11

Assembler les briques du LAN et du WAN

Jusqu'à présent, nous avons utilisé différentes technologies, les unes adaptées aux réseaux locaux, les autres aux réseaux étendus.

Arrive un moment où les deux mondes doivent se rencontrer puisque la vocation des réseaux est de relier des hommes, qu'ils fassent ou non partie de la même entreprise.

Les réseaux locaux ont de plus en plus tendance à s'étendre au-delà d'un simple site pour former un réseau de campus, repoussant ainsi la frontière qui les sépare des réseaux étendus.

Dans ce chapitre vous apprendrez :

- à étendre le réseau fédérateur jusqu'au campus ;
- à configurer des VLAN ;
- à établir le lien entre commutateurs LAN et routeurs WAN ;
- à élaborer un plan de routage ;
- à configurer les protocoles de routage.

Mettre en place un réseau fédérateur

Les données du problème

Le réseau de 800 postes dont nous avons décrit l'installation au chapitre 6 fonctionne parfaitement. Or, voici que l'ouverture d'un nouvel immeuble à proximité est annoncée. Elle implique un changement d'échelle, puisqu'il s'agit d'une tour de quinze étages, représentant environ 2 000 connexions, à raison de 130 par étage en tenant compte des postes de travail, des imprimantes, des serveurs, etc.

Le câblage a été conçu en fonction des besoins potentiels en matière d'architecture, incluant à la fois la téléphonie (la voix), le réseau local (les données), et la diffusion vidéo (l'image). Les principes sont ceux qui ont été étudiés au chapitre 5.

La démarche

Il semble tout d'abord évident qu'il faudra au moins un réseau local par étage, afin de contrôler les flux, et sans doute plus, car il faut toujours s'attendre à des besoins spécifiques pour une population de 1 500 utilisateurs. Il est donc sage de prévoir une quarantaine de réseaux.

Un constat s'impose : s'il faut descendre près de quinze réseaux en *collapse backbone*, les équipements fédérateurs doivent disposer d'une très grande capacité. De plus, un réseau redondant est absolument nécessaire pour assurer une bonne qualité de service. En effet, à une telle échelle, un problème survient nécessairement quelque part (en vertu d'un principe de probabilité).

Le point central de l'architecture concerne donc les caractéristiques du réseau fédérateur pour lequel nous nous posons les questions suivantes :

- Quelle technologie ?
- Quels équipements ?
- Routeurs ou commutateurs de niveau 3 ?

Quelle technologie ?

Nous avons ici le choix entre Ethernet et ATM, sujet que nous avions abordé au cours du chapitre précédent.

Bien qu'adapté aux réseaux WAN, les constructeurs nous proposent d'utiliser ATM également pour les réseaux locaux. Choix étrange, car l'utilisation de ce protocole pose un certain nombre de problèmes :

- Il faut mettre en place une mécanique complexe pour adapter un réseau multipoint tel qu'Ethernet à un réseau ne fonctionnant qu'avec des circuits virtuels point à point.
- Il faut mettre en place une mécanique non moins complexe pour adapter les VLAN Ethernet au monde ATM.
- Le débit d'ATM est aujourd'hui limité à 622 Mbit/s, 155 Mbit/s étant le débit le plus fréquemment rencontré dans les entreprises. Face au Gigabit Ethernet, l'argument est donc mince.

Un certain nombre de standards permettent d'effectuer cette intégration. Il s'agit de LANE (*LAN Emulation*) pour les VLAN, de MPOA (*Multi Protocol Over ATM*) pour le routage et l'interconnexion des ELAN (*Emulated LAN*) et de Classical IP pour la correspondance entre adresses ATM et adresses IP.

Face à cela, Ethernet nous offre la simplicité et une panoplie de solutions homogènes et évolutives, du 10 mégabits au Gigabit. Nous choisirons donc cette technologie pour l'ensemble de notre réseau local, du poste de travail au réseau fédérateur.

Quels équipements ?

Le réseau fédérateur concentre tous les flux entre les réseaux d'étage d'une part, et entre ces derniers et les ressources communes d'autre part. Cela suppose que la majorité des flux est émise entre les utilisateurs d'un étage donné.

Mais, de nos jours, la traditionnelle répartition 80/20 (80 % du trafic local sur le réseau et 20 % vers d'autres réseaux) n'est plus valable. Par exemple, la constitution de groupes de travail pluridisciplinaires amène des personnes dispersées au sein de l'immeuble à établir des liens de communication privilégiés entre elles.

Traduit en termes techniques, les flux réseau générés de manière privilégiée entre les postes de travail et les serveurs évolueront sans cesse.

En outre, la constitution de réseaux isolés regroupant des utilisateurs géographiquement dispersés pourrait s'avérer nécessaire.

La solution à ces besoins passe par les **VLAN** (*Virtual Local Area Network*). Cette technologie permet de définir des segments Ethernet logiques, indépendamment de la localisation géographique des postes de travail. Une trame émise au sein d'un VLAN ne sera diffusée qu'aux stations participant audit VLAN.

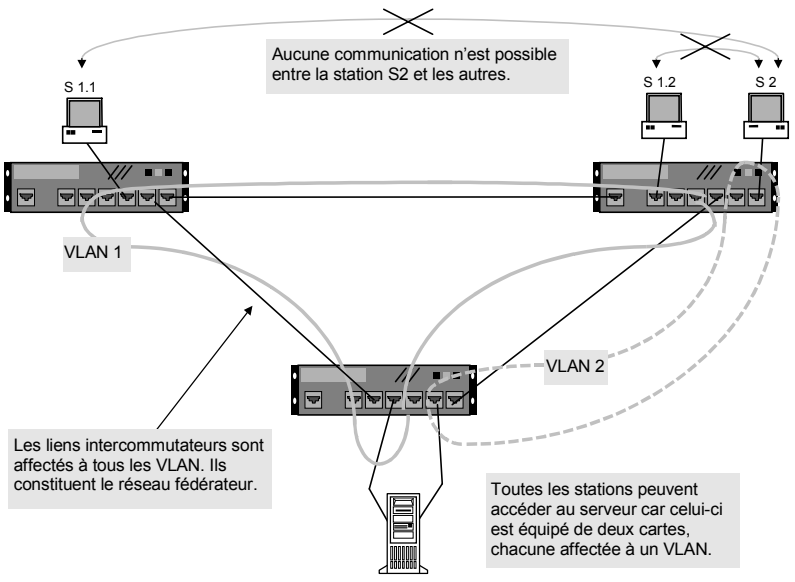


Figure 11-1.
Principe des VLAN.

Or, seuls les commutateurs permettent de créer des VLAN, c'est-à-dire de segmenter le réseau correspondant à des groupes d'utilisateurs indépendamment de leur localisation géographique.

De plus, nous avons vu au chapitre 6 que les commutateurs sont nécessaires pour des applications multimédias (téléphonie et visioconférence sur IP), pour de gros volumes de données et pour une question de fiabilité.

Si l'on veut répondre à tous ces besoins, il est donc nécessaire d'installer des commutateurs sur l'ensemble de notre réseau (eau et gaz à tous les étages, pour ainsi dire !). Le choix de ces équipements s'impose donc à la fois pour des questions de performances et d'architecture.

Nous choisissons donc la solution 100 % commutateurs.

Routeur ou commutateur de niveau 3 ?

L'intérêt de partitionner notre réseau en réseaux plus petits est de circonscrire localement les flux générés par un groupe d'utilisateurs, de créer des zones isolées, ou encore de réduire les flux générés par les broadcast (surtout pour les grands réseaux).

La constitution de réseaux distincts (constituant chacun un domaine de broadcast MAC) nécessite cependant de les interconnecter quelque part. En effet, même s'ils appartiennent à des groupes différents, les utilisateurs doivent, à un moment ou à un autre, accéder à des ressources communes (serveur d'annuaire, passerelles fax, base de données centrale, PABX, accès Internet, etc.).

Une fonction de **routage** est donc nécessaire pour interconnecter ces différents réseaux, qu'ils soient physiquement ou virtuellement constitués.

À ce niveau, nous avons le choix entre deux types d'équipements : les routeurs et les commutateurs de niveau 3.

Comparé au commutateur de niveau 3, le routeur présente un certain nombre de désavantages : il est nettement moins performant et, de ce fait, dispose rarement d'interfaces Gigabit. Par ailleurs, il ne sait pas gérer les VLAN.

Ce dernier point mentionné implique que le routeur dispose d'autant d'interfaces qu'il y a de VLAN (si ceux-ci sont créés par port), ou d'autant d'adresses IP sur une interface qu'il y a de VLAN créés par adresse IP.

Pour les petits et moyens réseaux (moins de 800 postes) sans liens gigabit, on peut envisager un routeur pour interconnecter quelques VLAN. Au-delà de ces restrictions, le commutateur de niveau 3 s'impose.

LES COMMUTATEURS DE NIVEAUX 2 ET 3

Un commutateur de niveau 2 agit au niveau des couches physique et logique (niveaux 1 et 2). Il ne traite que les trames MAC. On parle de commutation de niveau 2 ou layer 2 switching.

Un commutateur de niveau 3, quant à lui, agit au niveau de la couche réseau (niveau 3). Il ne traite que les paquets IP. C'est l'équivalent d'un routeur mais en beaucoup plus performant. On parle de commutation de niveau 3 ou layer 3 switching.

Quelle architecture ?

Nous voilà donc confortés dans le choix des commutateurs. Mais quelle architecture retenir ? Et à quel débit ?

On le voit, pour notre réseau de 1500 postes, de nouvelles considérations viennent compliquer notre tâche, de nouveaux paramètres influent sur le choix de l'architecture. En fait, tout tourne autour du fédérateur, pièce maîtresse du réseau. Résumons :

1. Une architecture basée uniquement sur des commutateurs de niveau 2 a le mérite de la simplicité. Elle a été étudiée au chapitre 6. Si l'on veut créer des réseaux séparés, il faut employer des VLAN par port ou par adresses MAC, ce qui augmente la complexité d'exploitation.
2. Une architecture basée uniquement sur des commutateurs de niveau 3 est plus coûteuse. Elle est cependant plus souple que la précédente, car on peut choisir les classes d'adresses IP et les combiner. L'architecture est identique à celle de la première solution, seule la technologie change.
3. Une architecture basée sur des routeurs est la moins performante de toutes et la moins souple (pas de VLAN possible). En fait, les routeurs sont plutôt destinés aux réseaux WAN.
4. Une architecture reposant sur un réseau fédérateur ATM est la plus complexe et la plus fragile, car elle impose une combinaison de plusieurs technologies. Son débit est, de plus, limité à 622 Mbit/s, ce qui est un handicap certain face au Gigabit Ethernet.

En fait, le routage n'est nécessaire qu'au niveau du réseau fédérateur, car tous les commutateurs d'étage y seront reliés.

En définitive, le choix se portera sur des commutateurs de niveau 2 pour les étages, et des commutateurs de niveau 3 pour le réseau fédérateur. Dans la pratique, ces derniers sont également des commutateurs de niveau 2 équipés de cartes de commutation de niveau 3.

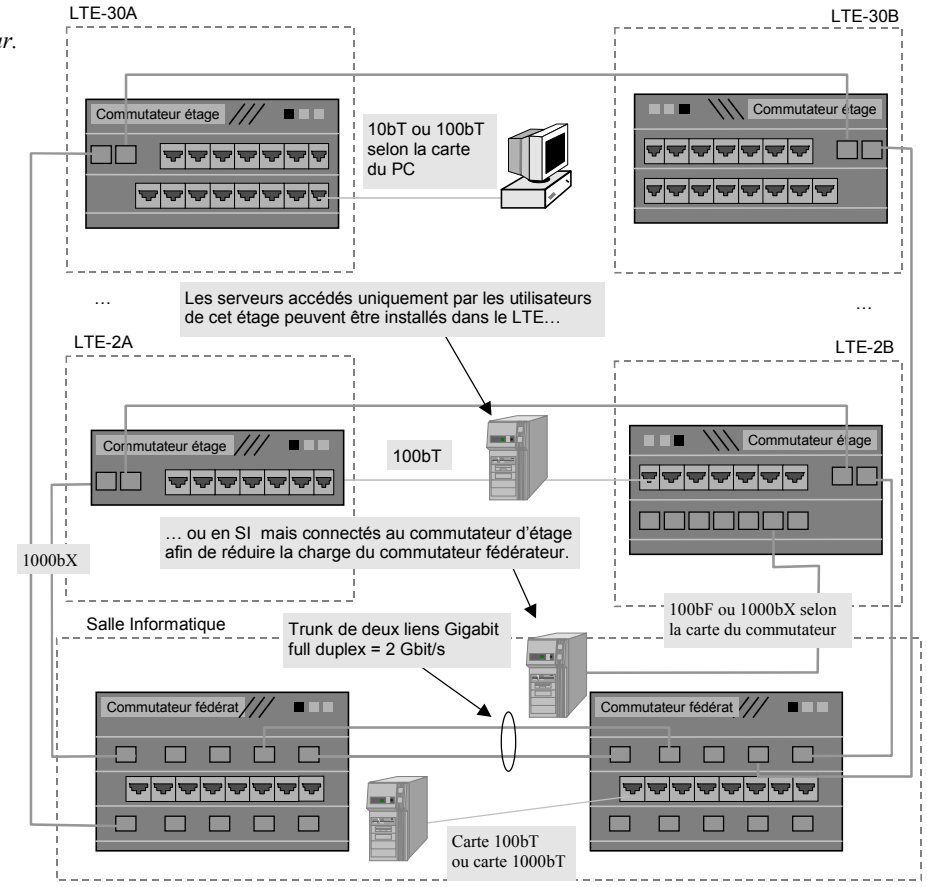
Pour le reste, nous appliquerons les recettes indiquées au chapitre 6.

Les commutateurs d'étage sont équipés de cartes 10/100bT ainsi que de deux ports uplink gigabit. Ils peuvent éventuellement être dotés de cartes 100bF ou de cartes gigabit pour connecter des serveurs délocalisés.

Les commutateurs fédérateurs sont principalement équipés de cartes gigabit pour être raccordés, d'une part, entre eux et, d'autre part, aux commutateurs d'étage. Ils peuvent éventuellement être dotés de cartes 10/100bT ou 1000bT, afin de connecter des serveurs situés dans des salles informatique. Les cartes 1000bT offrent, en effet, une plus grande densité de port que leurs équivalents en fibre optique.

Les cartes en fibre optique sont utilisées partout où les distances sont supérieures à 90 mètres. Leur emploi est cependant systématisé au niveau du réseau fédérateur, même en dessous de cette distance, afin de disposer de configurations homogènes.

Figure 11-2.
Réseau fédérateur.



Configurer les VLAN

Même si les commutateurs fédérateurs sont équipés de cartes de commutation niveau 3, tous assurent la commutation de niveau 2. Le spanning tree doit donc être configuré sur tous les commutateurs, comme indiqué au chapitre 7.

De la même manière, un VLAN doit être configuré sur tous les commutateurs, afin qu'il soit connu de tous. Sur nos équipements (de marque Cisco), la création d'un VLAN par port s'effectue de la façon suivante :

```
set vlan 100 name VLAN_principal
set vlan 100 2/1-48
```

Les ports 1 à 48 situés sur la carte n° 2 seront ainsi affectés au VLAN 100 que nous avons appelé " VLAN principal ".

L'opération suivante consiste à activer le protocole **802.1q** entre tous nos commutateurs, afin d'étendre la portée du VLAN à l'ensemble de notre réseau. Ce protocole ne doit être activé

que sur les ports qui raccordent des commutateurs entre eux, qu'on appellera des ports **trunk** (ports de liaison) :

```
set trunk 3/1 dot1q
set trunk 3/2 dot1q
```

Active le protocole 802.1q sur ces ports

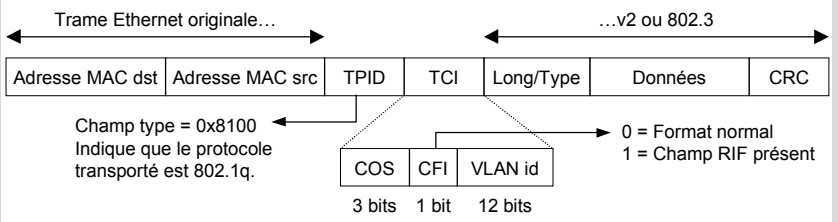
Chaque VLAN correspond à un domaine de broadcast et fonctionne donc avec un spanning tree indépendant. Il est donc possible de configurer ce protocole indépendamment pour chaque VLAN (voir chapitre 7) :

```
set spantree enable 100
set spantree fwddelay 15 100
set spantree hello 2 100
set spantree priority 16384 100
```

N° de VLAN

LE POINT SUR LES VLAN (IEEE 802.1Q)

La norme 802.1q consiste à ajouter un champ à l'en-tête de la trame Ethernet initiale (802.3) à la fois pour gérer les VLAN et pour gérer des classes de service (802.1p).



TPID = *Tag Protocol Identifier* : correspond au champ Type d'une trame Ethernet v2
TCI = *Tag Control Information* : le label 802.1q inséré dans la trame Ethernet v2
COS = *Class Of Service* : utilisé par la norme 802.1p (cf. chapitre 14)
CFI = *Common Format Identifier* : permet de transporter le champ RIF dans le cas d'un tunnel source routing
VLAN id = numéro de VLAN (4 096 possibilités)

Cette trame est véhiculée uniquement entre les commutateurs. Un VLAN peut donc être étendu à tout un réseau de commutateurs. Ces derniers ôtent le champ 802.1q lorsqu'ils transmettent la trame à un équipement terminal (PC, serveur, etc.) de manière que ces dernier retrouvent une trame conforme à la norme 802.3 ou Ethernet v2.
La constitution des VLAN dépend de l'implémentation qui en est faite au sein des commutateurs. Il est ainsi possible de créer des VLAN :

- **par port** : toute trame entrant par un port est affectée d'office à un VLAN ;
- **par adresse MAC** source : toute trame disposant d'une telle adresse est affectée à un VLAN ;
- **par protocole** : toute trame véhiculant de l'IP, par exemple, est affectée à un VLAN ;
- **par adresse IP** source : toute trame véhiculant un paquet IP avec une telle adresse est affectée à un VLAN.

Un processus **spanning tree** (802.1d) est créé par VLAN. Par conséquent, les trames de **broadcast** et de **multicast** MAC émises au sein d'un VLAN ne seront pas propagées aux autres VLAN. En outre, les stations d'un VLAN ne pourront pas communiquer avec celles appartenant à un autre VLAN. Pour permettre cette fonction, il faut interconnecter les VLAN à l'aide d'un **routeur** ou d'un **commutateur de niveau 3**.

Les ports gigabit peuvent, de plus, être configurés de manière à opérer un contrôle de flux. Cela consiste en un signal envoyé à un autre commutateur pour lui demander de ralentir temporairement l'envoi de trames :

```
set port flowcontrol send 0/0-1 on
set port flowcontrol receive 0/0-1 on
```

Envoie…

…et accepte les signaux de contrôle de flux.

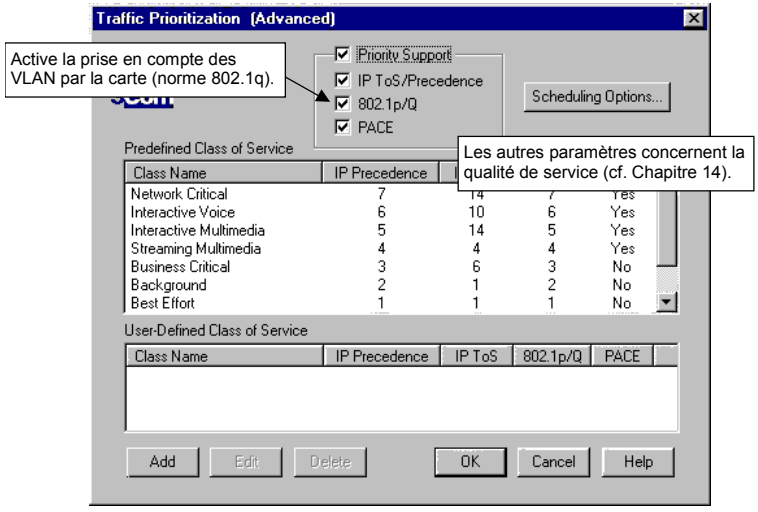
Dans notre architecture, il est également prévu d'agréger deux liens gigabit entre les deux commutateurs fédérateurs :

```
set vlan 100 2/0-1
set port channel 2/0-1 desirable
set trunk 2/0 desirable dot1q
```

Crée un groupe de 2 ports

L'activation du trunking 802.1q sur un port est appliquée à tout le groupe.

La configuration des VLAN au niveau des cartes Ethernet des PC et des serveurs est possible si elles supportent le protocole 802.1q, ce qui est le cas de nos cartes 3com.



Il est cependant préférable de ne pas utiliser cette facilité pour les raisons suivantes :

- cela ajouterait une complexité supplémentaire aux tâches d'administration : il faudrait configurer à distance toutes les cartes des PC ;
- les utilisateurs trouveraient toujours le moyen de modifier la configuration de leur carte de manière à changer de VLAN ;
- pour éviter cela, il faudrait réaliser un contrôle au niveau des commutateurs, ce qui induirait une double exploitation.

Nous pourrions également créer des VLAN dynamiques par adresse IP. Là encore, l'exploitation est délicate et les modifications de la part des utilisateurs sont toujours possibles. L'affectation des VLAN par port a le mérite d'être simple, de faire partie de la configuration normale des commutateurs et de maîtriser l'étendue du VLAN sur notre réseau.

Extension du réseau fédérateur

Nos deux commutateurs fédérateurs étaient suffisants pour accueillir les 15 réseaux d'étage. Maintenant, nous devons connecter un autre site situé à quelques centaines de mètres, puis deux autre situés à quelques kilomètres.

Si tous les bâtiments sont situés sur un terrain privé (par exemple, un campus universitaire), nous pouvons poser de la fibre optique comme nous l'entendons. Dans le cas contraire, soit un opérateur nous loue des câbles en fibre optique, soit nous devons obtenir une dérogation pour en poser entre nos bâtiments.

Dans tous les cas, nous supposerons donc que les bâtiments sont reliés entre eux par des câbles en fibre optique. Car l'enjeux est maintenant d'étendre notre réseau fédérateur pour en faire un réseau de campus.

On parle également de MAN (*Metropolitan Area Network*), bien qu'aucune technologie particulière, autre que celle utilisée en LAN, ne soit associée à ce type de réseau. Il s'agit simplement d'une dénomination conceptuelle.

Pour ceux qui en doutaient encore, le Gigabit Ethernet convient parfaitement à ce type de besoins. Certains opérateurs proposent même ce service sur quelques centaines de kilomètres. Tout dépend de la fibre optique utilisée.

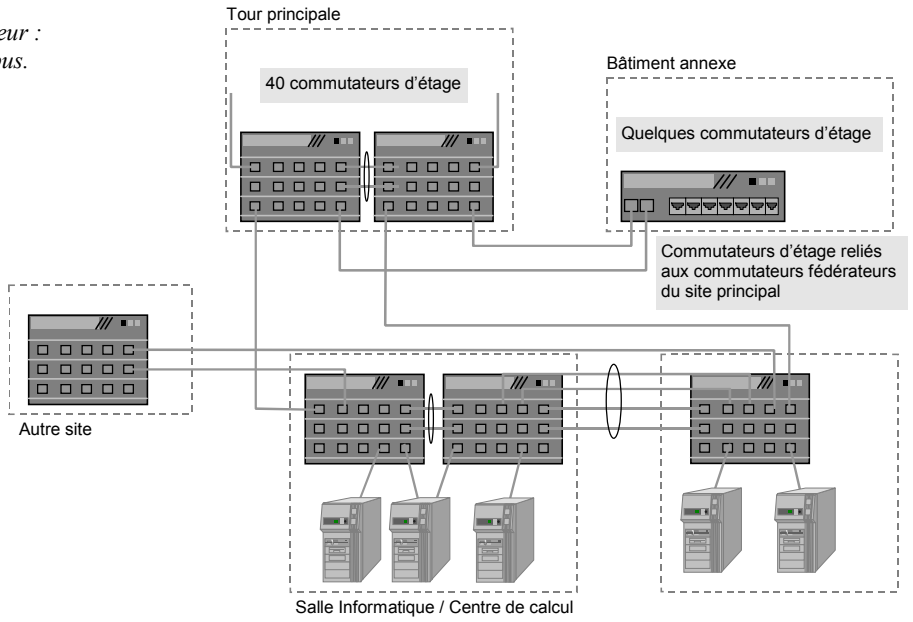
	Support de transmission	Distance
1000bSX	Fibre multimode 62,5µ	De 2 à 300 m
	Fibre multimode 50µ	De 2 à 550 m
1000bLX	Fibre multimode 62,5µ	De 2 à 550 m
	Fibre multimode 50µ	De 2 à 550 m
	Fibre monomode 9µ	De 2 à 3 000 m

Côté performances, le Gigabit Ethernet est à la hauteur des débits annoncés :

- un débit réel de 761 Mbit/s pour des trames de 64 octets (soit 1 488 095 paquets par seconde) ;
- un débit réel de 986 Mbit/s pour des trames de 1 518 octets (soit 81 274 paquets par seconde).

Confortés dans notre choix du gigabit, nous nous retrouvons avec plusieurs commutateurs fédérateurs à interconnecter.

Figure 11-3.
Extension du réseau fédérateur : le réseau de campus.



Nous aurions pu mailler tous les commutateurs fédérateurs de manière à offrir des routes multiples. Cela est envisageable si les serveurs sont disséminés dans différents bâtiments.

Quand cela est possible, il est cependant préférable de respecter les principes suivants :

- Choisir deux commutateurs fédérateurs de campus qui fédéreront également les autres commutateurs fédérateurs de site (ou en dédier deux autres), de manière à centraliser les flux intersites au sein d'un nombre réduit de matrices de commutation. Ces deux équipements peuvent être situés dans deux bâtiments différents.
- Relier les deux commutateurs fédérateurs de campus par un lien très haut débit, dans notre cas quatre liens gigabits.
- Connecter les fédérateurs de site aux fédérateurs de campus par deux liens distincts en partage de charge et en redondance, de préférence sur deux commutateurs distincts, de manière à pallier la défaillance d'un équipement.

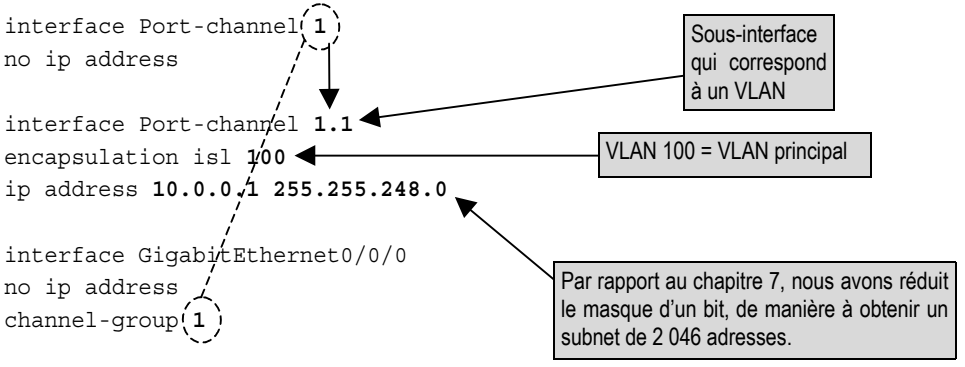
L'extension du réseau fédérateur se fait donc de manière très simple, sans remettre en cause les choix technologiques et l'architecture.

Il est à noter que, grâce aux VLAN, cette architecture permet à plusieurs sociétés de cohabiter sur la même infrastructure tout en étant isolées.

L'adressage et le routage IP

Nous avons décidé de créer un VLAN et d'affecter de manière statique les ports à ce VLAN, que nous avons appelé VLAN principal. D'autres VLAN peuvent être crées pour des réseaux dédiés.

Cela implique d'affecter un subnet IP à chaque VLAN, et par conséquent de configurer nos cartes de commutation niveau 3 (fonctionnellement équivalentes, rappelons-le, à des routeurs). Pour notre VLAN principal, cela est réalisé comme suit, conformément à notre plan d'adressage établi au chapitre 7 :



L'encapsulation " isl " (*Inter-Switch Link*) active le protocole propriétaire Cisco équivalent de la norme 802.1q. Dans ce cas particulier, nous ne pouvons faire autrement.

La carte de commutation de niveau 3 dispose ainsi d'un attachement sur le VLAN principal, qui est réalisé au niveau de la matrice de commutation. La création de tout autre VLAN sera réalisée sur le même modèle.

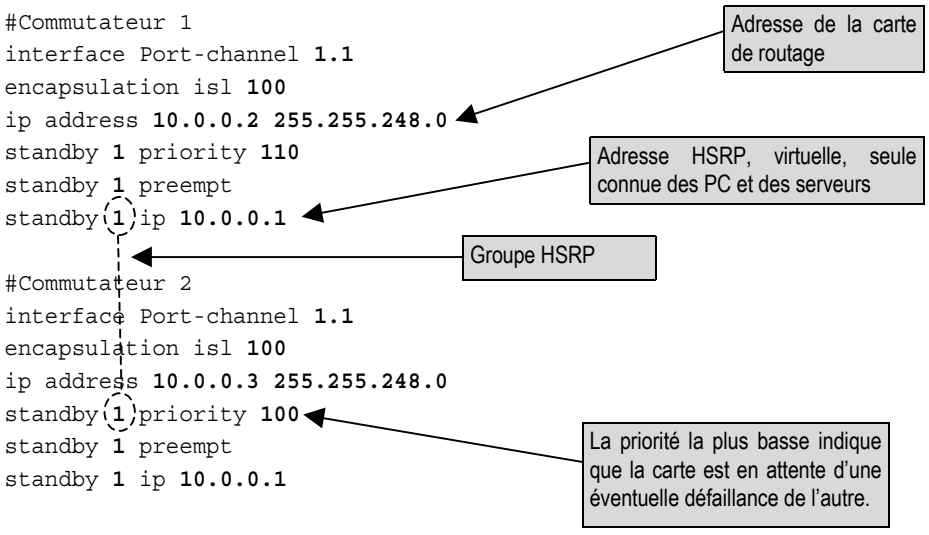
Le routage entre VLAN, et donc entre subnets IP, est effectué par la carte de commutation. Pour sortir du VLAN principal, les PC et les serveurs doivent connaître la route de sortie, c'est-à-dire la route par défaut (*default gateway*), comme cela était le cas au chapitre 8. Mais, cette fois, elle doit pointer sur l'adresse IP de la carte de commutation, à savoir 10.0.0.1.

Sur chaque VLAN, la passerelle par défaut des PC et des serveurs pointe donc sur l'adresse IP du commutateur de niveau 3 attaché audit VLAN.

Redondance du routage

Si, comme sur notre site parisien, nous disposons de deux commutateurs fédérateurs, chacun équipé d'une carte de commutation de niveau 3 (carte de routage), il est intéressant d'assurer la redondance de la route par défaut vis-à-vis des PC et des serveurs. Sur nos équipements, cela est réalisé grâce à la fonction HSRP (*Hot Standby Router Protocol*).

Le principe repose sur un groupe de *n* routeurs (ou cartes de commutation de niveau 3 dans notre cas), dont l'un est désigné actif. Comme d'habitude, chaque interface est associée à une adresse IP et à une adresse MAC. Mais le routeur actif reçoit en plus une adresse IP (définie comme route par défaut) associée à une adresse MAC qui seule répond au protocole de résolution d'adresses ARP (voir chapitre 7). En cas de défaillance du routeur actif, un nouveau routeur est élu parmi les N-1 restants en fonction des priorités affectées au sein du groupe HSRP, qui s'approprie les adresses HSRP (MAC et IP) :



La route par défaut configurée sur les PC et serveurs est celle de l'adresse HSRP, à savoir 10.0.0.1.

Ce principe peut être appliqué à chaque VLAN. Il est alors conseillé d'affecter les priorités de telle manière que chacune des cartes de commutation de niveau 3 soit active au moins pour un VLAN, et ce afin de répartir la charge de routage.

Certaines piles IP, comme celle de Windows NT, intègrent un mécanisme de détection de panne du routeur par défaut (*dead gateway detection*), tel que décrit dans le RFC 816. Si la station constate qu'elle ne parvient plus à joindre son routeur par défaut, elle en choisira un autre parmi une liste définie dans le menu des propriétés de TCP/IP, case " Avancé… ", section " Passerelle ". Pour activer ce mécanisme de détection, il faut positionner à " 1 " la clé de registre " HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters\EnableDeadGWDetect ".

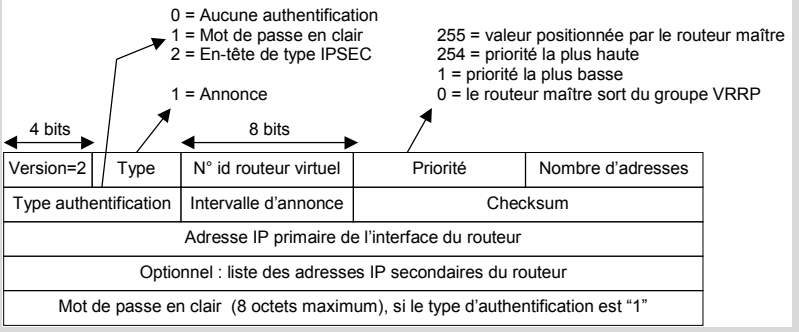
LE POINT SUR VRRP (RFC 2338)

Le protocole **VRRP** (*Virtual Router Redundancy Protocol*) reprend les principes du protocole **HSRP** (*Hot Standby Router Protocol* – RFC 2281) spécifié par Cisco. Les formats des paquets sont en revanche différents, ce qui rend ces deux protocoles incompatibles.

Le but est ici d'offrir une **redondance** de routeurs pour les machines (notamment les PC) utilisant le mécanisme de **passerelle par défaut** (*default gateway*). Même si plusieurs routeurs sont connectés à un segment Ethernet, la passerelle par défaut des PC pointera vers une seule adresse IP, celle d'un des routeurs choisis par l'administrateur du réseau.

Le protocole VRRP permet d'ajouter, en plus des adresses propres à chaque routeur, une **adresse IP virtuelle** vers laquelle les passerelles par défaut peuvent pointer. À un instant donné, seul le routeur désigné **maître** détiendra l'adresse virtuelle, et pourra assurer le traitement des paquets à destination de cette adresse.
Ainsi, lorsque la pile IP du PC devra résoudre, grâce à ARP, l'adresse de sa passerelle par défaut, seul le routeur maître répondra en indiquant l'**adresse MAC virtuelle**.

Le routeur maître envoie un paquet d'annonce à intervalle régulier. Si les autres routeurs n'en reçoivent plus au bout de l'intervalle de temps spécifié dans le dernier paquet reçu (par défaut une seconde), ils considèrent que le routeur maître est en panne et entrent alors dans un processus d'**élection** en envoyant des annonces. Celui dont la **priorité** est la plus élevée, devient alors maître et prend le contrôle de l'adresse virtuelle.



Les paquets VRRP disposent du numéro de **protocole 112**. Ils sont envoyés dans des paquets IP à destination de l'adresse multicast **224.0.0.18**, dont l'adresse source est la véritable adresse IP du routeur et dont le TTL est obligatoirement fixé à **255**. Le tout est envoyé dans une trame MAC d'adresse source **00-00-5E-00-01-xx** où "xx" représente le numéro d'identification du routeur virtuel (identique au champ "n° id routeur virtuel" du paquet VRRP).

De son coté, HSRP fonctionne au-dessus d'UDP avec l'adresse multicast 224.0.0.2 et un TTL fixé à 1. L'adresse MAC virtuelle utilisée est 00-00-0C-07-AC-xx.

Un même routeur peut participer à plusieurs groupes VRRP et plusieurs groupes VRRP peuvent cohabiter sur un LAN. Si, via le système des priorités, on s'arrange pour que chaque routeur d'un LAN soit maître pour un groupe, et, si on répartit les passerelles par défaut des PC sur chacune des adresses virtuelles, il est alors possible de **partager la charge** de routage entre les routeurs.

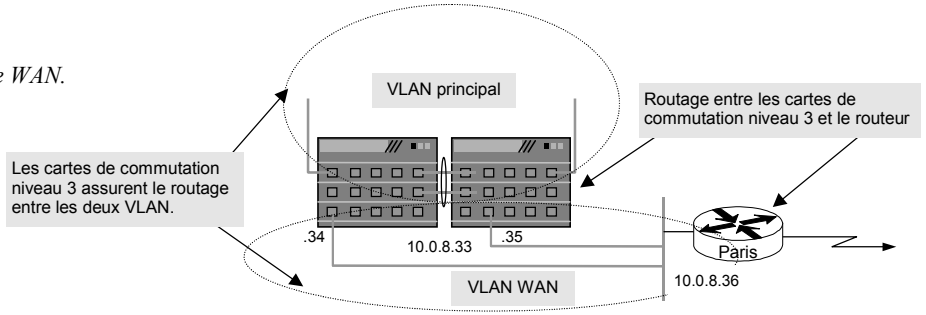
Les mécanismes classiques d'icmp-redirect et de proxy ARP sont toujours opérants.

La rencontre du LAN et du WAN

Les routeurs qui interconnectent notre site aux autres peuvent directement être raccordés au réseau fédérateur. Nous préférons cependant créer un VLAN spécifique, afin de marquer la frontière entre les deux mondes, ce qui procure certains avantages :

- tout changement d'architecture ou de configuration du LAN n'affecte pas le WAN, et inversement ;
- la gestion du WAN peut être centralisée à partir d'un autre site, ce dernier gardant son autonomie sur le LAN ;
- idem lorsque les routeurs sont gérés par un ou plusieurs opérateurs.

Figure 11-4.
La frontière entre le LAN et le WAN.



Comme pour le VLAN principal, il faut affecter un subnet IP au VLAN WAN, puis des adresses IP aux cartes de commutation et à HSRP. Un subnet de 30 adresses, pris dans notre plan d'adressage, sera largement suffisant :

```
interface Port-channel 1.2
encapsulation isl 5
ip address 10.0.9.34 255.255.255.224

standby 2 priority 110
standby 2 preempt
standby 2 ip 10.0.9.33
```

VLAN n° 5 = VLAN WAN

Groupe HSRP n° 2 affecté à ce VLAN

Il suffit ensuite de configurer le routage entre nos commutateurs LAN et le routeur WAN. La manière la plus simple de le faire est de définir des routes statiques, soit une par défaut, soit explicitement pour chaque site distant connu :

```
# route par défaut
ip route 0.0.0.0 0.0.0.0 10.0.9.36
# OU routes statiques explicites
ip route 10.4.0.0 255.252.0.0  10.0.9.36
ip route 10.8.0.0 255.252.0.0  10.0.9.36
```

Tout ce qui n'est pas connu est envoyé au routeur

Vers Toulouse et vers Strasbourg via le routeur

Inversement, nous indiquons au routeur comment joindre le VLAN principal du site parisien :

```
ip route 10.0.0.0 255.252.0.0 10.0.9.33

int e0
ip address 10.0.9.36 255.255.255.224
```

Sur Toulouse, nous avons deux routeurs WAN qui peuvent être redondants pour des liaisons Frame-Relay. Il est alors possible de configurer HSRP à la fois sur les routeurs WAN et sur les routeurs LAN (les cartes de commutation de niveau 3).

Le routage sur le WAN

Une fois arrivés sur le WAN, les paquets IP se trouvent face à de multiples routes allant vers la même destination.

Il est envisageable de programmer tous les routeurs avec des routes statiques, comme nous l'avons fait précédemment, mais cette tâche peut s'avérer complexe et fastidieuse, surtout s'il faut envisager des routes de secours.

Sur le WAN, le plus simple est d'utiliser un protocole de **routage dynamique**. Nous avons alors le choix entre RIP et OSPF (voir encadré). Ce dernier est cependant le plus performant et le plus répandu, même s'il est un peu plus complexe à programmer. Nous choisissons donc OSPF.

Configuration du routage

La première tâche est d'activer le routage OSPF. Sur nos routeurs Cisco, il faut attribuer un numéro de processus, car plusieurs instances d'OSPF peuvent fonctionner simultanément :

```
router ospf 1
```

Avec OSPF, la première tâche est de définir l'aire 0, appelée *backbone area*.

QU'EST-CE QU'UN PROTOCOLE DE ROUTAGE ?

Le routage est l'action de commuter les paquets d'un réseau IP à l'autre en fonction de leur adresse IP de destination. Le routeur se base sur des routes **statiques** (configurées par l'administrateur) et **dynamiques** (apprises par des protocoles de routage).

Le routeur maintient ainsi une base de données des coûts des routes associées, ce qui permet de calculer le meilleur chemin.

Afin de réduire le trafic réseau généré par les protocoles de routage, de réduire la taille des bases de données et de déléguer l'administration, les réseaux IP sont découpés en domaines appelés **systèmes autonomes** (AS, *Autonomous System*).

Les protocoles spécialisés dans le routage au sein d'un AS sont de type **IGP** (*Interior Gateway Protocol*). Les plus courants sont **RIP** (*Routing Information Protocol*) et **OSPF** (*Open Shortest Path First*). Les protocoles spécialisés dans le routage inter AS sont de type **EGP** (*Exterior Gateway Protocol*). Le plus répandu est **BGP** (*Border Gateway protocol*).

Au sein d'un AS, tous les routeurs disposent de la même base de données décrivant la topologie de l'AS.
Les IGP utilisent deux types d'algorithmes pour calculer les routes : celui à **vecteur de distance** (Bellman-Ford) utilisé par RIP, et celui de l'**arbre du plus court chemin** (Dijkstra), plus performant, qui est utilisé par OSPF.

Même si de multiples configurations sont possibles avec OSPF, il est cependant conseillé de respecter les règles suivantes :

- L'aire 0 doit couvrir toutes les interfaces WAN des routeurs (c'est-à-dire les interfaces série, Frame-Relay, ATM, LS, RNIS, etc.).
- Une aire doit être définie par site ou par groupe de sites fédérés autour d'un campus. L'intérêt est de pouvoir contrôler la diffusion des routes, par exemple, d'empêcher qu'un subnet parisien puisse être vu des autres sites.

Étant donné que notre plan d'adressage défini au chapitre 7 prévoit l'affectation d'un subnet complet à l'ensemble des liaisons WAN, une seule commande sur chaque routeur est nécessaire pour affecter l'aire 0 :

```
network 172.16.0.0 0.0.255.255 area 0.0.0.0
```

Tous les réseaux WAN sont dans l'aire backbone

L'aire OSPF est un numéro sur 32 bits qui peut être noté à la manière d'une adresse IP. La notation du masque associé au subnet à annoncer utilise, quant à elle, une convention inverse à celle utilisée pour les adresses IP (les bits à " 0 " indiquent la partie réseau).

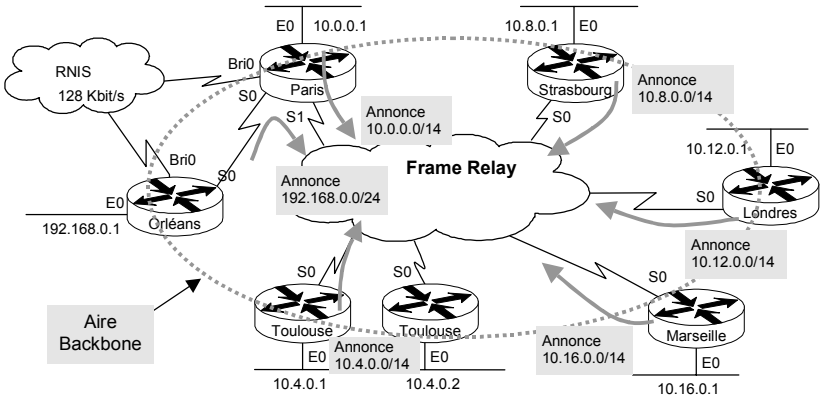


Figure 11-5.
Configuration OSPF.

Côté LAN, il n'y a pas de contrainte particulière à l'affectation d'une aire. Nous choisissons d'en affecter une par site (ou par campus) si cela se révélait nécessaire.

Aire OSPF	Site
0.0.0.1	Région parisienne
0.0.0.2	Région toulousaine
0.0.0.3	Strasbourg
Etc.	…

C'est justement le cas à Toulouse, car nous avons deux routeurs, connectés à l'aire 0 d'un côté et au même réseau local de l'autre. Afin que ces deux routeurs puissent échanger leurs tables de routage et se secourir mutuellement, il faut positionner leur interface locale dans une aire.

S'il n'y a que deux routeurs, le plus simple est de tout mettre dans l'aire 0. Si le réseau de Toulouse grandit au point d'intégrer plusieurs routeurs (ou cartes de commutation de niveau 3), on peut envisager de créer une aire sur ce site, afin de réduire le trafic sur le WAN et de mieux contrôler la diffusion des routes :

```
network 10.4.0.0 0.3.255.255 area 0.0.0.2
```

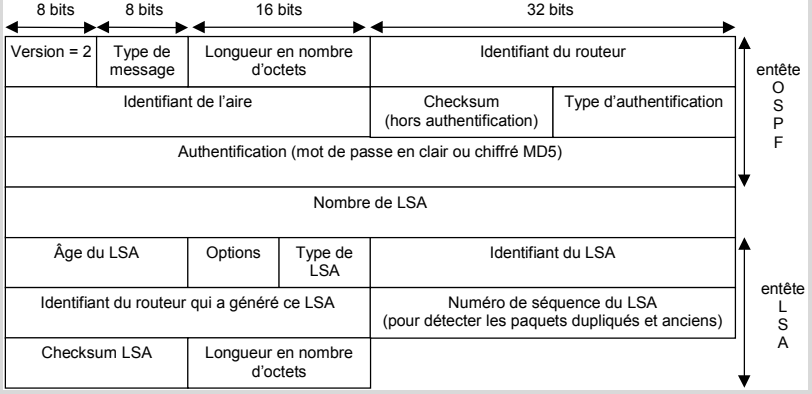
Redondance en cas de panne

En reprenant le réseau Frame-Relay que nous avons construit au chapitre 10, nous pouvons apercevoir que, en cas de panne du routeur de Strasbourg, l'aire backbone serait coupée en deux, empêchant toute diffusion des routes. Même s'il y a continuité du réseau local, l'aire backbone est séparée par l'aire 2.

LE POINT SUR OSPF (RFC 2328)

Le protocole OSPF (*Open Shortest Path First*) découpe l'AS (*Autonomous System*) en **aires**. Toutes les aires doivent être adjacentes à l'aire 0 (***backbone area***) qui doit être contiguë. Si elle ne l'est pas, un **lien virtuel** doit être configuré pour assurer sa continuité logique. Les paquets routés entre aires doivent tous passer par la backbone area *via* les **routeurs de bordure**.

Les routeurs diffusent régulièrement des messages d'annonce **LSA** (*Link State Advertisement*) pour indiquer quels réseaux leur sont directement attachés. Les LSA sont diffusés à tous les routeurs de l'aire ; ils permettent à chacun d'entre eux de disposer de la même base de données d'**état des liens** et de calculer l'**arbre du plus court chemin** dont il est la racine. Un routeur gère autant de bases de données et calcule autant d'arbres qu'il y a d'aires auxquelles il est connecté.



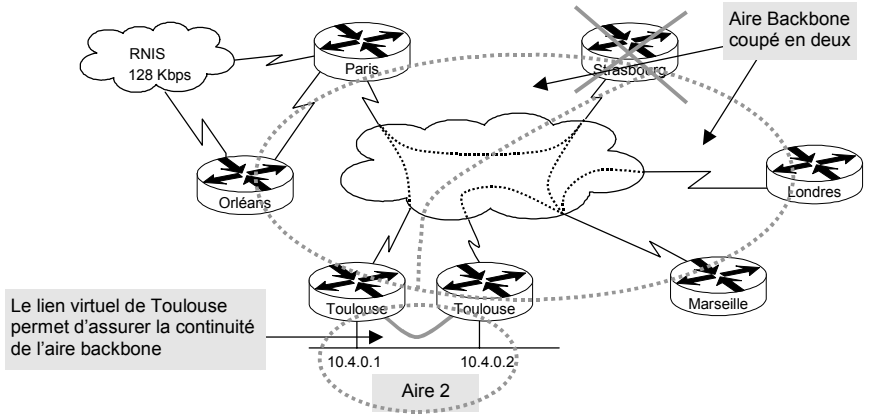
Les routeurs diffusent régulièrement des messages **Hello** afin d'annoncer leur présence à leurs voisins sur les réseaux multipoints supportantle broadcast (par exemple Ethernet). Celui dont la priorité est la plus grande est élu **routeur désigné** ; il a la charge d'inclure ce réseau dans ses LSA. Sur les réseaux Ethernet, les messages sont envoyés dans des paquets multicast 224.0.0.5.

Dans la *backbone area*, les routeurs de bordure s'échangent les bases de données des aires auxquelles ils sont rattachés. Ils calculent les meilleures routes qui sont ensuite diffusées aux routeurs intra-aire. Les routeurs intra-aire calculent la meilleure route pour sortir de l'aire (*via* un routeur de bordure). Le **routeur frontière** (qui peut être situé n'importe où dans l'AS) assure le même rôle pour les routes permettant de sortir de l'AS.
Quatre types de LSA sont échangés :

- *router-LSA* : émis par tous les routeurs d'une aire pour décrire l'état et indiquer le coût de leur interface.
- *network-LSA* : émis par les routeurs désignés pour annoncer les réseaux de type broadcast (Ethernet, par exemple) ;
- *summary-LSA* : émis par les routeurs de bordure ;
- *AS-external-LSA* :émis par les routeurs de frontière.

Aucun AS-external-LSA n'est envoyé dans les aires configurées en **stub area**. À la place, le routeur de bordure diffuse une route par défaut.

Figure 11-6.
Liens virtuels OSPF.



Il est à noter que ce cas de figure n'existerait pas si Toulouse ne disposait que d'un seul routeur ou si les interfaces Ethernet des deux routeurs étaient situées dans l'aire 0.

La solution à ce problème passe par la création d'un lien virtuel entre les deux routeurs de Toulouse :

```
# Routeur 1
area 2 virtual-link 10.4.0.2
# Routeur 2
area 2 virtual-link 10.4.0.1
```

Aire de transit commune aux deux routeurs

Adresse du routeur à l'autre bout du lien virtuel

Ce lien permet d'assurer la continuité de l'aire backbone *via* l'aire de transit de Toulouse.

Ajustement des paramètres

Diffuser les routes statiques

Certains sites peuvent comporter des routeurs configurés uniquement avec des routes statiques. Si ces routes doivent être connues des autres sites, il est alors impératif de les diffuser au processus OSPF, de manière que ce dernier les diffuse dynamiquement à ses voisins :

```
router ospf 1
redistribute static
```

Modifier le coût des routes

Pour calculer le coût des routes, et donc choisir la meilleure, OSPF se base sur la bande passante du lien. Sur nos routeurs, il est nécessaire de l'indiquer manuellement, par exemple 512 Kbit/s sur les routeurs de Toulouse :

```
int s0
bandwidth 512
```

Par défaut, le coût associé à l'interface est de 100 000 divisé par le débit exprimé en Kbit/s, ce qui donne, par exemple, un coût de 1 562 pour un débit de 64 Kbit/s. Il est néanmoins possible de le modifier, comme suit :

```
int s0
ip ospf cost 300
```

Valeur de 1 à 65 535

Limiter la diffusion des routes

Dans certains cas, il peut être intéressant de limiter la diffusion de certaines routes afin qu'elles ne soient pas connues d'autres site, par exemple pour des questions de confidentialité ou pour forcer le chemin à emprunter :

```
router ospf 1
distribute-list 11 out

access-list 11 deny 192.168.0.0 0  0.0.0.255
access-list 11 permit any
```

Le routeur parisien ne diffuse pas le réseau du site d'Orléans.

De la même manière, un routeur peut ne pas accepter une route si, par exemple, le site d'Orléans doit être caché uniquement à celui de Londres :

```
router ospf 1
distribute-list 11 in

access-list 11 deny 192.168.0.0 0  0.0.0.255
access-list 11 permit any
```

Le routeur de Londres filtre la route du site d'Orléans.

Modifier la fréquence des échanges

Les routeurs OSPF voisins s'échangent des paquets Hello selon une périodicité qu'il est possible de modifier :

```
int s0
ip ospf hello-interval 10
ip ospf dead-interval 40
```

Envoi un paquet Hello à ses voisins toutes les 10 secondes.

Le routeur voisin est déclaré absent au bout de 40 secondes (par défaut, 4 x le hello-interval).

Forcer l'élection du routeur désigné

Lorsque, comme cela est le cas à Toulouse, il existe deux routeurs sur le même réseau Ethernet, seul le routeur désigné va diffuser le subnet IP de l'aire n° 2. Est élu " désigné " le routeur dont la priorité est la plus haute ; en cas de niveau de priorité identique, c'est celui dont l'adresse IP est la plus haute :

```
ip ospf priority 1
```

Valeur par défaut

Les performances d'OSPF

Le RFC 1245 fournit quelques statistiques relevées sur les routeurs de l'Internet :

- Chaque entrée de la base d'états de liens est mise à jour toutes les 30 minutes en moyenne.
- Selon les cas, l'arbre du plus court chemin est recalculé toutes les 13 à 50 minutes.
- En moyenne, un paquet d'annonce contient trois LSA.
- Pour 2 000 entrées dans une base de données OSPF, la bande passante consommée par l'émission des LSA représente moins de 0,5 Kbit/s.

Type d'annonce	Taille moyenne dans les paquets	Mémoire routeur
External LSA	36 octets	64 octets
Router et Network LSA	108 octets	192 octets
Summary LSA	36 octets	64 octets
En-tête OSPF	24 octets	--
En-tête IP	20 octets	

Le temps CPU pour calculer l'arbre du plus court chemin (algorithme de Dijsktra) est de l'ordre de n*log(n) pour *N* routes et 200 routeurs, soit environ 15 millisecondes pour un processeur de 10 Mips. En découpant un système autonome en aires, la charge CPU est réduite, car il y a moins de routeurs à prendre en compte, le calcul SPF étant réalisé au sein d'une aire.