

7.3.1 L'architecture point à point

C'est la plus simple. La bande passante est entièrement dédiée aux deux équipements. C'est la première architecture qui a été mise en oeuvre.

Elle a un énorme défaut, lors de l'ajout de matériel, elle oblige à relier ce matériel avec tous les autres existants avec lesquels il aurait besoin de communiquer. Cette architecture n'a pratiquement plus lieu d'être.

7.3.2 L'architecture de type boucle arbitrée

On retrouve ici, le *Fibre Channel Arbitrated Loop* (FC-AL).

Ils est amusant de faire le parallèle entre les réseaux locaux et les réseaux SAN. Ils suivent à peu près les mêmes évolutions.

Cette architecture est basée sur une boucle, à laquelle tous les équipements sont reliés. Cette architecture est limitée à 127 points de connexion. On utilise parfois une double boucle pour permettre de doubler la bande passante, car les équipements sont équipés de deux ports de communications. Il existe aussi un système qui permet de supprimer de la boucle tout équipement qui tomberait en panne. Cela permet d'éviter que la boucle soit rompu.

7.3.3 Les commutateurs (ou Fabric)

C'est la dernière architecture. Elle assure au même titre qu'un réseau local commuté, une bande passante maximale entre deux équipements. Ils permettent ainsi une meilleure exploitation de la bande passante. De plus, ils permettent aussi la redondance des liens en cas de rupture de liaison.

Autrefois, limité par un prix prohibitif des commutateurs, le prix par port s'est effondré. Cette architecture se répand un peu partout, car elle permet la création de réseau de stockage SAN à la grandeur de l'entreprise.

Cette architecture est limitée théoriquement par le nombre d'adresses. Une adresse est codée sur 224 bits, soit environ 16 millions d'adresses.

Le rôle du commutateur est pratiquement identique à ceux utilisés dans les réseaux locaux. L'émetteur a juste besoin de créer une liaison point à point avec le commutateur.

8 Les interfaces

8.1 Le SCSI Parallèle

8.1.1 Vue globale

Le bus SCSI a, dès son origine, été développé afin d'offrir une interface entre les unités nécessitant de forts débits de données (disques durs par exemple) et les ordinateurs.

Différentes normes sont apparues depuis sa création. Dans ces normes, différents paramètres entrent en ligne de compte et conditionnent le débit, le nombre d'unités connectables ainsi que la longueur maximale des câbles.

- La fréquence du bus, qui varient de 5 à 160 Mhz ;
- la taille du bus, qui va de 8 à 32 bits ;
- le nombre de connecteurs (ou nombre de fils du câble) : entre 50 et 136 fils.

Ainsi, la longueur maximale des câbles varie entre 1,5 et 25m, sachant que celle-ci est limitée par l'augmentation du débit. Les débits rencontrés vont de 3 à 360 Mo/s selon la norme utilisée. Enfin, le nombre d'unités que l'on pourra connecter dépendra de la largeur du bus. En effet, un

bus 16 bits pourra accueillir 16 unités au maximum (incluant l'interface hôte).

La première norme SCSI (SCSI-1), mise sur le marché en 1986, fonctionnait à 5Mhz, sur un bus 8bits avec des napes de 50 fils. Les débits variaient entre 1,5 à 8 Mo/s selon le mode d'échange utilisé (synchrone ou asynchrone).

La dernière norme en date est l'Ultra-4 Wide SCSI (Ultra 320) qui fonctionne à 160Mhz, sur un bus 16bits avec des napes 68 fils. Le débit maximum pour cette norme est 320Mo/s. Les câbles ne doivent pas dépasser 1,5m.

8.1.2 Principe des échanges SCSI

Nous allons présenter ici, les séquences d'échanges :

Bus libre : état de repos du bus, il n'y a pas de signal.

Phase d'arbitrage : Les unités voulant émettre passent le bit correspondant à leur identifiant (ID) sur le bus à 1.

- l'identifiant détermine la priorité (ID petite = meilleure priorité)

Phase de sélection : - l'unité ayant le contrôle du bus (l'ID la plus petite ayant été passée à 1 sur le bus) passe à 1 l'ID de l'unité destination sur le bus.

- Un accusé de réception est alors renvoyé par cette unité lorsqu'elle est prête à recevoir.

Phase de transfert : Pour SCSI, il existe deux modes de transfert :

- Synchrone :
 - Une négociation préalable est effectuée sur le nombre de caractères à envoyer (cela sans acquitement)
 - On obtient des débits supérieurs à 10Mo/s
- Asynchrone
 - Un acquitement est envoyé pour chaque octet transmis
 - Les débits ne dépassent pas 5Mo/s

Phase de sélection : Cette phase permet à une unité cible de se déconnecter puis de se reconnecter le temps de traiter les données déjà reçues. Cela permet de libérer le bus et entraîne une meilleure utilisation de la bande passante. On s'aperçoit ici que l'arbitrage est une notion cruciale pour partager un bus avec de nombreuses unités.

Phase de commande : Elle sert, par exemple aux ordres d'entrée/sortie (E/S) sur une unité de stockage

Phase de données : C'est pendant cette phase que l'émetteur et le récepteur échangent des données.

Phase statut : Elle permet au destinataire d'envoyer à l'émetteur des messages de statut.

Phase de message : La phase de message est un système élaboré pour, entre autres, maîtriser la qualité de la communication, les acquitements...

Pour information, les commandes SCSI sont codées sur 10 bits.

8.2 Le SCSI Série

Celui-ci est une nouvelle évolution de la norme SCSI qui réduit considérablement les contraintes d'utilisation. En particulier, les câbles sont moins limités en longueur et les périphériques peuvent être chaînés les uns aux autres.

Il existe 3 normes de bus SCSI série :

IEEE-1394 (ou FireWire ou i-Link) : Elle est connue pour être largement utilisée dans le monde de la vidéo numérique. Elle permet des débits de 400Mo/s (bientôt de 800 Mo/s) sur seulement fils, la connexion à chaud et est disponible sur presque tous les ordinateurs neufs actuellement en vente (PC ou MAC). Cette norme a maintenant été implémentée dans SCSI-3.

Fibre Channel : dérivé des technologies réseau, Fibre Channel a été implémenté dans la norme SCSI-3. En fait, c'est une sous-catégorie un peu moins performante qui a été introduite : FC-AL (fibre Channel - Arbitrary Loop). Par ses performances, elle rentre un peu en concurrence avec l'IEEE-1394, qui a pour elle d'être plus polyvalente et soutenue par un marché grand public. Le Fibre Channel est le support qu'utilisent en majorité les systèmes SAN. On s'intéressera donc plus longuement à elle un peu plus loin.

SSA (Serial Storage Architecture) : comme les deux technologies précédentes, cette norme, d'origine IBM, a été agréée et intégrée (par le sous-comité X3T10.1 de l'ANSI) à la future norme SCSI-3. Cette interface série, comporte un module FULL DUPLEX à deux canaux, lui permettant donc de soutenir 4 transactions simultanées. Elle est capable d'un débit de 80 Mo/s (4 x 20 Mo/s). C'est une interface Plug And Play, acceptant même le Hot Plugging de 127 périphériques (dont la carte). On peut atteindre avec cette technologie des longueurs de 20 mètres en paire torsadée et 2500 mètres en fibre optique.

9 Fibre Channel

Le Fibre Channel (FC) est un standard mis au point en 1988 par le comité X3T11 de l'ANSI. Son but était de concevoir un standard permettant d'acheminer des données de façon très rapide.

Il existe deux types de communication logiques pour les données :

Canal Entrée/Sortie (E/S) (ex : SCSI, IDE...) : Ce type de communication permet une connexion directe ou commutée, il entraîne peu d'overhead (Temps mis par un ordre pour traverser toutes les couches d'un système, et qui peut être assez long). Cependant, sa limitation réside dans le fait que le nombre d'équipements est réduit par le petit nombre d'adresses fixes par contrôleur (2 pour IDE, 7 ou 15 pour SCSI).

Réseau (ex : pile TCP/IP) : La structure utilisée par ce type de communication consiste en une agrégation de noeuds avec plusieurs protocoles de communication. La communication réseau entraîne un overhead important car à chaque passage dans un noeud actif (ex : routeur), les trames doivent remonter plusieurs couches. Cependant, on peut utiliser avec ce type de communication un plus grand éventail de commandes. De plus, les environnements réseaux sont des environnements avec des communications pré-établies (la connexion nécessite une phase d'initialisation/négociation).

Le FC vise à reprendre le meilleur des deux types de communication. Il ne s'apparente purement à aucun des deux. Il utilise un système de communication intelligent : Fabric (qui prend souvent la forme d'un commutateur FC)

Quelques notions :

N_Port : point de connexion ;

NL_Port : Point de connexion dans une boucle arbitraire FC ;

Les N(L)_Ports : ils consistent en l'association sur un même équipement d'un circuit d'émission et d'un circuit de réception.

Afin d'améliorer les performances, les piles FC et les interfaces logiques sont directement gérés par des microcircuits (on appelle cela une implémentation matérielle).

FC est une liaison série hautes performances avec un protocole de communication propre (restant, via une interface, compatible SCSI, HIPPI, IPI... qui sont de niveau supérieur). Cette liaison

permet donc un transfert rapide d'une grosse quantité de données. La communication avec les LAN est possible grâce à un convertisseur.

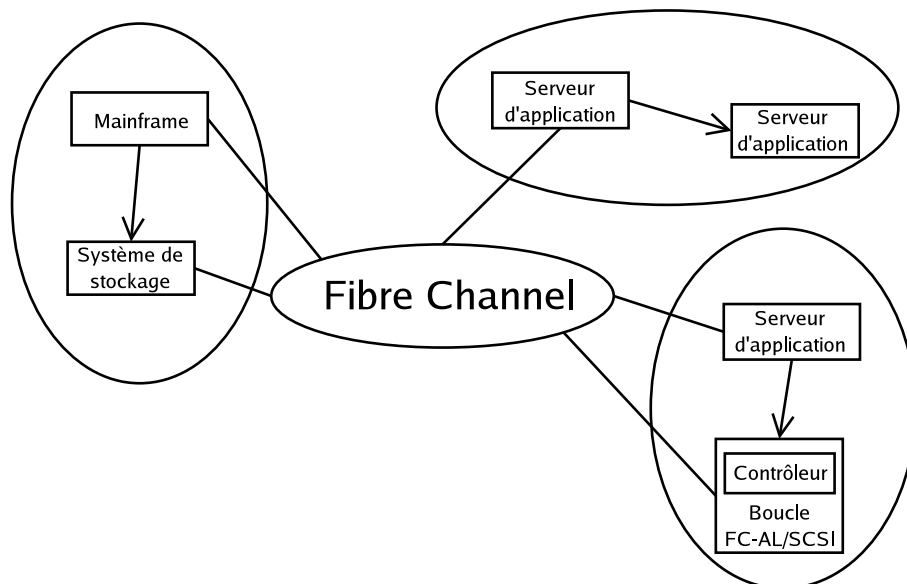


FIG. 5 – Schéma Fédération des réseaux de stockage multiprotocoles

9.1 La topologie FC

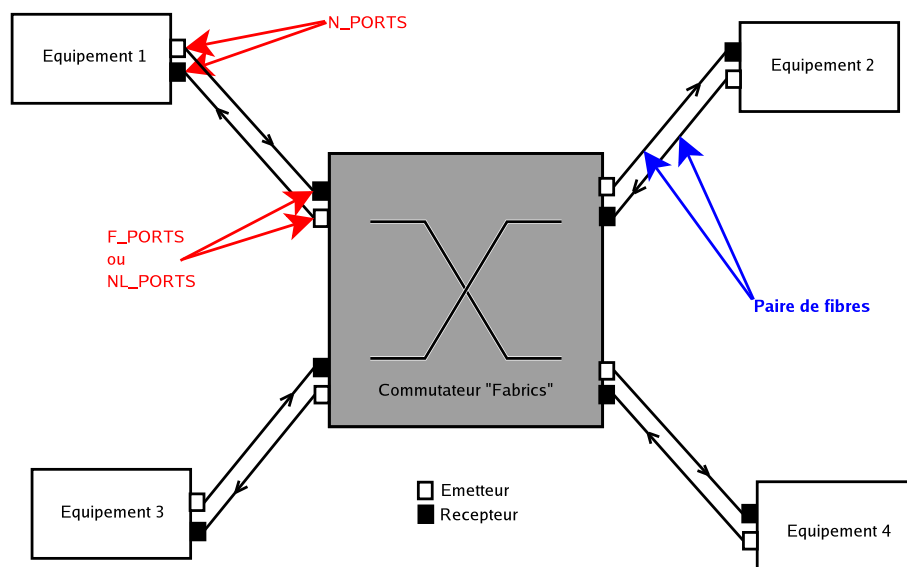


FIG. 6 – Schéma d'un commutateur Fibre Channel

La topologie des connexions FC repose sur :

- les commutateurs (cf. FIG.6) (même principe que pour les réseaux locaux) ;
- les concentrateurs (réunir plusieurs connexions réseaux sur un pipeline FC) ;
- les boucles arbitrées (FC Arbitrated Loop).

Le FC permet d'atteindre des vitesses allant de 133 Mbits/s à 10 Gb/s (pour comparaison, Ethernet : <1Gb/s, ATA : <1 Gb/s, serial-ATA : <1,2 Gb/s, SCSI parallèle : <2,9 Gb/s)

Le FC utilise les médias optiques ou électriques (cuivre) pour transiter, sachant que la solution optique est la plus performante et qu'elle se généralise.

9.2 Les couches de FC

Le Fibre Channel est composé de 5 cinq couches (cf. FIG.7) . Nous allons vous les décrire.

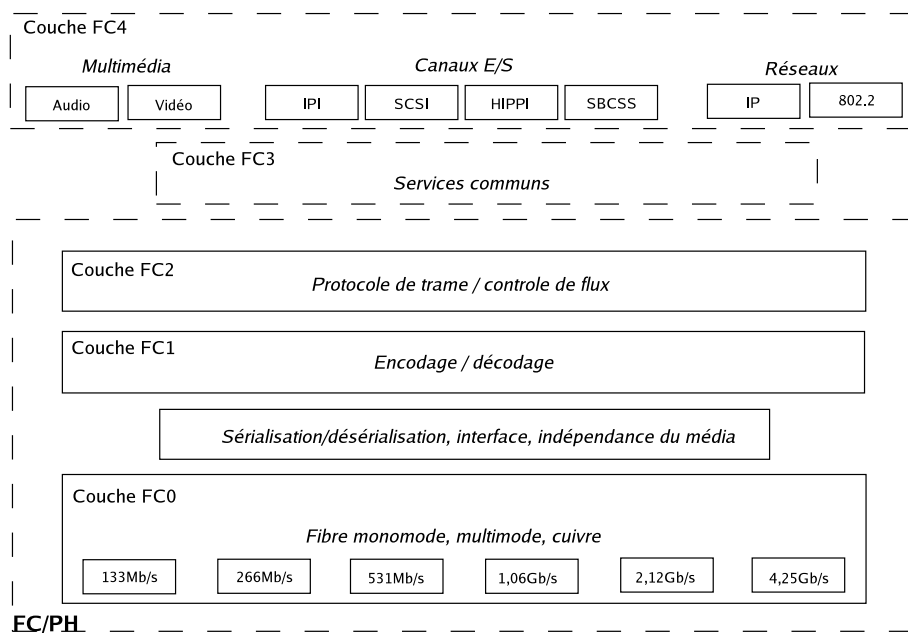


FIG. 7 – Les 5 différentes couches de Fibre Channel

9.2.1 FC-0 Layer

C'est la couche de liaison physique.

Elle est caractérisée par le type de média et les paramètres de liaison, qui conditionnent les taux de transfert. Cette couche définit la longueur et la vitesse du transfert (liés au média utilisé) La liaison optique MONOMODE permet de relier les sites distants (jusqu'à 100km). Les liaisons optique MULTIMODE et cuivre couvrent des distances plus courtes (dont les batiments).

Les interfaces médias utilisent le système BitErrorRate (BER) qui définit le taux d'erreur accepté our les trames, généralement $\langle 10^{-12}$ (cela veut dire qu'au maximum, une trame sera erronée sur $\langle 10^{12}$ transmises).

La couche physique est capable de supporter un grand nombre de technologies, ce qui rend le standard très compatible. En effet, une même communication pourra être supportée par différents types de supprots physiques.

Open Fibre Control (OFC) :

OFC est le mode opértoire utilisé par FC (cf. FIG.9).

Ce mode opératoire permet de se prémunir contre les dépassements de niveaux de puissance définis par les standards de sécurité sur les lasers.

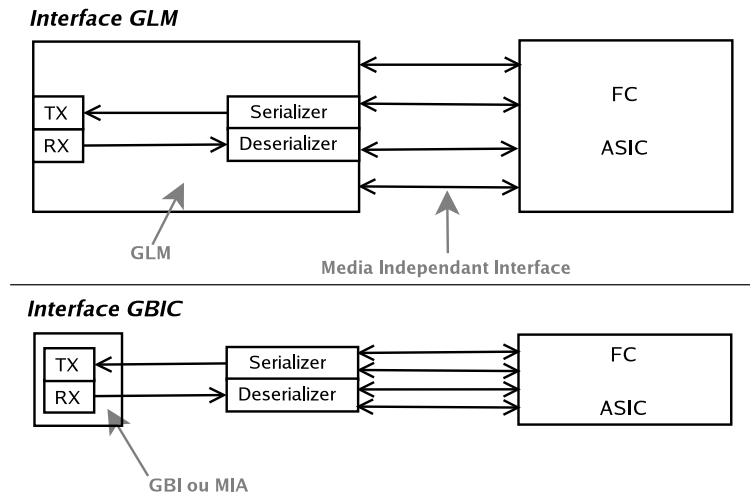


FIG. 8 – Schema des interfaces GLM et GBIC

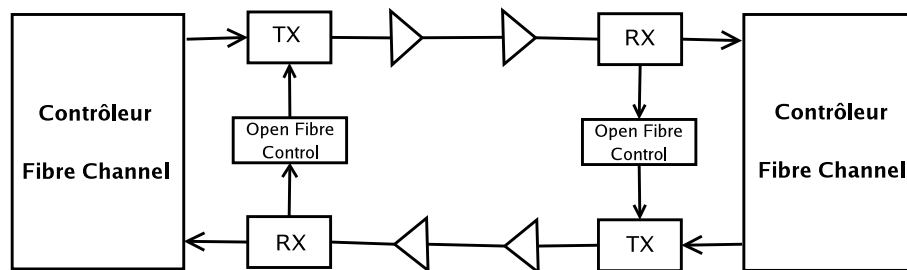


FIG. 9 – Open Fibre Control (OFC)

9.2.2 FC-1

Cette couche est celle du protocole de transmission.

Elle effectue les opérations suivantes :

- encodage/décodeage série ;
- traitement des caractères spéciaux ;
- contrôle des erreurs ;
- reconstitution de l'horloge de transmission
- synchronisation octets.

La transmission se fait en codage 8B/10B dont voici le principe global (cf. FIG.10).

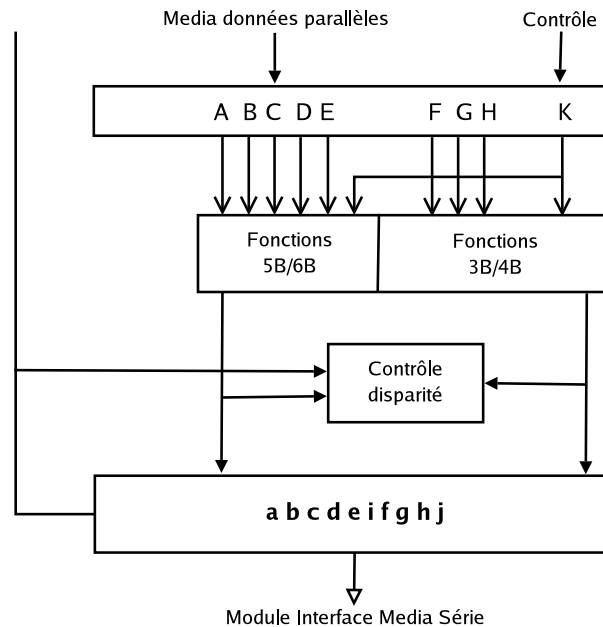


FIG. 10 – Principe du codage 8B/10B

Ce codage permet d'éviter les trop longues séquences de 0 ou de 1 qui ont tendance à désynchroniser l'horloge elle-même réglée sur le signal. De plus, la redondance de 25% apportée par ce codage améliore la vitesse de transmission.

9.2.3 FC-2

La troisième couche de FC gère la signalisation et les mécanismes de transport indépendamment des couches supérieures de FC.

Cette couche définit 5 blocs :

- Un jeu de commandes, qui organise les échanges (initialisation des circuits, séparation des trames, notification de la fin de communication...);
- des trames qui contiennent les informations de charge utile, les adresses source et destination et les contrôles de liaison. La taille des trames FC varie de 38 à 2148 octets ;
- une notion de séquence, qui est formée par une ou plusieurs trames transmises entre deux points A et B, numérotées par un compteur spécifique qui rend chaque trame unique ;
- un bloc d'échanges contenant un ou plusieurs séquences échangées (unidirectionnellement ou bidirectionnellement entre deux N_Ports). ;
- des protocoles, relatifs aux systèmes offerts par FC, spécifiques pour les services associés aux couches de haut niveau. FC fournit ses propres paramètres de protocole pour gérer les environnements opérationnels pour le transfert de données. FC-2 prend en charge indifféremment

les trois topologies vues précédemment.

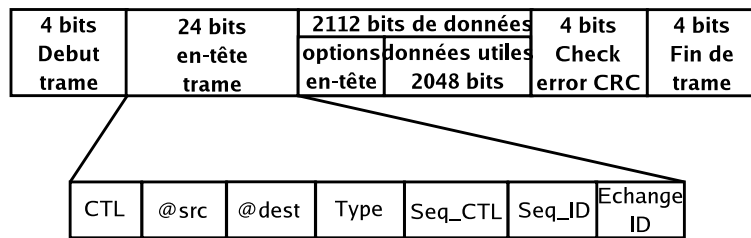


FIG. 11 – Une trame Fibre Channel

Le mécanisme de séquence permet de transmettre une information qui ne tient pas dans un seul champ Payload (cf. FIG.12).

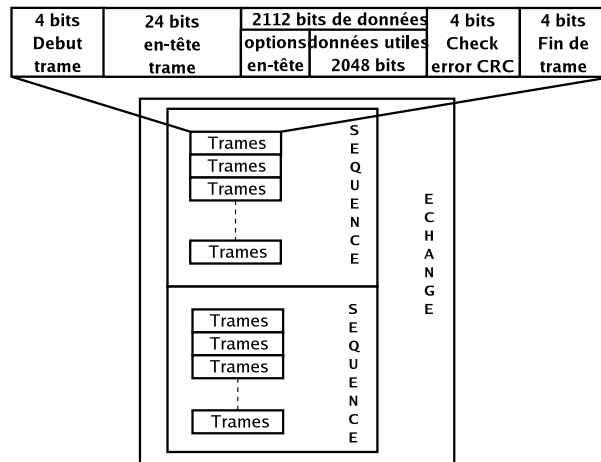


FIG. 12 – Le mécanisme de séquence

Contrôle de flux : Celui-ci assure le bon acheminement du flux entre deux N_Ports via un Fabric. Ce contrôle dépend des classes de services :

- 1** : contrôle de bout-en-bout ;
- 2** : contrôle basé sur une mémoire tampon ;
- 3** : utilisation simultanée des deux classes précédentes.

Le contrôle est géré à la fois par la source et la destination via un mécanisme de crédit et de crédit CNT (l'on limite le nombre d'acquitements de retard).

9.2.4 FC-3

Cette couche fournit les services communs entre agents pour les fonctionnalités comme :

- agréer plusieurs liaisons parallèles pour augmenter la bande passante pour une connexion spécifique ;
- la possibilité pour plusieurs ports d'émettre vers une même adresse ;
- à l'inverse, la possibilité d'émettre vers plusieurs adresses simultanément (multicast).

Cette couche n'est pas encore totalement standardisée.

9.2.5 FC-4

C'est la couche la plus élevée, elle définit l'interface avec l'application.

Cette couche définit une règle de mise en correspondance avec les protocoles de plus haut niveau. Elle peut transporter sans discrimination les données dans une topologie réseau ou canal. Elle est compatible avec de nombreux protocoles : SCSI, IPI, HIPPI, IP, ATM(ALL5), FC-LE, SBCCS, IEEE802.2...

Le Fibre Channel réunit les avantages des protocoles réseau et des protocoles E/S. Il permet de mettre en place des liaisons à très haut débit, bien adaptées au stockage. Son implémentation lui permet une grande compatibilité avec les équipements existants.

10 la virtualisation

Qu'est ce que la virtualisation ?

C'est la séparation entre les ressources physiques et les ressources logiques. Par exemple, nous avons un espace de stockage X. Et bien, cet espace de stockage peut être réparti physiquement sur un ensemble de disques.

10.1 La technologie RAID

Ce concept n'est pas nouveau. Depuis plusieurs années, il existe au sein du stockage de données simple (locale au système). La technologie RAID utilise ce concept. RAID est l'acronyme de *Redundant Array of Independent Disks*. Il existe différents niveaux dans RAID. Ils apportent chacun des améliorations grâce à une abstraction du support physique des données. Il y en a 5 à retenir :

Raid 0 : Cette approche est souvent utilisée pour améliorer les performances uniquement. On copie alternativement sur un nombre N de disques un flux de données. Chaque bout de fichier se retrouve éparpillé sur N disques. Cela permet ainsi d'avoir sur l'espace disque virtuel une bande passante N fois supérieure en comparaison d'un accès sur un espace sur un seul disque.

Raid 1 : Cette approche apporte uniquement de la redondance aux données. Elle n'apporte pas de performance supplémentaire. On obtient ainsi plusieurs copies des mêmes données sur des disques différents.

Raid 0+1 : Cette approche permet d'avoir la redondance des données et ainsi obtenir un gain de performance.

Raid 3 : Cette approche se base sur le fait qu'il y a un disque dédié à la parité. Il permet ainsi de fiabiliser l'ensemble des données, de ne pas perdre l'intégrité des données.

Raid 5 : Ici la parité et les données sont distribuées de manière séquentielle sur tous les disques composant la grappe. Cela assure une excellente protection de données.

10.2 La virtualisation dans les réseaux SAN

Ainsi comme dans le RAID, l'avantage de la virtualisation, est de pouvoir différencier l'aspect physique et l'aspect logique des données. Les serveurs d'applications n'ont que des données et rien de plus.

Il existe cinq grandes architectures de virtualisation dans le SAN :

SAN in the BOX : Cette approche repose sur un *appliance* qui joue le rôle de chef d'orchestre. Il dispose de sa propre technologie de virtualisation des données. Il est aussi à même de fournir du stockage virtuel aux serveurs d'applications sans modifier leur constitution

logicielle et matérielle.

Les serveurs de domaines : Ils s'intercalent entre les serveurs d'applications et les unités de stockage. Ils ne nécessitent pas de modification logicielle des serveurs d'applications. Cette approche est proche de la précédente mais, permet un support d'unités de stockage hétérogènes.

Les métaserveurs : Cette approche prend la forme d'un logiciel collaborateur installé sur les serveurs d'applications. Il est nécessaire d'installer cet agent sur l'ensemble des serveurs.

La virtualisation au niveau bloc : Dans ce cas, elle fait partie intégrante des couches logicielles des serveurs d'applications. Elle opère alors au niveau des systèmes de fichiers, qui sont virtualisés.

La virtualisation à l'échelle de l'entreprise : Dans ce cas, la virtualisation s'applique au niveau de l'entreprise.

11 Internet Fibre Channel Protocol (iFCP)

11.1 La description de iFCP

iFCP est un protocole de passerelle à passerelle pour l'implémentation de Fabric (FFC). Les éléments de routage et de communication IP remplacent FC. Cela permet l'attachement de produits de stockage FC déjà déployés à des réseaux IP, tout en prenant en charge les services de commutation de chacun des deux environnements.

iFCP s'appuie sur FC, en remplaçant la couche de transport par un réseau IP. Pour cela, il met en correspondance (mapping) la couche de transmission FC et de la pile TCP/IP.

On peut donc relier des réseaux SAN en mode point-à-point ou sans utiliser des équipements purement SAN (cf. FIG.13). Cela réduit considérablement les contraintes de distance posées par FC. De même, le coût d'un tel système est bien inférieur à la mise en place d'un réseau SAN complet.

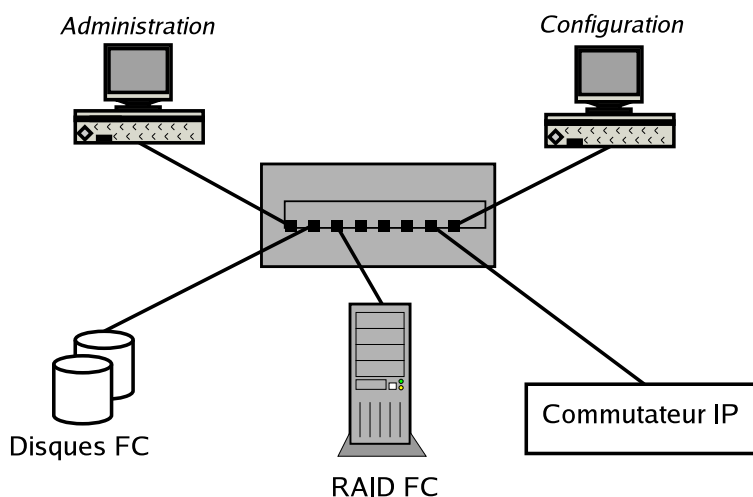


FIG. 13 – Un exemple de passerelle iFCP

On peut grâce à iFCP :

- interconnecter des réseaux de stockage distants en utilisant TCP/IP

- exploiter les réseaux TCP/IP en place pour bâtir un réseau FC. On conserve alors les bénéfices de l'infrastructure FC.
- relier simplement des équipements FC avec de l'IP local.

Le schéma (cf. FIG.14), montre un exemple d'utilisation de iFCP :

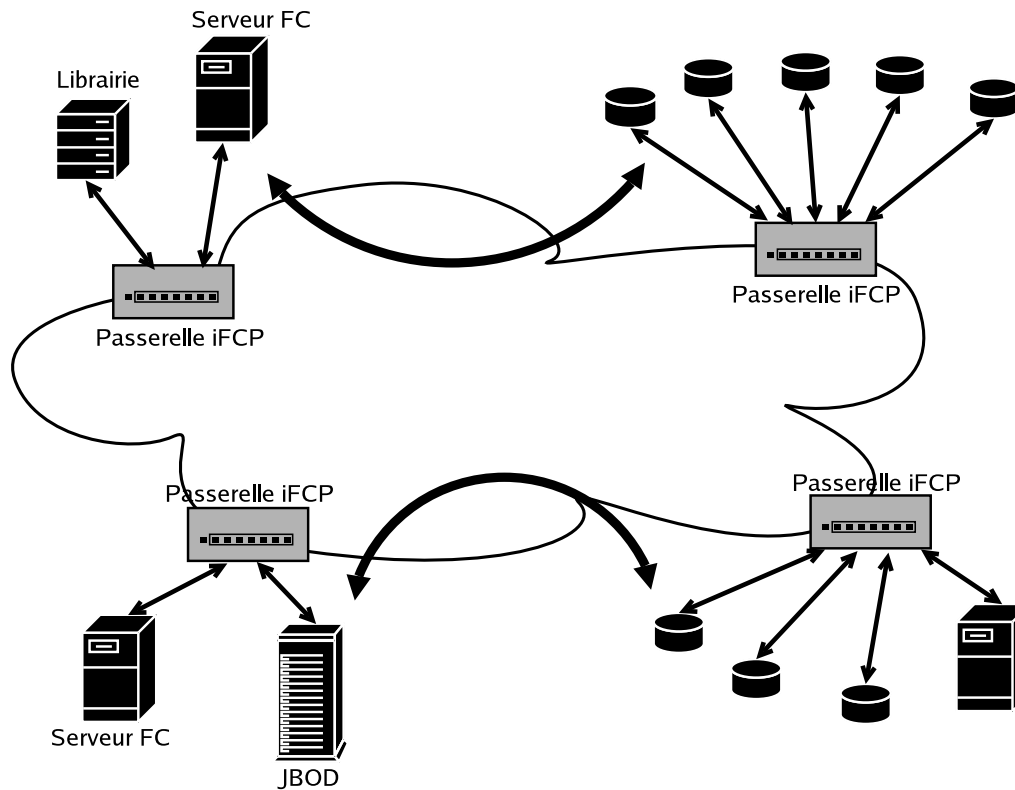


FIG. 14 – un exemple d'utilisation de iFCP

11.2 Architecture d'une passerelle iFCP

La passerelle iFCP sert d'interface entre les réseaux IP et les réseaux FC. De ce fait, elle assure, entre autres, la mise en correspondance des adresses. On peut voir sur le schéma (cf. FIG.15) la structure globale d'une passerelle iFCP.

Enfin, il est à signaler qu'il existe un mécanisme de translation FC vers iFCP et vice-versa, qui est assez complexe et ne sera pas traité ici car s'éloignant trop du sujet.

12 le iSCSI (ou le SCSI over IP)

Comme nous l'avons vu précédemment, SCSI fonctionne en mode Entrée/Sortie (E/S) entre plusieurs périphériques(internes ou externes). Il entraîne des contraintes liées à son support (distance et performances) qui ont engendré la nécessité du développement de protocoles tels que FC ou Gigabit Ethernet sur lesquels repose le SAN. Ainsi, les contraintes de distances sont repoussées.

Le partage du SCSI via des réseaux permet une grande flexibilité (transfert du type node bloc, ex : sauvegarde, clustering...). Il fonctionne selon un mode client (initiateur ou expéditeur) / serveur (resource ayant un rôle passif : serveur de fichiers).

Les commandes sont contenues dans un CDB (bloc de description de commande) délivré par un système hôte. iSCSI a pour mission première d'encapsuler et de faciliter les transactions CDB

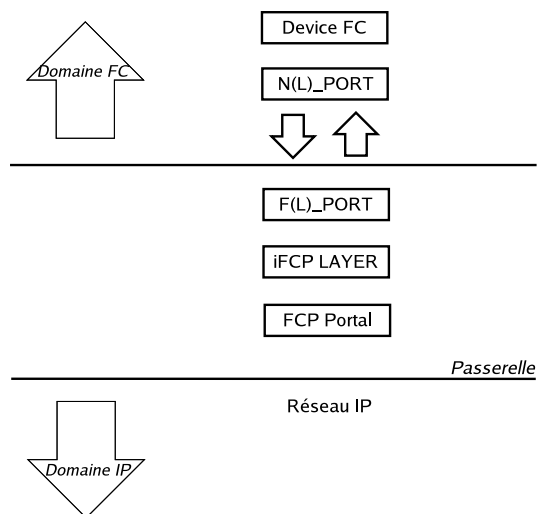


FIG. 15 – L'architecture d'une passerelle iFCP

entre destinataire et expéditeur en utilisant un réseau TCP/IP.

Ainsi, les infrastructures utilisant iSCSI peuvent avoir la forme du schéma (cf. FIG.16).

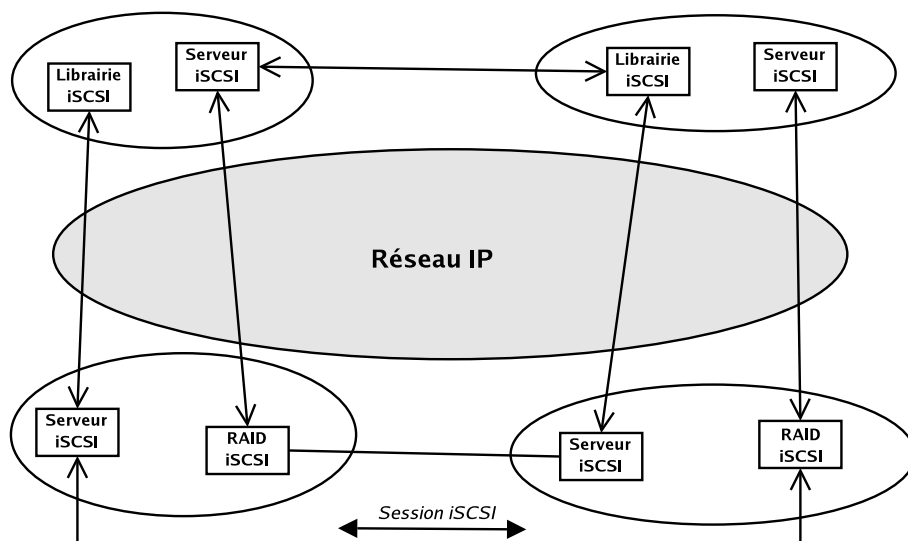


FIG. 16 – Un exemple de topologie iSCSI

iSCSI surveille seulement les transferts en mode bloc et valide les opérations de lecture-écriture, au travers d'une ou plusieurs connexions TCP entre cible et émetteur. Ainsi, ces informations traversent toutes les couches comme montré sur le schéma (cf. FIG.17) .

Ainsi, le iSCSI est en concurrence directe avec le LAN-NAS et HBA-FC. On peut voir sur le schéma (cf. FIG.18) qu'elle reprend la simplicité de HBA-FC tout en étant véhiculée par des paquets IP.

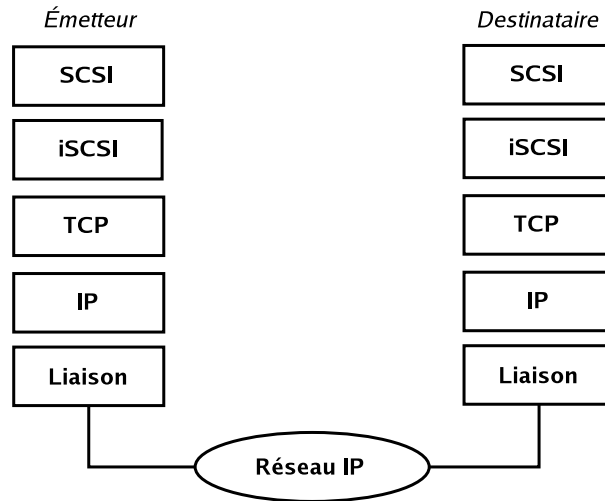


FIG. 17 – Le modèle du protocole iSCSI

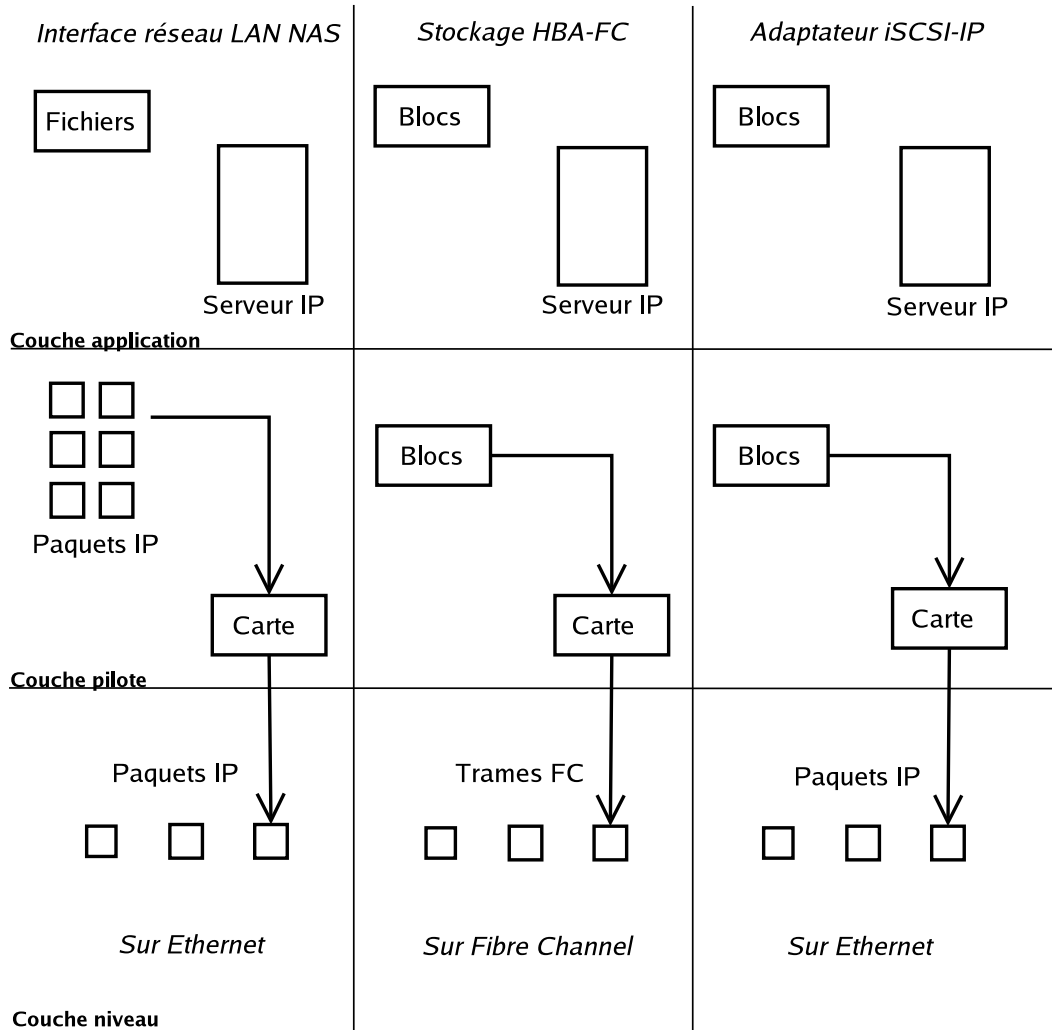


FIG. 18 – La comparaison du iSCSI aux autres types d'adaptateurs présents sur le marché

Cinquième partie

Conclusion

En conclusion, les réseaux SAN sont une solution très fiables, très évolutifs et très flexibles. Il est possible d'augmenter la capacité de stockage sans interrompre le fonctionnement et utiliser divers plans de sauvegarde qui n'encombrent pas votre réseau principal. Un réseau SAN peut gérer d'importants volumes de données réparties sur plusieurs disques. Une grande variété de commutateurs, routeurs et batteries de disques SAN sont commercialisés dans une fourchette de prix entre 7 600 et 304 000 euros. Les réseaux SAN peuvent également s'étendre physiquement sur la distance qu'autorise leur réseau, plusieurs kilomètres pour les systèmes en fibre optique.

Grand avantage, les réseaux SAN permettent de gérer de façon centralisée des unités de stockage disparates ; toutefois, les logiciels sont souvent propriétaires et leur intégration dans une installation n'est pas des plus simples.

Mais, les réseaux SAN sont encors assez chers, bien que les prix soient à la baisse. En outre, si le concept est éprouvé, il reste encore beaucoup à faire en matière d'interopérabilité entre les différents fabricants. Dernièrement, ces derniers se sont davantage poursuivis en justice au motif de violations de brevet, qu'ils n'ont cherché à concevoir des interfaces ouvertes. Certains signes montrent toutefois que cette folie est en train de leur passer ; il suffit de voir l'initiative Bluefin de l'association SNIA (Storage Network Industry Association). Une standardisation est en cours.

Sixième partie

Bibliographie et outils

Documentation :

- "*SAN et NAS, Solutions de stockage, sécurité, infrastructures*" de Xavier BOUCHET et de Henri GILLARES-CALLIAT
- "*SCSI mon amour*" <http://www.byc.ch/scsi/>

Outils :

- L^AT_EX pour le rapport.
- Dia pour les diagrammes.