



Gamme de produit MDS 9000

Fabric Switches et Directeurs Fibre Channel



1.	LA GAMME MDS9000 – INTRODUCTION	3
1.1	PRESENTATION GENERALE DES DIRECTEURS MDS9500	3
1.2	CONFIGURATIONS TYPES.....	5
1.3	PRESENTATION GENERALE DES COMMULATEURS MDS9216 ET MDS9216i.....	6
1.4	PRESENTATION GENERALE DE LA GAMME MDS9100.....	6
1.5	HOMOGENEITE DE L’OFFRE	7
2.	CARACTERISTIQUES MATERIELLES.....	8
2.1	EQUIPEMENT DE CLASSE « DIRECTEUR »	8
2.2	ARCHITECTURE INTERNE	8
2.2.1	<i>Commutation de type « Store-and-Forward »</i>	8
2.2.2	<i>Commutation centralisée</i>	9
2.3	PERENNITE DE L’INVESTISSEMENT	10
2.4	CARACTERISTIQUES DES CARTES LIGNE 16 ET 32 PORTS FC	10
2.5	CARACTERISTIQUES DES CARTES LIGNE 12, 24 ET 48 PORTS FC.....	11
2.6	CARACTERISTIQUES DE LA CARTE LIGNE 4-PORTS 10 GBPS FC.....	12
2.7	CARACTERISTIQUES DES CARTES LIGNE 4 ET 8 PORTS GBE.....	12
2.8	CARACTERISTIQUES DE LA CARTE HYBRIDE MPS	12
2.9	CARACTERISTIQUES DU MODULE DE SERVICE	13
2.10	AGREGATION DE LIENS ET REPARTITION DE CHARGE	13
2.11	EVOLUTION ET RECONFIGURATION SIMPLIFIEES DU SAN	14
2.12	SECURITE DE L’INFRASTRUCTURE.....	15
3.	MECANISMES RESEAUX EVOLUES.....	16
3.1	PARTITIONNEMENT LOGIQUE DE L’INFRASTRUCTURE MATERIELLE	16
3.1.1	<i>Description du mécanisme VSAN ou Virtual Fabric</i>	16
3.1.2	<i>Justification du partitionnement logique</i>	17
3.1.3	<i>Routage Inter-VSAN, IVR, FC-NAT et IVR Service Group</i>	18
3.1.4	<i>Les VSAN en environnement Mainframe FICON</i>	19
3.1.5	<i>Configuration des VSAN</i>	19
3.1.6	<i>Bénéfices de la Commutation FC et Routage intégrés</i>	20
3.2	DEFINITION DES ZONES PAR LUN.....	21
3.3	CONNECTIQUE IP ET DEPORT DE SAN.....	22
3.3.1	<i>Extension du SAN par encapsulation iSCSI</i>	22
3.3.2	<i>Déport de SAN par encapsulation FCIP</i>	24
3.3.3	<i>Interop mode</i>	26
3.3.4	<i>IPv6 sur les interfaces LAN</i>	26
3.3.5	<i>Multiplexage en longueur d’onde</i>	26
3.4	QUALITE DE SERVICE.....	26
3.5	CONTROLE DE FLUX	27
3.6	SERVICE INTELLIGENTS AU CŒUR DU RESEAU DE STOCKAGE.....	27
3.6.1	<i>Applications intégrées dans le Réseau</i>	28
3.6.2	<i>Applications assistées par le Réseau</i>	28
3.6.3	<i>Sauvegardes accélérées par le Réseau</i>	29
3.6.4	<i>Accélération des flux et statistiques LUN</i>	29
4.	ADMINISTRATION ET DEBUGGING	30
4.1	ADMINISTRATION PUISSANTE.....	30
4.2	CISCO FABRIC MANAGER SERVER	31
4.3	CISCO FABRIC SERVICES.....	31
4.4	SECURISATION OPTIMALE DE L’ADMINISTRATION	32
4.5	FIBRE CHANNEL SECURITY PROTOCOL	32
4.6	DIAGNOSTICS INTEGRES.....	33
5.	TABLEAU RECAPITULATIF DES CARACTERISTIQUES	34
6.	CONCLUSION	35
6.1.1	<i>Le Centre Informatique du futur</i>	35

1. LA GAMME MDS9000 – INTRODUCTION

La famille de Commutateurs et Directeurs de la gamme MDS9000 représente l'offre Cisco pour Réseaux de Stockage. Cette gamme a été annoncée en Août 2002 et est disponible depuis Janvier 2003. La famille MDS9000 comprend les éléments suivants :

- Directeurs FC : la famille MDS9500 est composée des châssis de classe Directeur universel MDS9506, MDS9509 et MDS9513 entièrement modulaires, multi-fonctions, pouvant offrir jusqu'à 528 ports FC 1, 2 ou 4 Gbps, du 10 Gbps FC ou un mélange de connectique FC, IP et de modules de service.
- Commutateur modulaire : la famille MDS9200 comprenant les MDS9216 et MDS9216i. Le MDS9216 dispose de 16 ports FC en standard, alors que le MDS9216i comprend 14 ports FC et 2 ports Gigabit Ethernet. Un slot d'extension est disponible sur ces deux modèles et il peut accueillir l'ensemble des cartes disponibles pour les directeurs FC MDS9500. Ces commutateurs peuvent évoluer jusqu'à 64 ports FC, ou une combinaison de ports FC et GbE par simple ajout de carte ligne dans le slot d'extension.
- Commutateurs fixes : la famille MDS9100 est composée de commutateurs fixes de 20 à 40 ports en connectique Fibre Channel 1, 2 ou 4 Gbps.
- Commutateurs Workgroup 4 Gbps : le MDS9020 dispose de 20 ports en connectique 4 Gbps.

Les gammes MDS9200 et MDS9500 sont disponibles depuis Décembre 2002. Ils ont subi des tests de compatibilité extrêmement rigoureux chez les principaux fournisseurs de baies de stockage, et sont maintenant au catalogue des fournisseurs que sont IBM, HP, EMC, HDS, SUN/STK et NetApp, ainsi qu'à l'ensemble de leurs revendeurs.

Le Cisco MDS9500 est un équipement de commutation Fibre Channel de classe Directeur, avec redondance complète de tous les composants actifs. Il offre ainsi un très haut niveau de disponibilité de l'infrastructure SAN.

1.1 Présentation générale des Directeurs MDS9500

Le Directeur MDS9500 est un équipement universel qui peut offrir dans une même unité, les protocoles Fibre Channel, FICON, FCIP et iSCSI, ainsi que des fonctions de Virtualisation intégrées.

Le directeur MDS9513 comprend 13 slots, le MDS9500 9 slots et le MDS9506 6 slots. Deux slots sont utilisés pour les matrices de commutation de type Cross-Bar. Les matrices de commutation supportent également la fonction de Supervision. Le MDS9500 peut recevoir les cartes lignes suivantes :

Cartes de première génération :

- Connectique Fibre Channel : carte ligne 16-ports FC ou 32-ports FC, 1 ou 2 Gbps,
- Carte SSM : Connectique Fibre Channel 32-ports FC 1 ou 2 Gbps avec ASIC pour Services Applicatifs,
- Connectique GbE : 4 ou 8-ports GbE pour le transport des flux de stockage SCSI encapsulés dans des protocoles TCP/IP (support de FCIP et/ou iSCSI),
- Connectique mixte : carte ligne MPS, 14 ports FC et 2 ports GbE (support de FCIP et/ou iSCSI).

Cartes de seconde génération (disponibles avec le code SAN_OS 3.0) :

- Connectique Fibre Channel : carte ligne 12, 24 ou 48 ports FC 1/2/4 Gbps,
- Connectique Fibre Channel : carte ligne 4 ports FC 10 Gbps,

Le directeur MDS9513 comprend 13 slots en face avant. Deux slots sont utilisés pour les fonctions de Supervision. Deux slots en face arrière permettent d'utiliser deux modules Fabric de commutation de type Cross-Bar. La défaillance d'un module Fabric (ou module Cross-bar) n'aura strictement aucun impact tant en terme de connectique (tous les ports accèdent à tous les ports) qu'en terme de performances, même débit par slot.

En configuration maximale, le MDS9513 va offrir 528 ports FC en carte ligne 48 ports FC. Les cartes lignes 1/2/4 Gbps 12, 24 et 48 ports présentent quelques différences au niveau de leurs caractéristiques et seront mixées dans le châssis MDS9500.

Le Directeur MDS9500 peut également recevoir le module de service Storage Service Module (SSM) pour la mise en œuvre de solutions logicielles directement en cœur de réseau. Le module SSM est un module destiné à recevoir des applications de gestion de l'espace disque, de réplication et de continuité de service. Le module SSM dispose de 32 ports FC 1/2 Gbps ainsi que d'ASIC additionnels pour le traitement d'applications de stockage et offre également des mécanismes d'accélération et d'optimisation des flux SCSI.

Cartes de première génération : Chaque port est « auto-sensing » 1 et 2 Gbps, va négocier en FC Boucle (FC-AL) ou FC Switch, et peut recevoir indifféremment des connecteurs optiques SFP de type SW (fibre multimode en local) et LW (fibre monomode pour l'interconnexion distante).

Cartes de seconde génération (12, 24 et 48 ports) : Chaque port est « auto-sensing » 1/2/4 Gbps, et peut recevoir indifféremment des connecteurs optiques SFP de type SW (fibre multimode en local) et LW (fibre monomode pour l'interconnexion distante).

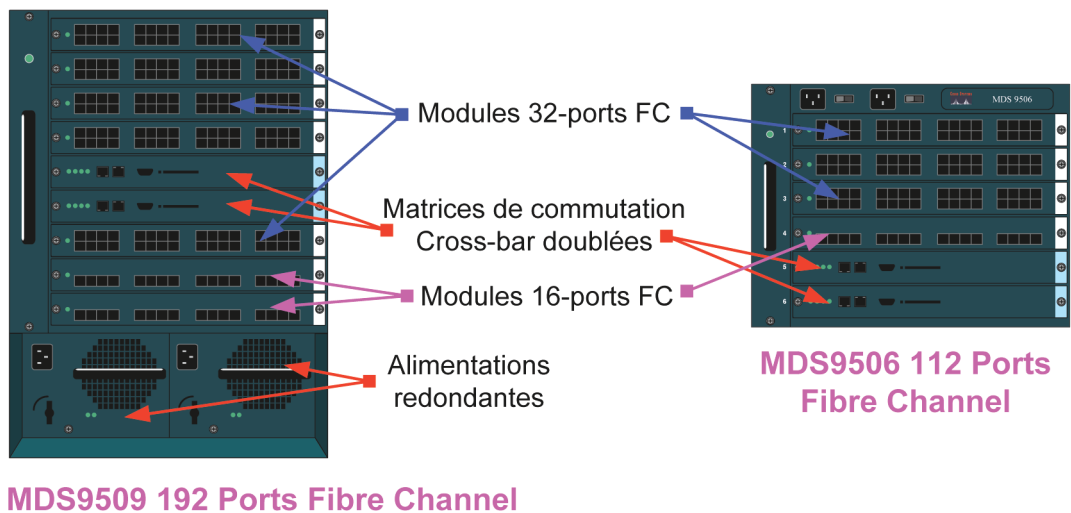
Tous les directeurs MDS9500, qu'ils soient équipés des cartes SUP-1 ou SUP-2, supportent la mixité des cartes de première et seconde génération (nécessite le code SAN_OS 3.0). Ceci est vrai également pour les MDS9216A et MDS9216i.

L'arrivée des nouveaux modules 4 Gbps et châssis MDS9513 a nécessité l'évolution du module de Supervision. Le module Supervisor-2 présente les mêmes caractéristiques que le module Supervisor-1 de première génération, avec la possibilité de gérer un nombre de ports physiques plus important puisqu'il passe de 252 ports géables avec la Sup-1 à 1024 ports sur la nouvelle carte Sup-2. La capacité maximale d'un châssis est aujourd'hui de 528 ports physiques avec le MDS9513, correspondant à 11 cartes 48 ports FC/FICON.

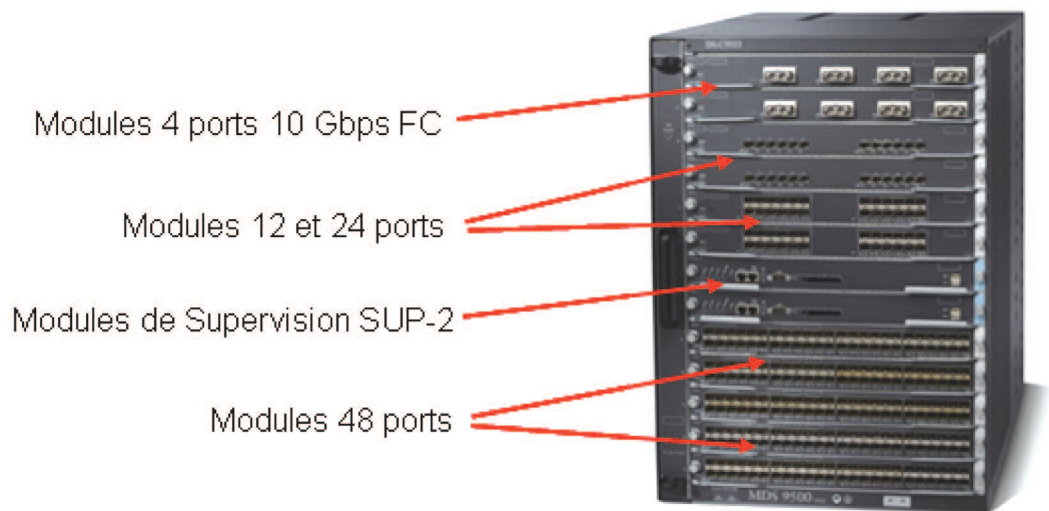
Le module Supervisor-2 est équipé d'une CPU Motorola PowerPC 7447A à 1.4 Ghz avec Marvell Discovery System MV64460, d'un port de management Ethernet 10/100/1000, d'un port COM1 et de deux ports USB. Il dispose d'une mémoire Flash interne d'une capacité de 512 MOctets, utilisée pour le stockage des microcodes utilisés lors du démarrage de la machine. Il est possible aussi de lui adjoindre une mémoire flash externe, extractible, d'une capacité de 512 MO. Cette mémoire peut servir à stocker différentes versions de microcode, différentes versions de configurations, à les sécuriser, et à les copier d'un MDS9500 vers un autre.

1.2 Configurations types

Voici quelques exemples de configuration des Directeurs MDS9500 :



Avec un MDS9513 :



Exemples de configuration des MDS9513 :

Les modules 48-ports seront utilisés en priorité pour la connexion des serveurs et robotique de sauvegarde. Pour la connectivité vers les baies de stockage et les liens ISL, plusieurs options sont possibles :

Fourniture de 4 Gbps garanti par port sur tous les ports en simultanément. La configuration est la suivante :

- 4 modules 12 ports FC et 7 modules 48 ports FC, soit un total de **384 ports FC**.

Fourniture de 4Gbps par port avec un certain niveau de mutualisation ce qui permet l'augmentation du nombre de ports. Dans ce cas entre 2 Gbps et 4 Gbps sont garantis sur tous les ports en simultanément. La configuration est la suivante :

- 4 modules 24 ports FC et 7 modules 48 ports FC, soit un total de **432 ports FC**.

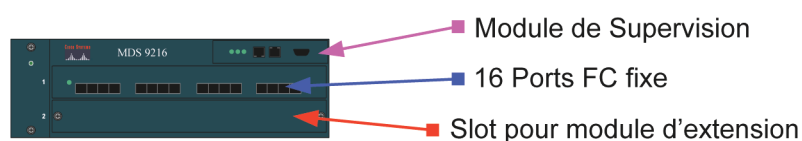
Il est à noter que les alimentations redondantes du MDS9513 et du MDS9506 sont situées à l'arrière du Directeur.

1.3 Présentation générale des Commutateurs MDS9216 et MDS9216i

Le commutateur MDS9216 dispose de 16 ports FC 1 et 2 Gbps en standard, ainsi que d'un slot d'extension. Les différentes cartes ligne disponibles autorisent des évolutions du MDS9216 vers 32 ports FC, 48 ports FC, 64 ports FC ou configuration mixte 16 ports FC + 8 ports Gigabit Ethernet.

Alimentations et ventilateurs sont doublés, fonctionnent en partage de charge et sont échangeables à chaud.

Le schéma ci-dessous présente le MDS9216 en configuration de base :



L'ajout d'un module additionnel se fera sans interruption de service du MDS9216 au niveau des 16 ports FC d'origine.

Le **commutateur MDS9216i** est basé sur la carte mixte 14+2. Il offre 14 ports FC équivalents à ceux du MDS9216, deux ports Gigabit Ethernet et il dispose également d'un slot d'extension. Ce slot d'extension permet d'augmenter le nombre de ports FC ou GbE ou de recevoir un module de service. Dans ce cas, le **MDS9216i se transforme en équipement universel supportant les protocoles FC, FICON, FCIP et iSCSI et disposant de fonctions de Virtualisation intégrées.**

Ces deux commutateurs modulaires tournent le même firmware SAN OS que les châssis MDS9500.

1.4 Présentation générale de la gamme MDS9100

La gamme MDS9100 est composée de trois commutateurs de configurations fixes. Ces trois commutateurs tournent le même firmware SAN OS que les châssis MDS9500.

La première génération de commutateurs MDS9100 comprend le MDS9120 et le MDS9140, tels que présentés ci-dessous :



Le MDS9120 offre 20 ports FC en 1 ou 2 Gbps et le MDS9140 offre 40 ports FC en 1 ou 2 Gbps, sur une hauteur de 1 RU. Ces deux commutateurs disposent de deux alimentations et deux ventilateurs, fonctionnant en partage de charge et pouvant être échangés à chaud.

La deuxième génération de commutateur MDS9100 est composée du MDS9124 :



Le MDS9124 dispose de 24 ports FC 1, 2 et 4 Gbps, sur 1 RU, avec une capacité de commutation globale de 192 Gbps, garantissant 4 Gbps (8 Gbps full-duplex) sur tous les ports en simultané. Il offre les fonctions suivantes :

- **Activation des ports par Licence** : la configuration de base offre 8 ports que l'on peut faire évoluer vers 16 et 24 ports par activation des ports via une simple clé logicielle. L'investissement de l'utilisateur sera donc en phase avec son besoin immédiat et l'augmentation du nombre de ports ne nécessitera pas le changement de commutateur,
- **Alimentations et ventilateurs redondants**, fonctionnant en partage de charge et échangeables à chaud,
- **Montée de niveau du firmware en dynamique**, sans interruption de service ni perturbation des flux en transit au sein du commutateur.

1.5 Homogénéité de l'offre

Le MDS9500 est un châssis modulaire de classe Directeur, avec redondance totale de chacun des composants. Le MDS9216 est un commutateur FC sans redondance de la fonction de Supervision. Les MDS9120 et MDS9140 sont des commutateurs fixes en connectique FC uniquement. Le MDS9124 est aussi un commutateur fixe en connectique FC 1, 2 et 4 Gbps avec licences d'activation des ports et montée de niveau de firmware en dynamique. Le Directeur MDS9513 offre une densité de port très importante, jusqu'à 528 ports FC. Le MDS9216 offre un niveau de modularité également très élevé avec la connectique IP possible, et l'intégration de modules de services dits «intelligents». En dehors de ces différences, le même microcode est utilisé pour l'ensemble des équipements de la gamme MDS9000, du MDS9124 de 8 ports FC au MDS9513 528 ports, et les différentes fonctions évoluées décrites tout au long de ce document, bien que présentées principalement par rapport au Directeur MDS9513, sont disponibles sur tous les modèles, sauf spécification explicite. L'homogénéité de la gamme MDS9000, ainsi que l'intégration des différents éléments de cette gamme dans un réseau de stockage, sont donc parfaitement garanties.

2. CARACTERISTIQUES MATERIELLES

Les paragraphes suivants vont décrire les caractéristiques matérielles de la gamme MDS9500, à savoir ses spécificités techniques «bas-niveaux».

2.1 Equipement de classe «Directeur»

Le MDS9500 est un équipement de classe Directeur.

A ce titre, la Haute Disponibilité est garantie.

Tous les composants internes sont doublés (alimentations, horloges, ventilateurs, matrices de commutation Cross-Bar, fonction de Supervision, canaux d'administration), fonctionnent en partage de charge et sont tous échangeables à chaud, sans interruption de service du Directeur MDS9500.

Une évolution du nombre de port du Directeur MDS9500 se fait également en dynamique. Il suffit pour cela d'ajouter une carte ligne dans un slot disponible, et ce nouveau module est automatiquement pris en compte par la fonction de supervision du Directeur, sans impact sur son fonctionnement.

Les montées de niveau du microcode (firmware) se font, là encore, en dynamique, sans interruption de service et sans impact sur les flux en transit au sein du Directeur.

2.2 Architecture interne

Les Directeurs MDS9500 sont architecturés autour de deux matrices de commutation actives de type Cross-Bar. La capacité de commutation totale au niveau d'un châssis MDS9513 est supérieure à 2,4 Tbps en full duplex. Chaque carte ligne est reliée aux deux matrices de commutation par des connecteurs redondants, 4 connecteurs par carte ligne, 2 vers chaque matrice.

La capacité maximale au niveau de chaque carte ligne est de 128 Gbps, soit 256 Gbps en Full-Duplex. En fonctionnement normal, la moitié de la bande passante par slot est utilisée afin de garantir l'aspect Directeur-Class. Ainsi, même en cas de défaillance d'une cross-bar, les 64 Gbps par slot seront toujours disponibles.

D'autre part, les matrices de commutation cross-bar sont dimensionnées afin que la défaillance de l'une d'entre elles n'entraîne aucune baisse de la capacité de commutation du châssis.

L'architecture interne des Directeurs MDS9500 met en œuvre les techniques de commutation de trames les plus sophistiquées et les plus performantes du marché à l'heure actuelle. Deux caractéristiques sont particulièrement pertinentes, commutation centralisée et de type «Store-and-Forward».

2.2.1 Commutation de type « Store-and-Forward »

Les équipements de commutation Fibre Channel traditionnels utilisent une commutation de type Cut-through, dans laquelle la trame est commutée dès le décodage de l'en-tête. Le Directeur MDS9500, équipement de commutation Fibre Channel plus récent, utilise une technique de commutation plus moderne de type Store-and-Forward, dans laquelle la trame est reçue entièrement avant d'être commutée. Ces deux types de commutation offrent une latence différente, supérieure pour le Store-and-Forward mais qui est restée négligeable sur un temps de réponse standard et complètement invisible sur le chemin de données. Voici les raisons pour lesquelles la commutation de type Store-and-Forward est incontournable :

- Les temps de réponse moyens sont aujourd'hui de quelques millisecondes à une dizaine de millisecondes, contre quelques micro secondes pour les deux types de commutation.

- La commutation Cut-Through n'est possible que si tous les chemins empruntés sont au même débit, ce qui était le cas des premières infrastructures FC à 1 Gbps. Dans une infrastructure mixte 1, 2, 4 ou 10 Gbps, les équipements vont basculer en Store-and-Forward. Dans une architecture comprenant des liens de débits différents, le Cut-Through n'est plus possible, sauf à réduire la capacité globale de l'infrastructure au débit des liens les plus faibles. Le cas typique concerne les liens inter sites en débit 2, 4 ou 10 Gbps alors qu'un certain nombre de HBA sont des HBA à 1 ou 2 Gbps.
- La mise en œuvre de services intelligents ne peut se faire qu'avec une technique de commutation de type Store-and-Forward. Sans citer tous les services concernés, il suffit d'étudier la Qualité de Service. Il n'y a pas d'autre moyen, pour prioriser les flux entre eux, que de bufferiser les moins prioritaires. La commutation de type Cut-Through pourra offrir une latence réduite mais elle empêchera la mise en œuvre de toute fonction intelligente.

Il est à noter que cette évolution du Cut-Through vers le Store-and-Forward est une évolution naturelle de la commutation réseau. Les premiers réseaux Ethernet fonctionnaient en Cut-Through, mais cette technique a complètement disparue de nos jours, depuis qu'il a fallu intégrer des équipements hétérogènes en terme de débit. Ces évolutions sont en cours au niveau des réseaux Fibre Channel, et le Directeur MDS9500, équipement disposant des dernières avancées en terme de commutation Fibre Channel adopte tout naturellement cette technique incontournable.

D'autre part, en fonctionnement de production, les flux Fibre Channel vont avoir des comportements très divers, dus à de fortes contraintes applicatives. Des phénomènes de congestion vont se produire, qui seront adressés, dans un équipement de type Directeur Cross-bar, par un mécanisme de «Virtual Output Queuing» (VOQ) afin de prévenir l'apparition du «Head-of-Line blocking» (blocage HoL). Enfin, le contrôle de la congestion sera assuré par le mécanisme de Buffer-to-Buffer credits.

Dans tous les cas, les équipements vont utiliser des tampons mémoire (buffers), qui vont annuler les caractéristiques de latence réduite de la commutation Cut-Through. En d'autres termes, «on ne commute plus, on bufferise». Ainsi donc, le temps de latence en Cut-Through devient une valeur de laboratoire, sans plus aucune signification dans une exploitation informatique en fonctionnement de production réel.

2.2.2 Commutation centralisée

Sur les équipements de type châssis, il existe deux possibilités de commutation :

- Commutation locale au niveau des cartes ligne,
- Commutation centralisée.

La commutation locale sur une carte ligne apporte la performance maximale à partir du moment où le port d'entrée et le port de sortie sont situés sur le même ASIC de la même carte. D'où un câblage extrêmement contraignant car tous les serveurs accédant à une interface d'une baie de disque doivent se situer sur le même ASIC du même module. Si un serveur accède à plusieurs volumes situés derrière des interfaces distinctes ou même sur des baies différentes, les différentes interfaces doivent également être sur la même carte ligne. Ce qui revient de fait à empiler des commutateurs 16 ou 32 ports au sein d'un rack. Aucun bénéfice n'est retiré de la consolidation. De plus, si la carte ligne a une défaillance, on perd l'accès à la fois côté serveur et côté baie. D'autre part, si certains des flux sont commutés en local et d'autres doivent passer d'une carte à une autre, la latence sera différente selon les cas et le châssis devient complètement imprédictible en terme de latence. Si on utilise en plus des lames de routage, les latences seront encore moins prédictibles car le routage va encore ajouter à la latence, outre le fait que la lame de routage devient un goulet d'étranglement.

Bien évidemment, le MDS9500 n'utilise pas ce type d'architecture. Il est basé sur une commutation des trames centralisée sur les cross-bar. Ainsi, la latence est fixe et garantie, que les ports d'entrée et de sortie soient situés sur les mêmes cartes ligne ou sur des cartes ligne différentes. Le MDS9500 est donc bien un équipement Directeur permettant aux utilisateurs d'envisager une véritable consolidation de

l'infrastructure SAN, disposant de ports banalisés sur l'ensemble des cartes ligne et garantissant une latence stable quelle que soit la position des ports en entrée et en sortie pour un flux donné.

Au sein de la gamme MDS, le partitionnement logique d'une infrastructure matérielle mutualisée est mis en œuvre par le mécanisme des VSAN, les SAN Virtuels. Chaque VSAN va tourner l'ensemble des Processus Fibre Channel, et va se présenter comme un Fabric à part entière. L'isolation entre chaque Fabric est donc totale, car les processus sont isolés les uns des autres. Une reconfiguration ou un dysfonctionnement dans une partition logique n'aura strictement aucun impact sur les ressources des autres partitions. Ce partitionnement logique permet de s'affranchir des problèmes d'interopérabilité, car les éléments FC non compatibles seront installés dans des partitions logiques distinctes et ne se verront donc absolument pas. Les mécanismes de partitionnement logique sont décrits plus loin dans ce document.

Le Directeur MDS9500 représente un seul nœud Fibre Channel d'un point de vue topologie réseau FC. Les cartes ligne sont reliées aux deux matrices de commutation et n'importe quel port du MDS9500 peut accéder à n'importe quel autre port via les deux matrices, même en configuration maximale de 528 ports sur un MDS9513. Les trames FC sont réparties entre les deux matrices de commutation. En cas de défaillance d'une matrice de commutation dans un Directeur, aucune perte de connexion des serveurs et périphériques, car tous les ports seront interconnectés par la deuxième matrice de commutation et cela, **sans interruption sur les flux en transit.**

2.3 Pérennité de l'investissement

La capacité de commutation d'un Directeur MDS9513 est supérieure à 2,4 Tbps. Les deux modules de commutation Cross-Bar sont actifs et fonctionnent en partage de charge. Un MDS9513 supporte jusqu'à 528 ports FC 4 Gbps. Chaque carte ligne dispose d'une électronique back-end banalisée et utilise 4 connecteurs en fond de panier vers les deux modules Cross-Bar.

Cette banalisation du fond de panier garantit la pérennité de l'investissement pour les commutateurs MDS9216 et les Directeurs MDS9500, car ils sont à même de mixer les cartes ligne de 1ère et 2ème génération sans remise en cause de l'investissement initial, pas de changement de châssis ni de matrice de commutation.

Ainsi les utilisateurs des premières générations de MDS9000 ont pu évoluer de façon transparente du 1 et 2 Gbps FC d'origine vers le 4 et 10 Gbps aujourd'hui. Ces mêmes utilisateurs pourront mettre en œuvre les prochaines générations de protocoles (par exemple le 8 Gbps FC) de façon tout aussi transparente, toujours sans aucune remise en cause de l'investissement initial. Le MDS9500 est «8 Gbps ready».

De la même façon que toutes les fonctions évoluées sont utilisées indifféremment en 1 et 2 Gbps, les technologies 4 et 10 Giga FC seront transparentes pour les couches hautes. Tous les mécanismes décrits dans ce document – VSAN, QoS, Zoning, Trunking, etc. – seront exploités aussi bien sur les liens en 1 Gbps, 2 Gbps, 4 Gbps et 10 Gbps.

2.4 Caractéristiques des cartes ligne 16 et 32 ports FC

Chaque interface Fibre Channel fonctionne en auto négociation pour le débit – 1 ou 2 Gbps – pour le mode – half-duplex ou full-duplex – et pour le type de port.

Chaque interface peut négocier indifféremment en E Port, F Port, FL Port (périphériques en FC-AL Public) ou TE Port, ce dernier type assurant le transport des VSAN ou partitions logiques, tel que décrit plus loin. Les modes TL-Port, pour les périphériques en FC-AL Privé, et SD Port, pour le port SPAN, fonction de debugging décrite plus loin, devront être définis spécifiquement. Selon le besoin, chaque interface peut être forcée dans un mode précis.

Le choix entre carte ligne 16 et 32 ports dépendra des besoins d'interconnexion distante.

La carte ligne 32 ports dispose de 12 Buffer Credit par port. Elle sera utilisée pour la connectique locale, serveurs et robotiques de sauvegarde. L'objectif de ce module est la densité de connexion.

La carte ligne 16 ports dispose de 255 Buffer Credit par port. Ce n'est pas un pool partagé, mais bien 255 Buffer Credit dédiés à chaque port.

Cette capacité en Buffer Credit permet de garantir le débit maximal de 2 Gbps entre des sites distants. En technologie 2 Gbps, il faut environ un BC par kilomètre pour garantir le débit maximal. Les Directeurs MDS9500 peuvent donc assurer le débit maximal sur environ 250 kilomètres. En technologie 1 Gbps, le débit maximal est assuré sur 500 kilomètres. Si les équipements DWDM inter-sites n'assurent que du 1 Gbps, le MDS s'adaptera à ce débit.

Encore une fois, il faut bien noter que **tous les ports d'un MDS9509 font partie du même nœud FC, avec un Domain-Id unique par VSAN.**

2.5 Caractéristiques des cartes ligne 12, 24 et 48 ports FC

La deuxième génération de cartes ligne dispose de 4 nouveaux modules, dont trois en connectique 1, 2 et 4 Gbps :

- Module 12 ports FC 1/2/4 Gbps, garantissant le débit de 4 Gbps sur tous les ports en simultané. Ce module sera utilisé pour des liens ISL et baies de stockage pour lesquels l'utilisateur veut garantir les 4 Gbps sur tous les ports,
- Module 24 ports FC 1/2/4 Gbps, offrant le meilleur ratio coût / performance, car les 24 ports vont se partager les 64 Gbps de fond de panier. Ce module sera aussi utilisé pour des liens ISL et baies de stockage pour lesquels l'utilisateur veut un certain nombre de lien à débit garanti de 4 Gbps – via le mécanisme de Port Bandwidth Reservation décrit ci-après - les autres ports se partageant la bande passante restante,
- Module 48 ports FC 1/2/4 Gbps, offrant le meilleur ratio coût / densité de ports, les 48 ports se partageant la bande passante de 64 Gbps de fond de panier. Ce module sera principalement utilisé pour la connectique Serveurs et Robotique, mais pourra parfaitement assurer la connectique ISL et baies de stockage en fonction des débits.

Ces modules disposent des caractéristiques suivantes :

- Mécanisme de Port Bandwidth Reservation : pour garantir la bande passante sur un ou plusieurs ports des modules 24 et 48-ports. Cette bande passante garantie peut-être de 2, 3 ou 4 Gbps. Cette bande passante sera réservée même dans le cas où le port ne l'utilise pas. Les autres ports du module se partageront la bande passante restante.
- 6000 Buffer Credit par carte ligne, partagé par l'ensemble des ports du module. Par défaut, un port va disposer de 16 BC lorsqu'il négocie en F_Port, ou de 255 lorsqu'il négocie en E_Port ou TE_Port. Il est possible d'allouer jusqu'à 4096 BC - du pool de 6000 - sur un seul port en cas de nécessité de connectique longue distance. 4096 BC garantissent un débit ligne jusqu'à 4000 kms en 2 Gbps ou 2000 kms en 4 Gbps,

C'est la valeur de Buffer Credit la plus importante disponible sur le marché à l'heure actuelle.

Chaque interface peut négocier indifféremment en E Port, F Port, FL Port (périphériques en FC-AL Public) ou TE Port, ce dernier type assurant le transport des VSAN ou partitions logiques, tel que décrit plus loin. Selon le besoin, chaque interface peut être forcée dans un mode précis.

Le choix entre carte ligne 12, 24 et 48 ports dépendra des besoins de bande passante par port et des besoins d'interconnexion distante et donc de la capacité en nombre de Buffer Credit par port.

2.6 Caractéristiques de la carte ligne 4-ports 10 Gbps FC

D'un point de vue pur protocolaire, il est intéressant de noter que les couches basses – FC0, FC1 - du protocole 10 Giga FC sont basées sur les couches basses du protocole 10 Giga Ethernet. De la même façon, le protocole Gigabit Ethernet est basé sur les couches basses – FC0, FC1 – du 1 Gbps FC, avec une fréquence d'horloge légèrement supérieure. Le standard 802.3ae de l'IEEE décrivant le 10 Giga Ethernet a été ratifié courant 2002, et des commutateurs Ethernet intégrant cette technologie sont disponibles chez Cisco depuis Septembre 2001. Cisco maîtrise la commutation de trame sur du 10 Gigabit natif depuis plus de cinq ans maintenant.

La connectique 10 Giga FC est utilisée principalement pour l'interconnexion des Directeurs MDS9500 de cœur de réseau. Elle utilise un encodage / décodage nettement plus efficace que le 8b / 10b des 1, 2 et 4 Gbps. La bande passante effective d'un lien 10 Gbps est d'environ 9,8 Gbps, contre 1,6 Gbps pour un lien 2 Gbps. Ainsi, deux liens 10 Giga FC entre deux Directeurs remplacent ainsi un Trunk de 12 liens en 2 Gbps, grâce à la différence d'efficacité de l'encodage / décodage.

Le module 4-ports 10 Gbps FC est une carte ligne de deuxième génération et dispose donc des mêmes caractéristiques que les modules 4 Gbps, notamment en ce qui concerne le nombre de Buffer Credit, soit 6000 pour les 4 ports 10 Gbps FC.

Au vu de la bande passante disponible sur chaque slot d'un châssis MDS9500, le débit de 10 Gbps FC est parfaitement garanti sur les 4 ports en simultané.

2.7 Caractéristiques des cartes ligne 4 et 8 ports GbE

Chaque interface Gigabit Ethernet est capable de supporter en même temps les protocoles FCIP pour l'encapsulation des trames FC dans IP, ou iSCSI pour l'encapsulation du protocole SCSI dans IP. Ces deux protocoles sont maintenant standardisés. Ils utilisent un port TCP réservé, 3225 pour FCIP et 3260 pour iSCSI. Ils seront généralement connectés sur une interface Gigabit optique d'un commutateur Ethernet du réseau d'entreprise.

En utilisation FCIP, l'accès sera fourni à un routeur d'accès WAN pour pouvoir établir le tunnel FCIP vers le site distant. FCIP est basé sur un modèle «Peer-to-Peer». Une fois établi, le tunnel FCIP est vu par les processus FC comme un simple lien ISL, indépendamment de la topologie et de la taille du réseau IP sous-jacent. Il sera donc possible d'agrèger plusieurs liens FCIP dans un PortChannel au même titre que les liens ISL FC. L'agrégation de liens est décrite au chapitre 2.7. Dans le cas où deux tunnels FCIP sont présentés au protocole de routage FSPF, les mécanismes de choix des routes mis en œuvre seront identiques à ceux utilisés pour les liens FC. Un lien FC de 2 Gbps dispose d'un poids de 500 pour FSPF. Un tunnel FCIP sur un seul lien à 1 Gigabit Ethernet aura un poids de 1000 et FSPF pourra faire de l'équilibrage de charge sur des chemins à coût égal, ces chemins pouvant emprunter indifféremment des liens FC pur et FCIP.

En utilisation iSCSI, l'accès sera fourni potentiellement à l'ensemble du réseau d'entreprise afin de pouvoir ventiler les interfaces GbE des MDS9000 vers un nombre important de serveurs de débit moindre, en bénéficiant des architectures de type «Core-Edge» des réseaux modernes. Il n'est pas souhaitable de connecter les serveurs directement sur les MDS car la mise en place du protocole iSCSI va se justifier par la consolidation de nombreux serveurs de débit faible à moyen, sur les cartes IP des MDS9000.

Une description plus précise de la mise en œuvre des protocoles d'encapsulation des flux de stockage dans IP est présentée au chapitre 3.3.

2.8 Caractéristiques de la carte hybride MPS

Le module MPS pour Multi-Protocol Service module dispose de 14 ports FC et 2 ports Gigabit Ethernet. Ce module peut s'insérer dans n'importe quel slot d'un Directeur MDS9500 ou MDS9216. Il sert également de base au nouveau MDS9216i et la description de ce module s'applique à l'identique au MDS9216i.

Cette carte MPS peut être vue comme une carte 16-ports FC sur laquelle les ports 15 et 16 auraient été remplacés par des ports GbE. Les 14 ports FC disposent donc des mêmes caractéristiques, par exemple les 255 Buffer Credit par port sur tous les ports.

Les deux interfaces Gigabit Ethernet apportent également plusieurs améliorations marquantes par rapport aux modules 4 et 8 ports IP comme décrit ci-après :

- Compression en hardware. Capacité de compression au niveau du lien Gigabit, en fonction du profil des données en entrée et du ratio de compression obtenu.
- Chiffrement des flux FCIP. Une fois compressé, le flux de données peut-être chiffré en IPSec par algorithme DES, 3DES ou AES sur clé jusqu'à 256 bits. Le chiffrement s'effectue également en hardware.

Les nouvelles possibilités offertes par ce module MPS seront exploitées essentiellement pour l'extension de SAN via FCIP. Mais au-delà de ces évolutions, la mise en œuvre en iSCSI et en FCIP est identique à la mise en œuvre sur carte ligne IPS-8.

2.9 Caractéristiques du module de Service

Le module SSM est basé sur une carte ligne 32-ports FC sur laquelle ont été implantés les ASIC pilotés par les logiciels appropriés. Ce module de service dispose donc de sa propre connectique externe. Les services intelligents supportés par ce module sont décrits au chapitre 3.6.

2.10 Agrégation de liens et répartition de charge

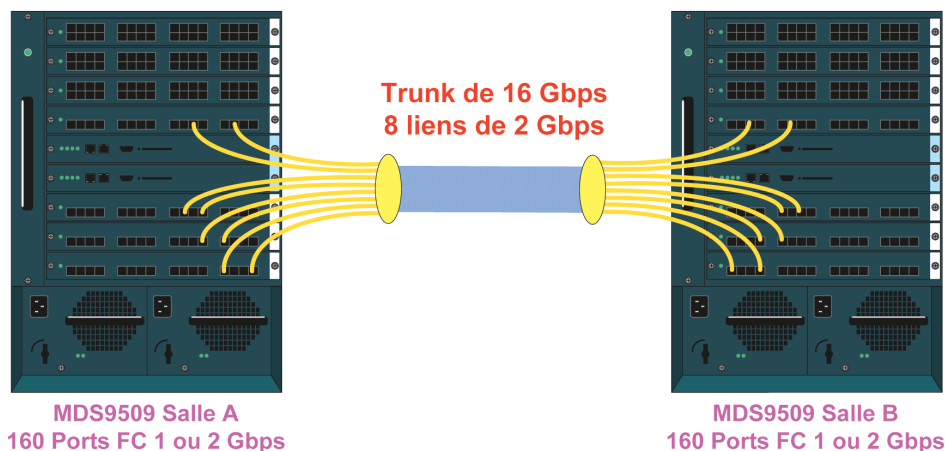
La fonction PortChannel – ou Trunking – permet d'agrèger plusieurs liens ISL physiques dans un lien logique de plus grande capacité et de meilleure résilience aussi bien pour les flux de données FC et FICON. Jusqu'à 16 liens peuvent être agrégés par Trunk pour une capacité max de 64 Gbps en connectique 4 Gbps. Une caractéristique unique sur le marché aujourd'hui est que ces liens peuvent être répartis sur plusieurs cartes ligne, améliorant d'autant la fiabilité. Avec un mécanisme de Trunk dans lequel tous les liens sont adjacents sur la même carte ligne, en cas de défaillance de cette carte ligne, le Trunk est perdu dans sa totalité. Ce n'est pas le cas sur le MDS9500. Les liens d'un même Trunk peuvent également passer via des chemins optiques différents, de longueur différente.

Il n'y a aucune restriction sur la combinaison des mécanismes. Le nombre élevé de Buffer Credit par carte ligne – jusqu'à 4000 sur un seul port – permet d'étendre le Trunk sur plusieurs milliers de kilomètres au débit ligne sans que la distance supportée baisse si le nombre de liens dans le Trunk augmente. D'autre part, si le besoin s'en fait sentir, l'administrateur réseau pourra très simplement augmenter la capacité de l'agrégation en intégrant un lien supplémentaire dans le PortChannel, sans aucune interruption de service de ce dernier.

Il n'y a pas de notion de lien Master dans un PortChannel Cisco. Tous les liens sont banalisés. La défaillance de n'importe quel lien n'aura pas d'autre impact que la baisse de capacité de l'agrégation de la valeur du lien.

La répartition de charge entre les différents liens d'un Trunk est basée sur un algorithme de Hashing entre Src-ID, Dst-ID et Exchange-ID. Ainsi, la cohérence d'une Entrée/Sortie sera garantie en envoyant toutes les trames par le même lien du Trunk, en respectant donc l'ordre d'émission des trames. Cette répartition de charge basée sur l'Exchange-ID permettra de bien équilibrer le flux d'une sauvegarde sur les différents liens du PortChannel.

Le schéma ci-après présente l'agrégation de lien de type PortChannel sur modules ligne multiples :



Sachant que la conception interne du MDS9500 - commutation centralisée - garantit également le respect de l'ordre d'émission des trames, il est tout simplement impossible dans une infrastructure à base de Directeur MDS9500 de récupérer des trames dans le désordre, et cela, même en interconnexion distante, via une agrégation de lien de type Trunk, comprenant des liens de longueurs inégales.

La configuration d'un PortChannel est facilitée par un mécanisme de détection de l'agrégation. Il suffit de configurer le PortChannel que d'un côté, sur un seul MDS9000, pour que l'autre MDS9000 le détecte et configure automatiquement le PortChannel avec les mêmes caractéristiques. Cela permet d'éviter tout risque d'erreur.

En plus de l'agrégation physique de type PortChannel, la gamme MDS utilise aussi la répartition de charge sur chemins à coût égal de type FSPF. L'algorithme de répartition de charge est identique à celui du PortChannel. Il est possible de combiner les deux mécanismes, pour avoir jusqu'à 128 PortChannel dans un Directeur, chacun pouvant recevoir jusqu'à 16 liens physiques à 2 ou 4 Gbps.

Les MDS9000 disposent de la fonction «Port Tracking». En cas de défaillance d'un lien ISL dans une configuration de type Dual-Fabric, ce mécanisme va permettre un basculement immédiat des flux vers le chemin alterné, sans déclencher les timers Fibre Channel. Il garantit ainsi une convergence FSPF immédiate.

La combinaison de l'agrégation de liens ISL en PortChannel avec le partitionnement logique en VSAN va apporter un bénéfice immédiat à l'utilisateur. Au lieu d'utiliser de multiples liens ISL dédiés à chacun des SAN, les différents VSAN seront transportés sur la même agrégation, ce qui va donc amener une meilleure utilisation de la bande passante et donc réduire le nombre de liens inter-sites nécessaires. Le bénéfice sera d'autant plus important si l'interconnexion des sites est assurée via des MUX optiques utilisant des transpondeurs pour les différents liens ISL. **L'agrégation de liens ISL apporte donc un bénéfice immédiat en terme de Retour sur Investissement.**

2.11 Evolution et reconfiguration simplifiées du SAN

Chaque entité Fibre Channel, périphérique ou serveur, comprend une adresse unique de type WWN (World Wide Name) ou WWPN (World Wide Port Name). Lors de la connexion de l'interface au réseau FC, le Service de Login de la Fabrique, sur réception du WWPN, va retourner une adresse physique FCID, ou Fibre Channel Identifier, qui sera utilisée pour le routage des trames Fibre Channel au sein du réseau FC. Ce FCID est constitué de trois octets :

- Le Domain-ID correspondant au Domain-ID du commutateur FC sur lequel est connectée l'interface FC,
- L'Area_ID utilisé en fonction des caractéristiques d'adressage du nœud considéré,
- Le Port_ID correspondant à l'adresse ALPA en FC-AL.

Le problème potentiel est que certains Operating Systems, tel que HP-UX ou AIX, utilisent l'adresse physique FCID pour identifier les volumes logiques. Si le FCID de la baie de stockage change, alors les systèmes de fichiers de tous les serveurs HP-UX ou AIX accédant à cette baie doivent être posés et remontés.

Contrairement à l'offre Fibre Channel disponible aujourd'hui sur le marché, dans laquelle l'adresse FCID est liée au port physique sur le commutateur, le Port_ID sur la gamme MDS n'est pas dépendant du port physique, mais assigné de façon séquentielle. Ainsi, toute reconfiguration de l'environnement FC n'aura plus d'impact au niveau des définitions système dans les Serveurs Unix. De la même façon, une baie de stockage peut changer de port sur le MDS9000, pour le remplacement d'une carte ligne par exemple, sans incidence pour les définitions système Unix.

L'affectation WWPN – FCID peut être définie en statique. Dans le cadre de migration de l'existant, il sera possible de reprendre les serveurs et baies de stockage en leur présentant le même adressage FC. Ainsi, les OS n'auront pas besoin d'être reconfigurés avec la nouvelle adresse FC des disques, il suffira de les connecter sur les MDS pour qu'ils retrouvent leur environnement.

2.12 Sécurité de l'infrastructure

La gamme MDS9000 dispose de plusieurs mécanismes garantissant la stabilité du réseau FC contre toute erreur de manipulation, reconfiguration ou malveillance. De plus, la gamme MDS9000 met en œuvre le protocole Fibre Channel Security Protocol (FC-SP) en cours de standardisation à l'ANSI T11.

Tout d'abord, il est possible de forcer chaque port à négocier par défaut en F Port ou FL Port. Ainsi, un commutateur malicieux ne sera pas en mesure de se substituer à un périphérique ou serveur, négocier en E Port, provoquer un Rebuild de la Fabrique et, avec la priorité appropriée, prendre le contrôle des processus FC. Le mode E Port lui sera interdit.

Pour les ports déjà en lien ISL, il est possible de restreindre les négociations de priorité uniquement au niveau de quelques commutateurs considérés comme extrêmement sécurisés. Cela est possible en bloquant les messages de Rebuild de la Fabrique émanant de commutateurs considérés comme moins sécurisés.

Pour les nœuds terminaux, un serveur ou périphérique va se retrouver par défaut dans une zone nulle dans laquelle il n'aura accès à aucune ressource. Seul l'administrateur pourra lui fournir un accès en le positionnant dans la Zone voulue. Là encore, un serveur malicieux n'aura aucune possibilité de perturbation sans que l'administrateur ne s'en aperçoive.

Enfin, le mécanisme Port Security va affecter un ou une liste de WWPN sur un port spécifique. Tout élément voulant se connecter dans la Fabrique via ce port, se verra interdire l'accès à la Fabrique s'il n'est pas spécifiquement listé dans la configuration du port concerné.

3. MECANISMES RESEAUX EVOLUES

La gamme des Directeurs MDS9500 de Cisco Systems dispose de nombreuses fonctions réseaux évoluées qui vont accroître de façon très importante les possibilités de l'infrastructure SAN et ouvrir de nouveaux horizons aux grands Centres Informatiques.

3.1 Partitionnement logique de l'infrastructure matérielle

Afin d'optimiser l'utilisation de ses équipements de commutation Fibre Channel, l'administrateur du Réseau de Stockage pourra envisager la consolidation de ses différents SAN sur une infrastructure mutualisée. Cette consolidation va apporter plus de souplesse dans le pilotage et l'administration de l'infrastructure SAN, des évolutions simplifiées, et une meilleure maîtrise des coûts par un taux d'occupation des ports supérieur ainsi que par une économie au niveau des liens ISL. En effet, les liens ou Trunks ISL seront également mutualisés entre les différentes partitions logiques avec un meilleur taux d'utilisation de la bande passante de ces liens, au lieu d'avoir des liens ISL dédiés par environnement physique. Cette consolidation devra cependant garantir l'isolation entre les différents environnements, Production, Sauvegarde, Développement, etc.

Dans le cadre de Réseau de Stockage devant héberger des entités autonomes, le partitionnement logique va apporter un niveau d'isolation correspondant à celui apporté par des environnements physiquement dédiés.

3.1.1 Description du mécanisme VSAN ou Virtual Fabric

Au sein de la gamme MDS, le partitionnement logique d'une infrastructure matérielle mutualisée est mis en œuvre par le mécanisme des VSAN, les SAN Virtuels. Chaque VSAN va tourner l'ensemble des Processus Fibre Channel, et va se présenter comme un Fabric à part entière. L'isolation entre chaque Fabric est donc totale, car les processus sont isolés les uns des autres. Une reconfiguration ou un dysfonctionnement dans une partition logique n'aura strictement aucun impact sur les ressources des autres partitions. Ce partitionnement logique permet de s'affranchir des problèmes d'interopérabilité, car les éléments FC non compatibles seront installés dans des partitions logiques distinctes et ne se verront donc absolument pas.

La mise en œuvre du partitionnement logique en VSAN s'effectue par l'imposition d'un label ou tag sur l'en-tête de trame FC au moment où celle-ci rentre dans le premier MDS9000 du réseau SAN. Ce champ additionnel dans l'en-tête de chaque trame FC va contenir, entre autres informations, l'identification du VSAN auquel appartient cette trame. Cette information sera utilisée tout au long du cheminement de la trame dans le réseau de stockage et sera retirée au niveau du dernier commutateur, afin de restituer une trame standard au serveur ou périphérique de stockage.

Extended Header du comité T11 de l'ANSI : le comité T11 de l'ANSI est l'organisme en charge de la standardisation du protocole Fibre Channel. Entre autres évolutions, le comité T11 a ajouté les en-têtes étendus (Extended Headers) dans la version 2 du standard FC-FS. Ces en-têtes sont au nombre de trois, VFT, IFR et Encapsulation Header. Le Virtual Fabric tag Header (VFT Header) est le plus avancé avec les spécifications de tagging de la trame FC, les recalculs de CRC, etc. Il comprend 8 octets qui sont ajoutés à l'en-tête de trame FC et c'est lui qui définit l'appartenance de la trame à un Fabric Virtuel via le VF_ID sur 12 bits. Ce VFT_Header comprend aussi 3 bits de QoS pour la priorisation des flux au sein du réseau. A noter que l'en-tête de type VFT apparaît également dans les standards FC-LS et FC-SW-4, en plus du FC-FS-2.

L'imposition de label de type VSAN dans une trame FC par le MDS9000 est identique à la recommandation de l'ANSI sur l'en-tête étendu de type VFT. La taille, la position et la signification des champs dans l'en-tête VSAN correspondent exactement au VFT Header tel que décrit dans les standards de l'ANSI. Nos équipements de commutation Fibre Channel disposent donc d'un mécanisme de partitionnement logique de l'infrastructure physique parfaitement standardisé.

Les VSAN sont transportés sur les différents MDS interconnectés entre eux, avec une granularité d'allocation dans une partition au niveau du port. Dans chaque partition, les définitions de Zones (Hard et Soft) seront effectuées de la même façon que dans un Fabric classique. Les Zones seront garanties en hardware, grâce à des Listes de Contrôle d'Accès (ACL), positionnées dans les ports en entrée.

Il faut bien comprendre que le partitionnement logique ne se substitue pas au Zoning. Ce dernier va masquer des ressources entre elles, mais ne va pas garantir l'isolation d'exploitation. Celle-ci ne peut être totalement garantie qu'avec des Fabriques physiquement isolés.

3.1.2 Justification du partitionnement logique

Prenons le cas d'un environnement de type SAP dans lequel on peut retrouver l'environnement de Production, l'environnement d'Intégration, le «Bac à sable» pour le développement, l'environnement de Sauvegarde, et, éventuellement les besoins de connectique pour la Réplication distante, soit potentiellement cinq environnements pour la même application. Chacun de ces environnements peut très bien être totalement isolé, et dispose de contraintes de production qui lui sont propres.

Détaillons les options possibles qui s'offrent à l'administrateur.

Option 1-) Il déploie une infrastructure physique dédiée à chaque environnement, avec un taux d'utilisation très faible des ports FC des commutateurs utilisés pour les environnements de réplication, d'intégration et Bac à sable. En effet, ces environnements sont assez réduits en nombre de ports et la redondance des chemins d'accès va augmenter le nombre de commutateurs de faible densité. De plus, l'administration d'un parc important de petits commutateurs va s'en trouver complexifié d'autant.

Option 2-) Il optimise les coûts, et déploie une infrastructure unique pour tous les environnements, avec des risques importants en terme d'intégration et de pilotage. Par exemple, Production et Sauvegarde ne font pas souvent bon ménage, à la fois pour des problèmes d'interopérabilité et de non maîtrise des flux. D'autre part, les environnements d'Intégration et le Bac à sable sont sujets à des reconfigurations incessantes, complètement antinomiques avec la stabilité nécessaire à l'environnement de Production. On comprend aisément que, même si les Zones masquent les ressources d'Intégration des ressources de Production, elles sont gérées par le même processus de Serveur de Zone, unique dans toute la Fabrique. A chaque reconfiguration de Zones dans l'un ou l'autre des environnements, la Production peut être impactée.

Les risques sont encore plus importants lorsque l'administrateur du Réseau de Stockage doit héberger des clients différents, ou groupes d'utilisateurs, soit des Productions distinctes. Le Zoning va bien masquer les ressources du Client A par rapport au Client B, mais lors d'une modification de Zone demandée par le Client A, rien ne garantit à Client B de ne pas être impacté, car l'administrateur va intervenir au niveau du processus Serveur de Zone gérant le Client A, mais aussi le Client B, qui lui, n'a rien demandé.

Option 3-) Le client déploie les différents environnements de son parc SAP sur une infrastructure à base de Directeurs MDS9500. Chaque environnement sera isolé dans sa propre partition logique et géré par un jeu de processus FC dédié. La partition de Production, la plus importante en taille, utilisera la majorité des ports de l'infrastructure. Les autres partitions disposeront chacune du nombre exact de ports requis et pas plus. L'environnement de réplication, par exemple, pourra très bien se satisfaire d'un VSAN à 4 ports répartis sur deux Directeurs MDS9500. La redondance sera garantie pour chaque partition car les VSAN seront définis sur chaque Directeur MDS9500. L'utilisation des liens ou Trunk ISL sera optimisée car ils transporteront les différents VSAN. Les reconfigurations incessantes dans les environnements d'Intégration et Bac à sable seront totalement transparentes et sans aucun risque pour la Production car gérées par des processus FC distincts.

Cisco a poussé très loin le développement de ce mécanisme de partitionnement logique car il est aussi possible de définir des profils administrateur ne disposant de droit que dans une partition donnée. C'est le mécanisme RBAC pour Role-Based Access Control. Ainsi, le Client A pourra complètement superviser sa ou ses partitions sans jamais avoir la moindre visibilité sur les partitions du Client B. L'administrateur du centre d'hébergement aura la visibilité sur la totalité de l'infrastructure matérielle, alors que chacun des Clients hébergés n'aura d'accès qu'à ses partitions logiques.

Les retours d'expérience sur la mise en œuvre d'une infrastructure mutualisée avec isolation des environnements par VSAN montrent une économie sur le nombre de ports de l'ordre de 10 à 30 %. **Le partitionnement logique en VSAN apporte donc un bénéfice immédiat en terme de Retour sur Investissement.**

3.1.3 Routage Inter-VSAN, IVR, FC-NAT et IVR Service Group

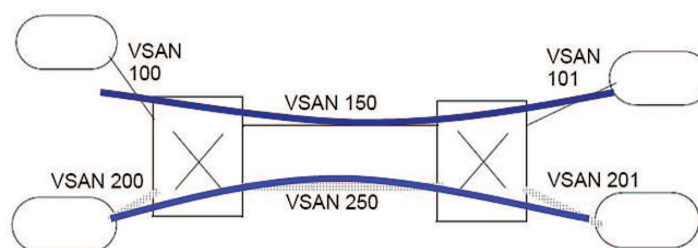
Il est possible de router entre différents VSAN. Une ressource dans un VSAN pourra communiquer avec une ressource située dans un autre VSAN qui peut être adjacent ou distant. Dans ce cas, un ou plusieurs VSAN dits de transit seront empruntés par les flux de données. Le routage inter-VSAN ne remet pas en cause les apports du partitionnement logique, à savoir l'isolation des ressources (celles qui ne sont pas concernées par le routage) et l'indépendance des processus FC entre chaque partition. Le routage s'effectue via les matrices de commutation cross-bar sans impact sur les performances.

Le routage inter-VSAN permet d'isoler certaines ressources dans une partition spécifique et de fournir l'accès de ces ressources à d'autres partitions. Par exemple, les ressources de sauvegarde peuvent être regroupées dans une partition dédiée, un « VSAN de Backup ». Il n'y a pas de ressource de sauvegarde dans les partitions de Production. Les flux de sauvegarde seront routés des partitions de Production vers la partition de Sauvegarde.

Le routage inter-VSAN apporte aussi l'isolation de certains tronçons dans le cadre de SAN étendu. Ainsi, si les Réseaux de Stockage de deux Centres Informatiques sont reliés via un réseau WAN IP, il est possible d'englober le lien WAN dans un VSAN de transit et de router entre les partitions de Production des deux Centres Informatiques via ce VSAN de transit. Toute perturbation au niveau du réseau WAN sera limitée au VSAN de transit et ne sera jamais propagée au sein des partitions de production.

Il est intéressant de noter que la mise en place de routage entre Fabriques ne nécessite pas de revoir l'adressage en place au niveau des Fabriques existantes. En effet, deux Fabriques distinctes peuvent très bien avoir des commutateurs FC configurés avec le même Domain-ID. Deux commutateurs avec le même Domain_ID ne peuvent pas se retrouver dans la même Fabrique, sinon la Fabrique va segmenter. Si on met en place un nœud de routage entre ces deux Fabriques, le MDS9000 dans lequel IVR est mis en œuvre va effectuer de la translation d'adresse (FC-NAT pour Fibre Channel Network Address Translation) au niveau des Domain_ID afin d'éviter les adresses dupliquées. Le routage IVR aussi bien que la translation d'adresse FC-NAT sont effectués en hardware au niveau des ASIC des cartes lignes et des Cross-bars. Il n'y aura donc pas d'impact sur la latence lors de la mise en œuvre du routage et de la translation d'adresse.

IVR Service Group : cette fonctionnalité permet de sélectionner un VSAN de transit spécifique pour l'interconnexion de deux VSAN distants. Dans le cadre d'une interconnexion de deux Centres Informatiques via un réseau FC ou IP, elle apporte une plus grande flexibilité dans le contrôle des chemins utilisés par les différents trafics.



Dans le schéma ci-dessus, deux Service Group seront définis :

SG1 comprenant les VSAN 100, 150 et 101

SG2 comprenant les VSAN 200, 250 et 201

Le routage inter-VSAN nécessite la licence Enterprise sauf dans le cas de routage via un VSAN de transit en FCIP, auquel cas l'IVR est inclus dans la licence SAN Extension.

3.1.4 Les VSAN en environnement Mainframe FICON

La gamme MDS9000 supporte le protocole FICON et dispose de l'ensemble des fonctions nécessaires au support de l'environnement Mainframe. Ces fonctions sont :

- Administration In-Band via la fonction CUP.
- Port et Fabric Binding afin de garantir la stabilité de la topologie SAN vue par le Mainframe.
- Switch Cascading pour pouvoir chaîner deux MDS9000 entre les ressources Mainframe.

L'intérêt du VSAN pour la mise en place du protocole FICON est évident. Les ressources FICON sont des ressources dédiées, qui ne seront jamais accédées par une ressource FC. Un VSAN FICON sera configuré dans lequel ne seront connectés que les canaux FICON des Mainframes et baies de stockage, à l'exclusion de tout autre HBA ou périphériques et recevra toutes les ressources Mainframe. Ce VSAN FICON sera géré par son jeu de processus FC dédié et sera complètement isolé de tout événement survenant dans les partitions Open voisines. Une reconfiguration de zones pour insérer un serveur Windows ne présentera strictement aucun risque pour l'environnement Mainframe.

Toutes les cartes ligne 2^{ème} génération ont été qualifiées par IBM en environnement FICON. Le MDS9500 est le Directeur FICON disposant de la plus grande densité de port sur le marché à l'heure actuelle avec jusqu'à 528 ports FICON sur un seul MDS9513.

3.1.5 Configuration des VSAN

Nous avons vu que la granularité d'affectation dans un VSAN est au niveau du port sur le MDS9000. Avant d'affecter les ports au VSAN, il faut d'abord définir le VSAN. Cela peut se faire en local sur chaque MDS9000, mais il faut une définition rigoureusement identique sur chaque MDS9000. On utilisera donc l'outil d'administration embarqué Cisco Fabric Manager. A partir du CFM, l'administrateur définit le VSAN puis sélectionne l'ensemble des MDS9000 sur lequel ce VSAN doit être déployé.

L'affectation des ports dans le VSAN peut se faire de façon manuelle, au niveau de chaque MDS9000. Mais il est aussi possible d'utiliser un mécanisme d'affectation automatique en fonction du WWPN du nœud qui se connecte sur le port. Ce mécanisme d'affectation est centralisé via la fonction Cisco Fabric Services. CFS est décrit plus loin dans le document. CFS est accédé en In-Band et contient une table de correspondance des WWPN avec les VSAN. Un port non utilisé d'un MDS9000 est dans un état dynamique et n'appartient à aucun VSAN. Lorsque un serveur ou périphérique est connecté sur ce port du MDS9000, ce dernier récupère son WWPN lors de la phase de négociation initiale. Il interroge alors la base de donnée de CFS pour apprendre à quel VSAN appartient ce WWPN. Il configure ensuite le port concerné dans le VSAN voulu et le serveur peut effectuer ses phases de Login dans la Fabrique SAN à laquelle il appartient.

3.1.6 Bénéfices de la Commutation FC et Routage intégrés

La mise en œuvre de routage entre Fabriques avec la gamme MDS9000 offre plusieurs avantages uniques sur le marché à l'heure actuelle. Le MDS9000 supporte la Commutation et le Routage des trames Fibre Channel de façon intégrée. Les bénéfices en sont les suivants :

- Pas de nécessité d'un équipement de routage externe,
- Pas de gaspillage de liens ISL entre les commutateurs FC et les routeurs externes,
- Pas besoin de sécuriser le nœud de routage lui-même lorsqu'il est mis en œuvre dans un châssis MDS9500 car il dispose automatiquement de la haute disponibilité propre à cet équipement de classe Directeur.
- Pas de saut (hop) supplémentaire dans le réseau de stockage SAN,
- Pas besoin de modifier les Domain-ID des commutateurs lorsqu'ils sont identiques grâce à la translation d'adresse FC-NAT,
- Routage et translation d'adresse supportés entre Fabriques hétérogènes,
- Lorsque le routage s'effectue au-dessus de liens FCIP vers des sites distants, la connectique IP peut également être intégrée directement dans le MDS9500 ou MDS9200,
- Administration de bout en bout avec le logiciel Cisco Fabric Manager, y compris dans le cas de Fabriques hétérogènes ou d'interconnexion via IP (tunnels FCIP).

Le routage FC faisant partie intégrante du MDS9500, la latence sera toujours stable quels que soient les cas de figures :

- Commutation de trames Fibre Channel au sein d'une même Fabrique,
- Commutation et Routage de trames Fibre Channel entre deux Fabriques distinctes,
- Commutation et Routage de trames Fibre Channel entre deux Fabriques distinctes avec translation d'adresse FC-NAT.

Au sein du Directeur MDS9500, les mêmes ASIC exécutent aussi bien la commutation des trames FC que le routage de ces trames. Les bénéfices pour l'utilisateur sont les suivants :

- Aucune dégradation de performance, ni de latence. Chaque slot offre 256 Gbps full-duplex de capacité de Commutation ET de Routage.
- Pas d'augmentation de la complexité, vu que le même ASIC effectue les deux tâches.
- Pas d'augmentation de la complexité de l'infrastructure et de la gestion de la topologie, vu qu'il n'y a aucun besoin d'appliance externe.

La gamme MDS9500 est capable de proposer cette intégration Commutation / Routage car ces mécanismes sont maîtrisés depuis plus de 10 ans par les équipes de développement de Cisco sur d'autres types de protocoles – Ethernet / IP. Le cheminement normal d'un équipementier en phase d'acquisition de ces compétences est le suivant :

1. Equipement de commutation monolithique, ni partitionnement, ni routage,
2. Routage entre équipement via routeur externe,
3. Intégration du routeur externe sous forme de lame dans le châssis, la lame devient un goulet d'étranglement dans le châssis et dégrade sa fiabilité,
4. Partitionnement du châssis, mais routage toujours via la lame avec les mêmes contraintes que ci-dessus,
5. Intégration commutation et routage banalisé.

Ce cheminement a été effectué par Cisco il y a 10 ans environ sur d'autres protocoles. Dès notre arrivée sur le SAN il y a 3 ans et de par la compétence acquise, nous avons été à même de passer à l'étape finale, étape 5, directement. L'intégration complète de ces mécanismes garantit leur simplicité et leur fiabilité car les mêmes ASIC vont effectuer aussi bien la commutation des trames que le routage inter-SAN. Ils ont été acceptés et mis en œuvre par pratiquement la totalité des utilisateurs de MDS9500.

3.2 Définition des Zones par LUN

Dans une partition logique, le Zoning va masquer des ressources les unes par rapport aux autres. Ce mécanisme fonctionne par entité FC, qu'elle soit identifiée par port physique, adresse FCID ou WWPN. Il n'est pas possible de découper plus finement, notamment au niveau de l'interface d'une baie de stockage sur laquelle plusieurs volumes logiques (LUN) sont accrochés. La seule possibilité est d'utiliser la fonction de LUN Masking de la baie, qui ne va montrer que la LUN au serveur qui lui a été affecté. Cette fonction est propre à chaque baie, en fonction des fournisseurs, et la définition n'est pas centralisée.

Aujourd'hui, afin de proposer une isolation efficace, il faut combiner les fonctions de Zoning dans la Fabrique, avec les fonctions de LUN Masking dans chaque baie, le tout sur des Fabriques physiquement isolées dans le cas où une isolation forte est demandée, cas d'un hébergeur. L'intégration de la solution ne s'en trouve pas facilitée.

En combinant le mécanisme des VSAN avec le Zoning par LUN, la gamme MDS répond aux deux contraintes énoncées ci-dessus, à savoir à la fois l'isolation forte par Fabrique distinct – c'est le partitionnement logique par VSAN - et la finesse d'isolation au niveau des volumes logiques attachés à une même interface baie de stockage – c'est le LUN Zoning. Ces deux fonctions étant exploitées au sein même de l'infrastructure réseau, l'administrateur va disposer d'un outil unique, centralisé, pour répondre à toute demande, partitionnement logique à grande échelle, ou finesse de découpage sur chaque interface baie de stockage.

La gamme MDS9000 dispose aussi de la fonction de Zones Read-Only. Quelle que soit la définition initiale, WWPN, Port physique ou LUN Zoning, cette zone peut être marquée en lecture uniquement. Les serveurs pourront consulter les données des volumes logiques auxquels la visibilité leur est donnée, mais ils ne pourront en aucun cas les modifier. Cela est particulièrement approprié dans le cas où une entité dans l'entreprise est responsable de la mise à disposition d'informations qui seront consultées, mais sans possibilité de les modifier, par les autres entités de l'entreprise.

Le Zoning est garanti en hardware, par liste de Contrôle d'accès positionnée directement dans les ASIC sur les ports en entrée dans les Directeurs MDS9500. Ces ACL sont positionnées au moment de l'activation du ZoneSet. Elles vont filtrer les trames sur l'ensemble des champs d'en-tête FC et SCSI, y compris le numéro de LUN, sans aucun impact sur les temps de commutation. Le nombre maximum d'entrées dans une ACL est de 20000. La gamme MDS9000 supporte les dernières évolutions des standards de l'Ansi, FC-GS-4 et FC-SW-4.

Sécurité accrue du réseau avec le Zoning Cisco. Les réseaux de type SAN à base du protocole FC restent aujourd'hui peu sujet aux attaques que l'on trouve dans les réseaux IP. Cependant, il est préférable de se tenir prêt et la gamme MDS9000 peut répondre à certains risques inhérents aux topologies réseaux, comme les attaques en déni de service (Denial of Service - DoS - attacks). Le filtrage des trames en fonction des règles de Zoning est effectué en entrée du réseau SAN. Une attaque de type DoS sera systématiquement bloquée en entrée et ne parviendra pas à inonder le réseau ni à saturer le port d'un périphérique. Au niveau d'une baie de stockage, le LUN Masking permet de masquer certains volumes logiques pour les serveurs non autorisés. Mais le LUN Masking n'est pas protégé contre les attaques de type DoS. Pour se prémunir contre ce risque, le LUN Zoning de la gamme MDS9000, unique sur le marché à l'heure actuelle, permet d'isoler les volumes logiques au sein d'une interface unique et de filtrer les trames indésirables dès leur entrée dans le réseau SAN. Elles ne pourront donc pas inonder le réseau et encore moins accéder à la baie de stockage pour perturber l'accès des autres serveurs autorisés.

3.3 Connectique IP et déport de SAN

La gamme MDS dispose d'un panel de possibilités très complet pour dépasser les limites de la technologie Fibre Channel. Il s'agit de la connectique IP intégrée pour les protocoles iSCSI et FCIP, ainsi que du multiplexage en longueur d'onde sur Fibre Noire. Il faut noter que la connectique IP n'est pas disponible pour la gamme MDS9100.

3.3.1 Extension du SAN par encapsulation iSCSI

Le besoin d'extension du Réseau SAN se fait de plus en plus pressant. Ce besoin est dicté par la volonté d'intégrer le maximum de serveurs de l'entreprise dans une infrastructure commune pour le transport du protocole SCSI vers des ressources mutualisées. Malheureusement, le ticket d'entrée de la technologie FC, technologie de fait pour le transport du protocole SCSI, est relativement élevé. Le ratio moyen entre HBA serveurs et interfaces baies de stockage est de 7 pour 1 en technologie 2 Gbps, soit un débit moyen de 30 MO/s là où le débit nominal de la technologie FC est de 100 ou 200 MO/s. Dans certains cas, le ratio est de 15 pour 1, soit un débit moyen d'environ 13 MO/S.

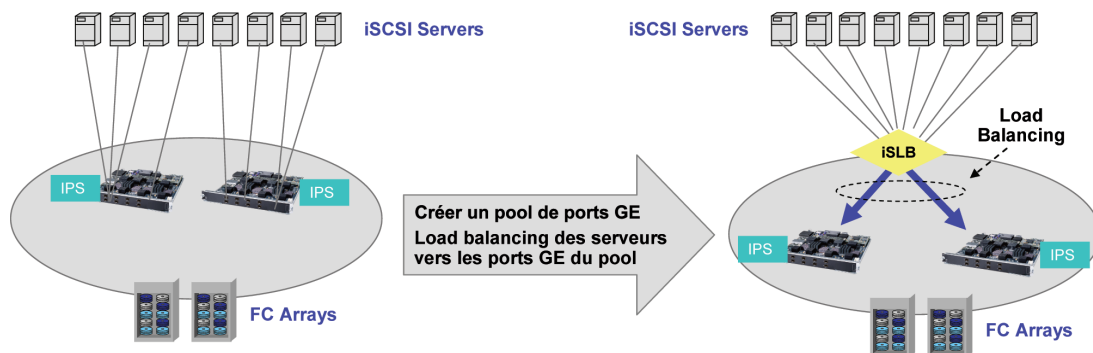
C'est le but du protocole iSCSI, ou SCSI sur IP, qui encapsule le jeu de commande et données SCSI dans des trames TCP/IP pour utiliser l'infrastructure IP existante, et offrant des débits en adéquation avec le besoin réel de chaque serveur. L'encapsulation iSCSI masque complètement le réseau IP utilisé pour le transport des flux SCSI. L'OS du serveur dispose d'un accès direct aux données, tout comme un serveur avec HBA FC.

Sur des débits de 10 à 15 MO/s de flux SCSI, le coût de l'infrastructure en SCSI sur IP est de l'ordre de 2 à 3 fois moindre par rapport au réseau FC. Mais le gain financier majeur se retrouve sur l'interface serveur d'un coût négligeable en technologie SCSI sur IP, alors qu'un HBA vaut environ 1500 Euros, prix public, en technologie SCSI sur FC, soit 3000 Euros si l'on veut la redondance des chemins d'accès. D'un point de vue topologie, les utilitaires de MultiPathing fonctionnent de la même façon en SCSI sur IP qu'en SCSI sur FC. Toute défaillance d'un élément du réseau sera traitée par la redondance des chemins d'accès, que l'élément défaillant soit sur l'infra IP / Ethernet ou l'infra Fibre Channel.

En environnement Wintel, les dernières versions de l'OS Windows intègrent le MultiPathing par MPIO pour iSCSI. Il n'est pas nécessaire de mettre en œuvre de MultiPathing en fonction de la baie, il est natif dans Windows.

Les modules IP supportent également le mécanisme VRRP et iSLB (iSCSI Server Load balancing) pour la haute disponibilité des passerelles d'encapsulation iSCSI.

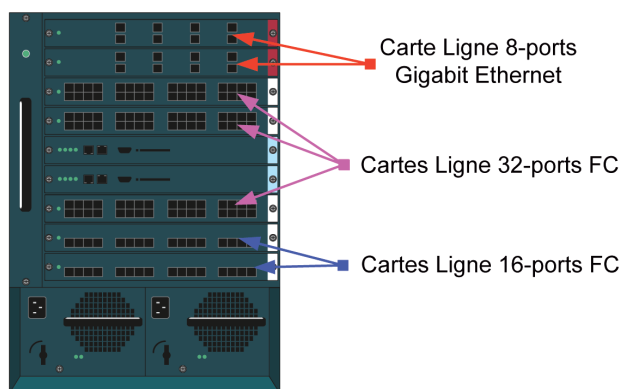
Cette fonctionnalité simplifie et optimise le déploiement d'un grand nombre de serveurs iSCSI. Combiné au mécanisme VRRP, elle accroît la sécurité d'accès et les performances à l'ensemble des serveurs iSCSI qui accèdent à des ressources de stockage sur le SAN. En effet, le service iSLB permet de répartir dynamiquement la charge des serveurs iSCSI vers un groupe de ports Gigabit Ethernet, tous actifs, disponibles sur les MDS9000. Il offre une meilleure utilisation des port GE et accroît la haute disponibilité des ressources.



Dans la version actuelle, VRRP impose un fonctionnement en mode active/esclave, avec iSLB / VRRP le système fonctionne en mode active/active pour partager la charge.

La gamme MDS9500 dispose des modules IPS-8 (8 GE) et MPS (14FC + 2GE) en interface Gigabit Ethernet afin d'offrir cette connectique IP en natif. Chaque interface GbE pourra transporter les flux de stockage sur IP, soit en encapsulation iSCSI, soit en encapsulation FCIP. Les mécanismes d'isolation et de sécurité IP seront intégrés aux configurations des partitions logiques Fibre Channel, les serveurs en IP disposeront d'une adresse FC virtuelle et ils seront gérés par le réseau FC de la même façon que les serveurs avec HBA FC. Le Zoning, par exemple, sera mis en œuvre de la même façon pour les serveurs en FC natif et pour les serveurs en iSCSI. Chacun de ces serveurs présentera une adresse de type WWPN, réelle ou virtuelle.

Dans le schéma de configuration ci-après, le Directeur MDS9509 dispose de 128 ports FC et 16 ports GbE. Il peut fournir l'accès aux baies de stockage à environ 100 serveurs en attachement HBA et 200 serveurs de moindre capacité en accès iSCSI. Avec redondance des chemins d'accès, il faut deux Directeurs MDS9509 équivalents, pour 100 serveurs avec double HBA et 200 serveurs de moindre capacité avec double accès Ethernet.



MDS9509 128 Ports FC + 16 Ports GbE

Il n'y a pas de restriction dans l'implantation des cartes ligne dans les Directeurs MDS9500, le nombre et la localisation des modules IPS-8 par rapport aux modules FC n'a aucune importance.

La mise en œuvre du protocole iSCSI peut se faire en deux modes, Transparent et Proxy :

- Mode Transparent : lorsque le module IP du MDS9000 reçoit une demande de connexion d'un serveur iSCSI, il va affecter un WWPN à ce serveur en connectique IP. Ce WWPN peut être affecté en dynamique ou prédéfini en statique. Ce WWPN est ensuite présenté sur la Fabrique pour les phases de Login. Le Serveur de Noms va enregistrer ce WWPN et lui affecter un FCID comme un serveur en attachement HBA classique. Le Zoning sera mis en œuvre pour chaque serveur iSCSI de la même façon qu'un serveur HBA. Ce WWPN est aussi utilisé par les baies de stockage pour les fonctions de LUN Masking.
- Mode Proxy : un WWPN est prédéfini et présenté aux processus FC de la Fabrique. Les différents serveurs iSCSI qui montent leur session iSCSI à destination du MDS9000 seront représentés par ce WWPN unique. Cela permet de faciliter l'ajout de serveur iSCSI sans avoir à intervenir au niveau du Zoning ou du LUN Masking de la baie de stockage.

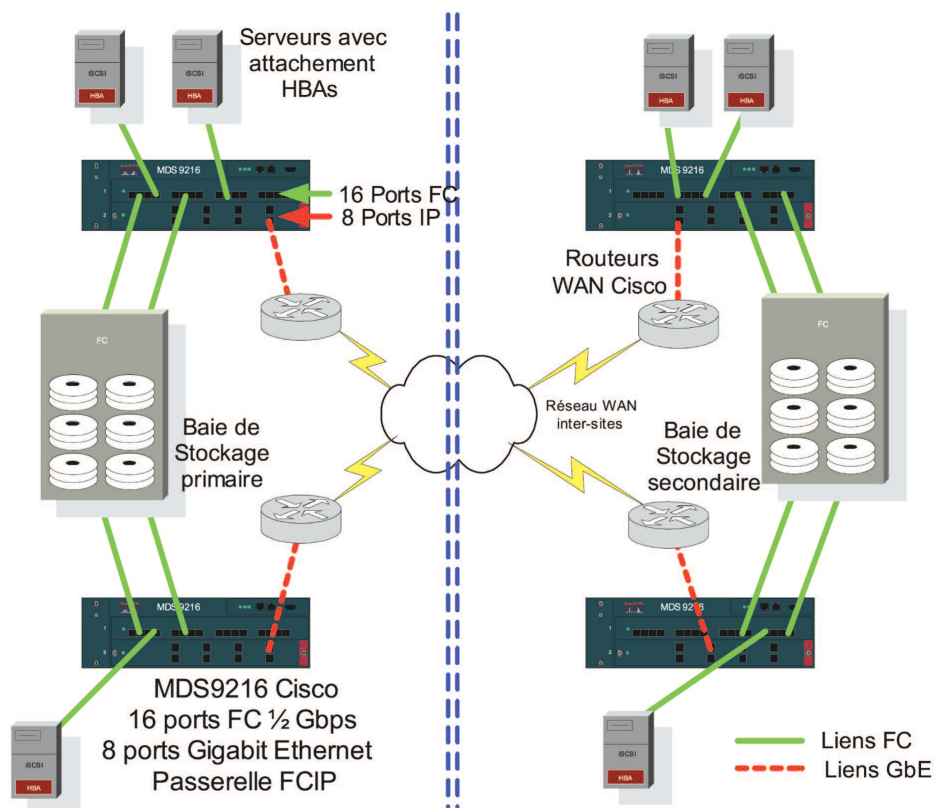
Chacune des interfaces des cartes IPS-4, IPS-8 et MPS peut gérer jusqu'à 200 serveurs iSCSI.

3.3.2 Déport de SAN par encapsulation FCIP

La nécessité d'étendre le Réseau SAN est aussi dictée par des besoins d'interconnexion de Centres Informatiques distants, Production et Secours, ou Hébergement et Secours. Plusieurs techniques sont possibles. Soit par Fibre Noire, lorsque les distances sont de l'ordre de la centaine de kilomètres, soit par Tunneling des trames FC dans IP au-delà, ou lorsque le coût de la fibre est prohibitif, ou que le Centre de Secours peut changer de localisation géographique – besoin d'un réseau d'accès maillé.

Pour aller au-delà des distances de la Fibre Noire, tels que décrit dans le chapitre suivant, ou pour bénéficier d'un réseau WAN IP existant, la gamme MDS peut utiliser le protocole FCIP qui va encapsuler les trames FC dans TCP/IP, et masquer complètement le réseau IP. Les deux environnements FC distants vont se connecter l'un à l'autre comme s'ils étaient adjacents, et les données de Production pourront être répliquées sur le site de Secours, comme si les ressources étaient locales.

Le schéma ci-dessous présente un exemple d'utilisation des modules IP intégrés dans les MDS pour assurer la réplication de données entre sites distants. Les baies de Stockage utilisent des interfaces dédiées pour leur interconnexion, et l'encapsulation FC dans IP est assurée directement dans les MDS9000 via le module IPS-8 :



Solution de Réplication sur IP avec Infrastructure Cisco

Les ports Gigabit Ethernet peuvent également être affectés dans des partitions logiques, au même titre que les interfaces Fibre Channel. Ainsi, on peut définir un VSAN réplication qui va isoler logiquement l'infrastructure FC et GbE nécessaire à la réplication, tout en utilisant la passerelle FC sur IP intégrée dans le Directeur MDS.

Les flux FCIP peuvent être compressés en software sur les cartes IPS et en hardware sur la carte MPS. Ils peuvent être chiffrés en IPSec sur la carte MPS. Afin d'optimiser le comportement des applications sans que la distance pénalise trop les performances, les modules IPS et MPS disposent des mécanismes d'accélération :

- Write Acceleration : ce mécanisme effectue un « Spoofing » du R_RDY afin de ne pas avoir de dégradation des performances lorsque la distance augmente. La fonction WA effectue un acquittement local de la demande de transfert de données ce qui a pour effet de démarrer le transfert de façon immédiate.
- Tape Acceleration : destiné à optimiser le comportement des périphériques de sauvegarde, la fonction TA travaille en conjonction avec la fonction WA et apporte un « Spoofing » de la phase Etat de fin d'I/O en local afin d'enchaîner les I/Os en continu

Les tunnels FCIP peuvent être ajustés au niveau des paramètres TCP, en fonction des caractéristiques de débit et de latence du lien WAN. Le fenêtrage TCP sera ajusté en fonction des valeurs fournies.

La gamme MDS9000 dispose également de la fonction SET (SAN Extension Tuner) qui va générer des I/O spécifiques pour évaluer le comportement d'une application donnée. En fonction du profil des I/Os, l'administrateur va générer un jeu de commande SCSI. Ensuite, il pourra vérifier le débit et la latence obtenue.

3.3.3 Interop mode

Pour faciliter la migration des infrastructures SAN existantes, les MDS9000 offre la possibilité de s'interconnecter aux switches FC Brocade et McData sans avoir à modifier leur configuration. Le mode interop-4 permet de connecter des MDS9000 à des commutateurs McData en mode natif sur une infrastructure SAN existante. Il vient compléter les modes d'interopérabilité 2 et 3 existants pour l'interopérabilité avec des commutateurs Brocade en mode natif.

Les modes d'interopérabilité se configurent par VSAN. Il n'y a pas de restriction à la mise en œuvre des différentes fonctions réseaux évoluées selon les différents modes d'Interop, à savoir VSAN, IVR, QoS, PortChannel, multi-protocoles, etc. Ainsi, des ressources connectées sur des commutateurs non-Cisco pourront communiquer entre elles via le routage intégré au sein du MDS9000. Chaque commutateur non-Cisco sera dans un VSAN différent avec, le cas échéant, un mode d'Interop différent. Les commutateurs non-Cisco n'auront pas connaissance l'un de l'autre car les VSAN n'auront pas mergé. On s'affranchit ainsi des problèmes de compatibilité.

3.3.4 IPv6 sur les interfaces LAN

Le MDS9000 supporte le protocole IPv6 sur les interfaces d'administration ainsi que sur les cartes lignes IPS et MPS disposant des interfaces Gigabit Ethernet pour ISCSI et FCIP. On peut ainsi établir des tunnels FCIP en IPv6 de bout en bout.

3.3.5 Multiplexage en longueur d'onde

En déport par Fibre Noire, Cisco dispose des équipements de la gamme ONS15454 qui apportent des capacités en multiplexage de type DWDM (Dense Wavelength Division Multiplexing).

L'ONS15454 est capable de multiplexer jusqu'à 32 longueurs d'onde de 10 Gigabit chacune, soit un débit total de 320 Gbps sur une fibre monomode. Sur chacune de ces longueurs d'onde, il est possible d'effectuer un premier niveau de multiplexage temporel, de type TDM. Ainsi, sur chaque longueur d'onde à 10 Gbps, on pourra avoir 8 canaux de type Fibre Channel, FICON ou Gigabit Ethernet, permettant de transporter jusqu'à 256 canaux - 8 canaux par longueur d'onde, 32 longueurs d'onde – de type Fibre Channel, FICON ou Gigabit Ethernet, sur des distances de plusieurs centaines de Kms sur une fibre monomode. L'ONS15454 peut recevoir également des modules d'agrégation de 8 ports ESCON sur une longueur d'onde à 2,5 Gigabit.

Dans le cas où les besoins en connectique sont moindres, il est possible d'utiliser le multiplexage en longueur d'onde de type CWDM. Les commutateurs de la gamme MDS9000 supportent des interfaces FC spécifiques à une longueur d'onde (SFP-CWDM). Avec ces multiplexeurs de type passif, on peut transporter jusqu'à 8 canaux de type Fibre Channel ou Gigabit Ethernet sur une fibre monomode, sur des distances jusqu'à 80 Kms.

3.4 Qualité de Service

L'objectif des mécanismes de QoS – Qualité de Service – de la gamme Cisco MDS9500 est de traiter les problèmes de congestion au sein du réseau. La QoS permet de différencier les flux de données et de les prioriser en cas de congestion, n'importe où dans le réseau de stockage. En fonction des niveaux de service – SLA – assigné par l'administrateur SAN aux différentes applications, le MDS9500 sera à même de prioriser des flux critiques tels que des flux transactionnels par rapport à d'autres flux moins critiques tels que des flux de sauvegarde ou de transfert de fichier.

Il faut définir les différentes classes de flux qui seront différenciés. Plusieurs possibilités sont offertes, la plus simple étant au niveau de la Zone. La hiérarchie de priorisation sera établie en fonction de l'appartenance des nœuds à telle ou telle zone. Il est possible de différencier les flux entre VSAN. La valeur de QoS pour traiter la priorité d'une trame lors de son transit dans le réseau sera spécifiée dans le champ d'en-tête du Virtual Fabric Tag tel que défini par le comité T11 de l'Ansi. La mise en œuvre de la Qualité de Service par les MDS9000 est conforme au standard en vigueur.

La QoS n'est pas un mécanisme pour gérer la congestion dans le MDS9500, ce dernier ne peut pas se retrouver en congestion car son architecture interne de type cross-bar est totalement non-bloquante. La QoS sera mise en œuvre en cas de congestion au sein du réseau SAN, sur un lien ISL ou une interface baie de stockage. Dans le cas de deux serveurs, S_A supportant des applications critiques et S_B du développement, voulant envoyer simultanément 2 Gbps de données sur un lien ISL à 2 Gbps, ce dernier ne pourra pas transporter les 4 Gbps de flux. Sans QoS et dans le meilleur des cas, chaque serveur aura 50 % de la bande passante disponible, soit 1 Gbps chacun. Avec les MDS9500 équipés de QoS, S_A pourra avoir 80 % des 2 Gbps du lien ISL et le flux de développement de S_B sera contraint à 20 %.

La QoS améliore le taux d'utilisation des ressources sans devoir faire grossir son infra, ce qui permet aux utilisateurs d'optimiser l'utilisation de leur Réseau de Stockage tout en retardant dans le temps les acquisitions de bande passante additionnelle. Un cas relativement fréquent est celui dans lequel les liens inter-sites sont utilisés pour transporter des flux de réplication et des flux de sauvegarde. Sans QoS, lorsque les flux de sauvegarde augmentent, ils vont pénaliser la réplication et donc impacter directement les temps de réponse. Sans QoS, la seule option est d'ajouter de la bande passante inter-site, ce qui va aussi nécessiter l'ajout de transpondeur sur les MUX optiques dans le cas de multiplexage en longueur d'onde. La QoS va permettre d'optimiser l'utilisation de la bande passante et donc d'éviter de faire grossir la capacité inter-sites. **La QoS apporte donc un bénéfice immédiat en terme de Retour sur Investissement.**

3.5 Contrôle de flux

La gamme MDS9000 supporte les Classes de Service 2, 3 et F. La Classe de Service 2 met en œuvre les mécanismes de contrôle de flux Buffer to Buffer Credit pour gérer la congestion dans le réseau, et End-to-End Buffer Credit, pour gérer la congestion de bout en bout, cas où le récepteur ne peut s'aligner sur le débit des trames envoyé par l'émetteur. Malheureusement, 99 % des implémentations ne supportent que la Classe de Service 3. Dans ce cas, le récepteur saturé va utiliser son mécanisme Buffer to Buffer Credit, mais il ne pourra que ralentir le commutateur adjacent, qui lui-même va ralentir son commutateur adjacent, et ainsi de suite, la congestion va remonter dans le réseau, et pénaliser d'autre flux que le flux en congestion.

En Classe de Service 3, la gamme MDS9000 dispose du mécanisme Fibre Channel Congestion Control, qui va apporter le contrôle de la congestion de bout en bout. Lorsqu'un récepteur est saturé, il va prévenir le MDS9000. Ce dernier, en examinant les trames en attente dans ses Buffers, va détecter l'origine de la congestion et la source du flux. Il va prévenir directement le MDS9000 en amont du flux, celui sur lequel est connecté l'émetteur, quelle que soit la topologie réseau entre les deux MDS. Le MDS9000 en amont va alors utiliser le mécanisme Buffer to Buffer credit pour ralentir l'émetteur. Le flux sera ainsi régulé au mieux des possibilités du récepteur, sans pénaliser les autres flux transportés au sein de l'infrastructure Fibre Channel. Dans le cas de topologie réseau se complexifiant, le mécanisme FCC apporte aux flux de données en Classe de Service 3, la plus utilisée, le contrôle de flux de type End-to-End Buffer Credit de la Classe de Service 2.

3.6 Service Intelligents au cœur du Réseau de Stockage

Les équipements modulaires de la gamme MDS, les Directeurs MDS9500 et commutateurs MDS9216 peuvent recevoir le module de service SSM. L'objectif est de déployer des utilitaires de stockage au cœur même du réseau de stockage. En tant qu'élément fédérateur, le réseau offre ainsi la meilleure visibilité à

de tels utilitaires pour accéder à l'ensemble des ressources serveurs et périphériques de stockage.

Ces services intelligents sont mis en œuvre grâce au module de service Storage Service Module (SSM). Le module SSM dispose de 32 ports FC en connectique 1 et 2 Gbps. Ces ports offrent les mêmes caractéristiques FC que ceux de la carte 32 ports.

Le module SSM permet de mettre en œuvre différents types de services intelligents.

3.6.1 Applications intégrées dans le Réseau

Le module SSM est conforme aux normes de l'API Fabric Assisted Interface Spécification (FAIS). Tout logiciel de stockage au format FAIS peut s'intégrer avec la SSM et permet de mettre en œuvre des utilitaires de stockage intégrés dans le réseau. En fonction des possibilités du logiciel pilotant la SSM, les services suivants vont pouvoir être mis en œuvre :

- Administration simplifiée,
- Migration de données en dynamique sans impact applicatif, entre baies de stockage hétérogènes,
- Regroupement de l'espace disque (pooling), création de pools entre baies de stockage hétérogènes,
- Provisionning,
- Support de copie Point-in-Time ou Snapshot,
- Services de continuité de service et de protection des données – Continuous Data Protection (CDP),
- Réplication synchrone ou asynchrone en environnement hétérogène.

Le logiciel Invista est un exemple de solution au format FAIS s'intégrant nativement sur le module de service SSM du MDS9000.

3.6.2 Applications assistées par le Réseau

Le module SSM dispose de la fonction SANTap. Ce mécanisme va dupliquer le flux de données commuté au sein du MDS pour traitement par appliance externe. La duplication du flux se fait sans impact de performance ni de latence sur le flux primaire entre le serveur et la baie. Les services intelligents seront ensuite mis en œuvre par l'appliance et seront donc fonction du type d'appliance. Les bénéfices de la fonction SANTap sont les suivants :

- Pas d'agent logiciel au niveau de chaque serveur applicatif (cas des solutions Out-of-Band),
- Pas d'impact sur les performances des flux primaires,
- Pas de point de passage obligé, pouvant affaiblir la haute disponibilité du réseau (cas des solutions In-Band),
- Peut être mis en œuvre de façon dynamique et sans impact,
- Supporte les environnements hétérogènes,

En fonction des appliances, SANTap peut être mis en œuvre en mode transparent dans lequel soit l'Initiator, soit le Target est connecté directement sur la SSM ou en mode Proxy dans lequel aucun des deux n'est connecté directement sur la SSM.

Les appliances supportant le protocole SANTap sont celles des fournisseurs tels que Kashya, Topio, Xiotech, Alacritus ou encore FalconStor.

3.6.3 Sauvegardes accélérées par le Réseau

Le module SSM s'interface également avec les logiciels de sauvegarde pour la mise en œuvre de sauvegardes accélérées. Dans une sauvegarde classique, le serveur de sauvegarde (Media Server) doit lire les données sur la baie de stockage et écrire ces données sur les périphériques de sauvegarde. Les flux de sauvegarde vont ainsi passer deux fois à travers le réseau de stockage et le serveur de sauvegarde devra avoir un nombre d'interface FC suffisant pour supporter le volume de données à sauvegarder. La SSM supporte les commandes de type XCOPY. Ainsi, le Media Server va simplement superviser la SSM. Il va envoyer les ordres de sauvegarde à la SSM et c'est elle qui va lire les données sur les baies à sauvegarder et les écrire directement sur les périphériques de sauvegarde. Il va donc y avoir une double économie, au niveau du ou des Media Servers, ainsi qu'au niveau de la charge des flux au sein du réseau SAN.

Les solutions supportées aujourd'hui sont celles de Computer Associates, Comvault et EMC/Legato.

3.6.4 Accélération des flux et statistiques LUN

Le module SSM offre également la fonction de Write Acceleration pour les flux FC natifs. De la même façon que le FCIP-WA, le FC-WA va effectuer un «spoofing» de la requête SCSI initiale et économiser le RTT correspondant au X-RDY. Dans le cadre d'interconnexion de Centres Informatiques distants en FC natif, via fibre noire ou multiplexage en longueur d'onde, DWDM ou CWDM, la dégradation du temps de réponse en fonction de la distance sera ainsi atténuée.

Enfin, le module SSM est à même de fournir des statistiques détaillée au niveau des accès aux différents volumes logiques. C'est la fonction SCSI Flow Statistics.

Les différentes fonctions évoluées décrites ci-dessus sont accessibles par licence logicielle :

- Une carte SSM sans licence offre 32 ports de base pour la connectique FC,
- La licence Enterprise est obligatoire pour les fonctions FC-WA et SCSI Flow Statistics,
- La licence Storage Service Enabler (SSE) est obligatoire pour le support des logiciels au format FAIS, de la mise en œuvre du protocole SANTap et du backup accéléré avec XCOPY.

La licence Enterprise s'applique au MDS9000 dans son ensemble, une seule est nécessaire par équipement alors que la licence SSE s'applique au module SSM lui-même, il en faut une par module.

A noter que l'ensemble des fonctions intelligentes décrites dans ce chapitre sont disponibles avec le niveau de firmware 2.1(2).

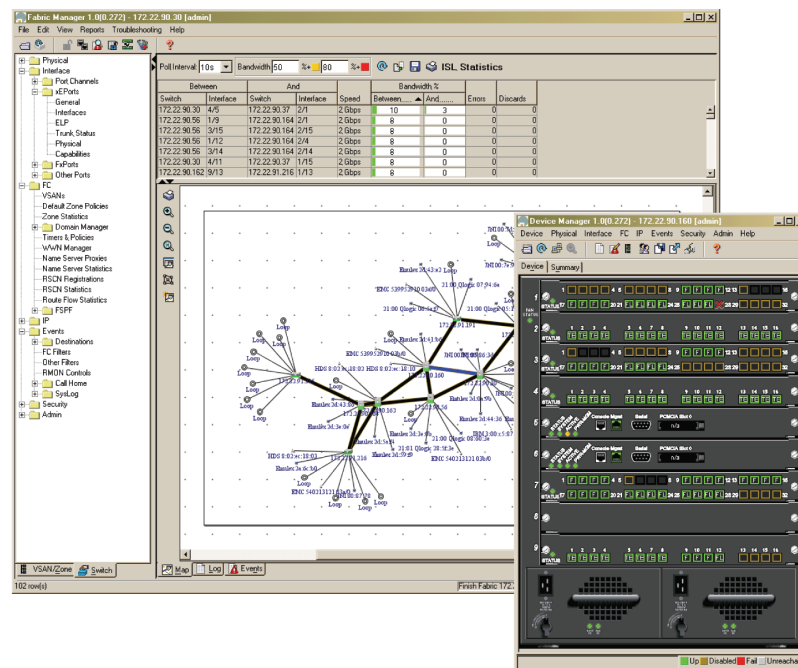
4. ADMINISTRATION ET DEBUGGING

Le pilotage du Réseau de Stockage devient extrêmement critique car ces environnements se complexifient en devenant de vrais réseaux, à topologie maillée, interconnexion distante, design de type Core-Edge, etc.

L'administrateur doit donc disposer d'un point focal pour une supervision globale, d'accès entièrement sécurisé, qu'elle que soit sa localisation géographique, et enfin d'outils d'analyse performants.

4.1 Administration puissante

Le Cisco Fabric Manager est un outil d'administration intégré dans chaque MDS. Il va assurer la découverte et la gestion de la topologie FC, la configuration et administration de chaque Directeur MDS9509, le diagnostic du réseau, la remontée d'alertes, la supervision et le «reporting» statistique et temps réel de l'état du réseau et de la charge des liens. L'accès au Fabric Manager est sécurisé par SSHv2 et SNMPv3. De plus, en mode console ou Telnet, le CLI est identique à celui de L'IOS Cisco. L'accès est possible en «Out-of-Band», par interface Ethernet, et en «In-Band», par IP sur FC (RFC2625). Les protocoles FC-GS-3 et FC-GS-4 vont apporter la gestion de toute la topologie FC, y compris les Serveurs et périphériques. Les MDS peuvent s'intégrer dans RME de CiscoWorks pour la gestion du parc réseau. Enfin, la fonction Call Home assure le relais vers les équipes de permanence.



Toutes les opérations relatives à la topologie seront effectuées dans la vue Fabric ci-dessus. Ces opérations seront, entre autres, la configuration des partitions logiques, la définition du Zoning, etc. Elles seront toujours réparties sur l'ensemble des nœuds FC.

Le paramétrage d'un MDS en particulier sera effectué à partir de la vue Device. Les opérations comme l'affectation des ports dans les VSAN, la création des PortChannels, etc. seront réalisées dans la vue Device.

En sélectionnant l'onglet Summary, le comportement d'un MDS9000 sera visualisé en temps réel, par la vue ci-dessous :

The screenshot shows the 'Device Manager 1.0(0.272) - 172.22.90.56 [admin]' window. The 'Summary' tab is active, showing a poll interval of 10s and thresholds for CPU (0%), Memory (7%), and Flash (23%). Below this, there are three tables:

xEPorts (Inter Switch Links)									
Port	Mode	Channel	Speed	VSAN(s)	Neighbor pW/WWN	Rx Utilization%	Tx Utilization%	Errors	Discards
3/4	TE		2 Gbps	300	20:4b CiscoMDS 00:26:9e (2/11)	88	68	0	0
1/12	TE		2 Gbps	300	20:44 CiscoMDS 00:25:5e (2/4)	80	85	0	0
1/9	TE		2 Gbps	300	20:41 CiscoMDS 00:25:5e (2/1)	80	2	0	0
1/10	TE		2 Gbps	300	20:42 CiscoMDS 00:25:5e (2/2)	88	49	0	0
1/11	TE		2 Gbps	300	20:43 CiscoMDS 00:25:5e (2/3)	85	60	0	0
3/9	TE		2 Gbps	300	20:49 CiscoMDS 00:26:9e (2/9)	70	18	0	0
3/12	TE		2 Gbps	300	20:4a CiscoMDS 00:26:9e (2/10)	66	67	0	0
3/11	TE		2 Gbps	300	20:4c CiscoMDS 00:26:9e (2/12)	57	31	0	0
1/13	TE		2 Gbps	400	20:cd CiscoMDS 00:29:9e (4/13)	55	61	0	0
1/15	TE		2 Gbps	400	20:cf CiscoMDS 00:29:9e (4/15)	52	91	0	0
1/14	TE		2 Gbps	400	20:ce CiscoMDS 00:29:9e (4/14)	49	52	0	0
3/16	TE		2 Gbps	400	20:50 CiscoMDS 00:25:5e (2/16)	46	50	0	0
3/7	TE		2 Gbps	200	20:c7 CiscoMDS 00:25:5e (4/7)	44	0	0	0

FxPorts (Switch Side)						NxPorts (Attached Hosts & Storage)				
Port	Speed	VSAN	Rx Utilization%	Tx Utilization%	Errors	Discards	Port	Type	Port WWN	Fold
8/1	2 Gbps	200	80	43	0	0	fc8/1		21:00 Qlogic 08:5c:3e	0x380000
8/2	2 Gbps	200	33	43	0	0	fc8/2		21:00 Qlogic 08:5d:3e	0x380200
8/3	2 Gbps	200	88	92	0	0	fc8/3		JNI 00:8d:80	0x380107
8/4	2 Gbps	200	16	4	0	0	fc8/4		JNI 00:8d:85	0x380106
8/5	2 Gbps	300	98	82	0	0	fc8/5		21:01 Qlogic 28:5f:3e	0x380000
8/6	1 Gbps	300	37	12	0	0	fc8/6		22:00 Seagate 5a:41:9b	0x3803b4
8/9	1 Gbps	300	39	34	0	0	fc8/6		22:00 Seagate 46:e7:21	0x3803b3
8/10	2 Gbps	300	57	37	0	0	fc8/6		22:00 Seagate 46:58:39	0x3803b2
8/13	2 Gbps	400	7	10	0	0	fc8/6		22:00 Seagate 5a:df:57	0x3803b1

Les outils disponibles dans le MDS vont faciliter l'administration des ressources Directeurs FC, Serveurs et baies répartis sur les différents sites, à partir d'un point focal. Il sera possible d'administrer les périphériques et serveurs avec le protocole FC-GS-3. Les anciens commutateurs 1 Gbps pourront également être administrés par ce moyen, selon leur niveau de firmware.

4.2 Cisco Fabric Manager Server

Le Cisco Fabric Manager Server, disponible sous la forme d'une licence additionnelle, va apporter le Reporting statistique. Chaque MDS9000 effectue la collecte des informations de charge et de type d'I/O au niveau de chaque port. Le Cisco FMS permet ensuite de générer des graphes de charge par lien, à la journée, à la semaine, au mois et à l'année.

Sur la base des informations collectées, le Cisco FMS permet également de déterminer des tendances sur les évolutions futures et effectuer des prédictions de performances ou «Capacity Planning».

Le Cisco FMS permet également de superviser plusieurs Fabriques distinctes à partir d'un point focal. Ainsi, dans le cas d'une architecture Dual-Fabric dans laquelle les deux réseaux SAN ne sont pas connectés entre eux, le Cisco FMS va détecter que ces deux Fabriques sont connectés à des périphériques communs et saura représenter cette topologie Dual-Fabric sur une vue consolidée unique.

4.3 Cisco Fabric Services

L'utilitaire CFS est un processus géré par le MDS9000 et dont le but est de centraliser les paramètres de configuration du Réseau de Stockage et d'automatiser la distribution de ces différents paramètres au sein des différents MDS9000 du réseau.

Plusieurs mécanismes des MDS9000 vont utiliser le service CFS pour la synchronisation des paramètres dans le réseau. Ces mécanismes sont : RADIUS, NTP, Syslog, CallHome, routage inter-VSAN, WWN-based VSAN, définitions des Rôles, etc.

4.4 Sécurisation optimale de l'administration

La sécurité de l'administration est un critère déterminant dans le choix d'équipement de Réseau de Stockage. Les équipements de commutation FC de la gamme MDS9000 vont demander une authentification forte lors de la phase de connexion à l'équipement à des fins d'administration, comme la majorité des équipements de commutation FC existants.

Mais la gamme MDS va bien au-delà. Qu'il s'agisse de l'administration en mode ligne de commande, en mode graphique, ou pour les transferts de fichiers, le Directeur MDS9500 va chiffrer les flux, grâce à l'utilisation exclusive de protocoles appropriés, SSHv2 pour le mode ligne de commande, SNMPv3 pour le mode graphique, et SFTP pour le transfert de fichiers système et de configuration. La confidentialité des données d'administration du réseau de stockage est donc parfaitement assurée. Et c'est un minimum, car s'authentifier n'est pas suffisant, si ensuite, il suffit d'installer une sonde au sein du réseau d'entreprise pour récupérer toutes les données relatives au réseau de stockage, sans que rien ne puisse prévenir ce genre d'intrusion, ni masquer les données proprement dites.

Un administrateur voulant se connecter aux unités MDS9000 devra s'authentifier soit auprès d'une base locale, soit par redirection vers un serveur RADIUS. Ce dernier vérifiera les droits de l'administrateur et pourra retourner le profil de l'administrateur, via VSA RADIUS, au MDS qui répondra ainsi aux requêtes de l'administrateur selon ses droits. Jusqu'à 64 profils administrateur peuvent être définis, y compris des profils par VSAN.

4.5 Fibre Channel Security Protocol

La gamme MDS9000 supporte la fonction DH-CHAP du protocole FC-SP. Ce protocole est basé sur des algorithmes largement éprouvés dans le monde IP, à savoir Diffie-Hellman, CHAP et DES.

1. Challenge Handshake Authentication Protocol ou CHAP a pour but d'authentifier les deux entités qui veulent dialoguer. Il est basé sur un nombre généré de façon aléatoire, le «Challenge», qui est échangé entre les deux parties et sert à coder le mot de passe d'authentification. Avec ce mécanisme, seul le challenge peut être intercepté, mais il ne permet aucun accès. Le mot de passe, lui, ne passera jamais en clair sur le réseau reliant entre les deux entités. Une authentification mutuelle des deux entités est également possible avec CHAP.
2. Diffie-Hellman est très largement répandu dans le monde IP et est une des briques de base du protocole IPSec. L'objectif de l'algorithme Diffie-Hellman est de générer un secret partagé entre les deux parties sans que jamais ce secret ne soit échangé sur le réseau. Il est aussi basé sur l'échange en claire de nombres générés de façon aléatoire, qui peuvent être interceptés mais ne pourront jamais suffire à dériver le secret partagé en cours d'élaboration. En tournant l'algorithme Diffie-Hellman sur ces nombres aléatoires, ainsi que sur d'autres nombres qui ne seront pas échangés et sont donc connus que d'une seule entité, chacune des deux parties va arriver au même résultat, le secret partagé. Ce secret partagé peut être ensuite utilisé par d'autres algorithmes, par exemple DES, 3DES ou AES, pour le chiffrement des flux de données à transmettre.
3. Data Encryption Service ou DES a pour objectif de chiffrer les flux de données afin de garantir la confidentialité lors de leur transit sur un réseau considéré comme non sûr. Il va utiliser un secret partagé généré par Diffie-Hellman. La longueur des clés de chiffrement de DES est de 56 bits. La clé de chiffrement en 3DES est de 168 bits. En AES, Advanced Encryption Services, elle peut varier de 128 ou 256 bits.

La gamme MDS9000 met en œuvre les fonctions CHAP pour l'authentification des deux parties en présence et Diffie-Hellman pour la génération des secrets partagés. Ces algorithmes sont traités par la fonction de supervision des MDS9000. Afin de connecter un commutateur FC vers un MDS9000 utilisant FC-SP, ce commutateur devra lui aussi mettre en œuvre FC-SP et les deux parties devront connaître le mot de passe requis afin que le lien ISL puisse monter et la Fabrique converger. FC-SP peut aussi être mis en œuvre sur un F-Port. Dans ce cas, le HBA du serveur devra supporter FC-SP et connaître le mot de passe du MDS9000 afin de pouvoir s'insérer dans le réseau SAN et s'enregistrer auprès des processus FC.

4.6 Diagnostics intégrés

Des mécanismes extrêmement novateurs sont mis en œuvre dans le Directeur MDS9509. Le MDS dispose d'outils d'enregistrement et d'analyse des flux Fibre Channel.

L'utilitaire Cisco Fabric Analyzer fonctionne au niveau de la fonction de Supervision. Il peut décoder tous les flux de commande et contrôle Fibre Channel. Il est basé sur l'utilitaire Ethereal.

Pour décoder les flux de données, la gamme MDS9000 met en œuvre la fonction de SPAN de port, qui permet de répliquer le flux d'un port, en entrée ou en sortie, vers un autre port sur lequel est connecté un Analyseur de trame FC. Cette fonction d'analyse peut être assurée par le Port Adapter Analyzer, un élément additionnel de l'offre MDS, qui sera connecté sur le port SPAN par son interface FC, va encapsuler les trames FC dans IP, et les envoyer vers le Cisco Fabric Analyzer via son port GbE.

La gamme MDS supporte aussi le mode RSPAN, dans lequel le flux peut être répliqué à partir d'un commutateur MDS, puis envoyé via le réseau FC vers le commutateur sur lequel est connecté le PAA. Il n'est donc pas nécessaire de le déplacer sur le MDS dont on veut analyser le flux, la totalité de la topologie étant accessible via un point focal.

La fonction FC Ping sera utilisée pour vérifier l'accès entre deux éléments périphériques du réseau SAN, via l'émulation d'une commande de type Process Login (PRLI). La fonction FC TraceRoute vérifiera le chemin emprunté entre deux nœuds du réseau pour déterminer tout point de blocage, et fournira le temps de latence de chacun des nœuds empruntés par le flux de données. Ces mécanismes seront très utiles dans les phases d'installation et reconfiguration. Dans une interconnexion de SAN distants via le protocole FCIP, en cas de dégradation du temps de réponse, la fonction TraceRoute permettra de déceler instantanément le point dans le réseau d'où provient la dégradation en vérifiant le temps de réponse de chaque nœud.

Enfin, le MDS dispose d'un utilitaire de découverte des volumes logiques, afin de vérifier, non seulement l'accessibilité à la baie de stockage, mais également les LUN visibles par chaque interface, avec leurs caractéristiques techniques et la taille de chaque volume logique.

Ces fonctions sont disponibles via le Cisco Fabric Manager. Ainsi, l'administrateur pourra les activer depuis sa station, grâce à la couverture globale de CFM, et en toute confidentialité, grâce au chiffrement de bout en bout des données d'administration. L'analyseur de trame intégré est un puissant outil de debugging, disponible en standard dans les MDS9509.

5. TABLEAU RECAPITULATIF DES CARACTERISTIQUES

Le tableau ci-dessous récapitule les principales caractéristiques des équipements de la gamme MDS9000 :

	9120/40	9124	9216/9216i	9506	9509	9513
Type de Configuration	Fixe	Fixe	Modulaire	Modulaire	Modulaire	Modulaire
Nombre de slots pour cartes ligne	N/A	N/A	1 + 1	4	7	11
Densité de port FC minimale	20/40	8	16	12	12	12
Densité de port FC maximale	Idem	24	64/62	192	336	528
Incrément	N/A	N/A	12, 16, 24, 32 ou 48 ports			
Débit FC	1/2 Gbps	1/2/4	1/2/4 et 10 Gbps			
Transceiver SFP SW, LW, CWDM	Oui	SW, MR (4 kms) et LW (10 kms)				
Emulation Ports F, FL et TL	Oui sauf TL sur cartes ligne 2ème génération					
Emulation inter-switch E et TE	Oui					
Buffer Credit maximum par port	255	64	Jusqu'à 4000 par port sur module 2ème génération			
Agrégation max (PortChannel)	4/8 liens	16 liens				
Débit de l'agrégation	8/16 G	64 Gbps sur module 2ème génération				
Distance max de l'agrégation en 1 Gbps / 2 Gbps / 4 Gbps	500 km / 250 km	128/64/32 km	Avec modules 2ème génération : 8000/4000/2000 kms			
Alimentations redondantes	Oui					
Ventilateurs redondants	Oui					
Cross-bar et SUP redondantes	Non	Non	Non	Oui		
Nombre de VSAN maximum (partitions logiques)	256	16	256			
Zoning par interface	Oui					
Zoning par WWPN, fWWN, FCID	Oui					
LUNZoning, Zones Read-Only	Oui					
Zoning protégé par ACL	Oui					
Qualité de Service (QoS)	Non	Oui avec carte ligne 2ème génération				
Connectique iSCSI	Non			Oui		
Connectique FCIP	Non			Oui		
Densité de port max iSCSI / FCIP	N/A		10	24	48	80
Modules de services évolués	Non			Oui		
Admin ligne de commande	Oui					
Admin graphique embarquée	Oui					
RADIUS, SNMPv3, SSHv2, RBAC	Oui					
SNMP traps, Call Home, Syslog	Oui					

6. CONCLUSION

La gamme MDS multi protocoles de Cisco, comprenant des commutateurs fixes d'entrée de gamme, des commutateurs modulaires et des équipements de classe Directeur offre un choix très large pour adresser les besoins d'évolution vers les réseaux de stockage des grands Centres Informatiques, aussi bien que des petites et moyennes entreprises, ou des environnements départementaux.

La gamme MDS apporte des évolutions majeures à la technologie Fibre Channel. Ces évolutions fortes permettent enfin d'envisager des réseaux de stockage de grande taille, grâce aux fonctions de conversion de protocole intégrés, et à l'utilisation de technologies appropriées au besoin réel des applications, que ce soit en terme de débit, de temps de réponse ou de haute disponibilité.

Les mécanismes d'agrégation de liens, de Buffer Credit de grande capacité, de partitionnement logique et routage FC intégré, de Qualité de Service ou de Zoning par LUN sont sans équivalent sur le marché de la commutation FC à l'heure actuelle et ouvrent de nouveaux horizons pour répondre de manière satisfaisante aux nouvelles contraintes de réseaux de stockage. En effet, on assiste aujourd'hui à une banalisation de l'espace disque, et les fonctions intelligentes vont se concentrer dans le réseau, endroit le plus approprié car seul véritable point focal pour la centralisation de services évolués aux données.

6.1.1 Le Centre Informatique du futur

Avec les MDS9500 de Cisco, les utilisateurs ne seront plus limités dans leurs choix par les contraintes des équipements SAN. Ils pourront maintenant concevoir leur SAN avec n'importe quel protocole et débit, avec de la commutation et / ou du routage FC, avec des utilitaires de Virtualisation embarqués, avec connectivité FC native ou Extension sur IP, avec du partitionnement logique et de la ségrégation des flux applicatifs, avec des niveaux de SLA définis par application, etc. grâce au MDS9500. Il sera également possible d'envisager la prise en compte des protocoles et services de prochaine génération quand ils seront disponibles, sans remise en compte de l'investissement initial ni remplacement des Directeurs.

Le MDS9500 apporte aux Centres Informatiques la Consolidation et Virtualisation de l'infrastructure. L'Automatisation et le Provisioning sont en train d'être finalisés. Les utilisateurs peuvent donc envisager l'évolution de leurs Centres Informatiques avec des services intelligents entre les applications et les ressources physiques qu'elles utilisent. Une fois l'ensemble des ressources physiques consolidé et la virtualisation en place, les utilitaires d'automatisation et de Provisioning vont assigner des ressources logiques aux applications sans plus aucune connexion directe entre une application donnée et la ressource physique qu'elle utilise. Une nouvelle application pourra basculer en Production sans qu'il soit nécessaire pour les différentes équipes du Centre Informatique de définir et assigner des ressources physiques dédiées à cette application. Les Centres Informatiques vont pouvoir évoluer du modèle applicatif en silo vers un environnement nettement plus flexible, basé sur une décorrélation totale entre applications et ressources physiques, qui leur permettra au final de répondre de façon immédiate aux évolutions des besoins métiers de l'Entreprise. Le MDS9500 est un élément clé pour permettre le déploiement de Centres Informatiques de prochaine génération. C'est aussi une composante à part entière de la stratégie Cisco autour de ces Centres Informatiques.



Siège social Mondial
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
Etats-Unis
www.cisco.com
Tél. : 408 526-4000
800 553 NETS (6387)
Fax : 408 526-4100

Siège social France
Cisco Systems France
11 rue Camilles Desmoulins
92782 Issy Les Moulineaux
Cédex 9
France
www.cisco.fr
Tél. : 33 1 58 04 6000
Fax : 33 1 58 04 6100


Siège social Amérique
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
Etats-Unis
www.cisco.com
Tél. : 408 526-7660
Fax : 408 527-0883

Siège social Asie Pacifique
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapour 068912
www.cisco.com
Tél. : +65 317 7777
Fax : +65 317 7799

Cisco Systems possède plus de 200 bureaux dans les pays et les régions suivantes. Vous trouverez les adresses, les numéros de téléphone et de télécopie à l'adresse suivante :

www.cisco.com/go/offices

Afrique du Sud • Allemagne • Arabie saoudite • Argentine • Australie • Autriche • Belgique • Brésil • Bulgarie • Canada • Chili • Colombie • Corée
Costa Rica • Croatie • Danemark • Dubaï, Emirats arabes unis • Ecosse • Espagne • Etats-Unis • Finlande • France • Grèce • Hong Kong SAR
Hongrie • Inde • Indonésie • Irlande • Israël • Italie • Japon • Luxembourg • Malaisie • Mexique • Nouvelle Zélande • Norvège • Pays-Bas
Pérou • Philippines • Pologne • Portugal • Porto Rico • République tchèque • Roumanie • Royaume-Uni • République populaire de Chine
Russie • Singapour • Slovaquie • Slovénie • Suède • Suisse • Taiwan • Thaïlande • Turquie • Ukraine • Venezuela • Vietnam • Zimbabwe

 Copyright©2006 Cisco Systems, Inc. Tous droits réservés. CCSP, CCVP, le logo Cisco Square Bridge, Follow Me Browsing et StackWise sont des marques de Cisco Systems, Inc. ; Changing the Way We Work, Live, Play, and Learn, et iQuick Study sont des marques de service de Cisco Systems, Inc. ; et Access Registrar, Aironet, ASIST, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, Cisco, le logo Cisco Certified Internetwork Expert, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, le logo Cisco Systems, Cisco Unity, Empowering the Internet Generation, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, FormShare, GigaDrive, GigaStack, HomeLink, Internet Quotient, IOS, IP/TV, iQ Expertise, le logo iQ, iQ Net Readiness Scorecard, LightStream, Linksys, MeetingPlace, MGX, le logo Networkers, Networking Academy, Network Registrar, Packet, PIX, Post-Routing, Pre-Routing, ProConnect, RateMUX, ScriptShare, SlideCast, SMARTnet, StrataView Plus, TeleRouter, The Fastest Way to Increase Your Internet Quotient et TransPath sont des marques déposées de Cisco Systems, Inc. et/ou de ses filiales aux Etats-Unis et dans d'autres pays.

Toutes les autres marques mentionnées dans ce document ou sur le site Web appartiennent à leurs propriétaires respectifs. L'emploi du mot partenaire n'implique pas nécessairement une relation de partenariat entre Cisco et une autre société. (0202R)